

Special Paper

방송공학회논문지 제30권 제7호, 2025년 12월 (JBE Vol. 30, No. 7, December 2025)

<https://doi.org/10.5909/JBE.2025.30.7.1155>

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

A VQ-HEVC Hybrid Compression Method for Learned 3DGS (3D Gaussian Splatting) Models

Dong-Ha Kim^{a)}, Byung-Yoon Choi^{a)}, Kwan-Jung Oh^{b)}, Gwangsoon Lee^{b)}, and Jae-Gon Kim^{a)†}

Abstract

Recently, 3DGS (3D Gaussian Splatting) has attracted significant attention as an efficient technology for representing immersive 3D scenes. In response, MPEG is currently conducting an exploration phase toward standardizing the compression of learned 3DGS models, referred to as Gaussian Splatting Coding (GSC). As part of this activity, the GSC group is investigating an approach that compresses the attribute maps of learned 3DGS models using High Efficiency Video Coding (HEVC). Although Parallel Assignment Linear Sorting (PLAS) is employed to enhance the spatial correlation of attribute maps, the compression efficiency of 2D video codecs remains limited, particularly for AC spherical harmonics (SH) coefficients, due to their inherently low spatial correlation. To overcome this limitation, this paper proposes a VQ-HEVC hybrid compression method for 3DGS models, where two-stage Vector Quantization (VQ) is applied to the AC SH coefficients, while the DC SH coefficients and other attributes are compressed using HEVC. The proposed two-stage VQ scheme combines Residual Quantization (RQ) and Product Quantization (PQ) to improve quantization efficiency, and introduces a zero-mask-based residual VQ method to enhance the entropy coding efficiency of VQ indices. Experimental results demonstrate that the proposed VQ-HEVC hybrid method achieves BD-rate reductions of 21.28% and 23.15% compared to the GSC anchor on the Bartender and Cinema datasets, respectively, under the GSC Common Test Conditions (CTC). These results show the potential of the proposed VQ-HEVC hybrid approach as a promising candidate for the MPEG GSC compression standard.

Keyword : 3D Gaussian Splatting (3DGS), Gaussian Splatting Coding (GSC), HEVC, Vector Quantization (VQ), Hybrid compression

a) Department of Smart Air Mobility, Korea Aerospace University

b) Electronics and Telecommunications Research Institute (ETRI)

† Corresponding Author : Jae-Gon Kim

E-mail: jgkim@kau.ac.kr

Tel: +82-2-300-0414

ORCID: <https://orcid.org/0000-0003-3686-4786>

※ This work was supported equally by the IITP grant (No. 2018-0-00207, Immersive Media Research Laboratory) funded by the Ministry of Science and ICT, and the KOCCA grant (No. RS-2024-00395886, Development of AI-based image expansion and service technology for high-resolution (8K/16K) service of performance contents) funded by the Ministry of Culture, Sports and Tourism.

• Manuscript October 20, 2025; Revised October 23, 2025; Accepted October 24, 2025.

1. Introduction

3D Gaussian Splatting (3DGS) has attracted significant attention as an efficient method for representing immersive 3D scenes^[1]. The approach first extracts an initial point cloud using Structure-from-Motion (SfM) and then derives attributes of each 3D Gaussian through learning, including volume, opacity, rotation, and color represented by spherical harmonics (SH) coefficients. During training, all attrib-

utes—including point positions—are jointly optimized to minimize the difference between rendered and multi-view images. This method effectively represents immersive 6DoF video content^[2] and enables fast rendering through an explicit 3D representation.

A trained 3DGS model consists of optimized 3D Gaussians, each represented by 59 attributes stored as 32-bit floating-point numbers. Due to this high-dimensional structure, trained 3DGS models require substantial storage capacity, making them inefficient for storage and transmission. Various compaction or compression methods have been proposed, but most fail to preserve the original structure. They typically reduce attributes, replace components with MLP-based representations^{[3]-[5]}, or apply pruning and retraining^{[6]-[8]}. Consequently, models are often re-trained from scratch or redesigned with alternative architectures, making unified compression frameworks difficult to establish.

Meanwhile, to achieve efficient compression for 6DoF video, MPEG initially explored neural scene representation techniques based on Neural Radiance Fields (NeRF)^[9]. During this period, several hybrid NeRF-based approaches were also investigated, where both explicit and implicit representations were jointly used to achieve neural volumetric video compression with various codecs^[10]. However, despite these advances, NeRF-based methods showed limitations in rendering speed and reconstruction quality for practical 6DoF applications, leading MPEG to shift its focus toward 3DGS, which provides real-time rendering and higher visual fidelity. As a result, MPEG established the Gaussian Splatting Coding (GSC) Ad-Hoc Group^[11] to explore the standardization potential of 3DGS-based compression. Within this activity, MPEG's Working Group 7 (3D Graphics Coding) and Working Group 4 (Video Coding) jointly initiated the Joint Exploration Experiment 6 (JEE 6)^[12] to investigate candidate technologies for both static and dynamic 3DGS compression.

The GSC group currently explores two tracks: I-3DGS focuses on compressing models following the INRIA format^[13], while A-3DGS targets alternative model formats^{[14]-[16]}. Among the investigated approaches, the video-based compression framework applies Parallel Assignment Linear Sorting (PLAS)^[17] to convert 3DGS attributes into attribute-wise 2D grids, which are then compressed using 2D video codecs such as High Efficiency Video Coding (HEVC)^[18]. Since this approach compresses trained 3DGS models without altering their structure, it preserves both rendering consistency and attribute integrity.

PLAS enhances spatial correlation within each 2D grid by sorting 3D Gaussians according to spatial proximity, under the assumption that neighboring Gaussians share similar attribute values. However, because each Gaussian has 59 attributes, position-based sorting does not ensure spatial correlation for all attributes. AC (Higher-order) SH coefficients, in particular, retain noisy spatial distributions even after PLAS, limiting the coding efficiency of video-based compression.

To address this limitation, we propose a hybrid 3DGS compression method that applies Vector Quantization (VQ) to AC SH coefficients while using HEVC to compress the remaining attributes. The proposed two-stage VQ combines residual and product quantization to reduce quantization errors and employs a zero-masked residual VQ scheme to improve entropy coding efficiency by concentrating the index distribution at zero.

Experimental results show that the proposed method achieves Bjøntegaard Delta-rate (BD-rate) reductions of 21.28% and 23.15% on the Bartender and Cinema datasets^[19], respectively, compared with the GSC anchor^[20], which applies PLAS and HEVC to all attributes. These results demonstrate that the proposed method provides an efficient and practical framework for compressing trained 3DGS models.

II. Introduction to 3DGS compression in GSC

Understanding the structure of a trained 3DGS model is essential for designing efficient compression methods. This section outlines key attributes of trained 3DGS models and summarizes how the GSC group applies PLAS with video codecs for compression.

1. Attributes of trained 3DGS

A trained 3DGS model represents a 3D scene as a collection of structured 3D Gaussians with spatial and appearance attributes. Each 3D Gaussian contains learnable attributes jointly optimized to minimize the photometric error between rendered and ground-truth multi-view images. During training and rendering, each Gaussian is projected onto the 2D image plane based on its spatial attributes, while appearance attributes determine the resulting color.

The spatial attributes consist of 3D position (x, y, z) , anisotropic scale (s_x, s_y, s_z) , and orientation represented using a unit quaternion (q_i, q_j, q_k, q_r) , which define the Gaussian's location, volume, and rotation. For appearance, each Gaussian uses SH coefficients: the DC (zeroth-order) SH coefficients f_{dc} comprise three RGB values, while AC SH coefficients f_{ac} capture view-dependent color details with 15 coefficients per color channel (45 in total). An opacity scalar α controls transparency. In total, each Gaussian has 59 attributes optimized to reproduce the scene's appearance from arbitrary viewpoints.

2. PLAS for 3DGS Compression

Each Gaussian in a 3DGS model is defined by a

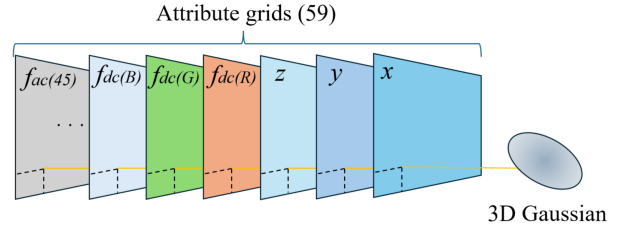


Fig. 1. 2D grid mapping of each attribute of a 3D Gaussian

high-dimensional attribute vector composed of spatial and appearance properties. Rendering depends only on visible positions, independent of Gaussian ordering. Thus, reordering Gaussians does not affect rendering results. Leveraging this property, PLAS enhances spatial correlation within each 2D attribute grid by reordering and mapping attribute vectors onto 2D grids, improving compression efficiency.

Each Gaussian has 59 attributes, each mapped to a separate grid, resulting in 59 grids aligned such that attributes of the same Gaussian share identical coordinates. The grid resolution is determined by fitting all Gaussians into a square layout, with low-opacity Gaussians pruned if capacity is exceeded.

PLAS, based on Fast Linear Assignment Sorting (FLAS)^[21], enhances spatial correlation by mapping millions of Gaussian attributes onto 2D grids. PLAS sorts 3D Gaussians on 2D grids using 6D vectors (position and DC SH coefficients), iteratively minimizing distance to enhance spatial correlation. Once the permutation is determined, it is uniformly applied to all attribute grids. For example, if the attribute at (3,71) moves to (21,37) in one grid, the same permutation applies to all grids.

Figure 2 shows attribute grids after PLAS-based reordering. Although sorting relies only on 3D positions and DC SH coefficients, most attributes show improved spatial correlation, confirming that PLAS enhances compressibility.

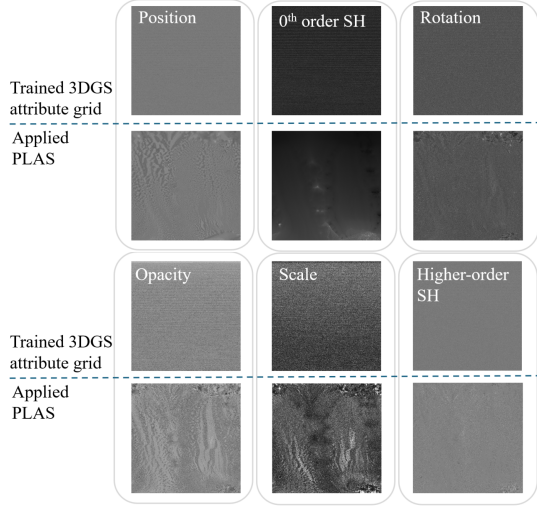


Fig. 2. Comparison of 2D attribute grids before and after PLAS reordering

3. Trained 3DGS compression with HEVC

The GSC group combines PLAS with a 2D video codec as an anchor for video-based 3DGS compression. PLAS first enhances spatial correlation of attribute 2D grids, which are then compressed using HEVC. This anchor serves as the baseline for evaluating other methods.

Figure 3 illustrates the pipeline of the video-based GSC anchor^[20]. PLAS reorders Gaussians based on position

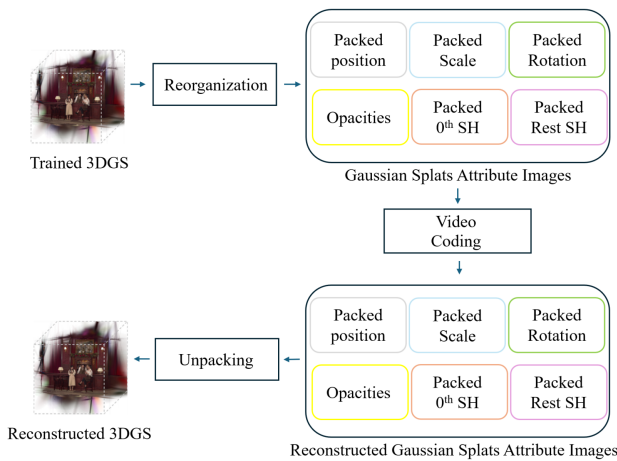


Fig. 3. Pipeline for video-based GSC anchor generation

and DC SH coefficients. The resulting permutation is applied to all 59 attributes, producing 59 reordered 2D grids. These grids are packed into RGB images for HEVC compression.

Position attributes are quantized to 16-bit, while others are min-max normalized and quantized to 8-bit. Attributes are grouped in sets of three and assigned to R, G, and B channels—for example, position attributes (x, y, z) form one RGB image. Using this scheme, 57 attributes form 19 RGB images. The remaining opacity and fourth quaternion component are stored as grayscale images. Thus, all attributes are represented as 21 images (19 RGB and 2 grayscale) compressed with HEVC.

During compression, the position image is encoded losslessly to preserve geometry, while other attributes are compressed using adaptive Quantization Parameters (QPs) depending on importance: lower QPs for high-importance attributes and higher QPs for less critical ones. GSC defines attribute-specific QP sets to support multiple bitrate points. For reconstruction, the compressed bitstreams are decoded, and attribute grids are unpacked by reversing the reorganization. The restored grids are then converted back into Gaussians to reconstruct the trained 3DGS for rendering.

III. Proposed VQ-HEVC hybrid method

Figure 2 compares the 2D attribute grids before and after PLAS reordering. Using position and DC SH coefficients as sorting criteria produces grids with high spatial correlation, while other attributes such as rotation, scale, and opacity show moderate improvement. However, AC SH coefficients still exhibit noisy distributions, indicating that PLAS alone cannot ensure spatial coherence across all attributes, thus limiting compression efficiency.

To overcome this limitation, we propose a 3DGS compression method that applies VQ to AC SH coefficients,

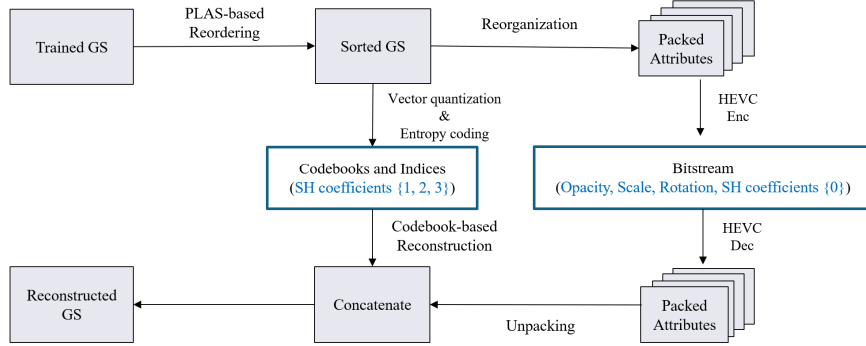


Fig. 4. The proposed VQ-HEVC hybrid 3DGS compression pipeline

which typically exhibit low spatial correlation even after PLAS. Instead of individually compressing each 2D grid, VQ clusters coefficient vectors in the attribute space based on value similarity, exploiting inter-vector correlation for efficient compression.

As shown in Figure 4, AC SH coefficients are compressed using VQ, while the remaining attributes follow the GSC video-based anchor pipeline. The transmitted data consists of the HEVC-compressed bitstreams from 2D attribute grids, along with VQ-generated indices and codebooks.

A single-stage VQ applied to the 45-dimensional AC SH

coefficients can cause large quantization errors, even with a large codebook. Therefore, we adopt a two-stage composite VQ that integrates Residual Vector Quantization (RVQ)^[22] and Product Quantization (PQ)^[23] to effectively reduce quantization errors. Furthermore, a zero-masked residual VQ is introduced to set small residuals to zero, concentrating indices near zero and improving both entropy coding efficiency and rate - distortion performance.

1. Two-stage VQ

Figure 5 illustrates the processing pipeline of the pro-

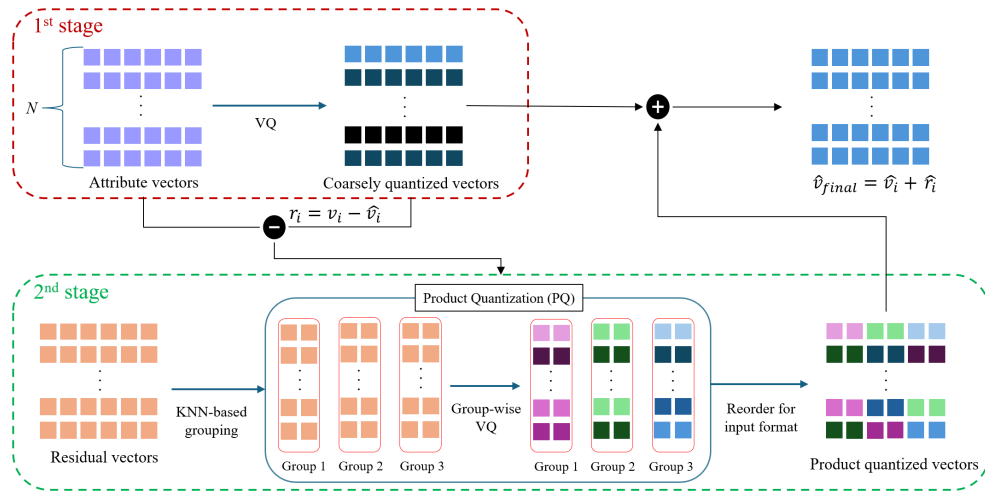


Fig. 5. Proposed two-stage VQ process, combining coarse quantization of AC SH coefficients with residual vector quantization using KNN-based grouping and PQ

posed VQ process, including PQ. Prior to quantization, the attribute vectors composed of a Gaussian's AC SH coefficients are normalized on a per-attribute basis. In the first stage of VQ, the i -th 45-dimensional normalized attribute vector v_i is quantized to the nearest code vector \hat{v}_i in a coarse codebook, which is constructed via K-Means clustering. A residual vector r_i is then computed by taking the difference between v_i and \hat{v}_i .

In the second stage, PQ is applied to the residual vectors to further reduce quantization errors. PQ divides each residual vector into multiple sub-vectors and performs vector quantization independently on each. Specifically, coefficient attributes with similar statistical properties are grouped into the same sub-vector using K-Nearest Neighbors (KNN) groups. After PQ, the quantized sub-vectors are reordered to match the original attribute order.

The final output vector \hat{v}_{final} is reconstructed by combining the coarsely quantized vector from the first stage with the refined residual vector from the second stage. For reconstruction at the decoder, the compressed bitstreams

must contain the following information:

- the coarse quantization codebook and indices from the first stage VQ,
- group-specific codebooks and indices from PQ,
- a grouping map that specifies the attribute-to-sub-vectors assignment, and
- the min - max values used for normalization

2. Zero-masked residual VQ

In PQ, the number of sub-vector groups and the codebook size for each group should be balanced to trade off quantization error against the bit rate required to transmit the codebooks and indices. Entropy coding efficiency depends on the quantization index distribution. A small codebook reduces index data size but increases quantization error and yields a more uniform index distribution, which degrades coding efficiency.

To address this, we introduce zero-masked residual VQ, which improves coding efficiency by masking small-magnitude residual vectors to zero. As shown in Figure 6, when the magnitude of a residual vector falls below a predefined

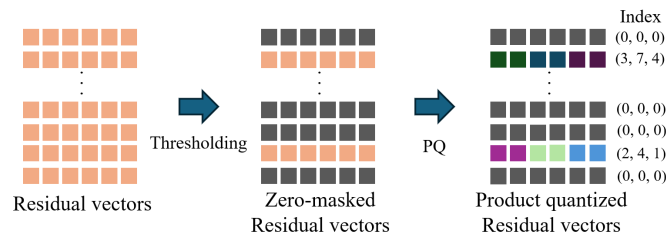


Fig. 6. Process of zero-masked residual VQ

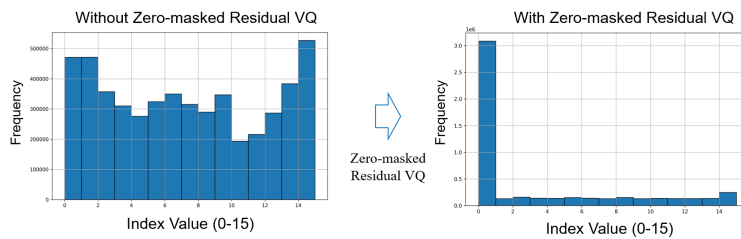


Fig. 7. Effect of zero-masked VQ on the index distribution

threshold, it is explicitly masked to zero before clustering and excluded from the codebook construction. Due to the natural characteristics of residual vectors, the index distribution is heavily biased toward zero, thereby maximizing entropy coding efficiency. Residual vectors with component magnitudes smaller than a predefined threshold are replaced with zero vectors (shown in black). The remaining non-zero residual vectors are product-quantized (PQ) into sub-vector codebook indices. Each index tuple (x, y, z) represents the quantization codebook indices assigned to the corresponding sub-vectors.

Figure 7 shows histograms comparing index distributions for residual vectors that are quantized with a codebook size of 16, with and without zero-masked VQ. The zero-masked approach produces a clear concentration of indices at zero, resulting in a notable improvement in entropy coding efficiency. The zero-masking threshold—defined as the percentage of residual vectors to be masked—controls the trade-off between bit-rate and quantization error. For example, a threshold of 0.3 masks the lowest 30% of residual vectors. Higher thresholds further reduce bit-rate at the cost of increased quantization error.

IV. Experimental results

The MPEG GSC currently uses two datasets, Bartender and Cinema^[17], in JEE 6. Each dataset consists of 20 mul-

ti-view images at Full HD resolution. In JEE 6, a trained 3DGS model is provided for each dataset, and exploration experiments are conducted to evaluate and compare compression methods using these models. To comprehensively evaluate compression performance, rate - distortion (RD) metrics are calculated, with PSNR, Structural Similarity Index Measure (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS)^[24] used as distortion measures. The model size and rendered image quality in terms of PSNR, SSIM, and LPIPS for the provided trained 3DGS models are summarized in Table 1.

We applied the proposed compression method to the provided 3DGS models and evaluated its performance in terms of compressed model size and the above quality metrics. For fair comparison with the GSC anchor, results are reported at rate points (RPs) corresponding to similar target bitrates used for the GSC anchor.

1. Anchor for GSC

The GSC anchor compresses each attribute grid of the PLAS-reordered 3DGS model using HEVC in FFmpeg^[25]. Different Quantization Parameters (QPs) are assigned to each attribute according to its importance, while position attributes are encoded losslessly to preserve geometry. Predefined QP sets for various rate points (RPs) are summarized in Table 2, with lower QPs for critical attributes and higher QPs for less sensitive ones.

Table 1. Model size and quality of trained 3DGS models used in GSC

Dataset	Size (MB)	PSNR (dB)	SSIM	LPIPS
Bartender	134	35.16	0.927	0.197
Cinema	100	34.51	0.928	0.208

Table 2. QP configurations for each attribute in the GSC anchor across different rate points

RP	Position	Opacity	Rotation	Scale	SH0	SH1	SH2	SH3
4	Lossless	4	4	4	4	4	4	4
3		4	10	10	4	16	22	28
2		10	16	16	10	22	28	34
1		16	22	22	16	28	34	40

2. Proposed two-stage VQ

As illustrated in Figure 4, the proposed method compresses AC SH coefficients using VQ, while the remaining attributes are encoded with HEVC as in the GSC anchor. The QPs listed in Table 2 are applied to the remaining attributes to obtain compression results for each RP. For the proposed zero-masked residual VQ, masking thresholds of $\{0, 0.6, 0.9, 0.98\}$ were experimentally determined for RPs 1–4. These thresholds were not optimized through a specific input-dependent procedure but were empirically selected to achieve rendering quality comparable to that of the anchor (in terms of PSNR).

As described in Section III, the proposed VQ employs a two-stage structure: (1) coarse quantization of the 45-dimensional vector using a 1024-entry codebook (10 bits), and (2) sub-vector quantization of residuals divided into 15 groups of 3 dimensions, each encoded with a 64-entry codebook (6 bits). Thus, each 3D Gaussian is represented by one 10-bit coarse index and fifteen 6-bit sub-vector indices, along with their codebooks.

3. Coding performance

Table 3 compares the compression performance of the GSC anchor and the proposed method on the Bartender and Cinema datasets, with Figure 8 showing the corresponding RD curves. The top and bottom rows represent Bartender and Cinema, respectively, while red dashed lines denote the

quality of the original trained 3DGS models as upper bounds.

For Bartender, the proposed method consistently achieves comparable or superior visual quality while significantly reducing model size across all RPs. At RP4, it maintains nearly identical PSNR and SSIM to the anchor with about 44% smaller size and minimal LPIPS degradation. As bitrate decreases (RP3–RP1), the method yields higher quality across all metrics. A similar pattern is observed for Cinema, where the proposed method achieves comparable quality at RP4 with about 38% smaller size and greater gains at lower bitrates.

These gains stem from the design of the proposed compression pipeline. RP4 uses only two-stage VQ, demonstrating that it compresses AC SH coefficients more efficiently than video codec-based methods. From RP3 downward, the addition of zero-masked residual VQ yields further model size reductions and improved entropy coding efficiency.

Since PLAS-based reordering does not guarantee spatial locality across all attributes, compressing only the AC SH coefficients with VQ is both feasible and effective. Similar to varying QPs in video codecs, different bitrates are achieved by adjusting the zero-masked ratio. Consequently, the proposed hybrid approach—VQ for AC SH coefficients and HEVC for other attributes—achieves more efficient compression across a wide bitrate range. The BD-rate analysis from Figure 8 indicates bitrate savings of 21.28% and 23.15% for the Bartender and Cinema datasets, respectively,

Table 3. Comparison of model size and rendering quality between the GSC anchor and the proposed method across RPs

RP	Bartender								Cinema							
	Anchor for GSC				Proposed method				Anchor for GSC				Proposed method			
	Size (MB)	PSNR (dB)	SSIM	LPIPS	Size (MB)	PSNR (dB)	SSIM	LPIPS	Size (MB)	PSNR (dB)	SSIM	LPIPS	Size (MB)	PSNR (dB)	SSIM	LPIPS
4	24.74	34.83	0.923	0.203	13.88	34.70	0.923	0.204	16.91	34.15	0.924	0.214	10.46	34.07	0.923	0.215
3	12.26	33.71	0.913	0.213	11.01	33.93	0.916	0.210	8.59	32.47	0.913	0.228	8.33	33.45	0.919	0.221
2	8.66	32.40	0.899	0.231	7.75	32.85	0.903	0.226	6.12	30.65	0.898	0.244	5.89	31.92	0.905	0.237
1	5.95	30.11	0.869	0.264	5.97	30.81	0.876	0.254	4.34	28.24	0.867	0.279	4.56	29.67	0.877	0.269

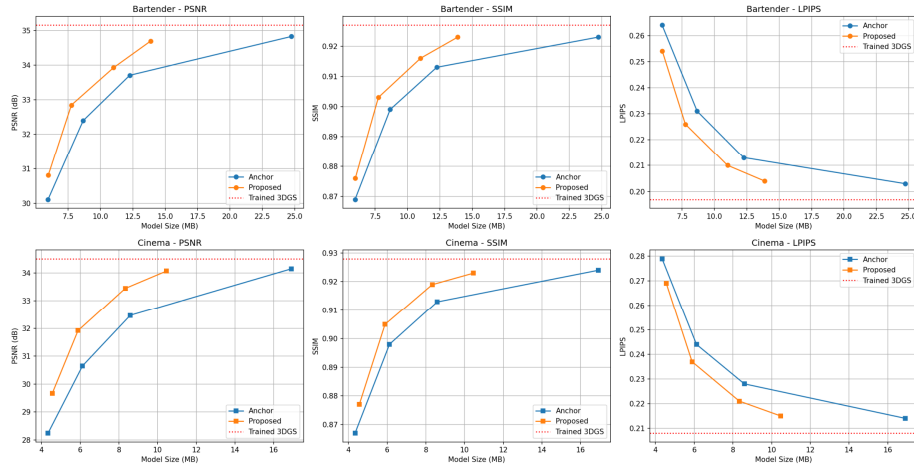


Fig. 8. Comparison of rendering qualities (PSNR, SSIM, and LPIPS) versus model sizes between the GSC anchor and the proposed method

compared with the GSC anchor.

4. Subjective quality

Figure 9 presents rendering quality comparisons under RP1 for the Cinema and Bartender datasets. To assess the perceptual impact of the proposed method, we conducted

a subjective evaluation under low-bitrate conditions where compression artifacts are more likely to appear. Since compressing AC SH coefficients mainly affects color fidelity rather than geometry, visual differences are less apparent at high bitrates. Therefore, the analysis focuses on RP1, where color degradation and visual artifacts are most noticeable.

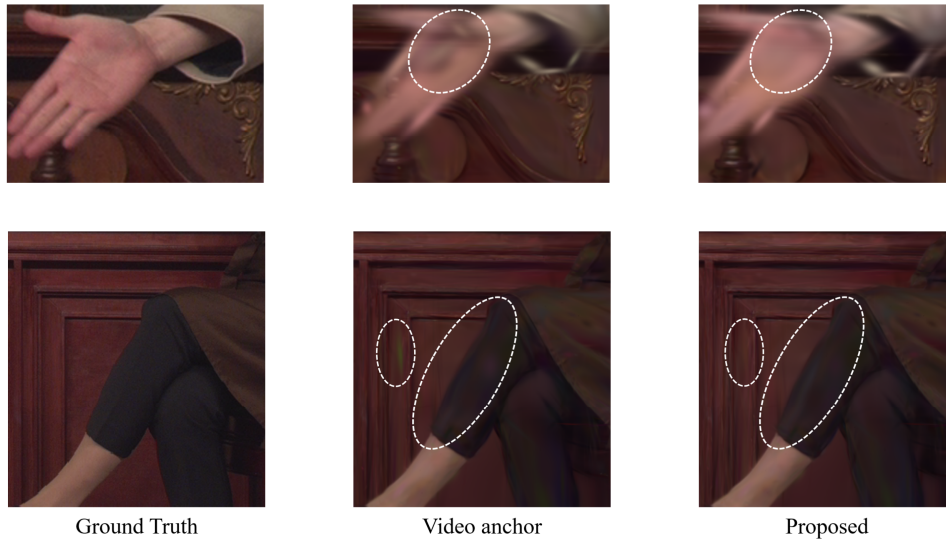


Fig. 9. Visual quality comparison of rendered novel views at RP1 for the Cinema dataset (top row) and Bartender dataset (bottom row). From left to right: Ground truth, GSC video-based anchor, and the proposed method

In the Cinema dataset, the anchor shows blurring and ghosting artifacts around objects such as the hand, while the proposed method effectively suppresses these distortions. In the Bartender dataset, the anchor output exhibits color distortion in dark background regions, where wooden textures appear unnaturally green. In contrast, the proposed method restores natural wood tones and provides smoother shading in clothing folds, yielding more stable color transitions overall.

These observations confirm that even when compression is applied to only a subset of attributes, the proposed method enhances perceptual quality at low bitrates by preserving color consistency and reducing visual distortion.

IV. Conclusion

This paper presents a hybrid framework for efficient compression of trained 3DGS models by combining VQ with HEVC. The current video-based anchor of MPEG GSC compresses each attribute of 3D Gaussian using HEVC after PLAS-based reordering to enhance spatial correlation. However, for certain attributes—such as AC SH coefficients—spatial locality is not sufficiently preserved even after reordering, making them difficult to compress efficiently with a 2D video codec.

To address this limitation, the proposed method applies a two-stage VQ to the AC SH coefficients, while the remaining attributes are compressed using HEVC as in the GSC anchor. In the first stage, coarse quantization is performed using a global codebook. In the second stage, product quantization (PQ) is applied to residual vectors by partitioning them into multiple sub-groups to improve quantization efficiency. Furthermore, a zero-masked residual VQ scheme is introduced to enhance entropy coding efficiency by masking small residuals to zero, thereby concentrating the index distribution.

Experimental results on the GSC benchmark datasets,

Bartender and Cinema, show that the proposed method achieves BD-rate reductions of 21.28% and 23.15%, respectively, compared to the GSC anchor. Beyond these objective gains, subjective evaluations confirmed that the proposed method preserves color fidelity and reduces perceptual artifacts, particularly in low-bitrate conditions. These findings demonstrated that the proposed method can significantly reduce model size while maintaining visual quality across a wide range of rate points. Therefore, the proposed method is a promising and efficient solution for compressing trained 3DGS and is expected to be a strong candidate technology for future MPEG GSC standardization.

References

- [1] B. Kerbl, G. Kopanas, T. Leimkühler, and T. Ritschel, “3D Gaussian Splatting for Real-Time Radiance Field Rendering,” *ACM Trans. Graphics*, vol. 42, no. 4, pp. 47:1 – 47:14, July 2023. doi: <https://doi.org/10.1145/3592106>
- [2] S. Rossi, I. Viola, L. Toni, and P. Cesar, “Extending 3 DoF Metrics to Model User Behaviour Similarity in 6 DoF Immersive Applications,” in *Proc. ACM Multimedia Syst. Conf. (MMSys)*, Vancouver, Canada, pp. 39 – 50, Jun. 2023. doi: <https://doi.org/10.1145/3587819.3590976>
- [3] Z. Tang, J. Zhang, H. Chen, and X. Li, “NeuralGS: Bridging Neural Fields and 3D Gaussian Splatting for Compact 3D Representations,” *arXiv preprint, arXiv:2503.23162*, Mar. 2025, [Online]. Available: <https://arxiv.org/abs/2503.23162>
- [4] K. L. Navaneet, K. P. Meibodi, S. A. Koohpayegani, and H. Pirsiavash, “CompGS: Smaller and Faster Gaussian Splatting with Vector Quantization,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2024, [Online]. Available: <https://arxiv.org/abs/2403.08550>
- [5] T. Lu, H. Lin, K. Ma, and Y. Zhang, “Scaffold-GS: Structured 3D Gaussians for View-Adaptive Rendering,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recogn. (CVPR)*, Seattle, USA, pp. 20654 – 20664, Jun. 2024. doi: <https://doi.org/10.1109/CVPR52733.2024.01952>
- [6] Z. Fan, K. Wang, K. Wen, Z. Zhu, D. Xu, and Z. Wang, “LightGaussian: Unbounded 3D Gaussian Compression with 15× Reduction and 200+ FPS,” *arXiv preprint, arXiv:2311.17245*, Nov. 2023, [Online]. Available: <https://arxiv.org/abs/2311.17245>
- [7] S. Xie, W. Zhang, C. Tang, Y. Bai, R. Lu, S. Ge, and Z. Wang, “MesonGS: Post-Training Compression of 3D Gaussians via Efficient Attribute Transformation,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Cham, Switzerland, pp. 434 – 452, Oct. 2024. doi: https://doi.org/10.1007/978-3-031-73414-4_25

- [8] M. S. Ali, M. Qamar, S.-H. Bae, and E. Tartaglione, "Trimming the Fat: Efficient Compression of 3D Gaussian Splats through Pruning," arXiv preprint, arXiv:2406.18214, Jun. 2024, [Online]. Available: <https://arxiv.org/abs/2406.18214>.
- [9] 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," in Proc. Eur. Conf. Comput. Vis. (ECCV), Glasgow, UK, pp. 405 – 421, Aug. 2020. doi: https://doi.org/10.1007/978-3-030-58452-8_24
- [10] J. Shin, J. Lee, G. Bang, and J. Kang, "Neural Volumetric Video Coding With Hierarchical Coded Representation of Dynamic Volume," IEEE Trans. Multimedia, vol. 27, pp. 4412 – 4426, Jan. 2025. doi: <https://doi.org/10.1109/TMM.2025.3544415>
- [11] MPEG WG 7 & WG 4, "Establishment of Gaussian Splatting Coding (GSC) Ad-Hoc Group and Launch of JEE 6," MPEG Meeting Documents, MPEG 150, Apr. 2025.
- [12] MPEG WG 7 & WG 4, "Description of JEE 6.3 on New Technologies," ISO/IEC JTC 1/SC 29/WG 04 Technical Report N0586, Kemer, Turkey, Nov. 2024.
- [13] V. Ye, R. Li, J. Kerr, M. Turkulainen, B. Yi, Z. Pan, O. Seiskari, J. Ye, J. Hu, M. Tancik, and A. Kanazawa, "Gsplat: An Open-Source Library for Gaussian Splatting," arXiv preprint, arXiv:2409.06765, Sep. 2024. Available: <https://arxiv.org/abs/2409.06765>
- [14] Z. Li, Z. Chen, Z. Li, and Y. Xu, "Spacetime Gaussian Feature Splatting for Real-Time Dynamic View Synthesis," arXiv preprint, arXiv:2312.16812, Dec. 2023; revised Apr. 2024, [Online]. Available: <https://arxiv.org/abs/2312.16812>.
- [15] G. Wu, T. Yi, J. Fang, L. Xie, X. Zhang, W. Wei, W. Liu, Q. Tian, and X. Wang, "4D Gaussian Splatting for Real-Time Dynamic Scene Rendering," arXiv preprint, arXiv:2310.08528, Oct. 2023; revised Jul. 2024, [Online]. Available: <https://arxiv.org/abs/2310.08528>.
- [16] H. Li, S. Li, X. Gao, A. Batuer, L. Yu, and Y. Liao, "GIFStream: 4D Gaussian-Based Immersive Video with Feature Stream," arXiv preprint, arXiv:2505.07539, May 2025, [Online]. Available: <https://arxiv.org/abs/2505.07539>.
- [17] W. Morgenstern, F. Barthel, A. Hilsman, and P. Eisert, "Compact 3D Scene Representation via Self-Organizing Gaussian Grids," in Proc. Eur. Conf. Comput. Vis. (ECCV), Cham, Switzerland, pp. 18 – 34, 2024. doi: https://doi.org/10.1007/978-3-031-73013-9_2
- [18] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1649 – 1668, Dec. 2012. doi: <https://doi.org/10.1109/TCSVT.2012.2221191>
- [19] MPEG WG 7 & WG 4, "Common Test Conditions on Gaussian Splat Coding," ISO/IEC JTC 1/SC 29/WG 04 Technical Report N0655, Online Meeting, Apr. 2025.
- [20] MPEG WG 7 & WG 4, "JEE 6.7 on Coordinating the 1F-Vid Track," ISO/IEC JTC 1/SC 29/WG 7 Technical Report N1278, Daejeon, Korea, Jul. 2025.
- [21] K. U. Barthel, N. Hezel, K. Jung, and K. Schall, "Improved Evaluation and Generation of Grid Layouts using Distance Preservation Quality and Linear Assignment Sorting," Comput. Graph. Forum, vol. 42, no. 2, pp. 123 – 136, May. 2023. doi: <https://doi.org/10.1111/cgf.14718>
- [22] R. A. Gray and Y. Linde, "Vector Quantization and Predictive Quantizers for Gauss-Markov Sources," IEEE Trans. Comm., vol. 30, no. 2, pp. 381 – 389, Feb. 1982.
- [23] H. Jégou, M. Douze, and C. Schmid, "Product Quantization for Nearest Neighbor Search," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 1, pp. 117 – 128, Jan. 2011. doi: <https://doi.org/10.1109/TPAMI.2010.57>
- [24] Z. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recogni. (CVPR), Salt Lake City, USA, pp. 586 – 595, Jun. 2018. doi: <https://doi.org/10.1109/CVPR.2018.00068>
- [25] MulticoreWare Inc., "x265: Open Source HEVC Encoder," x265 Project, 2024, [Online]. Available: <https://www.videolan.org/developers/x265.html>

Introduction Authors

Dong-Ha Kim



- 2021. 8 : B.S., School of Electronics and Information Engineering, Korea Aerospace University
- 2023. 8 : M.S., Dept. Electronics and Information Engineering, Korea Aerospace University
- 2023. 9 ~ Present : Pursuing Ph.D., Dept. Smart Air Mobility, Korea Aerospace University
- ORCID : <https://orcid.org/0009-0007-9918-571X>
- Research interests : AI-based media signal processing, Immersive video, 3D Gaussian splat

Introduction Authors



Byung-Yoon Choi

- 2023. 8 : B.S., School of Electronics and Information Engineering, Korea Aerospace University
- 2025. 3 ~ Present : Pursuing M.S., Dept. Smart Air Mobility, Korea Aerospace University
- ORCID : <https://orcid.org/0009-0001-5686-2051>
- Research interests : 3D Gaussian splat, Neural network compression



Kwan-Jung Oh

- 2002 : B.S., Dept. Electronic Computer Engineering, Chonnam University
- 2005 : M.S., Dept. Information and Communications Engineering, GIST
- 2010 : Ph.D., Dept. Information and Communications Engineering, GIST
- 2010 ~ 2013 : Samsung Advanced Institute of Technology (SAIT)
- 2013 : Principle Researcher, Electronics and Telecommunications Research Institute (ETRI)
- ORCID : <https://orcid.org/0000-0001-7676-1883>
- Research interests : Immersive video processing and coding, holography, 3D Gaussian splat coding



Gwangsoon Lee

- 1993 : B.S., Dept. Electronic Engineering, Kyungpook National University
- 1995 : M.S., Dept. Electronic Engineering, Kyungpook National University
- 2004 : Ph.D., Dept. Electronic Engineering, Kyungpook National University
- 2001 ~ Present : Principle Researcher/Team Leader, Electronics and Telecommunications Research Institute (ETRI)
- ORCID : <https://orcid.org/0000-0001-6981-2099>
- Research interests : Immersive video processing and coding, Learning-based 3D representation and coding



Jae-Gon Kim

- 1990. 2 : B.S., Dept. Electronic Engineering, Kyungpook National University
- 1992. 2 : M.S., Dept. Electrical and Electronic Engineering, KAIST
- 2005. 2 : Ph.D., Dept. Electrical Engineering and Computer Science, KAIST
- 1992. 3 ~ 2007. 2 : Senior Researcher / Team Leader, Broadcasting Media Research Group, ETRI
- 2001. 9 ~ 2002. 11 : Staff Associate, Dept. Electrical Engineering, Columbia University, NY
- 2014. 12 ~ 2016. 1 : Visiting Scholar, Video Signal Processing Lab., UC San Diego
- 2007. 9 ~ Present : Professor, College of AI Convergence, Korea Aerospace University
- ORCID : <https://orcid.org/0000-0003-3686-4786>
- Research interests : Video compression, Video signal processing, Immersive video