MARRYING CAPACITY AND PRECISION FOR EFFICIENT 2D GAUSSIAN IMAGE REPRESENTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

2D Gaussian representations provide an explicit and efficient way to represent images, but their rate-distortion performance under limited budgets depends critically on how Gaussian capacity and parameter precision are allocated. Existing methods typically optimize these two resources separately, overlooking content-dependent capacity demands and heterogeneous quantization sensitivity. We present CP-GS, a capacity-precision coupled Gaussian splatting framework for budget-aware 2D Gaussian image representation that jointly optimizes capacity allocation and attribute precision under a shared rate constraint. CP-GS first constructs a content-adaptive Gaussian representation by dynamically reallocating a fixed Gaussian budget from redundant regions to areas with higher representation demand. It then introduces hierarchical distortion-aware representation quantization, using efficient rendering-based calibration to allocate precision across Gaussian attributes and further refine it within heterogeneous attribute groups. Finally, it jointly optimizes Gaussian capacity and attribute precision by constructing candidate representations with different Gaussian counts and precision configurations and selecting operating points according to decoded distortion and actual serialized size. Experiments on DIV2K and Kodak show that CP-GS achieves improved rate-distortion performance over representative 2D Gaussian and implicit neural representations across a range of budgets, while maintaining a fixed overall representation budget and low optimization cost. These results demonstrate the benefit of jointly optimizing spatial capacity and parameter precision for compact Gaussian image representation.

1 INTRODUCTION

Efficient image representation aims to encode visual content compactly while preserving the information necessary for faithful reconstruction and downstream processing (Zhang et al., 2025a; Kao et al., 2025). Such representations serve as a fundamental building block for a wide range of visual tasks, including image reconstruction (Tian et al., 2025), super-resolution (Chen et al., 2021), restoration (Jiang et al., 2025), editing (Ye et al., 2026), and generation (Xiao et al., 2025). Building on the success of Gaussian primitives as an explicit representation for 3D scenes, 2D Gaussian splatting has recently emerged as a promising paradigm for image representation. It models visual content with a finite set of learnable anisotropic Gaussian primitives rather than dense pixels or implicit neural fields (Hu et al., 2025). Its explicit and differentiable formulation enables compact parameterization, efficient rendering, and flexible primitive-level manipulation, making it particularly attractive for efficient image representation. However, representation efficiency critically depends on how the model allocates Gaussian capacity and parameter precision under a constraint budget.

Existing approaches improve the efficiency of 2D Gaussian image representation primarily along two separate dimensions: representation capacity and parameter precision. As illustrated in Fig. 1(a), representation-oriented methods (Zhang et al., 2024; Li et al., 2026), such as Image-GS (Zhang et al., 2025d), focus on constructing and optimizing Gaussian primitives to match image content and reconstruction demands better. In parallel, Gaussian compression methods (Liang et al., 2025; Chen et al., 2026c) reduce storage costs through uniform or manually designed quantization schemes, assigning predefined precisions to position, geometry, and color attributes. However, representation capacity and attribute precision are typically optimized separately, with the Gaussian capacity determined before attribute quantization. This decoupled treatment overlooks the fact that

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

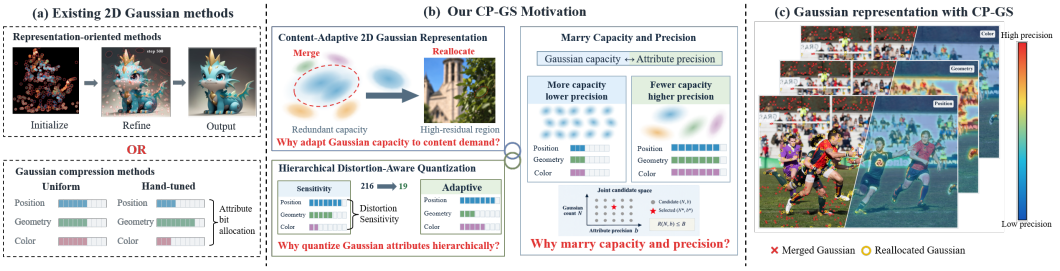

Figure 1: **CP-GS at a glance.** (a) Existing methods optimize Gaussian capacity and attribute precision separately. (b) We identify capacity mismatch, heterogeneous precision sensitivity, and capacity-precision coupling under limited budgets. (c) CP-GS performs content-adaptive capacity reallocation, hierarchical distortion-aware quantization, and joint capacity-precision optimization.

both resources compete for the same coding budget. Under a fixed rate, a representation may favor either more low-precision Gaussians for broader spatial coverage or fewer high-precision Gaussians for more faithful per-primitive encoding. Therefore, the preferred capacity and precision are inherently interdependent and should be jointly optimized rather than determined in isolation.

Figure 1(b) highlights three key observations that motivate CP-GS. First, **Gaussian capacity may not match content demand.** Redundant Gaussians can accumulate in over-represented regions, while high-residual regions remain under-represented. This suggests that a fixed Gaussian budget should remain reallocatable based on content demand. Second, **Gaussian attributes exhibit heterogeneous quantization sensitivity.** Their sensitivity varies across attributes and within attributes, making uniform precision inefficient and motivating hierarchical distortion-aware quantization without exhaustive search. Third, **capacity and precision are coupled under a shared budget.** Increasing Gaussian capacity improves spatial coverage but reduces precision available to each primitive, whereas higher precision improves attribute fidelity at the cost of fewer Gaussians. This inherent trade-off motivates jointly optimizing capacity and precision in a unified rate-distortion space.

Motivated by these observations, we propose **CP-GS**, a **C**apacity-**P**recision coupled **G**aussian **S**platting framework for budget-aware 2D Gaussian image representation. CP-GS first constructs a content-adaptive representation by reallocating a fixed Gaussian budget from redundant primitives to high-demand regions. It then performs hierarchical distortion-aware quantization to allocate precision according to heterogeneous quantization sensitivity at both the attribute and intra-attribute levels. Finally, CP-GS jointly optimizes Gaussian capacity and attribute precision as competing representation resources under a shared budget, adaptively balancing spatial coverage and per-primitive fidelity. As illustrated in Fig. 1(c), CP-GS therefore adapts both the spatial distribution and precision of Gaussian primitives to image content and representation budget. Extensive experiments on DIV2K and Kodak demonstrate that CP-GS achieves improved rate-distortion performance across a wide range of representation budgets. Our main contributions are summarized as follows:

- We develop a content-adaptive 2D Gaussian representation that reallocates a fixed Gaussian budget from redundant primitives to high-demand regions, improving capacity utilization without increasing the initial primitive budget.
- We introduce hierarchical distortion-aware representation quantization that models heterogeneous quantization sensitivity at both the attribute and intra-attribute levels, enabling efficient mixed-precision allocation without exhaustive search.
- We propose capacity-precision coupled optimization that jointly optimizes Gaussian capacity and attribute precision as competing representation resources, enabling adaptive trade-offs between spatial coverage and per-primitive fidelity under a shared budget.

2 Related Work

2.1 2D Gaussian Image Representation

Explicit 2D Gaussian representations have recently emerged as an efficient alternative to INR-based image codecs, such as SIREN Sitzmann et al. (2020) and COIN Dupont et al. (2021). Gaussian-

based methods represent images with position, covariance, and color parameters, and support fast differentiable rasterization (Zhang et al., 2024). Subsequent works have extended this paradigm along two main directions: continuous and arbitrary-scale signal representation (Hu et al., 2025; Peng et al., 2026), and compact or structured encoding (Pan et al., 2026). Collectively, these works demonstrate that 2D Gaussians provide an explicit and highly adaptable representation space. Our work builds on this perspective but focuses on how limited representation resources should be distributed between Gaussian capacity and parameter precision.

2.2 Adaptive Gaussian Modeling

A central problem in 2D Gaussian image representation is how to allocate a finite number of primitives according to spatially varying image complexity. Image-GS Zhang et al. (2025d) performs content-adaptive allocation by progressively placing anisotropic Gaussians according to image content. GaussianImage++ Li et al. (2026) uses distortion-driven densification to allocate primitives according to signal intensity. Instant GaussianImage Zeng et al. (2025) predicts Gaussian parameters in a feed-forward manner, enabling self-adaptive representation without per-image optimization. LocoADC Chen et al. (2026b) exploits locality in both image and Gaussian spaces through region-wise densification and similarity-driven merging. From a frequency-domain perspective, Spectrum-Adaptive Distribution Wan et al. (2025) distributes Gaussians according to spectral characteristics, while GPSToken Zhang et al. (2026) uses entropy-driven partitioning to achieve non-uniform tokenization. At the budget level, SPARE-GS Chen et al. (2026d) formulates structural evolution as a global budget-constrained optimization problem. In the 3DGS domain, GDAGS Zhou et al. (2026) introduces gradient-direction-aware density control. Collectively, these methods establish adaptive density control as an effective mechanism for improving spatial capacity utilization. However, they primarily determine *where* primitives should be created, retained, merged, or removed, while attribute precision is typically fixed or optimized separately.

2.3 Gaussian Compression and Quantization

Gaussian compression reduces representation cost through quantization, pruning, and entropy coding of primitive attributes. Structure-Guided Allocation Liang et al. (2025) couples image structure with adaptive bitwidth quantization, while CGVQ Cheng et al. (2026) applies K-means-guided codebook quantization. In the 3DGS domain, RDO-Gaussian Wang et al. (2024) formulates end-to-end rate-distortion optimized compression with dynamic pruning and entropy-constrained quantization. CAT-3DGS Zhan et al. (2025) employs context-adaptive triplane autoregressive coding to capture inter- and intra-primitive correlations. CSGaussian Tseng et al. (2026) combines compression and segmentation using an INR-based hyperprior, while HAC Chen et al. (2024) introduces hash-grid-assisted context for entropy coding. PCGS Chen et al. (2026c) studies progressive compression for scalable decoding, and HilComp Huang et al. (2025) uses Hilbert curve-based balanced clustering for compact 3D Gaussian representation. Collectively, these studies show that Gaussian attributes have heterogeneous coding requirements and that representation structure affects rate-distortion behavior. However, existing methods typically optimize precision under a fixed Gaussian capacity or treat capacity and precision as sequentially optimized resources. In contrast, CP-GS treats adaptive Gaussian modeling as one dimension of representation allocation and jointly optimizes attribute precision and Gaussian capacity under a shared budget.

3 Preliminary

Following prior 2D Gaussian image representations (Zhang et al., 2024; 2025d), we represent an image using a set of anisotropic Gaussian primitives defined directly on the image plane. Given an image $I \in [0,1]^{H\times W\times 3}$, its representation with N primitives is:

$$\mathcal{G}_N = \{(\boldsymbol{\mu}_i, \boldsymbol{\gamma}_i, \mathbf{c}_i)\}_{i=1}^{N}, \tag{1}$$

where $\boldsymbol{\mu}_i$, $\boldsymbol{\gamma}_i$, and $\mathbf{c}_i$ denote position, geometry, and color, respectively. The geometry parameters $\boldsymbol{\gamma}_i$ define a positive-definite covariance $\boldsymbol{\Sigma}_i$, such that primitive i provides an anisotropic spatial basis $G_i(\mathbf{x}) = \exp[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)]$. A differentiable splatting renderer $\mathcal{R}$ aggregates the colored Gaussian responses to reconstruct the image.

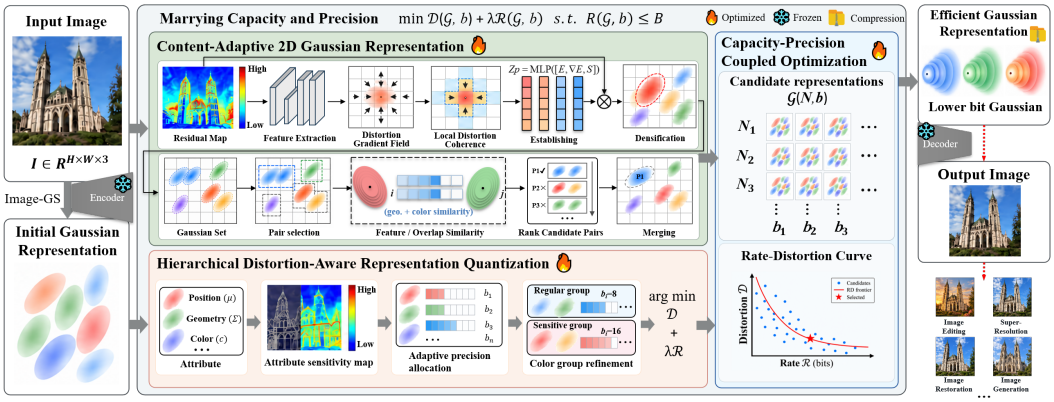

Figure 2: Illustration of our CP-GS framework. Given an input image, an initial Gaussian representation is first obtained via Image-GS. Two parallel branches, content-adaptive 2D Gaussian representation and hierarchical distortion-aware representation quantization, are then jointly optimized under a shared budget, yielding a capacity–precision coupled Gaussian representation that supports efficient decoding and downstream tasks.

Let $\mathbf{b}$ denote the attribute-precision configuration, yielding the quantized reconstruction $\hat{I}_{N,\mathbf{b}} = \mathcal{R}(Q_{\mathbf{b}}(\mathcal{G}_N))$. We characterize each representation by its distortion $D(N, \mathbf{b})$ and formulate budget-aware Gaussian representation as

$$\min_{N,\mathbf{b}} D(N, \mathbf{b}) \qquad \text{s.t.} \qquad R(N, \mathbf{b}) \le B, \tag{2}$$

where N controls representation capacity, $\mathbf{b}$ controls attribute precision, and B denotes the target budget. Further details are provided in *Appendix*.

4 Methodology

4.1 Overview

As illustrated in Fig. 2, given an input image, CP-GS first obtains an initial Gaussian representation via Image-GS (Zhang et al., 2025d). It then applies two parallel branches: a content-adaptive 2D Gaussian representation that reallocates redundant capacity toward regions with higher reconstruction demand and a hierarchical distortion-aware representation quantization that assigns attribute precision according to rendering sensitivity and further refines heterogeneous color groups. Under a shared budget, capacity and precision configurations are jointly evaluated in a shared candidate space to select the best rate-distortion operating point, yielding an efficient and compact low-bit Gaussian representation that can be decoded for image reconstruction and downstream applications such as image editing, super-resolution, image restoration, and image generation.

4.2 Content-Adaptive 2D Gaussian Representation

Building on the content-aware initialization of Image-GS, CP-GS improves Gaussian representation efficiency under limited budgets by merging locally redundant Gaussians and reallocating the released capacity to regions with higher representation demand.

4.2.1 Demand-Guided Content Selection

We estimate content demand from the reconstruction before reallocation. Let $e(\mathbf{x}) = \|I(\mathbf{x}) - \mathcal{R}(\mathcal{G}_N)(\mathbf{x})\|_2$ denote the residual magnitude. To suppress isolated errors, we measure the local coherence of residual-gradient directions:

$$\mathbf{u}(\mathbf{x}) = \frac{\nabla e(\mathbf{x})}{\|\nabla e(\mathbf{x})\|_2 + \varepsilon}, \qquad s_{\text{dem}}(\mathbf{x}) = e(\mathbf{x}) \left\| \frac{1}{|\mathcal{N}_{\mathbf{x}}|} \sum_{\mathbf{y}\in\mathcal{N}_{\mathbf{x}}} \mathbf{u}(\mathbf{y}) \right\|_2 . \tag{3}$$

Here, ∇ denotes the Sobel gradient and $\mathcal{N}_{\mathbf{x}}$ is a 7×7 neighborhood. We smooth the demand map and select high-scoring local maxima using radius-three non-maximum suppression, yielding spatially separated recipient locations. Computing the demand map before merging ensures that the selected recipients reflect the original reconstruction error rather than artifacts introduced by donor removal.

4.2.2 Redundant Gaussian Merging

To release capacity, we identify locally stable primitives with similar appearance and overlapping spatial support. Candidate pairs are further evaluated in the rendered domain, and low-cost disjoint pairs are merged to limit reconstruction degradation. The released capacity is then reassigned to the selected high-demand locations while preserving the total number of primitives:

$$|\mathcal{G}_N^+| = |\mathcal{G}_N|. \tag{4}$$

The resulting representation is subsequently refined through recovery optimization. Thus, CP-GS redistributes existing representation capacity rather than increasing the Gaussian budget. Detailed merging operations are provided in the *Appendix*.

4.3 Hierarchical Distortion-Aware Representation Quantization

Quantization affects Gaussian attributes through different rendering mechanisms: position and geometry determine spatial support, whereas color controls appearance. CP-GS therefore performs rendered-domain distortion calibration and hierarchically allocates precision across attributes and color groups.

4.3.1 Structured Attribute Calibration

Let $\mathbf{b} = (b_P, b_G, b_C)$ denote the precision configuration for position, geometry, and color, respectively, where $b_a \in \mathcal{B} = \{4, 6, 8, 10, 12, 16\}$. Exhaustive calibration requires evaluating $|\mathcal{B}|^3 = 216$ configurations, introducing substantial rendering overhead. Instead, CP-GS estimates attribute-wise distortion responses using a compact calibration procedure.

Let D_{ref} denote the distortion when all attributes are represented at 16-bit precision, and let $D_a(b)$ denote the distortion when only one attribute a is quantized to b bits. We approximate the joint distortion response as

$$\widetilde{D}(\mathbf{b}) = D_{\text{ref}} + \sum_{a\in\{P,G,C\}} \left[D_a(b_a) - D_{\text{ref}}\right]. \tag{5}$$

Structured-19 uses 19 logical calibration requests, corresponding to 16 unique precision configurations after caching, to efficiently approximate the finite precision search space. The selected configuration is then validated using the actual serialized representation.

4.3.2 Hierarchical Color Precision Refinement

Color primitives exhibit additional quantization heterogeneity within the color attribute. We therefore partition the color representation into two groups using robust channel-wise color statistics, separating primitives with atypical color responses from the dominant population. Each group is assigned its own quantization range and precision, while the selected position and geometry precisions remain fixed.

For each attribute-level operating point, we evaluate group-specific color precision candidates under the same serialized-rate constraint and select the pair that minimizes the predicted rendering distortion. The physical costs of the group map and group-specific quantization parameters are included in the final serialized representation, ensuring that the additional group structure is not treated as free side information. This hierarchical refinement captures intra-color quantization heterogeneity beyond global color precision allocation.

4.4 Capacity-Precision Coupled Optimization

Gaussian capacity and attribute precision are inherently coupled: the preferred precision depends on the retained Gaussian capacity, while the optimal capacity may change after precision adaptation.

CP-GS therefore jointly optimizes these two resources within a unified rate-distortion candidate space.

4.4.1 BUDGET-CONDITIONED CANDIDATE CONSTRUCTION

Given a fitted representation with N_0 primitives, we construct candidate capacities using retained ratios $r \in \{1, 0.8, 0.6, 0.4\}$. Primitives are selected according to reconstruction residuals, local gradients, and redundancy. Precision is then recalibrated for each retained count using initial 6-, 8-, 10-, and 12-bit configurations, followed by hierarchical color refinement.

Let $\mathcal{N}$ denote the candidate Gaussian counts and $\mathcal{A}_N$ the corresponding precision configurations. The joint candidate space is

$$\mathcal{X} = \bigcup_{N \in \mathcal{N}} \{(\mathcal{G}_N, \mathbf{b}) : \mathbf{b} \in \mathcal{A}_N\}. \tag{6}$$

Each candidate $x \in \mathcal{X}$ is evaluated using its decoded distortion D_x and actual serialized size S_x, after which dominated candidates are removed. Candidate construction and pruning details are provided in the *Appendix*.

4.4.2 JOINT CAPACITY—PRECISION OPTIMIZATION

The resulting candidate pool contains representations with different combinations of Gaussian capacity and parameter precision. Here, joint optimization refers to constructing and evaluating a shared candidate space over Gaussian capacity and attribute precision, followed by rate-distortion-based selection using decoded distortion and actual serialized size. Since each candidate is physically serialized and decoded before final evaluation, selection is based on measured distortion and actual storage cost rather than estimated parameter counts.

Given a target byte budget B, CP-GS selects the lowest-distortion feasible representation:

$$x_B^\star = \arg \min_{x \in \mathcal{X}:\, S_x \le B} D_x. \tag{7}$$

For continuous rate control, we further use the Lagrangian formulation.

$$x_\lambda^\star = \arg \min_{x \in \mathcal{X}} \left[D_x + \lambda R_x \right], \qquad R_x = \frac{8 S_x}{HW}, \quad \lambda > 0. \tag{8}$$

Both objectives search the same capacity-precision space but address different rate-control requirements. The hard-budget formulation selects the best representation satisfying a storage constraint while retaining non-dominated operating points missed by a linear rate penalty. More importantly, the search allows the preferred Gaussian count to change after precision adaptation: at a given rate, a representation may favor more primitives with coarser attributes or fewer with finer precision. CP-GS therefore treats capacity and precision as exchangeable resources and jointly optimizes them rather than committing to either in advance.

5 EXPERIMENTS

5.1 EXPERIMENTAL SETUP

Datasets and metrics. We evaluate Gaussian representation fidelity and rate-distortion performance on DIV2K Agustsson & Timofte (2017) and Kodak using a unified protocol with fixed hyperparameters. Coding ablations are conducted on the DIV2K validation set and Kodak24. We report PSNR, SSIM Wang et al. (2004), MS-SSIM Wang et al. (2003), and LPIPS Zhang et al. (2018), with the bitrate computed from complete serialized representations. The BD-rateHanhart & Ebrahimi (2014) is computed over valid overlapping quality intervals, and the number of valid images is reported for each comparison.

Implementation details. Experiments use 4 NVIDIA RTX 4090 GPUs. Unless otherwise stated, optimization time includes capacity reallocation, recovery optimization, precision calibration, serialization, decoding, and candidate evaluation. Content reallocation starts from a fixed Image-GS

Table 1: Comparison with representative INR- and GS-based image representation methods on Kodak and DIV2K. GS-based methods are evaluated at 5K/10K Gaussians on Kodak and 10K/30K on DIV2K. We report PSNR, MS-SSIM, parameter count, and optimization time. Optimization time denotes the single-image wall-clock time required to obtain the final representation.

Methods	Kodak				DIV2K			
	PSNR	MS-SSIM	Params(M)	Optim. Time (s)	PSNR	MS-SSIM	Params(M)	Optim. Time (s)
INR-based								
SIREN Sitzmann et al. (2020)	26.50	0.8705	3.74	889	30.01	0.9630	0.42	2242
SIREN+T_{sym} Zhang et al. (2025b)	32.80	0.9820	0.10	166	29.61	0.9712	0.10	261
EVOS Zhang et al. (2025c)	33.64	0.9809	0.10	106	30.38	0.9687	0.10	189
GS-based	Number of Gaussians (5K)				Number of Gaussians (10K)			
GaussianImage Zhang et al. (2024)	29.94	0.9632	0.04	234.47	29.71	0.9727	0.08	260.12
EA-GI Chen et al. (2026a)	30.84	0.9614	0.04	247.80	30.23	0.9653	0.08	301.31
GaussianImage++ Li et al. (2026)	31.90	0.9614	0.04	247.75	32.05	0.9742	0.08	283.28
Image-GS Zhang et al. (2025d)	36.50	0.9940	0.04	52.49	31.61	0.9684	0.08	74.37
CP-GS (Ours)	**37.41**	**0.9950**	0.04	**46.15**	**32.47**	**0.9771**	0.08	**66.21**
	Number of Gaussians (10K)				Number of Gaussians (30K)			
GaussianImage Zhang et al. (2024)	32.58	0.9824	0.08	258.11	34.64	0.9926	0.24	287.15
EA-GI Chen et al. (2026a)	33.87	0.9806	0.08	267.21	35.59	0.9896	0.24	346.78
GaussianImage++ Li et al. (2026)	35.41	0.9832	0.08	283.20	38.30	0.9940	0.24	290.85
Image-GS Zhang et al. (2025d)	39.31	0.9955	0.08	71.31	35.97	0.9913	0.24	112.43
CP-GS (Ours)	**40.79**	**0.9958**	0.08	**54.90**	**38.85**	**0.9943**	0.24	**95.70**

Table 2: Cumulative ablation of CP-GS on Kodak24. Quality metrics are averaged over all 24 images, with changes relative to the Spatial baseline shown in parentheses. BD-rate is measured relative to Spatial.

CP-GS Components				BD-Rate ↓ vs. Spatial (%)	Quality Metrics		
Spatial Realloc.	Attr. RD	Hier. Color	Joint Capa.–Prec.		PSNR ↑	SSIM ↑	LPIPS ↓
✓				–	19.310	0.6457	0.3699
✓	✓			−14.21	22.740 (+3.430)	0.6951 (+0.0494)	0.3639 (−0.0060)
✓	✓	✓		−23.71	23.225 (+3.915)	0.7053 (+0.0596)	0.3371 (−0.0328)
✓	✓	✓	✓	−44.93	**23.781** (+4.471)	**0.7147** (+0.0690)	**0.3308** (−0.0392)

representation with 5,000 Gaussians after 50,000 optimization steps, followed by one reallocation and 10,000 recovery steps. Coding ablations use representations optimized for 120,000 iterations with at most 5,000 Gaussians. Precision is selected from $\{4, 6, 8, 10, 12, 16\}$ under four payload budgets corresponding to uniform 6-, 8-, 10-, and 12-bit configurations, while count ablations retain 100%, 80%, 60%, or 40% of the fitted Gaussian set. Further details are provided in the *Appendix*.

5.2 Comparison with Existing 2D Image Representations

We compare CP-GS with representative INR- and GS-based image representation methods on Kodak and DIV2K. As shown in Table 1, CP-GS achieves the best reconstruction quality among the compared methods at the evaluated Gaussian counts while maintaining compact parameterization and low optimization cost. Compared with its Image-GS backbone, CP-GS improves PSNR by 0.91 and 1.48 dB on Kodak at 5K and 10K Gaussians, respectively, while reducing the corresponding single-image optimization time from 52.49 to 46.15 s and from 71.31 to 54.90 s. Similar gains are observed on DIV2K: CP-GS improves PSNR by 0.86 dB at 10K Gaussians and 2.88 dB at 30K, with higher MS-SSIM and lower optimization time than Image-GS. These results demonstrate that jointly optimizing Gaussian capacity and attribute precision can improve both representation fidelity and optimization efficiency.

5.3 Ablation Studies

We systematically analyze the three stages of CP-GS: content-adaptive 2D Gaussian representation, hierarchical distortion-aware representation quantization, and capacity-precision coupled optimization. Table 2 reports their cumulative contributions, while Table 3 provides targeted analyses of

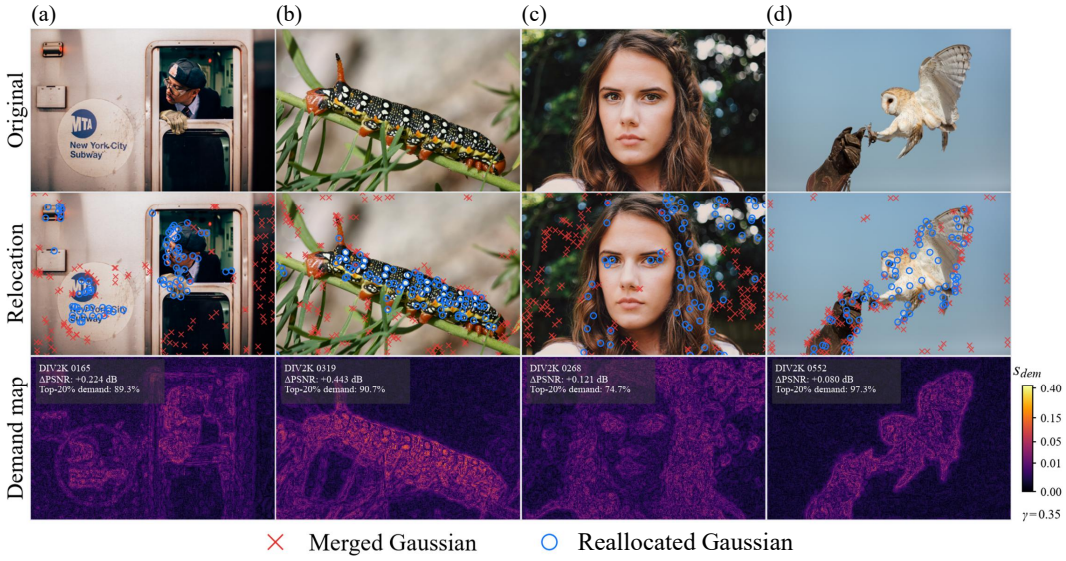

Figure 3: Visualization of content-adaptive Gaussian capacity reallocation on four representative images (a)–(d). Rows show the original image, Gaussian reallocation, and the pre-reallocation demand map. Red crosses denote merged Gaussians, and blue circles denote newly reallocated Gaussians.

attribute precision allocation, hierarchical color refinement, and capacity-precision coupling. Further ablation details are provided in the *Appendix*.

5.3.1 Effect of Each Component

Table 2 shows complementary gains from the three stages of CP-GS. Starting from the content-adaptive representation baseline, attribute-aware quantization reduces BD-rate by 14.21% while improving PSNR by 3.43 dB. Hierarchical color refinement further increases the BD-rate reduction to 23.71%, while capacity-precision coupled optimization achieves the best overall quality and a 44.93% BD-rate reduction. These progressive gains demonstrate that content-adaptive capacity utilization, hierarchical precision allocation, and their joint optimization provide complementary improvements in Gaussian representation efficiency.

5.3.2 Effect of Content-Adaptive Gaussian Reallocation

We evaluate whether demand-guided reallocation improves the utilization of a fixed Gaussian budget. All controls use the same merging and recovery settings and differ only in capacity placement. At $K = 150$, demand-guided reallocation achieves 24.520 dB PSNR, compared with 24.275 dB for random placement and 24.211 dB without reallocation, while placing 106 of the 150 newly allocated Gaussians in high-demand regions. Fig. 3(a)–(d) further shows that the reallocated capacity consistently follows high-demand structures. These results demonstrate that demand-guided reallocation improves capacity utilization without increasing the Gaussian budget.

5.3.3 Ablation of Hierarchical Distortion-Aware Quantization

We investigate attribute-level and intra-color precision allocation in Table 3(a,b,d). Oracle-216 exhaustively evaluates the finite precision space and therefore provides an empirical upper bound for precision allocation. Structured-19 achieves physical BD-rate reductions of 17.00% and 14.30% on DIV2K and Kodak, respectively, closely approaching this upper bound with mean PSNR gaps of only 0.038 and 0.032 dB. It requires approximately $11\times$ fewer logical calibration evaluations than Oracle-216 (19 vs. 216). Hierarchical K2 further reduces BD-rate by 7.09% and 10.68% over Global on DIV2K and Kodak, respectively, and consistently outperforms random grouping. These results show that distortion-aware precision allocation effectively captures quantization heterogeneity both across attributes and within color groups.

Table 3: Ablation analysis of precision allocation and capacity-precision coupling. Physical BD-rate (%, lower is better) is measured relative to Uniform in (a), Global in (b), and the fixed-count uniform baseline in (c). In (d), Structured-19 is compared with Oracle-216 under matched nominal payload targets; PSNR and BD-rate gaps are reported relative to Oracle-216.

(a) Attribute Precision Allocation

Method	Eval.	DIV2K (%)	Kodak (%)
Uniform (baseline)	0	0.00	0.00
Fixed mixed	0	+21.35	+20.65
Greedy	–	-5.84	-13.11
Random	16	-8.47	-3.19
Oracle	216	-17.94	-14.33
Structured (Ours)	**19**	**-17.00**	**-14.30**

(b) Hierarchical Color Refinement

Method	Groups	DIV2K (%)	Kodak (%)
Global	1	0.00	0.00
Shared	2	-0.80	-0.75
Random	2	-1.31	-1.62
Hierarchical (Ours)	2	**-7.09**	**-10.68**

(c) Capacity-Precision Coupling

Method	DIV2K (%)	Kodak (%)
Baseline	0.00	0.00
Hier. Precision	-7.18	-7.59
Count-only	-32.91	-48.12
Precision-only	-17.00	-14.30
Joint RD	-45.01	-50.14
Joint Hier. RD (Ours)	**-47.82**	**-53.35**

(d) Structured-19 (Ours) vs. Oracle-216

Metric	DIV2K	Kodak
Calibration evals. (S/O)	19 / 216	19 / 216
Config. match (%)	87.50	93.75
Mean PSNR gap (dB)	0.038	0.032
BD-rate gap (%)	0.127	0.013

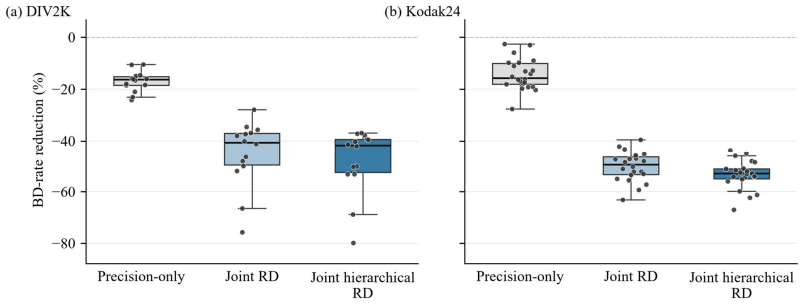

Figure 4: Image-level BD-rate distributions of Precision-only, Joint RD, and Joint Hierarchical RD on (a) DIV2K and (b) Kodak24. Each point denotes one image, and the boxes summarize the corresponding distributions. A more negative BD-rate indicates a better rate-distortion efficiency.

5.3.4 Ablation of Capacity—Precision Coupling

We further examine joint capacity—precision allocation under a shared budget. As shown in Table 3(c), optimizing capacity or precision alone improves rate-distortion efficiency, while Joint RD achieves larger BD-rate reductions of 45.01% and 50.14% on DIV2K and Kodak, respectively. Adding hierarchical precision further improves these reductions to 47.82% and 53.35%, respectively. Fig. 4 shows the same overall trend across individual images, demonstrating that capacity and precision are complementary resources that benefit from joint optimization.

6 Conclusion

We presented CP-GS, a capacity-precision coupled framework for budget-aware 2D Gaussian image representation. CP-GS improves Gaussian capacity utilization through content-adaptive reallocation, assigns precision according to hierarchical distortion sensitivity, and jointly optimizes capacity and precision under a shared budget. Experiments on DIV2K and Kodak demonstrate strong rate-distortion performance, validating the benefit of jointly considering where Gaussian capacity is allocated and how precisely it is encoded.

Broader Implications and Limitations. The capacity-precision coupling explored in CP-GS may generalize to other resource-constrained explicit representations. However, CP-GS currently relies on image-specific calibration and a predefined candidate space, while its evaluation is limited to static 2D images. Future work will explore learned allocation strategies, more efficient rate-distortion search, and broader visual representation tasks.

AI USE STATEMENT

In this work, Generative AI tools were used for language polishing and grammar checking. The authors take full responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

We provide the experimental protocols and implementation settings required to reproduce our results in Sec. 5.1, including datasets, Gaussian budgets, optimization schedules, precision candidates, and evaluation metrics. Detailed implementations of content-adaptive Gaussian reallocation, distortion-aware quantization, hierarchical color refinement, and capacity-precision optimization are provided in the *Appendix*. All reported bitrate results are computed from complete serialized representations, and additional ablations, rate-distortion statistics, and evaluation details are included in the supplementary material. We use fixed experimental settings across datasets and report the hardware configuration used for evaluation.

REFERENCES

Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

Jiacong Chen, Qingyu Mao, Shuai Liu, Chao Li, Jierun Lin, Xiandong Meng, Fanyang Meng, and Yongsheng Liang. Entropy-aware image representation via 2d gaussian splatting. *Signal Processing*, 244:110521, 2026a. ISSN 0165-1684.

Jiacong Chen, Qingyu Mao, Xiandong Meng, Shuai Liu, Chao Li, Fanyang Meng, Youneng Bao, and Yongsheng Liang. Locality-aware density control for efficient gaussian-based image representation, 2026b.

Yihang Chen, Qianyi Wu, Weiyao Lin, Mehrtash Harandi, and Jianfei Cai. HAC: Hash-grid assisted context for 3D Gaussian splatting compression. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 422–438, 2024.

Yihang Chen, Mengyao Li, Qianyi Wu, Weiyao Lin, Mehrtash Harandi, and Jianfei Cai. PCGS: Progressive compression of 3D Gaussian splatting. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 40, pp. 3111–3119, 2026c.

Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8628–8638, 2021.

Zhang Chen, Shuai Wan, Fuzheng Yang, Jiazhi Xia, Weiyao Lin, and Junhui Hou. Sparegs: Structural parsimony and resource efficiency for 3d gaussian splatting. *arXiv preprint arXiv:2607.16624*, 2026d.

Runze Cheng, Yicheng Zhan, Josef Spjut, and Kaan Akşit. Clustered codebook quantization for 2d gaussian-based image compression. *arXiv preprint arXiv:2607.05667*, 2026.

Emilien Dupont, Adam Goliński, Milad Alizadeh, Yee Whye Teh, and Arnaud Doucet. Coin: Compression with implicit neural representations. *arXiv preprint arXiv:2103.03123*, 2021.

Philippe Hanhart and Touradj Ebrahimi. Calculation of average coding efficiency based on subjective quality scores. *Journal of Visual Communication and Image Representation*, 25(3):555–564, 2014. ISSN 1047-3203. QoE in 2D/3D Video Systems.

Jintong Hu, Bin Xia, Bin Chen, Wenming Yang, and Lei Zhang. GaussianSR: High fidelity 2D Gaussian splatting for arbitrary-scale image super-resolution. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 39, pp. 3554–3562, 2025.

Yu Huang, Xiaowen Fu, Ling Ma, Zhengqi He, Jinbao Wang, and Xueliang Li. HilComp: Hilbert curve-based balanced clustering for 3D Gaussian splatting compression. In *Proceedings of the International Conference on Intelligent Computing (ICIC)*, pp. 421–432, 2025.

Junjun Jiang, Zengyuan Zuo, Gang Wu, Kui Jiang, and Xianming Liu. A survey on all-in-one image restoration: Taxonomy, evaluation and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(12):11892–11911, 2025.

Chia-Hao Kao, Cheng Chien, Yu-Jen Tseng, Yi-Hsin Chen, Alessandro Gnutti, Shao-Yuan Lo, Wen-Hsiao Peng, and Riccardo Leonardi. Bridging compressed image latents and multimodal large language models. In *Proceedings of the International Conference on Learning Representations (ICLR)*, volume 2025, pp. 7966–7986, 2025.

Tiantian Li, Xinjie Zhang, Xingtong Ge, Tongda Xu, Dailan He, Jun Zhang, and Yan Wang. GaussianImage++: Boosted image representation and compression with 2D Gaussian splatting. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 40, pp. 6442–6449, 2026.

Huanxiong Liang, Yunuo Chen, Yicheng Pan, Sixian Wang, Jincheng Dai, Guo Lu, and Wenjun Zhang. Structure-guided allocation of 2D Gaussians for image representation and compression. *arXiv preprint arXiv:2512.24018*, 2025.

Zixuan Pan, Kaiyuan Tang, Jun Xia, Yifan Qin, Lin Gu, Chaoli Wang, Jianxu Chen, and Yiyu Shi. Sgi: Structured 2d gaussians for efficient and compact large image representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12162–12172, June 2026.

Long Peng, Anran Wu, Wenbo Li, Peizhe Xia, Xueyuan Dai, Xinjie Zhang, Xin Di, Haoze Sun, Renjing Pei, Yang Wang, Yang Cao, and Zheng-Jun Zha. Pixel to gaussian: Ultra-fast continuous super-resolution with 2d gaussian modeling. In *International Conference on Learning Representations (ICLR)*, volume 2026, pp. 92172–92199, 2026.

Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, volume 33, pp. 7462–7473, 2020.

Xiangpeng Tian, Xiangyu Liao, Xiao Liu, Meng Li, and Chao Ren. Degradation-aware feature perturbation for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28165–28175. IEEE, 2025.

Yu-Jen Tseng, Chia-Hao Kao, Jing-Zhong Chen, Alessandro Gnutti, Shao-Yuan Lo, Yen-Yu Lin, and Wen-Hsiao Peng. CSGaussian: Progressive rate-distortion compression and segmentation for 3D Gaussian splatting. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 6883–6892, 2026.

Zunian Wan, Jiancheng Zhao, Yepeng Ding, Lingfeng Zhang, Hiroyuki Sato, and Takefumi Ogawa. Spectrum-adaptive distribution of 2D Gaussians for image representation and compression. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, 2025.

Henan Wang, Hanxin Zhu, Tianyu He, Runsen Feng, Jiajun Deng, Jiang Bian, and Zhibo Chen. End-to-end rate-distortion optimized 3D Gaussian representation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 76–92, 2024.

Z. Wang, E.P. Simoncelli, and A.C. Bovik. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pp. 1398–1402 Vol.2, 2003.

Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

Shitao Xiao, Yueze Wang, Junjie Zhou, Huaying Yuan, Xingrun Xing, Ruiran Yan, Chaofan Li, Shuting Wang, Tiejun Huang, and Zheng Liu. Omnigen: Unified image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13294–13304. IEEE, 2025.

Yang Ye, Xianyi He, Zongjian Li, Shenghai Yuan, Zhiyuan Yan, Bohan Hou, Li Yuan, et al. ImgEdit: A unified image editing dataset and benchmark. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, volume 38, 2026.

Zhaojie Zeng, Yuesong Wang, Tao Guan, Chao Yang, and Lili Ju. Instant GaussianImage: A generalizable and self-adaptive image representation via 2D Gaussian splatting. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 27896–27905, 2025.

Yu-Ting Zhan, Cheng-Yuan Ho, He-Bi Yang, Yi-Hsin Chen, Jui-Chiu Chiang, Yu-Lun Liu, and Wen-Hsiao Peng. CAT-3DGS: A context-adaptive triplane approach to rate-distortion-optimized 3DGS compression. In *Proceedings of the International Conference on Learning Representations (ICLR)*, volume 2025, pp. 26312–26334, 2025.

Jingwei Zhang, Anh Tien Nguyen, Xi Han, Vincent Quoc-Huy Trinh, Hong Qin, Dimitris Samaras, and Mahdi S Hosseini. 2dmamba: Efficient state space model for image representation with applications on giga-pixel whole slide image classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3583–3592. IEEE, 2025a.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

Weixiang Zhang, Shuzhao Xie, Chengwei Ren, Shijia Ge, Mingzi Wang, and Zhi Wang. Enhancing implicit neural representations via symmetric power transformation. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 39(10):10157–10165, 2025b.

Weixiang Zhang, Shuzhao Xie, Chengwei Ren, Siyi Xie, Chen Tang, Shijia Ge, Mingzi Wang, and Zhi Wang. Evos: Efficient implicit neural training via evolutionary selector. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 30472–30482, June 2025c.

Xinjie Zhang, Xingtong Ge, Tongda Xu, Dailan He, Yan Wang, Hongwei Qin, Guo Lu, Jing Geng, and Jun Zhang. Gaussianimage: 1000 fps image representation and compression by 2d gaussian splatting. In *European Conference on Computer Vision (ECCV)*, pp. 327–345. Springer Nature Switzerland, 2024.

Yunxiang Zhang, Bingxuan Li, Alexandr Kuznetsov, Akshay Jindal, Stavros Diolatzis, Kenneth Chen, Anton Sochenov, Anton Kaplanyan, and Qi Sun. Image-gs: Content-adaptive image representation via 2d gaussians. In *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Papers*, pp. 102:1–102:11. Association for Computing Machinery, 2025d.

Zhengqiang Zhang, Rongyuan Wu, Lingchen Sun, and Lei Zhang. GPSToken: Gaussian parameterized spatially-adaptive tokenization for image representation and generation. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, volume 38, pp. 11218–11249, 2026.

Zheng Zhou, Yu-Jie Xiong, Jia-Chen Zhang, Chun-Ming Xia, Xihe Qiu, and Hongjian Zhan. Gradient-direction-aware density control for 3D Gaussian splatting. In *Proceedings of the International Conference on Learning Representations (ICLR)*, volume 2026, pp. 21415–21437, 2026.