Geometry-Aware Style Transfer in 3D Gaussian Splatting
– *Supplementary Material* –

Algorithm A Geometry-aware style transfer in 3DGS

Input: Content images $\mathbb{I}^{\mathcal{C}} = \{\mathbf{I}_i^{\mathcal{C}}\}_{i=1}^{M}$, style image $\mathbf{I}^{\mathcal{S}}$
Output: Stylized Gaussians $\mathbb{G}^{\mathrm{s}} = \{\boldsymbol{\Theta}^{\mathrm{s}}, \boldsymbol{\Phi}^{\mathrm{s}}\}$
$\mathbb{G}^{\mathrm{init}} \leftarrow \textsc{Recon3DGS}(\mathbb{I}^{\mathrm{c}})$
$\mathbb{G}^{\mathrm{c}} = \{\boldsymbol{\Theta}^0, \boldsymbol{\Phi}^0\} \leftarrow \textsc{ColorMatching}(\mathbb{G}^{\mathrm{init}}, \mathbb{I}^{\mathcal{C}}, \mathbf{I}^{\mathcal{S}})$
for $k = 1, 2, \ldots, K$ **do** ▷ outer cycle
  $\boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta}^{k-1}$
  $\boldsymbol{\Phi} \leftarrow \boldsymbol{\Phi}^{k-1}$
  // **Color optimization phase**
  for $i = 1, 2, \ldots, N_{\mathrm{c}}$ **do**
    $v \leftarrow \textsc{RandomView}()$
    $\mathbf{I}^{\mathcal{R}} \leftarrow \textsc{Render}(v, \{\boldsymbol{\Theta}, \boldsymbol{\Phi}\})$ ▷ to compute $\mathcal{L}$
    $\mathcal{L} \leftarrow \lambda_{\mathrm{GC}}\mathcal{L}_{\mathrm{GC}} + \lambda_{\mathrm{TV}}\mathcal{L}_{\mathrm{TV}} + \lambda_{\mathrm{cont}}\mathcal{L}_{\mathrm{cont}}$
    $\boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta} - \eta\,\nabla_{\boldsymbol{\Theta}}\mathcal{L}$ ▷ $\boldsymbol{\Phi}$ is not updated
  end for
  $\boldsymbol{\Theta}^k \leftarrow \boldsymbol{\Theta}$
  // **Geometry optimization phase**
  for $j = 1, 2, \ldots, N_{\mathrm{g}}$ **do**
    $v \leftarrow \textsc{RandomView}()$
    $\mathbf{I}^{\mathcal{R}} \leftarrow \textsc{Render}(v, \{\boldsymbol{\Theta}, \boldsymbol{\Phi}\})$ ▷ to compute $\mathcal{L}$
    $\mathcal{L} \leftarrow \lambda_{\mathrm{GC}}\mathcal{L}_{\mathrm{GC}} + \lambda_{\mathrm{TV}}\mathcal{L}_{\mathrm{TV}} + \lambda_{\mathrm{cont}}\mathcal{L}_{\mathrm{cont}} + \mathcal{L}_{\mathrm{reg}}$
    $\boldsymbol{\Phi} \leftarrow \boldsymbol{\Phi} - \eta\,\nabla_{\boldsymbol{\Phi}}\mathcal{L}$ ▷ $\boldsymbol{\Theta}$ is not updated
  end for
  $\boldsymbol{\Phi}^k \leftarrow \boldsymbol{\Phi}$
end for
return $\mathbb{G}^K = \{\boldsymbol{\Theta}^K, \boldsymbol{\Phi}^K\}$

A Additional Complexity Analysis

Table A summarizes the number of Gaussian primitives, the average stylization time, and GPU memory usage across different scenes. The stylization process takes approximately 9–18 minutes per scene depending on scene complexity and rendering configurations. While the stylization time does not show a clear correlation with the number of Gaussians, it can vary depending on scene characteristics and rendering configurations. In contrast, GPU memory usage tends to increase with the number of Gaussian primitives.

B User Study Details

Recruitment and procedure. We conducted an online user study with 37 participants to evaluate the perceptual quality of our method. The survey was

Table A: Computational efficiency analysis. Summary of the average number of Gaussians, stylization time, and GPU memory usage for each scene.

Scenes	# of Gaussians (M)	Stylization Time	GPU Memory (GB)
trex	0.54	17m 54s	20.49
flower	0.54	17m 00s	16.63
horns	0.81	18m 54s	22.93
fern	0.82	18m 31s	15.41
garden	4.20	16m 23s	38.94
kitchen	1.53	7m 25s	20.06
train	1.07	8m 24s	25.69
truck	2.59	9m 16s	27.21

hosted using an online survey platform, and participants were recruited through our university's online community and bulletin boards.

Each participant completed evaluation trials comprising four scenes and four style images, resulting in a total of 16 scene–style combinations. In each trial, participants compared the stylized results from our method against four methods: StyleGaussian [21], G-Style [18], SGSST [4], and CLIPGaussian [8]. All results were presented anonymously and in a randomized order to mitigate potential bias.

Survey interface. Fig. A illustrates the typical interface presented to participants on the online survey platform. For each trial, participants were provided with the original scene, the target style image, and five anonymized stylized results. They were then asked to answer the following three questions.

Survey questions. Participants were asked the following three questions for each set of results:

Q1. Style Similarity (Ranking) "Please rank the following results according to their similarity to the reference style (1=Most Similar, 5=Least Similar)."
Note: This was a ranking question where each method received a unique rank from 1 to 5.

Q2. Visual Appeal (Rating) "On a scale from 'Very Low' to 'Very High', how would you rate the visual appeal of each creation?"
Note: This was a 5-point Likert scale (1=Very Low, 5=Very High) applied independently to each result.

Q3. Content Recognizability (Rating) "How well can the shape and structure of the original object be recognized in each result?"
Note: This was a 5-point Likert scale (e.g., 1=Very Difficult, 5=Very Easy) applied independently to each result.

Aggregated results. The aggregated results from all 37 participants are presented both graphically in the main paper and numerically in this supplement. Fig. 6 in the main paper visualizes the detailed distribution of participant responses for each question. For Q1, it depicts the cumulative vote count for each rank, while for Q2 and Q3, it shows the full distribution of 5-point Likert-scale

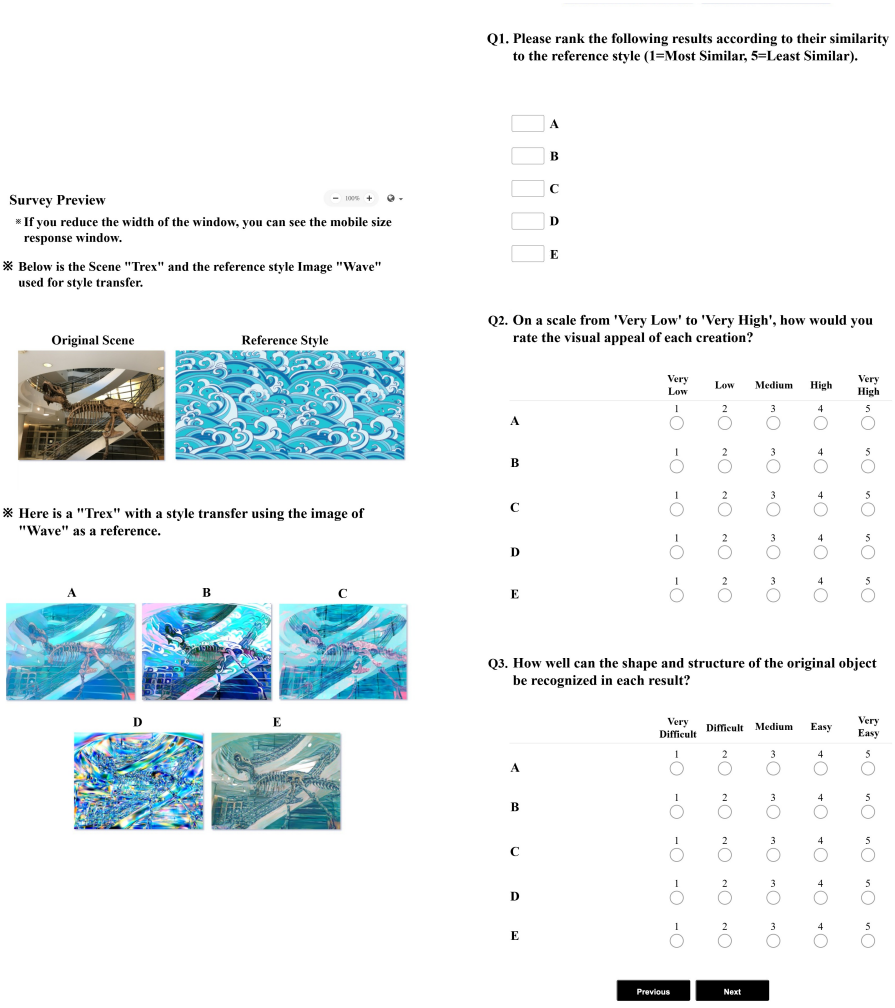

Fig. A: A screenshot of the user study interface on the online survey platform.

ratings. To provide a concise statistical summary of these distributions, we report the mean scores for all three questions in Table B. Together with the distributions shown in Fig. 6, these results demonstrate that our method was consistently preferred by participants, achieving the highest ratings across all evaluation criteria.

C Additional Ablation Studies

Additional ablation study on optimization strategies. In Fig. B, we compare four optimization strategies for 3D style transfer: only geometry, only color, joint-step, and our decoupled optimization. Only geometry optimization focuses on structural deformation, effectively embedding geometric traits of the target style but lacking color variation, which results in visually incomplete stylization. On the other hand, only color optimization preserves scene geometry but limits style transfer to surface appearance, resulting in weak spatial integration of the style. The joint-step approach updates geometry and color simultaneously, yielding stronger stylization effects but often overfitting to the style appearance. Our decoupled optimization achieves a balanced integration of color and geometry, enabling geometry-aware stylization that effectively reflects geometric characteristics of the target style while maintaining appropriate color stylization.

Fig. C additionally illustrates the results of only geometry optimization. Although it lacks color expressiveness, it still captures structural cues characteristic of the target style, producing distinctive surface reliefs and geometry-aware stylization effects. This observation suggests that geometry-based modulation can serve as a useful component for structural enhancement in 3D style transfer frameworks.

Ablation study on scale control. In Fig. D, we investigate the effect of the target texture scale. Here, the scale is defined relative to the default target texture resolution of 256×256: $0.5\times$, $1\times$, and $2\times$ correspond to resized resolutions of 128×128, 256×256, and 512×512, respectively. This parameter controls the strength and granularity of the transferred texture patterns, thereby affecting the balance between structural stability and stylistic deformation.

At a smaller scale (*e.g.*, $0.5\times$), the target texture is downsampled to 128×128, which suppresses excessive geometric changes and helps preserve the original scene structure. The default scale ($1\times$) provides a balanced result between geometry preservation and texture transfer. In contrast, a larger scale (*e.g.*, $2\times$) uses a higher-resolution target texture of 512×512, allowing more pronounced and expressive geometric patterns to appear in the stylized result. This qualitative analysis shows that the target texture scale provides a flexible control mechanism for adaptively balancing structural fidelity and stylistic expressiveness.

Ablation study on the geometry/color update ratio. We investigate the impact of varying the ratio between geometry and color updates within each

Table B: User study results. Average scores of 2,368 responses from 37 participants: Q1 – *Style Similarity* (↓), Q2 – *Visual Appeal* (↑), and Q3 – *Content Recognizability* (↑).

Methods	Q1↓	Q2↑	Q3↑
StyleGaussian [21]	4.78	1.95	1.90
G-Style [18]	2.64	3.65	3.83
SGSST [4]	<u>2.26</u>	<u>3.74</u>	3.70
CLIPGaussian [8]	3.55	3.45	<u>4.19</u>
Ours	**1.75**	**4.07**	**4.32**

Table C: Ablation study on the geometry/color update ratio. Comparison of SIFID scores across different N_g/N_c values.

N_g	0	30	60	90 (ours)	100
SIFID↓	1.5244	1.3272	1.2526	**1.1736**	<u>1.1885</u>

optimization cycle. Each cycle consists of 100 iterations, where N_g and N_c denote the number of geometry and color update steps, respectively, such that $N_g + N_c = 100$. Table C presents the SIFID scores across different update ratios. Our configuration ($N_g = 90$, $N_c = 10$) achieves the best perceptual score, indicating a balanced trade-off between stylization strength and geometric stability. In Fig. E, we present qualitative results showing the effect of varying the geometry/color update ratio. A higher geometry ratio (*e.g.*, $N_g = 100$) results in overly rigid structures with limited color adaptation, while a lower ratio (*e.g.*, $N_g = 30$) weakens geometric stylization and leads to flatter structures. Our setting ($N_g = 90$) yields vivid and spatially consistent stylization, demonstrating that a geometry-dominant update schedule is essential for stable and expressive 3D style transfer.

Ablation study on GCFM hyperparameter. In Table D, we evaluate the sensitivity of our framework to the weighting factor λ_{GC} in GCFM. The SIFID scores remain largely consistent across a wide range of values, indicating that GCFM is robust to moderate variations in its weighting. When the term is not used (*i.e.*, $\lambda_{GC} = 0$), the performance drops notably, suggesting that the GCFM plays a crucial role in guiding consistent stylization. Across a wide range of λ_{GC} values, our method produces stable and visually coherent stylization results. We adopt $\lambda_{GC} = 2.0$ in all experiments, as it consistently provides visually appealing and structurally coherent results across diverse scenes. This result demonstrates that our method is robust to variations in λ_{GC}, producing consistent stylization quality with stable optimization behavior.

D Comparison with StylizedGS

In Fig. F, we qualitatively compare our method with StylizedGS [36]. While StylizedGS achieves higher quantitative consistency (LPIPS and RMSE) under both short- and long-range evaluations, this can be attributed to its overly

Table D: Ablation study on GCFM hyperparameter. SIFID remains stable across various weights, confirming that our method is robust to the choice of λ_{GC}.

λ_{GC}	0.0	1.0	2.0 (ours)	5.0	10.0
SIFID↓	1.5718	**1.1683**	<u>1.1736</u>	1.1841	1.1878

smoothed and desaturated renderings rather than genuine structural coherence. As illustrated in the figure, StylizedGS tends to lose geometric details and fails to express distinctive style characteristics, yielding relatively flat and texture-suppressed outputs with diminished 3D structure. In contrast, our method preserves spatial geometry and reproduces vivid, structurally consistent stylization patterns. These results suggest that the high consistency scores of StylizedGS may not directly reflect superior visual quality, but rather a loss of stylization richness—whereas our framework achieves a balance between geometric coherence and expressive style fidelity.

E Additional Qualitative Results

This section presents additional qualitative stylization results and comparative analyses.

Scene-wise stylization results. Figs. G–L show stylization outcomes across multiple 3D scenes, including *Garden*, *Flower*, *Train*, *Truck*, *Kitchen*, and *Fern*. Figs. G and H present stylized results along with the corresponding depth maps, demonstrating that the geometric structure is effectively adapted to the target style while preserving the original scene content. Figs. I-L further illustrate stylizations using all exemplar styles adopted in our framework, showing consistent geometry–appearance adaptation across diverse environments. Overall, our method successfully preserves the structural integrity of the scene while transferring both color and geometric characteristics of the target style.

Comparisons with other methods. Figs. M–P provide additional comparisons with existing stylization approaches across diverse scenes and styles. Even under varying scene layouts and style characteristics, our method yields more stable and expressive stylizations than competing methods, maintaining geometric clarity and spatially coherent style patterns throughout the 3D space. These results complement the quantitative analyses in the main paper and further validate the robustness and visual fidelity of our proposed framework.

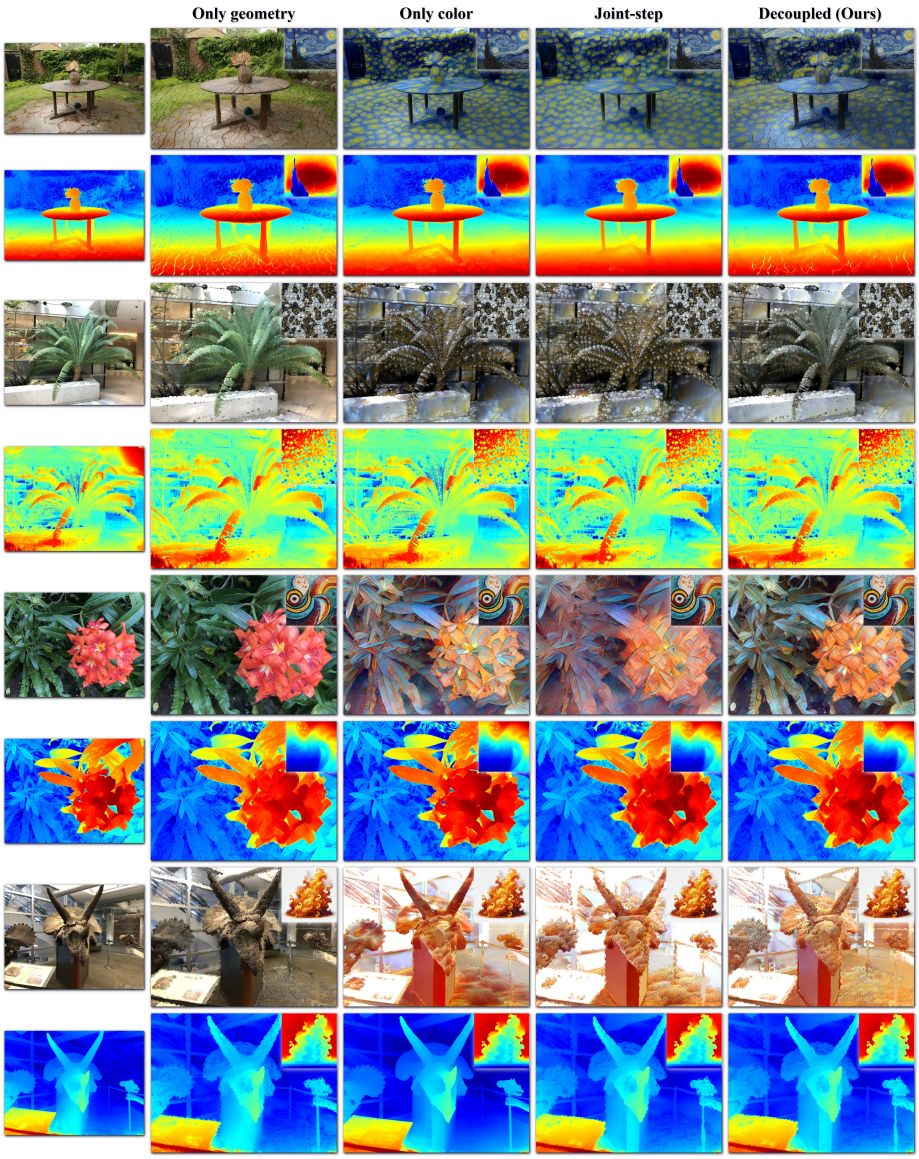

Fig. B: Qualitative comparison of optimization strategies.

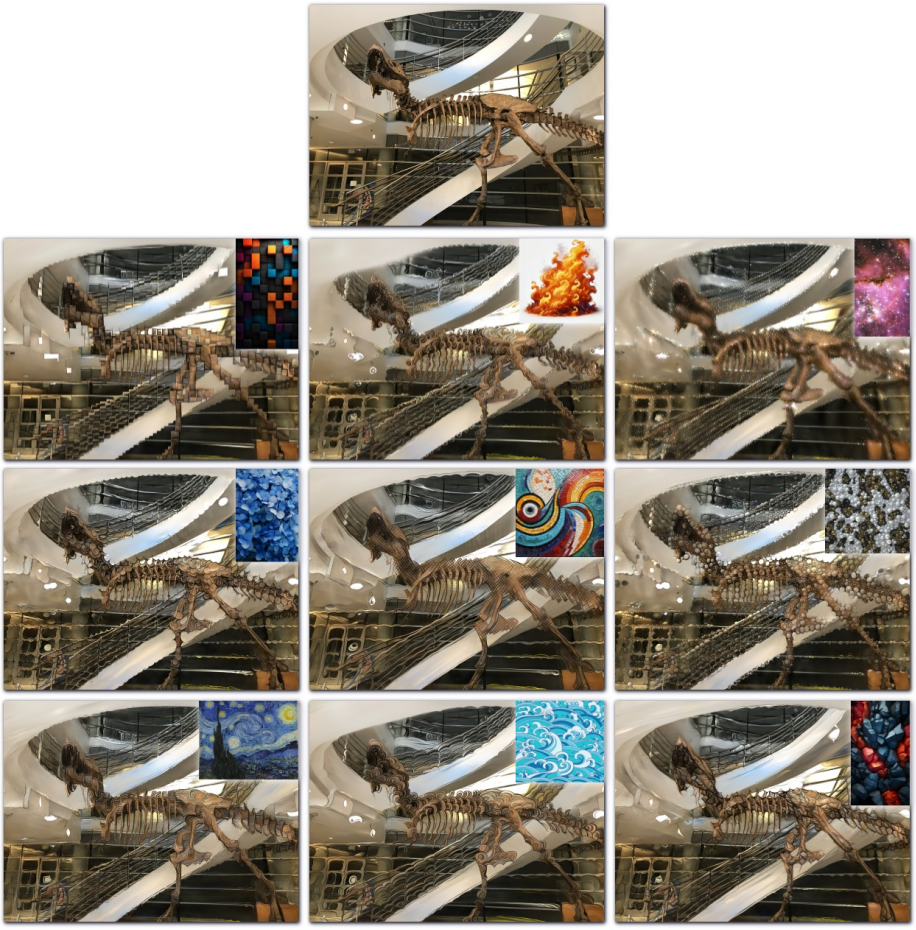

Fig. C: Qualitative results of only geometry optimization.

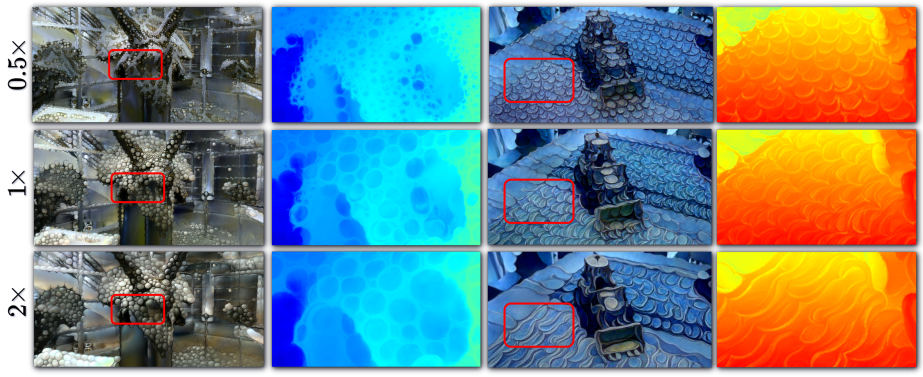

Fig. D: Qualitative results of target texture scale control. The scales 0.5×, 1×, and 2× correspond to target texture resolutions of 128 × 128, 256 × 256, and 512 × 512, respectively.

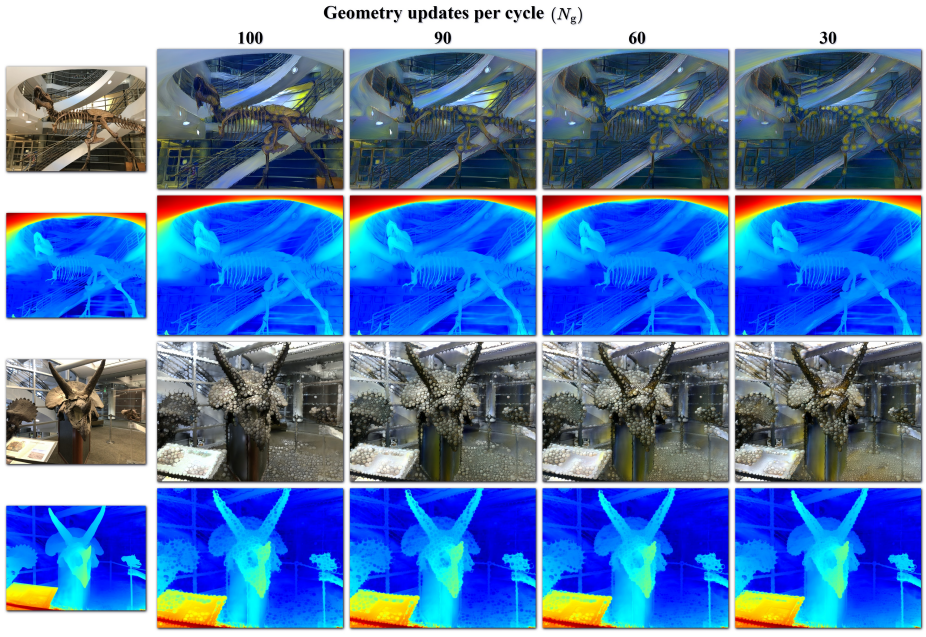

Fig. E: Qualitative results of the ablation study on the geometry/color update ratio.

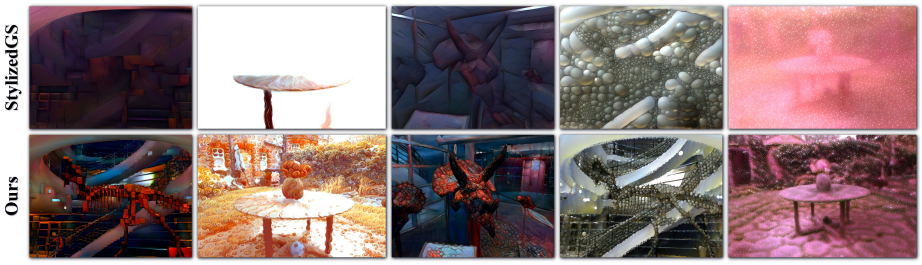

Fig. F: Qualitative comparison with StylizedGS.

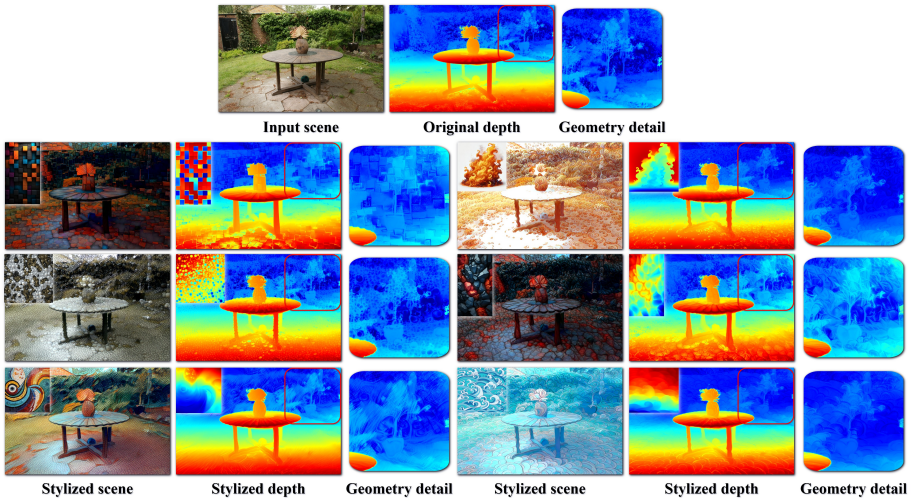

Fig. G: Stylization results for the *Garden* scene, including corresponding depth maps.

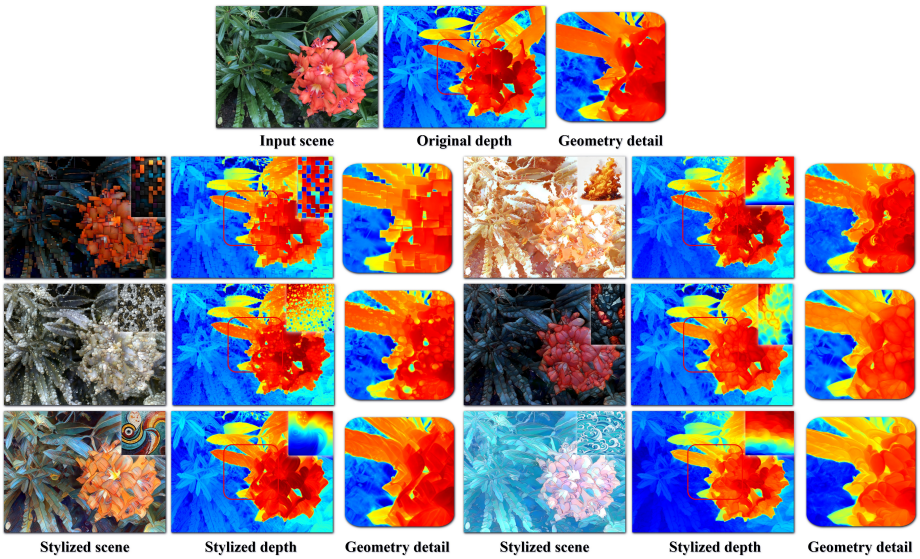

Fig. H: Stylization results for the *Flower* scene, including corresponding depth maps.

Fig. I: *Train* scene stylization results with nine exemplar styles.

Fig. J: *Truck* scene stylization results with nine exemplar styles.

Fig. K: *Kitchen* scene stylization results with nine exemplar styles.

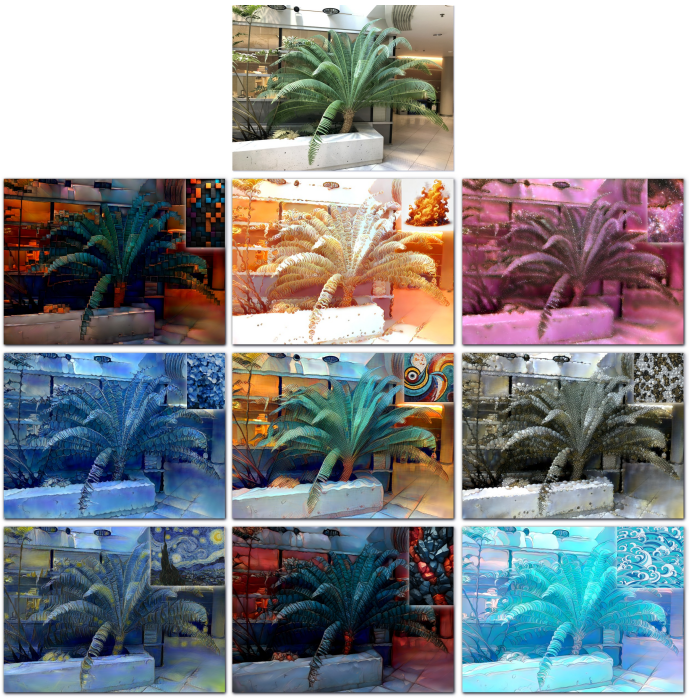

Fig. L: *Fern* scene stylization results with nine exemplar styles.

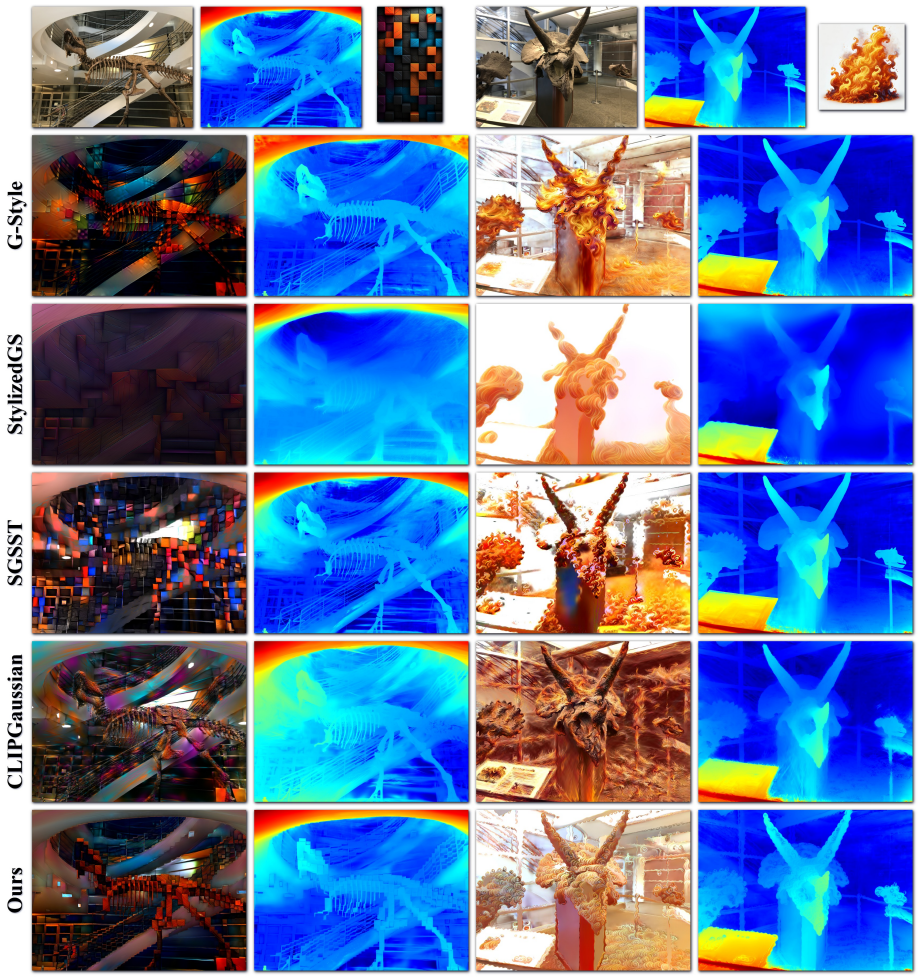

Fig. M: Qualitative comparison of the *Trex* and *Horns* scenes.

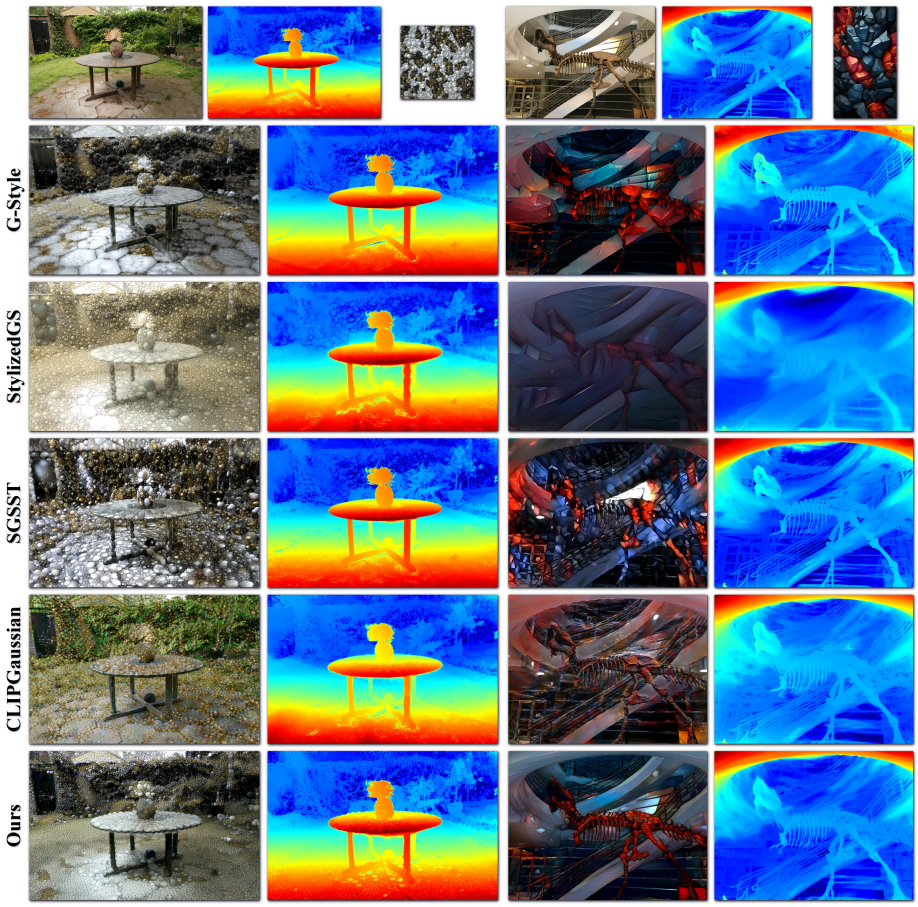

Fig. N: Qualitative comparison of the *Garden* and *Trex* scenes.

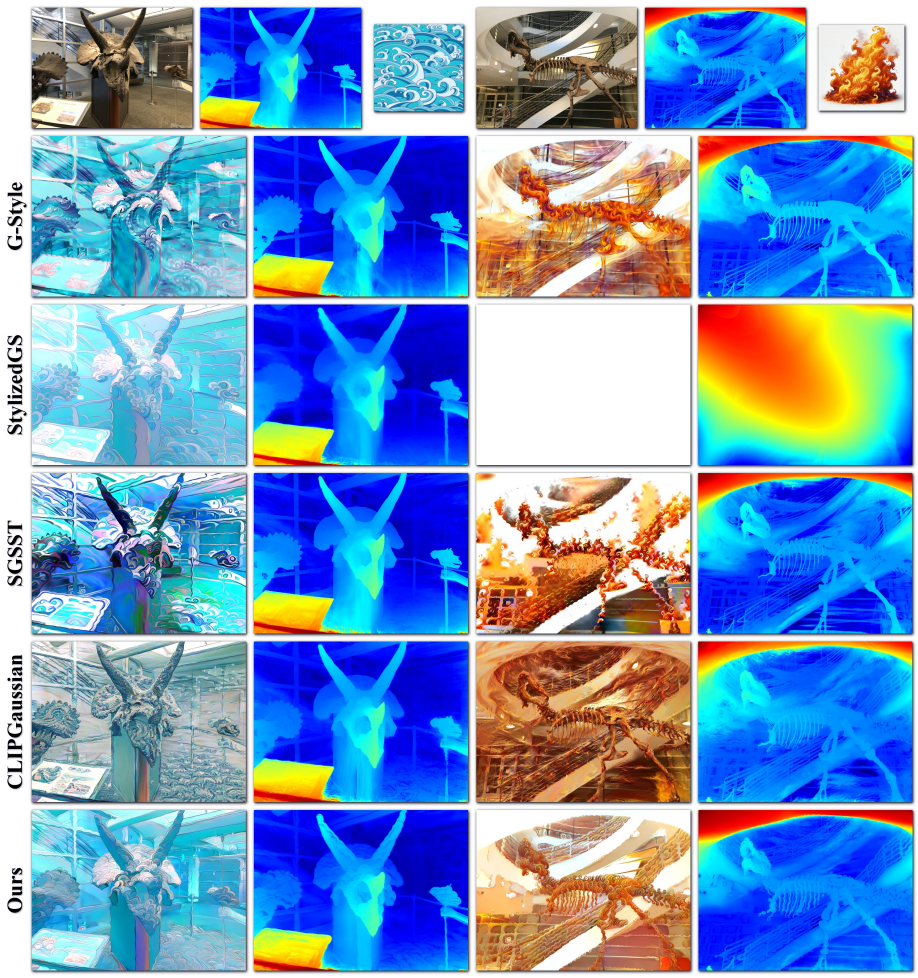

Fig. O: Qualitative comparison of the *Horns* and *Trex* scenes.

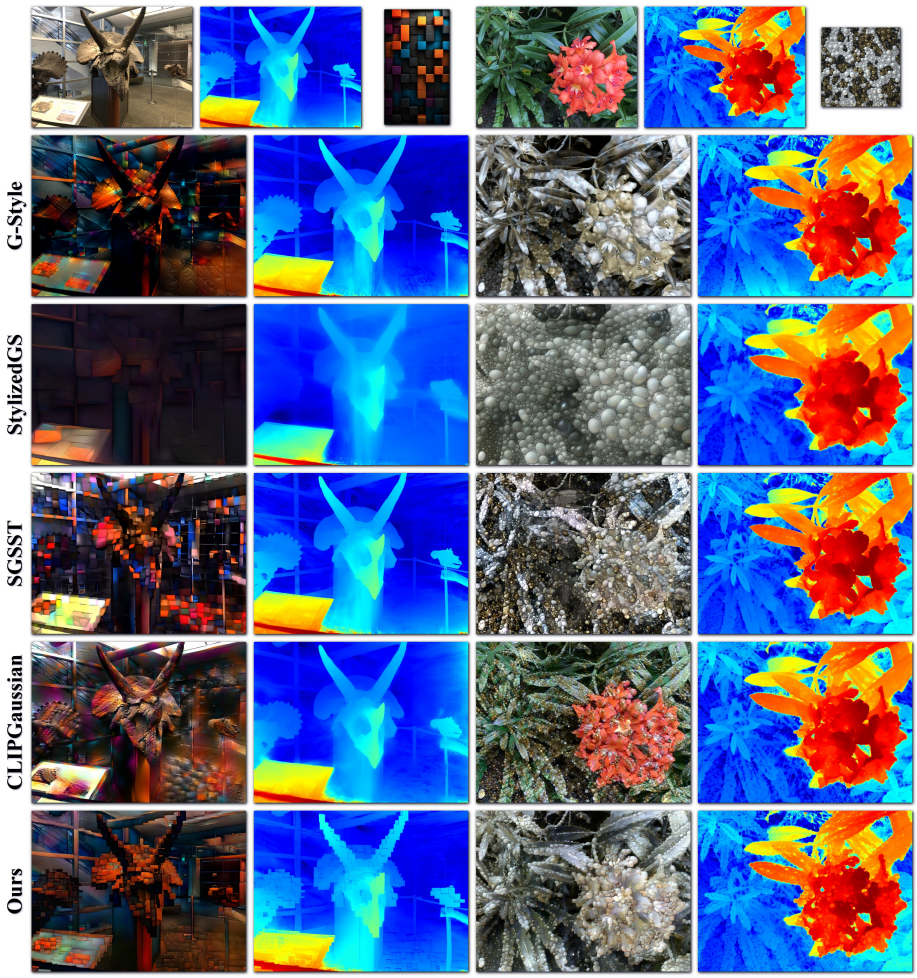

Fig. P: Qualitative comparison of the *Horns* and *Flower* scenes.