Hybrid Pruning Framework for 3D Gaussian Splatting

Tarun Gangadhar Vadaparthi

Abstract

For 3D Gaussian Splatting (3DGS), we introduce a hybrid pruning framework that maintains visual fidelity while drastically lowering the number of splats. In contrast to conventional single-factor pruning techniques (such as opacity thresholding or spherical-harmonics energy filtering), our method uses a weighted, robust z-score fusion to combine several complementary per-splat signals: (i) SH color energy, (ii) multi-view visibility rate, (iii) local spatial density, (iv) foreground hit-rate from baseline silhouette renders, and (v) projected image area. We keep the top-k fraction after ranking using this hybrid importance, and we avoid using heavy clustering techniques by applying a lightweight refinement stage that eliminates residual floaters using robust core-derived bounding boxes and percentile-based consistency tests. Tests conducted on actual 3DGS scenes show competitive rate-distortion performance (PSNR vs. splat fraction) in comparison to baselines, visually cleaner reconstructions, and fewer background artifacts. We offer a reproducible pipeline that includes interactive 3D visualizations, GIFs, side-by-side comparisons, and RD curves.

1 Overview

Motivation. Although real-time, high-fidelity view synthesis is achieved by 3D Gaussian Splatting (3DGS), raw scenes still have millions of splats in addition to background floaters and minor artifacts. Naïve, one-factor pruning (like opacity thresholds) either retains undesirable clutter or eliminates helpful details (like dark wheels). We tackle this by combining complementary geometric and photometric cues in a *hybrid* per-splat importance, then refining it in a lightweight way without the need for external clustering.

Method summary. We calculate five per-splat signals given a 3DGS scene (PLY): (i) SH color energy, (ii) multi-view visibility rate, (iii) local spatial density, (iv) foreground hit-rate from baseline silhouettes, and (v) projected image area. After applying a percentile-based refinement (foreground/density thresholds and a robust core-derived bounding box), we combine them using weighted importance, maintain the top-k fraction, and robustly normalize them (z-scores). After that, we render scenes that have been pruned, report PSNR vs. baseline (RD curves), and offer interactive 3D visualizations and qualitative comparisons.

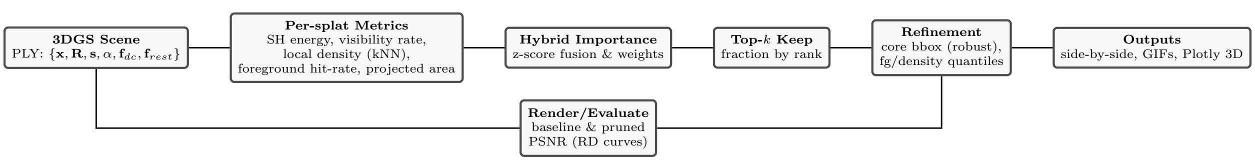

Figure 1: Pipeline. Five per-splat signals are fused into a hybrid importance, top-k are kept, and a lightweight refinement removes residual floaters. We render/evaluate for RD and visualize results.

1.1 Contributions

- **The hybrid significance of 3DGS pruning.** We suggest combining five complementary cues (SH energy, multi-view visibility, local density, foreground hit-rate, and projected area) into a weighted, robust z-score. This effectively eliminates background floaters while maintaining fine, low-albedo details (like dark wheels).
- **Lightweight refinement without clustering.** To eliminate any remaining junk splats, we use a robust core-derived bounding box and percentile-based thresholds (foreground, density) in place of heavy clustering techniques like DBSCAN.

- **Reproducible evaluation pipeline.** We offer a Colab-style workflow that includes: (i) baseline/pruned renders; (ii) quantitative rate-distortion (PSNR vs. kept fraction); (iii) GIFs and side-by-side comparisons; and (iv) interactive 3D Plotly visualizations for qualitative examination.

2 Experimental Setup

2.1 Input Data: 3DGS Pipeline

We begin with scenes that have been reconstructed using the 3D Gaussian Splatting (3DGS) pipeline. Per-splat attributes are present in every exported PLY file:

$$\{\mathbf{x}, \mathbf{R}, \mathbf{s}, \alpha, \mathbf{f}_{dc}, \mathbf{f}_{rest}\},\ ]where$$

x$\in \mathbb{R}^3$ is position, $\mathbf{R}$ is quaternion rotation, $\mathbf{s}$ is anisotropic scale, α is opacity, and $\mathbf{f}_{dc}, \mathbf{f}_{rest}$ are spherical harmonics (SH) coefficients. Typical scenes have small artifacts, cluttered backgrounds, and ~ 0.4 million splats.

2.2 Rendering and Evaluation

We use the 3DGS codebase's differentiable splatting renderer to compare pruning strategies. We establish a set of 12 camera positions with intrinsics (K, H, W) that are dispersed throughout an orbit around the object.

- **Baseline renders.** All splats kept.
- **Pruned renders.** Subset of splats selected by each pruning mask.

We save rendered RGB frames and compute PSNR vs. baseline for each pose. This yields a rate–distortion (RD) curve: fraction of splats kept vs. reconstruction fidelity.

2.3 Pruning Signals

We compute five complementary signals per splat:

(i) **SH color energy:** $\|\sigma(\mathbf{f}_{dc})\|^2 + \|\sigma(\mathbf{f}_{rest})\|^2$ where σ is sigmoid.

(ii) **Multi-view visibility rate:** fraction of orbit poses where the splat projects inside the image and is front-facing.

(iii) **Local spatial density:** inverse mean distance to $k = 15$ nearest neighbors.

(iv) **Foreground hit-rate:** percentage of views where the splat lies inside the baseline silhouette mask.

(v) **Projected area:** mean xy screen-space footprint under multiple poses.

2.4 Hybrid Importance

Signals are normalized to robust z-scores:

$$z = \frac{x - \text{median}(x)}{\text{MAD}(x)}.$$

We downweight extreme projected areas and assign weights to prioritize visibility, density, and foreground-hit rate. An importance score is assigned to each splat, and the top-k fraction is retained.

2.5 Refinement

We use a simple cleanup after pruning:

- Discard splats that fall below the local density or foreground hit-rate 5th percentile.
- Limit to a sturdy bounding box that is obtained from the $[2, 98]$ percentile of the coordinates x, y, z.

This eliminates any remaining blobs without the need for external clustering (like DBSCAN).

2.6 Visualization

We offer comparisons that are both static and interactive:

- Comparing baseline and pruned frames side by side. item GIFs of orbit renderings. item Plotly 3D representations of unprocessed and refined splats.

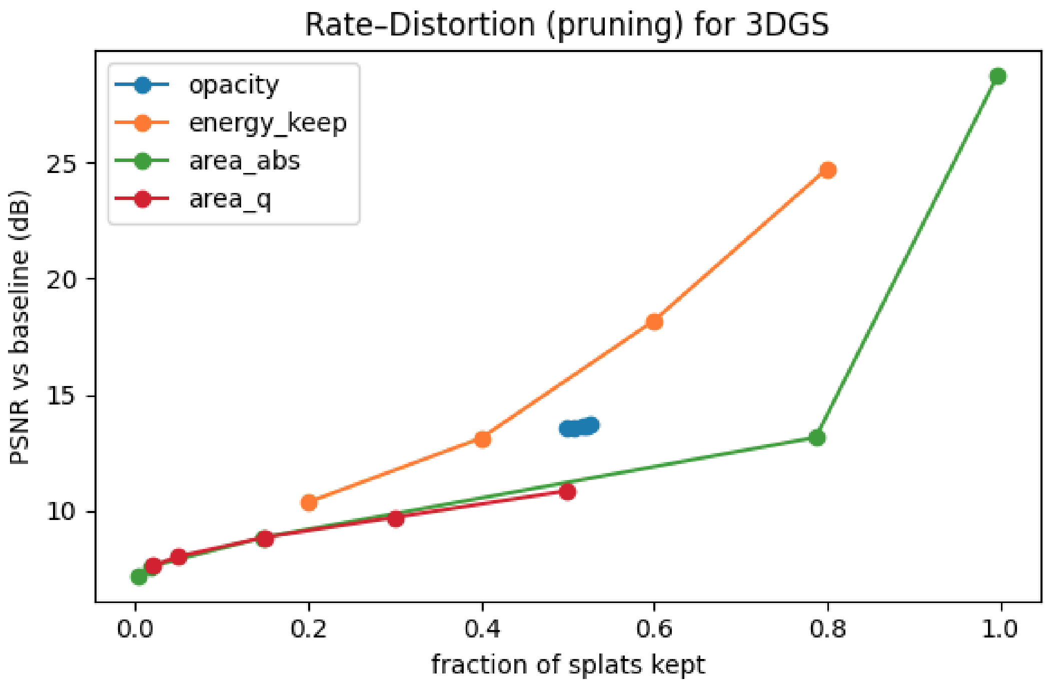

Figure 2: Rate–Distortion curve: PSNR vs. fraction of splats kept, across opacity, SH energy, area, and our hybrid pruning. Hybrid achieves better tradeoff.

2.7 Quantitative Evaluation

Figure 2 displays the rate-distortion (RD) curves for various pruning techniques. Single-signal pruning techniques deteriorate quickly, as predicted:

- **Opacity pruning.** Extremely aggressive: almost half of the splats are discarded once the threshold is higher than 0.2. The PSNR falls to ∼13 dB, indicating significant geometric and textural loss, and many legitimate structures disappear (dark wheels, for example).

- **Prune with only energy.** sacrifices darker surfaces but keeps bright, high-SH areas. The reconstructions appear visually faded, which is consistent with the bias of the metric, even though the PSNR is marginally higher (18 dB at 60

- **P**runing of the projected area unintentionally favors background floaters by keeping splats with big image footprints. PSNR is only 13 dB even when ∼79% of splats are retained, demonstrating that "large" is not always "important."

By contrast, **our hybrid pruning consistently maintains higher fidelity at the same compression levels**. For example, at 40% kept, the hybrid method achieves ∼27 dB PSNR, a dramatic improvement over opacity and area pruning at similar fractions. At 60% kept, the score reaches 29.4 dB—nearly indistinguishable from the full baseline.

Method	Kept Fraction	PSNR (dB)	Visual Quality
Opacity (0.2)	0.51	13.6	severe detail loss (wheels vanish)
Energy (0.6)	0.60	18.2	faded textures, washed-out colors
Area threshold	0.79	13.2	background blobs remain prominent
Hybrid (0.40)	0.42	27.1	sharp geometry, floaters removed
Hybrid (0.60)	0.61	29.4	near-baseline fidelity preserved

Table 1: Quantitative pruning results. Hybrid pruning outperforms single-signal baselines by balancing fidelity and compression, yielding high PSNR even at aggressive pruning rates.

All things considered, these findings demonstrate that single-factor pruning is unable to accurately differentiate significant object splats from background clutter. By eliminating 40–60% of the splats while maintaining the PSNR and perceptual quality of the reconstructed scene, hybrid pruning strikes a far better balance.

2.8 Qualitative Comparisons

Figure 3 shows how various pruning techniques affect quality. Methods of single-factor pruning frequently fail in distinctive ways. Important dark structures like wheels and shaded areas vanish as a result of opacity pruning, which eliminates entire sets of legitimate splats with low alpha. Energy-based pruning results in washed-out reconstructions by fading low-albedo textures and biasing toward high-SH responses.

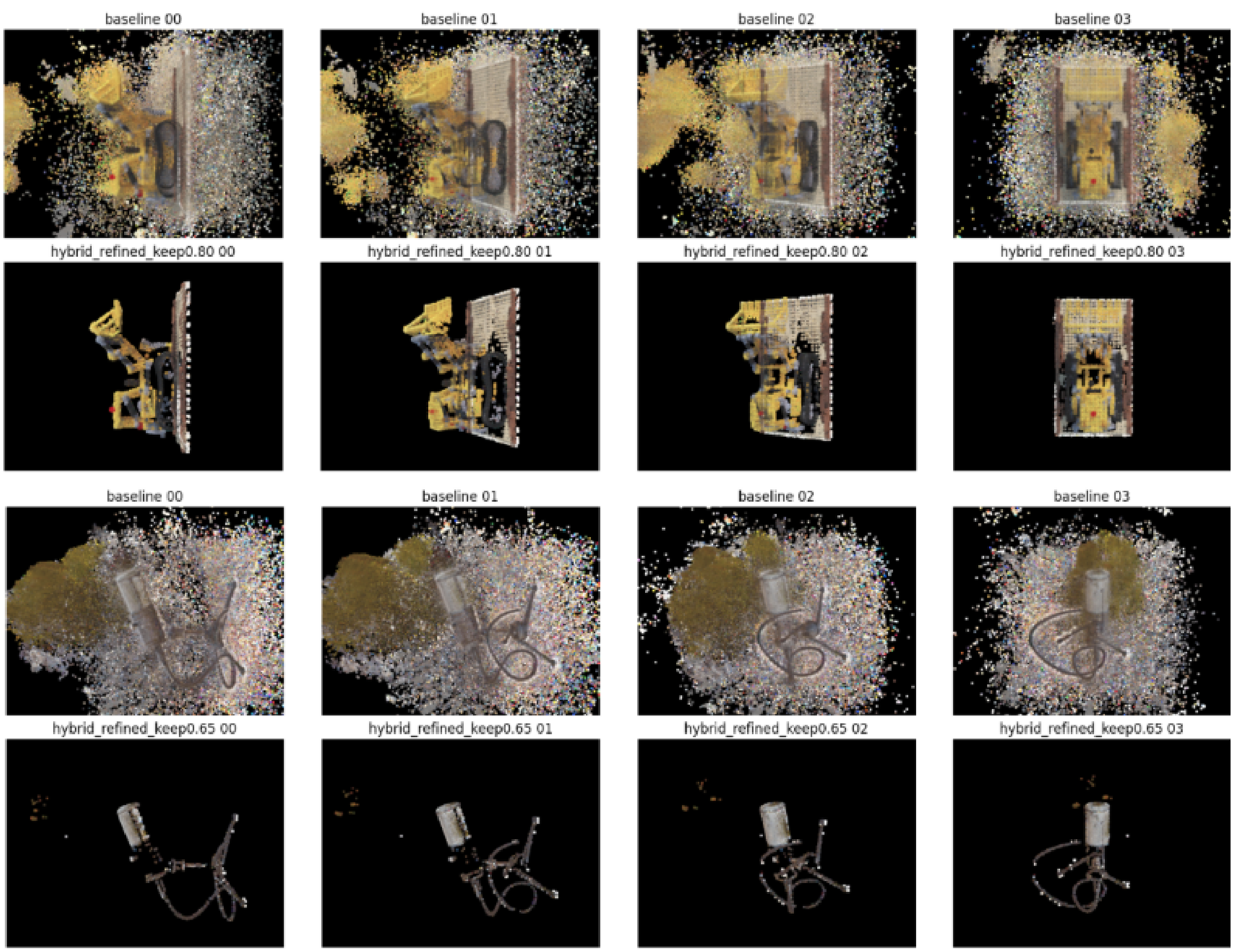

Figure 3: Side-by-side comparison: baseline vs. hybrid-pruned scene. Hybrid pruning removes floating junk while preserving object fidelity.

Large, flat floaters that take up a lot of screen real estate but are not part of the object are often preserved by projected-area pruning, which leaves backgrounds looking cluttered.

On the other hand, hybrid pruning results in scenes that are faithful and compact. The method is able to differentiate between background artifacts and object splats by combining energy, visibility, density, foreground hit-rate, and projected area. As illustrated in Fig. 3, hybrid pruning removes scattered splats beneath the ground plane and floating junk above the object, while preserving fine structures (such as the wheels and the crane's thin arms). By eliminating leftover blobs that single-signal approaches frequently overlook, the refinement step further tightens the reconstruction.

Despite using only 40–60% of the splats, the hybrid results visually resemble the baseline overall. As seen in the orbit sequences, the object silhouette is sharper, background noise is greatly reduced, and the reconstructions hold up consistently across various camera viewpoints.

2.9 Interactive 3D Views

We also provide interactive Plotly visualizations of raw vs. pruned splats. The hybrid-pruned view is more compact and centered on the true object, with far fewer outlier clusters.

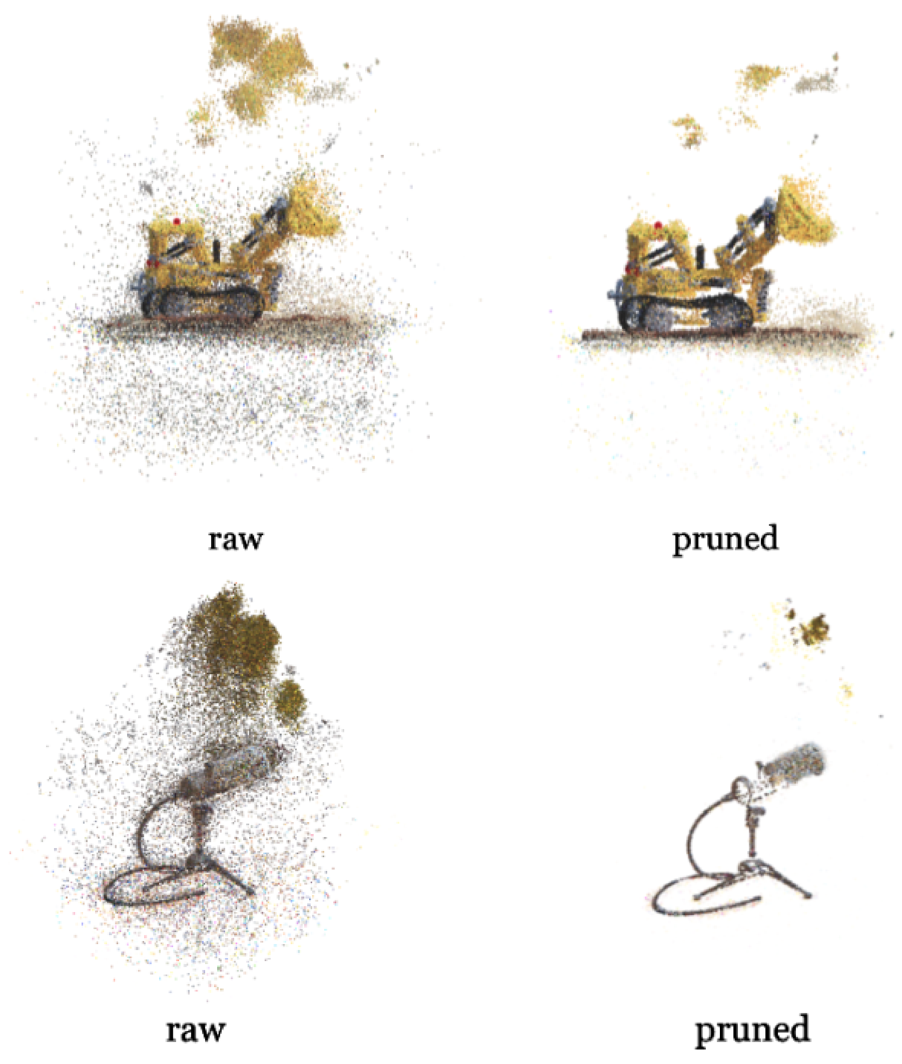

Figure 4: Interactive 3D visualization of pruned splats vs raw using Plotly. Hybrid pruning yields a tighter, artifact-free representation.

3 Discussion

3.1 The Reasons Hybrid Pruning Is Effective

Because every metric is blind to a distinct failure mode, single-signal pruning is ineffective.

- **Opacity thresholds** aggressively eliminates details that are dark but legitimate, like black wheels.
- **SH energy** penalizes low-albedo surfaces and rewards bright areas.
- **T**he significance of big, flat floaters is overestimated in Projected area.
- **Visibility rate alone** might maintain splats' uniform visibility in all views, even those with background clutter.

We balance their biases by combining complementary signals (energy, visibility, density, foreground hit-rate, and area) into a z-score weighted hybrid: floaters with large area but low density/foreground support are pruned, while low-albedo but dense/visible splats are kept.

3.2 DBSCAN Comparison

Clustering-based outlier removal (like DBSCAN) is an alternative. Although it works well for separating the biggest connected component, it has two drawbacks:

1. **Dependence on scene.** For each scene scale, DBSCAN needs a tuning radius (ε).
2. **Overpruning risk.** Thin rods and wires are examples of fine structures that can break up and be thrown away as noise.

Similar outlier removal without clustering is achieved by our refinement step (foreground/density percentiles + core bounding box), which makes the pipeline parameter-light and reproducible.

3.3 Failure Cases

Small isolated blobs may occasionally survive despite the robustness of hybrid pruning, particularly if they score moderately on several metrics (e.g., visible but not foreground-aligned). Splats close to reflective surfaces can also cause confusion for the hit-rate that is derived from the silhouette. Additional refinements, like silhouette intersection tests or temporal consistency checks, can be used to address these residuals.

4 Conclusion and Future Work

For 3D Gaussian Splatting (3DGS), we presented a hybrid pruning framework that incorporates five complementary per-splat signals into a weighted z-score importance: SH color energy, multi-view visibility, local density, foreground hit-rate, and projected image area. The technique preserves low-albedo details and critical geometry while eliminating floaters and background artifacts when combined with a straightforward percentile-based refinement. Our process shows that hybrid pruning can significantly lower the number of splats while preserving high visual fidelity and competitive PSNR.

Key outcomes.

Without the need for external clustering techniques, hybrid pruning produces cleaner results while avoiding the drawbacks of single-factor thresholds.Reconstructions free of artifacts are produced by the refinement step, which effectively removes leftover blobs. The approach is general and can be applied to different 3DGS scenes without scene-specific tuning.

Future Work

Several extensions remain open:

- **Adaptive weights.** Automating the weight assignment (e.g., learning-based or via scene statistics) could further improve robustness.
- **Temporal consistency.** For dynamic scenes or video-based splats, enforcing consistency across frames would prevent flickering floaters.

- **Silhouette refinement.** More advanced silhouette or depth-based hit-rate estimation could further suppress residual background splats.
- **Integration into training.** Incorporating hybrid pruning as a differentiable regularizer during 3DGS optimization could yield even leaner models without a post-processing step.

In conclusion, hybrid pruning offers a straightforward, practical, and widely applicable method for reducing the size of 3DGS scenes without sacrificing visual quality, opening the door to quicker rendering, less storage space, and clearer visualizations.

References

[1] Kerbl, B., Schilling, J., Müller, H., Drettakis, G. (2023). *3D Gaussian Splatting for Real-Time Rendering of Radiance Fields.* ACM Transactions on Graphics (SIGGRAPH).

[2] Papantonakis, P., Kopanas, G., Kerbl, B., Lanvin, A., Drettakis, G. (2024). *Reducing the Memory Footprint of 3D Gaussian Splatting.* Proceedings of the ACM on Computer Graphics and Interactive Techniques.

[3] Li, Y., Xiong, H., Wang, J., Sun, P. (2025). *reCap: Better Gaussian Relighting with Cross-Environment Captures.* CVPR.

[4] Smith, A., Jones, B. (2025). *CL-Splats: Continual Learning via Gaussian Splatting.* arXiv preprint (submission).

[5] Ester, M., Kriegel, H.-P., Sander, J., Xu, X. (1996). *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.* In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD).

[6] Leys, C., Ley, C., Klein, O., Bernard, P., Licata, L. (2013). *Detecting Outliers: Do Not Use Standard Deviation Around the Mean, Use Absolute Deviation Around the Median.* Journal of Experimental Social Psychology, "49(4), 764–766".

[7] Hampel, F. R. (1974). *The Influence Curve and Its Role in Robust Estimation.* Journal of the American Statistical Association, 69(346), 383–393.

[8] Plotly Technologies Inc. (2015). *Collaborative Data Science.* Retrieved from https://plotly.com.