

Light-SQ: Structure-aware Shape Abstraction with Superquadrics for Generated Meshes

YUHAN WANG, S-Lab, Nanyang Technological University, Singapore

WEIKAI CHEN*, ZEYU HU, RUNZE ZHANG, YINGDA YIN, RUOYU WU, KEYANG LUO, SHENGJU QIAN, YIYAN MA, HONGYI LI, YUAN GAO, YUHUAN ZHOU, HAO LUO, WAN WANG, XIAOBIN SHEN, ZHAOWEI LI, KUIXIN ZHU, CHUANLANG HONG, YUEYUE WANG, LIJIE FENG, and XIN WANG*

CHEN CHANGE LOY, S-Lab, Nanyang Technological University, Singapore

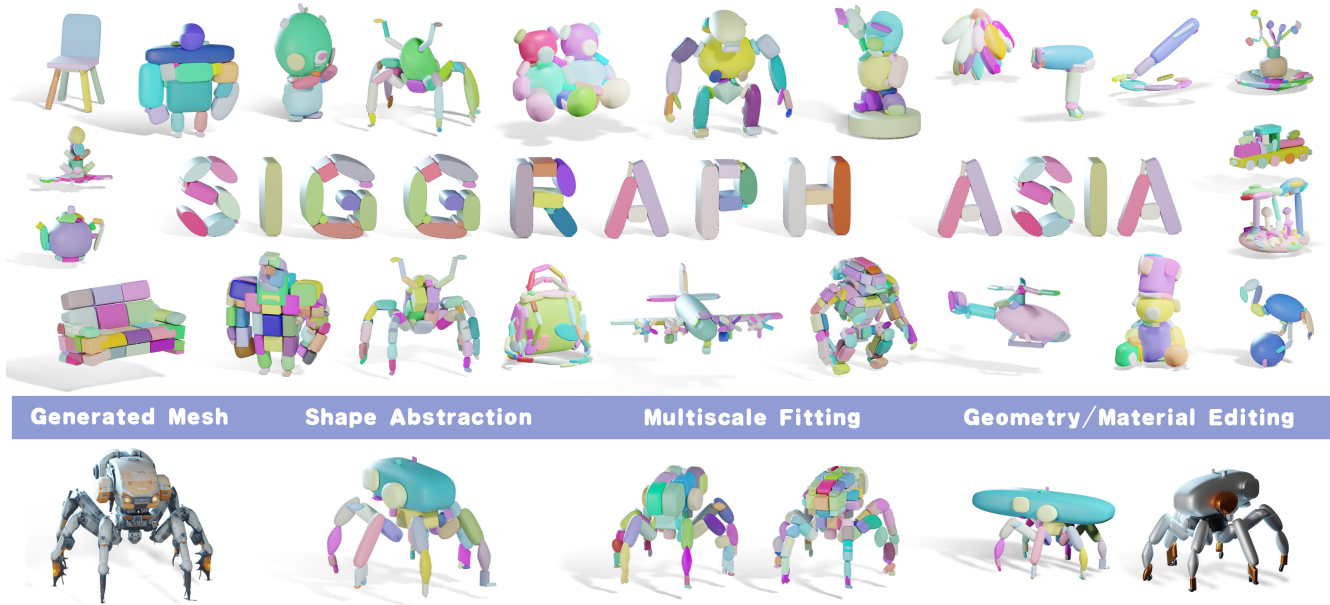


Fig. 1. We introduce **Light-SQ**, a superquadrics-based shape abstraction method tailored for generated 3D meshes and well-suited for UGC scenarios.

In user-generated-content (UGC) applications, non-expert users often rely on image-to-3D generative models to create 3D assets. In this context, primitive-based shape abstraction offers a promising solution for UGC scenarios by compressing high-resolution meshes into compact, editable representations. Towards this end, effective shape abstraction must therefore be structure-aware, characterized by low overlap between primitives, part-aware alignment, and primitive compactness. We present Light-SQ, a novel superquadric-based optimization framework that explicitly emphasizes structure-awareness from three aspects. (a) We introduce SDF carving

to iteratively update the target signed distance field, discouraging overlap between primitives. (b) We propose a block-regrow-fill strategy guided by structure-aware volumetric decomposition, enabling structural partitioning to drive primitive placement. (c) We implement adaptive residual pruning based on SDF update history to suppress over-segmentation and ensure compact results. In addition, Light-SQ supports multiscale fitting, enabling localized refinement to preserve fine geometric details. To evaluate our method, we introduce 3DGen-Prim, a benchmark extending 3DGen-Bench with new metrics for both reconstruction quality and primitive-level editability. Extensive experiments demonstrate that Light-SQ enables efficient, high-fidelity, and editable shape abstraction with superquadrics for complex generated geometry, advancing the feasibility of 3D UGC creation. *Project Page: <https://johann.wang/Light-SQ/>*.

CCS Concepts: • **Computing methodologies** → **Shape analysis**.

Additional Key Words and Phrases: Superquadrics, Shape Abstraction, Primitive Decomposition, Signed Distance Field

ACM Reference Format:

Yuhan Wang, Weikai Chen, Zeyu Hu, Runze Zhang, Yingda Yin, Ruoyu Wu, Keyang Luo, Shengju Qian, Yiyang Ma, Hongyi Li, Yuan Gao, Yuhuan Zhou, Hao Luo, Wan Wang, Xiaobin Shen, Zhaowei Li, Kuixin Zhu, Chuanlang Hong, Yueyue Wang, Lijie Feng, Xin Wang, and Chen Change Loy. 2025.

*Corresponding authors

Authors' Contact Information: Yuhan Wang, S-Lab, Nanyang Technological University, Singapore; Weikai Chen; Zeyu Hu; Runze Zhang; Yingda Yin; Ruoyu Wu; Keyang Luo; Shengju Qian; Yiyang Ma; Hongyi Li; Yuan Gao; Yuhuan Zhou; Hao Luo; Wan Wang; Xiaobin Shen; Zhaowei Li; Kuixin Zhu; Chuanlang Hong; Yueyue Wang; Lijie Feng; Xin Wang, LIGHTSPEED; Chen Change Loy, S-Lab, Nanyang Technological University, Singapore.



This work is licensed under a Creative Commons Attribution 4.0 International License. *SA Conference Papers '25, Hong Kong, Hong Kong*
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2137-3/2025/12
<https://doi.org/10.1145/3757377.3763835>

Light-SQ: Structure-aware Shape Abstraction with Superquadrics for Generated Meshes. In *SIGGRAPH Asia 2025 Conference Papers (SA Conference Papers '25)*, December 15–18, 2025, Hong Kong, Hong Kong. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3757377.3763835>

1 Introduction

Customizable 3D asset creation is a fundamental requirement in user-generated-content (UGC) platforms, yet remains largely inaccessible to non-expert creators. Recent advances in image-to-3D generation [Chen et al. 2025; Lai et al. 2025; Li et al. 2025b; Yang et al. 2024b; Zhang et al. 2024; Zhao et al. 2025] have significantly lowered the entry barrier, allowing users to generate high-quality 3D meshes from simple image prompts. However, the resulting meshes are typically over-tessellated, structurally unorganized, and difficult to edit, posing critical challenges in downstream applications such as animation, rigging, and interactive content creation.

To address these limitations, primitive-based shape abstraction has emerged as a promising solution. By converting bulky triangle meshes into a compact set of analytic primitives, this approach offers two key benefits: 1) substantial storage reduction, from megabytes to kilobytes; and 2) improved editability, as each primitive serves as an intuitive manipulation handle. Our objective is to extend these benefits to generated 3D assets, transforming coarse, unstructured meshes into compact, editable primitive representations, for a robust image-to-primitives pipeline.

For UGC scenarios, an ideal primitive abstraction should satisfy two high-level goals: *fidelity* and *editability*. While “fidelity” refers to the preservation of salient shape structures, “editability” relies on satisfying three structure-aware criteria: (a) *Low overlap*. Each primitive should occupy a distinct, low-overlapping spatial region; (b) *Structure awareness*. Primitives should conform to coherent volumetric partitions, avoiding cross-structure primitive placement; (c) *Compactness*. Excessive segments would hinder downstream tasks such as collision detection, rigging, and editing. Fig. 2 shows results that do not meet these criteria, along with an ideal abstraction.

Despite extensive research, existing abstraction techniques struggle with the irregularities and noise characteristic of generative geometry. Learning-based approaches [Li et al. 2024a; Paschalidou et al. 2020; Smirnov et al. 2020; Yang and Chen 2021; Ye et al. 2025] directly regress primitive parameters from the inputs, but typically generalize poorly outside the training domain, which is often limited to clean, curated datasets such as ShapeNet. Rule-based pipelines [Li et al. 2025a, 2024b; Lin et al. 2020] first segment the shape and then fit primitives to each segment. However, segmentation derived from generative meshes tend to be noisy – either due to being trained on artist-annotated assets with clean boundaries or adapted from 2D image priors lacking 3D awareness. Furthermore, even semantically meaningful segments may exhibit geometric irregularities that prevent accurate approximation by a single primitive, ultimately resulting in inconsistent or fragmented decomposition. Optimization-based techniques [Fedele et al. 2025; Liu et al. 2022, 2023; Monnier et al. 2023; Wu et al. 2022] avoid these pitfalls by iteratively fitting superquadrics to the geometry. While expressive, they often produce highly overlapped primitives, which compromises editability by eliminating clear ownership over spatial regions.

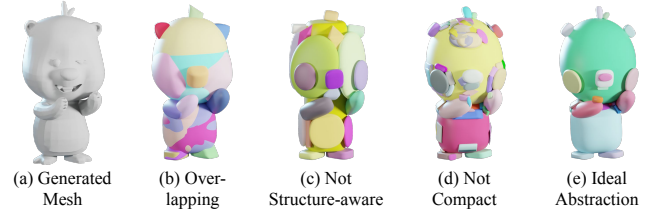


Fig. 2. Given the input mesh, (b), (c), and (d) exhibit, respectively, excessive primitive overlap, structural inconsistency, and excessive fragmentation. In (c), the centric primitive spans the head and the body.

To address these issues, we present Light-SQ, a structure-aware superquadrics decomposition framework tailored for generated geometry in UGC settings. We leverage superquadric for their compact representation and closed-form expressiveness over novel objects. Inspired by Marching Primitives [Liu et al. 2023], we adopt truncated signed distance fields (TSDFs) as the fitting target to exploit rich geometric information. To ensure structure awareness, we introduce three key components. First, we propose SDF carving, a volumetric exclusion mechanism that imposes explicit penalty over overlapping primitives to encourage spatial separation. Second, we present a structure-aware alignment framework that leverages geometric-feature-driven volumetric partitioning and a block-regrow-fill strategy to guide superquadric fitting to conform to structural boundaries. Third, to improve compactness, we track the SDF update history to classify residual primitives based on their geometric significance. Primitives below category-specific thresholds are discarded, preserving salient details while minimizing redundancy.

In addition, Light-SQ supports *multiscale fitting*, allowing recursive refinement of coarse primitives to capture fine-grained geometry. This allows flexible balancing between abstraction detail and reconstruction quality, which prior methods lack. Apart from the Light-SQ framework, we present 3DGen-Prim, a benchmarking suite for evaluating primitive decomposition on generated 3D geometry. Unlike existing protocols based on ShapeNet [Chang et al. 2015], which features clean, curated meshes, 3DGen-Prim extends 3DGen-Bench [Zhang et al. 2025] with outputs from recent image-to-3D methods [Li et al. 2025b; Zhao et al. 2025], and introduces metrics that assess both reconstruction fidelity and structural editability.

To summarize, our main contributions are as follows:

- We present Light-SQ, a superquadric decomposition algorithm that achieves high fidelity and structural awareness on generated geometry for the UGC scenarios.
- We introduce key algorithmic components, including SDF carving, structure-aware alignment, and adaptive residual pruning, which collectively promote structure-aware alignment and representation compactness.
- Extensive experiments on 3DGen-Prim dataset demonstrate our significant advantages in fitting fidelity and editability.

2 Related Work

Optimization-based Shape Abstraction. Optimization-based approaches reconstruct 3D shapes by optimizing primitive parameters, typically superquadrics. State-of-the-art efforts [Liu et al. 2022; Wu et al. 2022] began by fitting primitives to point clouds. They treat the

point cloud as an observation sampled from a probabilistic model upon superquadrics, and apply Maximum Likelihood Estimation (MLE) [Liu et al. 2022] and Non-parametric Bayesian Inference [Wu et al. 2022] respectively to solve their parameters. Subsequent works explored fitting on other modalities. Marching-Primitives [Liu et al. 2023] adopts the truncated signed distance field (TSDF) to avoid geometric ambiguities [Yang and Chen 2021]. Monnier et al. [2023] incorporates superquadrics into a NeRF-like [Mildenhall et al. 2020] reconstruction pipeline to fit multiview images.

Although existing methods have made significant progress in the accuracy of fitting, they lacked efforts to meet “structure-aware requirements”, as also noted in the concurrent work [Ye et al. 2025]. Light-SQ is the first superquadrics optimization algorithm that formally defines and explicitly emphasizes structural awareness.

Learning-based Shape Abstraction. Learning-based approaches aim to employ a single neural network to predict the corresponding primitive representation from 2D or 3D inputs. Tulsiani et al. [2017] pioneered using a convolutional neural network for cuboid representation prediction with volume input. 3D-PRNN [Zou et al. 2017] takes a depth map as input and uses recurrent neural networks. Paschalidou et al. [2019] extends the primitive type to superquadrics, and introduces point cloud reconstruction as the supervision. Subsequent works investigated the abstraction of other inputs such as distance fields [Smirnov et al. 2020] and the single-view image [Niu et al. 2018; Paschalidou et al. 2020], and explored the integration of frontier techniques, including unsupervised point cloud segmentation [Yang and Chen 2021] and autoregressive transformers [Li et al. 2024a]. Despite promising results in their data, the training source of these methods is typically confined to ShapeNet [Chang et al. 2015] (in some cases, only to a single subclass), resulting in poor generalization of the generated geometry.

Recently, SuperDec [Fedele et al. 2025] improved feed-forward predictions with an optimization refinement step. In parallel, PrimitiveAnything [Ye et al. 2025] constructed an unprecedented 120K-scale shape abstraction dataset and trained an autoregressive network on point cloud inputs, achieving remarkable generalization on generated geometry. However, it falls short of the UGC standards outlined in Sec. 1 and does not support multiscale decomposition.

Rule-based Shape Abstraction. Rule-based methods typically bind each primitive with a segmentation part. SEG-MAT [Lin et al. 2020] uses the medial axis transform to partition a mesh into skeletal regions and represent each with a cuboid. LMP [Li et al. 2024b] introduces Shared Latent Membership, where deformable superquadrics serve as both segmentation priors and abstraction targets. More recently, AISSR [Li et al. 2025a] aligns instance and semantic sparse representations to derive repeatable primitive templates. However, when applied to generated geometry, these segmentation schemes often fail to achieve high-fidelity shape fitting.

Approximate Convex Decomposition. Approximate convex decomposition (ACD) splits 3D geometry into a minimal set of pseudo-convex parts. CoACD [Wei et al. 2022] introduces a split-and-merge decomposition scheme. Andrews [2024] proposed navigation-driven ACD, which allows more flexible decomposition requirements. In this paper, we developed an adaptive ACD algorithm tailored for structure-aware shape abstraction.

3 Method

In this section, we first introduce how to fit a signed distance field with superquadrics (Sec. 3.1). Next, we describe how Light-SQ achieves structure-aware superquadric fitting from three perspectives: low-overlap (Sec. 3.2), structure-aware alignment (Sec. 3.3), and compactness (Sec. 3.4). Fig. 3-(a) shows our fitting pipeline. Finally, we demonstrate the multiscale fitting capability enabled by our structure-aware framework (Sec. 3.5).

3.1 Preliminaries: Superquadrics Fitting on TSDF

A signed distance function (SDF) returns the distance from a given point to the closest surface, which is negative when it is inside:

$$\phi(\mathbf{x}) = SDF(\mathbf{x}) = s : \mathbf{x} \in \mathbb{R}^3, s \in \mathbb{R} \quad (1)$$

Existing optimization-based approaches [Liu et al. 2023] use the truncated SDF (TSDF) $\phi_\tau(\mathbf{x})$ as the fitting target:

$$\phi_\tau(\mathbf{x}) = TSDF(\mathbf{x}; \tau) = \text{clamp}(SDF(\mathbf{x}), -\tau, +\tau) \quad (2)$$

The truncation parameter τ is typically set to the voxel edge length. Consequently, this field behaves almost like an **occupancy grid**, only providing local curvature information in regions adjacent to the surface. We find this setting reasonable, as it prevents overpenalty in the boundary region. We visualize it in Fig. 3-(b).

An axis-aligned superquadric can be represented by shape parameters $\epsilon_1, \epsilon_2 \in [0, 2]$ and scale parameters $a_x, a_y, a_z \in \mathbb{R}^+$ as

$$f(\mathbf{x}) = \left(\left(\frac{x}{a_x} \right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_y} \right)^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}} + \left(\frac{z}{a_z} \right)^{\frac{2}{\epsilon_1}} = 1, \quad (3)$$

where an arbitrarily positioned superquadric requires another three euler angles $\mathbf{e} \in \mathbb{R}^3$ and a translation vector $\mathbf{t} \in \mathbb{R}^3$ to describe the Euclidean transformation g . Given a superquadric parameterized by $\theta = [\epsilon_1, \epsilon_2, a_x, a_y, a_z, g(\mathbf{e}, \mathbf{t})] \in \mathbb{R}^{11}$ and a 3D point $\mathbf{x}$, we can use the signed radial distance function to approximate its SDF:

$$\phi_\theta(\mathbf{x}) = \left(1 - f^{-\frac{\epsilon_1}{2}}(g^{-1} \circ \mathbf{x}) \right) \|g^{-1} \circ \mathbf{x}\|_2 \quad (4)$$

Superquadric fitting is performed on each SDF connected component. After a new superquadric is fitted, the connected component is usually subdivided into several smaller segments. To fit this new superquadric, existing methods start from an initial superquadric θ , and, at each iteration, optimize its parameters by fitting to the TSDF field within its local neighborhood $V_a(\theta)$, then repeat this process with the updated superquadric parameter:

$$\hat{\theta} = \arg \min_{\theta} \sum_{\mathbf{x} \in V_a(\theta)} \lambda_{\mathbf{x}} \|\phi_{\hat{\theta}, \tau}(\mathbf{x}) - \phi_\tau(\mathbf{x})\|_2^2 \quad (5)$$

Here $\lambda_{\mathbf{x}}$ controls the penalty weight applied to each voxel.

3.2 Low-overlapping Superquadrics Fitting

In previous works [Liu et al. 2023], after a superquadric is fitted, the voxels inside it will no longer be penalized during the subsequent fitting. This allows superquadrics to overlap with each other, providing more flexibility for filling the remained regions. However, when applied to generated geometry, the complex curvature distribution causes this strategy to produce a large number of highly overlapping

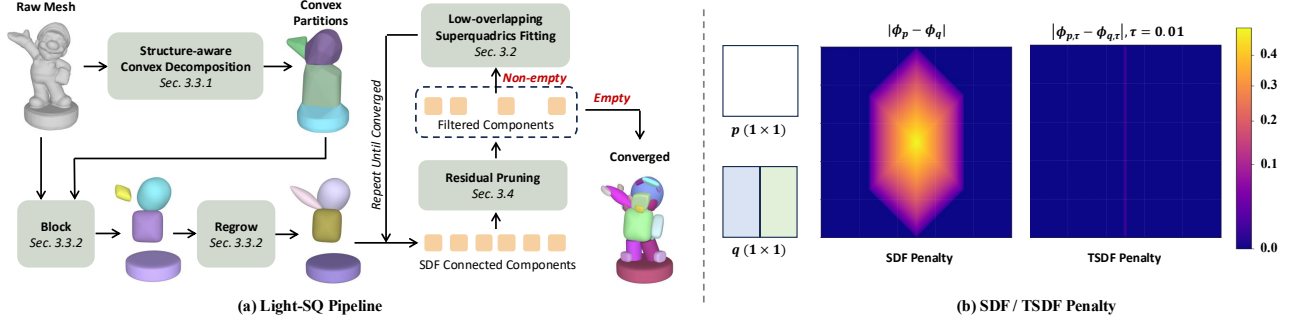


Fig. 3. **Overview.** (a) *Light-SQ pipeline.* Our method first employs block and regrow operations to satisfy the structural guidance of convex partitions, and then iteratively fills the SDF connected components until convergence, yielding a high-fidelity shape abstraction. (b) *Visualization of SDF and TDSF difference.* If a closed shape (p) is abstracted in the form of two primitives (q), around the boundary between them, the SDF values deviate significantly from the ground truth, incurring an over-penalty that hinders the necessary decomposition. This issue is mitigated in the TDSF field.

superquadrics, which not only violates structural compactness but also hinders downstream editing and interaction in UGC scenarios.

We revisited the optimization process of a single superquadric. Previous works derive the penalty weights through maximum likelihood estimation. We find that the computational framework of λ_x can be relaxed to the following form:

$$\lambda_x = \frac{P(\phi_\tau(\mathbf{x})|\phi_{\theta,\tau}(\mathbf{x}))}{\underbrace{(\phi_\tau(\mathbf{x}) < 0) \cdot C \cdot (1 - w)/w}_{\text{inside-voxel decay term}} + \underbrace{P(\phi_\tau(\mathbf{x})|\phi_{\theta,\tau}(\mathbf{x}))}_{\text{TDSF matching term}}} \quad (6)$$

Here w denotes the prior probability of any voxel covered by the current superquadric, which can be approximated as $1/\tilde{N}$ where $\tilde{N}$ is the expected number of superquadrics. C is a weighting constant, and $P(\phi_\tau(\mathbf{x})|\phi_{\theta,\tau}(\mathbf{x}))$ follows a normal distribution. The TDSF matching term ranges between 0 and 1, while the inside-voxel decay term is much larger. The motivation of this framework is to assign a much higher penalty to exterior voxels than to interior ones, thereby preventing a superquadric from expanding beyond the surface while still accommodating partial fitting within the interior regions.

Based on this property, we found that we can fulfill the low overlap requirement by iteratively updating the target SDF field. We convert the interior region of each fitted superquadric into the target shape’s exterior, thereby preventing subsequent superquadrics from encroaching on previously fitted regions. We call this operation *SDF carving*, denoted by $\phi \setminus \phi_\theta$. Specifically, $\phi \setminus \phi_\theta$ updates the original SDF field of the target shape as:

$$\phi(\mathbf{x}) = \begin{cases} -\phi_\theta(\mathbf{x}), & \phi(\mathbf{x}) < 0 \wedge \phi_\theta(\mathbf{x}) < 0, \\ \max(-\phi_\theta(\mathbf{x}), \phi(\mathbf{x})), & \phi(\mathbf{x}) < 0 \wedge \phi_\theta(\mathbf{x}) > 0, \\ \phi(\mathbf{x}), & \phi(\mathbf{x}) > 0 \end{cases} \quad (7)$$

SDF carving does not alter the SDF values of voxels outside the surface. The SDF field within the superquadric can be directly replaced, while those still inside the surface after carving requires a comparison and update. After this, we re-clamp $\phi(\mathbf{x})$ as Eq. (2) to obtain the updated TDSF field $\phi_\tau(\mathbf{x})$.

3.3 Structure-Aware Alignment

Compared to low overlap, achieving structure-aware alignment is substantially more challenging due to two primary issues. First,

“structure awareness” lacks a precise and consistent definition. A natural approach is to define semantics through segmentation; however, most existing methods only produce surface-level labels from rendered images [Liu et al. 2025; Tang et al. 2024; Yang et al. 2024a], which do not translate meaningfully to voxels within the volume. In contrast, direct volumetric segmentation approaches [Liu et al. 2025; Yang et al. 2025] are much less common and often suffer from unstable results. Second, even with a plausible structural segmentation, how to effectively leverage it for shape abstraction remains unclear. Semantic regions identified by 3D segmentation models may not align well with the geometric assumptions of primitive fitting – it is not guaranteed that each semantic part can be faithfully approximated by a single superquadric. This fundamental misalignment is also a key limitation for many rule-based abstraction pipelines.

Instead of relying on semantic learning, we adopt geometrical analysis to provide structural guidance. Our approach first introduces a structure-aware volumetric decomposition, followed by a three-phase convex segment-guided fitting strategy.

3.3.1 Structure-Aware Convex Decomposition. To enable structure-aware volumetric decomposition, we build upon CoACD [Wei et al. 2022], which formulates decomposition as a sequential decision process via MCTS. While effective, CoACD lacks structural guidance in plane sampling and relies on heuristic merging, often resulting in misaligned cuts and inconsistent part grouping. We address these issues by introducing structure-aware splitting and adaptive merging strategies to enhance structure alignment and geometric fidelity.

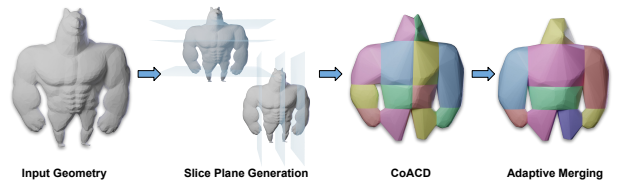


Fig. 4. Structure-aware convex decomposition pipeline.

Structure-Aware Slice Plane Generation. Instead of randomly sampling splitting planes, we precompute a fixed set of geometrically meaningful candidate planes using SDF-based volumetric analysis. Our method analyzes cross-sectional area variation and surface connectivity to select axis-aligned planes that are more likely to align with intrinsic geometric structures. In particular, we evaluate

each axis-aligned slice i based on two geometric cues: (1) the second-order difference of the cross-sectional area A_i , and (2) the variation in the number of connected components N_i on the slice.

The per-slice area $A_i = \sum_{j,k} \mathbf{1}[\text{SDF}(i, j, k) < 0]$, where $\mathbf{1}[\cdot]$ returns 1 if the condition holds, and 0 otherwise. We define its second-order difference (window size 3) to capture abrupt area transitions:

$$M_i = \sum_{j=1}^3 (A_{i-j} + A_{i+j}) - 2(A_{i-1} + A_i + A_{i+1}) \quad (8)$$

Let N_i denote the number of connected components formed by interior voxels on slice i . We compute the discrete component variation $\Delta N_i = |N_i - N_{i-1}|$. The final structural saliency score is:

$$S_i = \alpha \cdot \tilde{M}_i + (1 - \alpha) \cdot \tilde{\Delta N}_i, \quad (9)$$

where $\tilde{M}_i$ and $\tilde{\Delta N}_i$ are normalized scores, and $\alpha \in [0, 1]$ balances area change and connectivity jumps. We then select the top- K scoring slice indices as candidate planes, with a minimum spacing constraint δ to ensure diversity.

Adaptive Merging based on Geometric Continuity. To ensure meaningful part decomposition and avoid over-segmentation and erroneous merging, we propose an adaptive merging strategy based on geometric continuity. This strategy jointly considers curvature similarity and volumetric alignment between adjacent convex parts.

To compute curvature continuity, we denote Γ as the shared interface between two convex parts C_1 and C_2 . For each point $p \in \Gamma$, let $H_1(p)$ and $H_2(p)$ represent the mean curvature evaluated on the respective sides. The curvature continuity score is defined as:

$$S_{\text{curv}}(C_1, C_2) = 1 - \frac{1}{|\Gamma|} \sum_{p \in \Gamma} \frac{|H_1(p) - H_2(p)|}{H_{\max}}, \quad (10)$$

where $H_{\max}$ is a normalization constant representing the maximum expected curvature difference. For volumetric alignment, we quantify volumetric overlap via a volumetric IoU score:

$$S_{\text{vol}}(C_1, C_2) = \frac{\text{Vol}(C_1) + \text{Vol}(C_2)}{\text{Vol}(\text{CH}(C_1 \cup C_2))}, \quad (11)$$

where $\text{CH}(C_1 \cup C_2)$ denotes the convex hull of the merged region. A higher S_{vol} indicates better volumetric consistency between the parts. The merging score is computed as a weighted combination:

$$S(C_1, C_2) = \beta \cdot S_{\text{curv}}(C_1, C_2) + \gamma \cdot S_{\text{vol}}(C_1, C_2), \quad (12)$$

where β and γ control the relative importance of curvature and volume. Two parts are merged if $S(C_1, C_2) > \tau_m$, τ_m is a threshold.

3.3.2 Block-Regrow-Fill. Our structurally guided shape abstraction starts with a key observation that each volumetric partition is an occupancy grid – the voxels with one label as interior and all others as exterior. This yields the 3D partition for that label, and we can compute its corresponding SDF field.

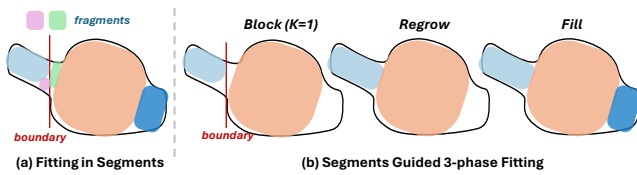


Fig. 5. A 2D illustration of direct fitting in segments and “block-regrow-fill”. The gaps around the boundary are filled through the regrow-stage.

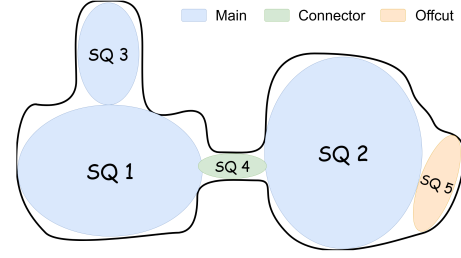


Fig. 6. A 2D illustration of superquadric classification used in adaptive residual pruning. SQ 1–2 are fitted during the ‘block’ and ‘regrow’ stages, while SQ 3–5 are fitted during the ‘fill’ stage.

After extracting the shape for each region, a straightforward strategy is to fit superquadrics independently to each partition. Although this achieves structure-aware alignment, the axis-aligned cutting planes may form unnatural boundaries for superquadrics. More specifically, the algorithm will generate cluttered and meaningless superquadrics fragments along these boundaries to fill the shape, as shown in Fig. 5-(a), undermining the goal of shape abstraction. Therefore, we proposed a three-stage fitting strategy to provide greater flexibility at structural boundaries:

- 1. Block.** Fit at most K superquadrics to each segment.
- 2. Regrow.** Insert all fitted superquadrics $\theta_1, \dots, \theta_N$ into the original shape. For each superquadric θ_i , use it as the initialization and run a second optimization, to obtain the final state $\hat{\theta}_i$. The target SDF field is computed from a series of SDF carving:

$$\phi_{\text{target}} = \phi \setminus \phi_{\theta_1} \setminus \dots \setminus \phi_{\hat{\theta}_{i-1}} \setminus \phi_{\theta_{i+1}} \setminus \dots \setminus \phi_{\theta_N} \quad (13)$$

Ideally, superquadrics in other regions can form flexible boundaries and allow “regrow” to fill the gaps around them.

- 3. Fill.** Fill under-fitted regions with additional superquadrics.

The effectiveness of this algorithm depends on the proper choice of K , and whether the placeholder superquadrics can reliably guard the boundaries. Since the structural partitions from Sec. 3.3.1 exhibit strong convexity, we found that $K = 1$ already yields ideal results.

3.4 Adaptive Residual Pruning

To enable more effective editing of primitive-based models after generation, a compact representation is essential. For instance, in a humanoid model, editing the “head” becomes much easier if it is represented by a single superquadric rather than multiple.

Achieving such compactness hinges on deciding when to stop fitting superquadrics on the TSDF. In other words, we must determine whether to fit a new superquadric on each connected component. This is particularly important for generated meshes, which often lack the clean structural organization of artist-designed models. To preserve key geometric features while minimizing redundancy, small but meaningful components (e.g., the propeller of a plane, as shown in Fig. 1) should be retained, while large but structurally insignificant artifacts (e.g., residuals from fitting a superquadric to a slightly tilted sphere) should be discarded.

The challenge then turns to how we can classify these components. An intuitive idea is that the superquadrics fitted during the “block” and “regrow” stages should be retained, since they each correspond to a distinct and meaningful structural partition. Therefore,

Table 1. **Quantitative Comparison.** We highlight the best value in blue, and the second-best value in green. “Optim.”, “Learn.”, and “Rule” stand for optimization-based, learning-based, and rule-based approaches, respectively. We do not highlight the best $\bar{N}$ because its comparison only makes sense when two methods achieve similar fitting performance; otherwise, a small $\bar{N}$ simply indicates underfitting to the input geometry.

Method	Type	Hunyuan3D-2.0					TripoSG				
		CD ↓	EMD ↓	Voxel-IoU ↑	OR ↓	$\bar{N}$	CD ↓	EMD ↓	Voxel-IoU ↑	OR ↓	$\bar{N}$
EMS [Liu et al. 2022]	Optim.	0.2345	0.2036	0.466	1.524	3.44	0.2472	0.2168	0.436	1.540	3.63
Marching-Primitives [Liu et al. 2023]	Optim.	0.0396	0.0544	0.868	4.201	67.7	0.0403	0.0537	0.860	3.778	62.5
AISSR [Li et al. 2025a]	Rule	0.1128	0.0918	0.403	1.051	7.94	0.1132	0.0915	0.394	1.056	7.81
PrimitiveAnything [Ye et al. 2025]	Learn.	0.1366	0.0986	0.442	1.845	82.7	0.1299	0.0936	0.457	1.870	79.9
Light-SQ (Ours)	Optim.	0.0388	0.0531	0.861	1.015	61.0	0.0385	0.0538	0.864	1.016	62.1

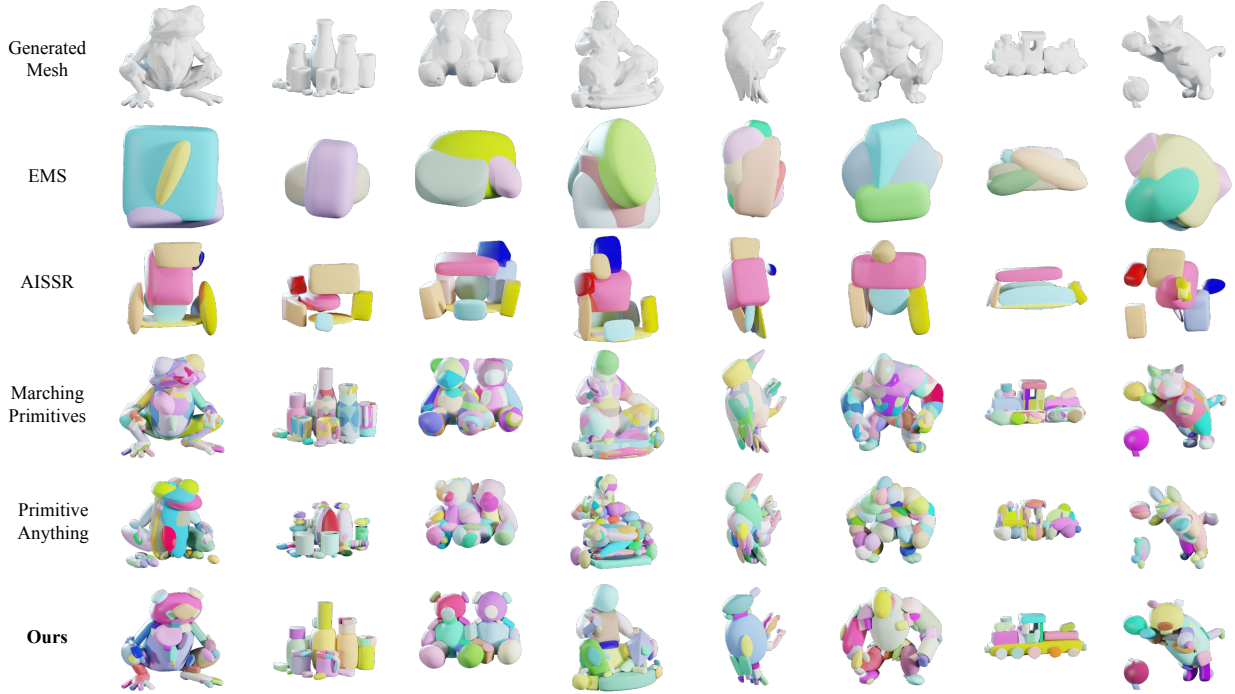


Fig. 8. **Qualitative Comparison.** Our method is the only one that simultaneously achieves faithful fitting to the original shape and strong editability.

an average of four superquadrics, predictably making the results difficult to edit. AISSR produces highly editable results, but it consistently fits only a small number of primitives, limiting its ability to represent complex shapes, which is a common shortcoming of rule-based methods. PrimitiveAnything, despite being trained on a 120K-sample dataset, still shows clear generalization issues on our generative data, offering no advantage in fitting quality or editability. Our method achieves fitting quality on par with Marching-Primitives, significantly outperforms other approaches, and simultaneously attains the lowest overlap—ensuring superior editability.

Qualitative Results. Fig. 8 presents the qualitative comparison. The generated mesh is randomly sampled from our 3DGen-Prim dataset. EMS and AISSR exhibit significant generalization challenges in fitting quality. Marching-Primitives achieves high reconstruction fidelity to the input shape, but the primitives overlap excessively and fail to reflect the underlying structure, resulting in poor editability. As a learning-based method, PrimitiveAnything demonstrates

Table 2. **Editability User Study.** We show the average ranking of each method in every metric. “PrimAny” stands for PrimitiveAnything [Ye et al. 2025]. “MPS” stands for Marching-Primitives [Liu et al. 2023].

Metric	PrimAny	MPS	Ours
Geometry Editability	2.43	2.50	1.07
Geometry Editing Efficiency	2.46	2.51	1.03
Texture Editability	2.32	2.61	1.06
Texture Editing Efficiency	2.32	2.62	1.06
Animation Friendliness	2.34	2.61	1.04

remarkable generalization, but its fitting quality is unstable, and on more complex generated shapes, it tends to do crude stacking, which harms editability. Our method achieves stable, high-precision fitting to the input geometry, while effectively ensuring low overlap and structure-aware alignment for editability.

Table 3. **Ablation on Low-overlapping.** Results are shown in OR (N). Here OR and N stand for overlap rate and average number of primitives, respectively. τ is the TSDF truncation.

	$w = 0.50$	$w = 0.10$	$w = 0.02$	$w = 0.01$
$C = 0.1$	1.019 (45.2)	1.018 (50.1)	1.016 (51.4)	1.016 (55.8)
$C = 1$	1.025 (51.1)	1.016 (59.5)	1.014 (61.7)	1.010 (79.4)
$C = 1/\tau = 50$	1.012 (75.2)	1.011 (91.0)	1.012 (113.8)	1.012 (117.8)

Table 4. User study results comparing our method with the ablated version and PrimitiveAnything, in terms of structure-aware alignment.

Method	Rank-1	Rank-2	Rank-3	Avg. Rank
PrimitiveAnything	6%	12%	82%	2.76
w/o. structure-aware	15%	74%	10%	1.96
Ours	75%	16%	9%	1.35

Table 5. **Block-Regrow-Fill Ablation.** Ablating the necessity of our 3-phase fitting strategy. “CH” stands for convex hull.

Method	CD ↓	EMD ↓	Voxel-IoU ↑	OR ↓	$\bar{N}$
One-per-CH	0.0613	0.0634	0.735	1.018	20.9
Ours	0.0388	0.0531	0.861	1.015	61.0

User Study on Editability. We conducted an extended user study for two baselines, Marching-Primitives and PrimitiveAnything, and our method on five metrics: Geometry Editability, Geometry Editing Efficiency, Texture Editability, Texture Editing Efficiency, and Animation Friendliness. In our user study, 25 participants received eight sets of uncurated shape abstraction results and were instructed to rank each set according to the five aforementioned criteria. The backgrounds of the participants included game development engineers and asset creation artists. They were provided with clear explanations and examples of each metric before evaluation, ensuring a consistent interpretation across raters. The explanations are detailed in *supplementary* Sec. B. Their ratings showed good agreement, suggesting that the metrics are interpretable and meaningful from a user’s perspective. Tab. 2 reports the average ranking of different methods on each metric. Our method exhibits a comprehensive, across-the-board advantage.

Efficiency Comparison. We benchmarked the speed of all methods on a single-GPU (14592 CUDA cores, 96GB VRAM) workstation with AMD EPYC 9K84 96-core CPU. Our method completes in 25.98 seconds per shape, including preprocessing, which is over 10× faster than Marching-Primitives (339.59s) and even slightly faster than PrimitiveAnything (29.10s), when accounting for its point-cloud scanning overhead. AISSR is very fast (0.06s), but its fitting fidelity is clearly insufficient (see Fig. 8 and Fig. 12). EMS provides no advantage in either speed (19.05s) or fitting quality.

5.2 Ablation Study and Discussions

Low overlapping. While our method does not guarantee zero overlap, it significantly reduces it via SDF carving and the penalty. Here, we provide an ablation on w and C in Eq. (6). We conducted a grid test for these two parameters on 51 uncurated Hunyuan3D-2.0 [Zhao et al. 2025] samples in our 3DGen-Prim dataset, comparing

Table 6. **Residual Pruning Ablation.** Pruning has almost no impact on fitting accuracy, yet it significantly reduces the number of primitives.

Hunyuan3D-2.0	CD ↓	EMD ↓	Voxel-IoU ↑	OR ↓	$\bar{N}$
w/o. pruning	0.0367	0.0533	0.874	1.018	95.4
Ours	0.0388	0.0531	0.861	1.015	61.0
TripoSg	CD ↓	EMD ↓	Voxel-IoU ↑	OR ↓	$\bar{N}$
w/o. pruning	0.0371	0.0526	0.872	1.018	92.2
Ours	0.0385	0.0538	0.864	1.016	62.1

the overlap rate and the average number of primitives. Tab. 3 shows that lower w and higher C reduce overlap rate (OR) but increase fragmentation. The paper setting is highlighted.

Structure-aware Alignment. We compare our method with an ablated version without structure-aware convex decomposition and “block-regrow-fill” strategy. Fig. 10 demonstrates how these two modules help our algorithm perceive and align with the input structure. Since this objective is difficult to quantify, we conducted a user study in which 28 participants were asked to rank 11 sets of randomly selected shape abstraction results. Results from PrimitiveAnything are also included. The participants were instructed to pick the best and worst primitive-decomposition results based on whether the structural partitioning was clear and reasonable. The comparisons are anonymized, and the order is randomized. The results shown in Tab. 4 further demonstrate the significant impact of our convex decomposition method and fitting strategy on enhancing structure-aware performance.

Block-Regrow-Fill. As mentioned in Sec. 3.3.2, given the structural guidance of a set of pseudo convex hulls, a straightforward approach is to fit superquadrics independently to each partition. To show the necessity of our three-phase fitting strategy, we compare our results with fitting a single superquadric to each pseudoconvex hull. As is shown in Tab. 5, since not all convex components can be well represented by a single superquadric, the ablated variant exhibits a pronounced drop in the fitting metrics.

Residual Pruning. Fig. 11 ablates the effect of adaptive residual pruning. The quantitative comparison in Tab. 6 further shows that, while the other metrics remain identical, pruning reduces the primitive count by an average of 30, thus reducing surface fragmentation.

Multiscale Fitting. Qualitative results are shown in Fig. 9. Multiscale fitting supports spatial upsampling and shape refinement, enabling the capture of finer local structure.

6 Conclusion

We introduce Light-SQ, a structure-aware superquadric decomposition framework for generated 3D geometry in UGC scenarios. Using SDF carving, structure-aware alignment, and adaptive residual pruning, Light-SQ yields abstractions that are simultaneously compact, editable, and faithful to the original geometry. We perform extensive experiments using our constructed 3DGen-Prim dataset to benchmark against prior methods, demonstrating consistent advantages in both fidelity and structural usability.

Acknowledgement. This work is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG2-PhD-2022-01-030).



Fig. 9. **Multiscale Fitting.** Based on the results of shape abstraction, multiscale fitting can upsample certain spatial regions and capture finer local curvature features in a block-stacking fashion. Users can refine the fitting of a specific region by selecting the corresponding superquadric.

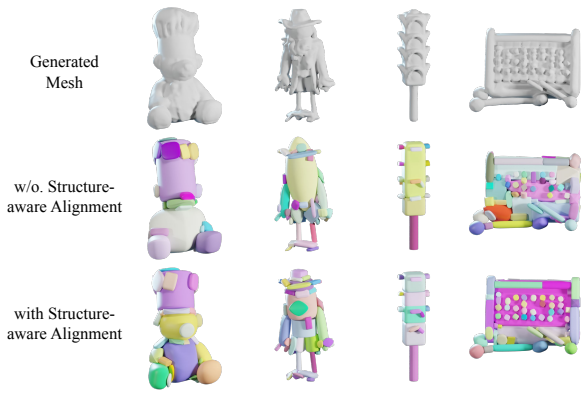


Fig. 10. **Ablation on Structure-aware Alignment.** Without structure-aware alignment, superquadrics fitted in early stages may span multiple structural partitions, resulting in degraded visual quality and editability.



Fig. 11. **Ablation on Residual Pruning.** Since optimization-based fitting algorithms always try to fully occupy the interior of the input geometry, they tend to generate many fragmented superquadrics around the surface. Adaptive residual pruning effectively removes these fragments.

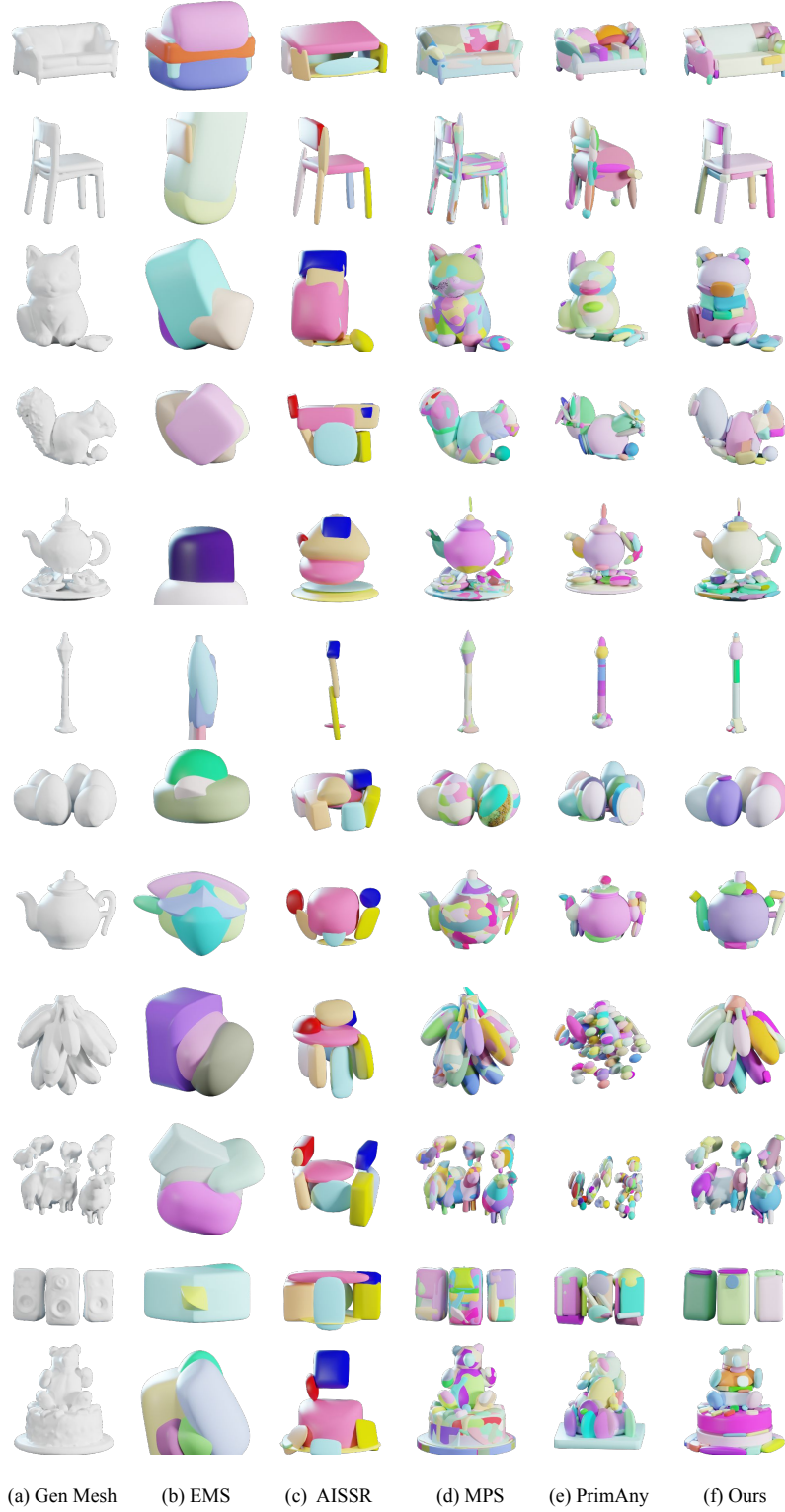


Fig. 12. Additional results on our 3DGen-Prim dataset, comparing (b) EMS [Liu et al. 2022], (c) AISSR [Li et al. 2025a], (d) Marching-Primitives [Liu et al. 2023], (e) PrimitiveAnything [Ye et al. 2025], and (f) Our Light-SQ.

References

- James Andrews. 2024. Navigation-Driven Approximate Convex Decomposition. In *ACM Trans. Graph. (SIGGRAPH)*.
- Angel X. Chang, Thomas Funkhouser, Leonidas J. Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. 2015. ShapeNet: An information-rich 3D model repository. *arXiv preprint arXiv:1512.03012* (2015).
- Jinnan Chen, Lingting Zhu, Zeyu Hu, Shengju Qian, Yugang Chen, Xin Wang, and Gim Hee Lee. 2025. MAR-3D: Progressive Masked Auto-regressor for High-Resolution 3D Generation. In *CVPR*.
- Elisabetta Fedele, Boyang Sun, Leonidas Guibas, Marc Pollefeys, and Francis Engelmann. 2025. SUPERDEC: 3D Scene Decomposition with Superquadric Primitives. *arXiv preprint arXiv:2504.00992* (2025).
- Zejiang Lai, Yunfei Zhao, Zibo Zhao, Haolin Liu, Fuyun Wang, Huiwen Shi, Xianghui Yang, Qinxian Lin, Jinwei Huang, Yuhong Liu, et al. 2025. Unleashing Vecset Diffusion Model for Fast Shape Generation. *arXiv preprint arXiv:2503.16302* (2025).
- Jiaxin Li, Hongxing Wang, Jiawei Tan, Zhilong Ou, and Junsong Yuan. 2025a. Aligning Instance-Semantic Sparse Representation towards Unsupervised Object Segmentation and Shape Abstraction with Repeatable Primitives. *IEEE. T. Vis. Comput. Gr.* (2025).
- Jiaxin Li, Hongxing Wang, Jiawei Tan, and Junsong Yuan. 2024b. Shared latent membership enables joint shape abstraction and segmentation with deformable superquadrics. *IEEE Trans. Image Process.* (2024).
- Songlin Li, Despoina Paschalidou, and Leonidas Guibas. 2024a. PASTA: Controllable Part-Aware Shape Generation with Autoregressive Transformers. *arXiv preprint arXiv:2407.13677* (2024).
- Yangguang Li, Zi-Xin Zou, Zexiang Liu, Dehu Wang, Yuan Liang, Zhipeng Yu, Xingchao Liu, Yuan-Chen Guo, Ding Liang, Wanli Ouyang, et al. 2025b. TripoSG: High-Fidelity 3D Shape Synthesis using Large-Scale Rectified Flow Models. *arXiv preprint arXiv:2502.06608* (2025).
- Cheng Lin, Lingjie Liu, Changjian Li, Leif Kobbelt, Bin Wang, Shiqing Xin, and Wenping Wang. 2020. SEG-MAT: 3D shape segmentation using medial axis transform. *IEEE. T. Vis. Comput. Gr.* 28, 6 (2020).
- Minghua Liu, Mikaela Angelina Uy, Donglai Xiang, Hao Su, Sanja Fidler, Nicholas Sharp, and Jun Gao. 2025. PARTFIELD: Learning 3D Feature Fields for Part Segmentation and Beyond. In *ICCV*.
- Weixiao Liu, Yuwei Wu, Sipu Ruan, and Gregory S Chirikjian. 2022. Robust and accurate superquadric recovery: A probabilistic approach. In *CVPR*.
- Weixiao Liu, Yuwei Wu, Sipu Ruan, and Gregory S Chirikjian. 2023. Marching-Primitives: Shape abstraction from signed distance function. In *CVPR*.
- Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *ECCV*.
- Tom Monnier, Jake Austin, Angjoo Kanazawa, Alexei Efros, and Mathieu Aubry. 2023. Differentiable Blocks World: Qualitative 3D decomposition by rendering primitives. In *NeurIPS*.
- Chengjie Niu, Jun Li, and Kai Xu. 2018. Im2Struct: Recovering 3D shape structure from a single RGB image. In *CVPR*.
- Despoina Paschalidou, Luc Van Gool, and Andreas Geiger. 2020. Learning unsupervised hierarchical part decomposition of 3D objects from a single RGB image. In *CVPR*.
- Despoina Paschalidou, Ali Osman Ulusoy, and Andreas Geiger. 2019. Superquadrics revisited: Learning 3D shape parsing beyond cuboids. In *CVPR*.
- Dmitriy Smirnov, Matthew Fisher, Vladimir G Kim, Richard Zhang, and Justin Solomon. 2020. Deep parametric shape predictions using distance fields. In *CVPR*.
- George Tang, William Zhao, Logan Ford, David Benham, and Paul Zhang. 2024. Segment Any Mesh: Zero-shot mesh part segmentation via lifting segment anything 2 to 3D. *arXiv preprint arXiv:2408.13679* (2024).
- Shubham Tulsiani, Hao Su, Leonidas J Guibas, Alexei A Efros, and Jitendra Malik. 2017. Learning shape abstractions by assembling volumetric primitives. In *CVPR*.
- Xinyue Wei, Minghua Liu, Zhan Ling, and Hao Su. 2022. Approximate convex decomposition for 3d meshes with collision-aware concavity and tree search. *ACM Trans. Graph.* 41, 4 (2022).
- Yuwei Wu, Weixiao Liu, Sipu Ruan, and Gregory S Chirikjian. 2022. Primitive-based shape abstraction via nonparametric bayesian inference. In *ECCV*.
- Kaizhi Yang and Xuejin Chen. 2021. Unsupervised learning for cuboid shape abstraction via joint segmentation from point clouds. *ACM Trans. Graph. (SIGGRAPH)* 40, 4 (2021).
- Xianghui Yang, Huiwen Shi, Bowen Zhang, Fan Yang, Jiacheng Wang, Hongxu Zhao, Xinhai Liu, Xinzhou Wang, Qingxiang Lin, Jiao Yu, et al. 2024b. Hunyuan3D 1.0: A Unified Framework for Text-to-3D and Image-to-3D Generation. *arXiv preprint arXiv:2411.02293* (2024).
- Yunhan Yang, Yuan-Chen Guo, Yukun Huang, Zi-Xin Zou, Zhipeng Yu, Yangguang Li, Yan-Pei Cao, and Xihui Liu. 2025. HoloPart: Generative 3D Part Amodal Segmentation. *arXiv preprint arXiv:2504.07943* (2025).
- Yunhan Yang, Yukun Huang, Yuan-Chen Guo, Liangjun Lu, Xiaoyang Wu, Edmund Y Lam, Yan-Pei Cao, and Xihui Liu. 2024a. SAMPart3D: Segment any part in 3D objects. *arXiv preprint arXiv:2411.07184* (2024).
- Jingwen Ye, Yuze He, Yanning Zhou, Yiqin Zhu, Kaiwen Xiao, Yong-Jin Liu, Wei Yang, and Xiao Han. 2025. PrimitiveAnything: Human-Crafted 3D Primitive Assembly Generation with Auto-Regressive Transformer. *ACM Trans. Graph. (SIGGRAPH)* (2025).
- Longwen Zhang, Ziyu Wang, Qixuan Zhang, Qiwei Qiu, Anqi Pang, Haoran Jiang, Wei Yang, Lan Xu, and Jingyi Yu. 2024. CLAY: A Controllable Large-scale Generative Model for Creating High-quality 3D Assets. *ACM Trans. Graph. (SIGGRAPH)* 43, 4 (2024).
- Yuhan Zhang, Mengchen Zhang, Tong Wu, Tengfei Wang, Gordon Wetzstein, Dahua Lin, and Ziwei Liu. 2025. 3DGen-Bench: Comprehensive Benchmark Suite for 3D Generative Models. *arXiv preprint arXiv:2503.21745* (2025).
- Zibo Zhao, Zejiang Lai, Qingxiang Lin, Yunfei Zhao, Haolin Liu, Shuhui Yang, Yifei Feng, Mingxin Yang, Sheng Zhang, Xianghui Yang, et al. 2025. Hunyuan3D 2.0: Scaling diffusion models for high resolution textured 3D assets generation. *arXiv preprint arXiv:2501.12202* (2025).
- Chuhang Zou, Ersin Yumer, Jimei Yang, Duygu Ceylan, and Derek Hoiem. 2017. 3D-PRNN: Generating shape primitives with recurrent neural networks. In *ICCV*.

A Implementation Details

In this section, we summarize the implementation details of our structure-aware superquadrics fitting algorithm.

Non-overlapping Fitting. The TSDF matching term follows a normal distribution as:

$$P(\phi_\tau(\mathbf{x})|\phi_{\theta,\tau}(\mathbf{x})) = \mathcal{N}(\phi_\tau(\mathbf{x})|\phi_{\theta,\tau}(\mathbf{x}), \sigma^2) \quad (17)$$

The σ^2 parameter is updated in closed form through the EM optimization process of the superquadric fitting. We further apply a truncation for numerical stability:

$$\sigma^2 = \max\left(\frac{\sum_{\mathbf{x} \in V_a(\theta)} \lambda_{\mathbf{x}} \|\phi_{\theta,\tau}(\mathbf{x}) - \phi_\tau(\mathbf{x})\|_2^2}{\sum_{\mathbf{x} \in V_a(\theta)} \lambda_{\mathbf{x}}}, \tau^2\right) \quad (18)$$

For the $\lambda_{\mathbf{x}}$ computation scheme, we set the primitive number prior $\tilde{N} = 50$, which yields $w = 1/\tilde{N} = 0.02$. The weighting constant C is set to 1. Compared to previous methods [Liu et al. 2023], our inside-voxel decay term is much smaller (yet still much larger than the TSDF matching term), allowing more flexible fitting around the generated shape surface, since the generated surface is not as clean as the artist-made ones.

Structure-aware Convex Decomposition. For the slice-plane generation, we operate on the axis-aligned setting, thus given a 100^3 grid, there are 300 potential cutting planes in total. We set $\alpha = 0.7$ for the structural saliency score, $K = 6$ and $\delta = 0.1$ to finalize the candidates. This setting also generalizes to other grid resolutions, given that the input geometry is normalized to $[-1, 1]^3$.

For adaptive merging, we set $\beta = 0.4$ and $\gamma = 0.6$ to balance between curvature and volumetric score. The merging threshold τ_m is set to 0.7.

Block-Regrow-Fill. During “block” stage, we set $K = 1$. The optimization parameters remain the same through all three stages as the setting above.

Adaptive Residual Pruning. In practice, we find that simply setting the classification percentages $P_M\% = P_C\% = P_O\% = 50\%$ yields robust classification. For the pruning threshold, we normalize all generated geometry into $[-1, 1]^3$ and set thresholds $T_M = V = 0.02$, $T_C = 1.5V = 0.03$, and $T_O = 2.5V = 0.05$, where $V = 2/100 = 0.02$ is the TSDF voxel size.

B User Study Settings

Here we detail the explanation of the five metrics in main paper Table 2, which are also provided to the participants.

- (1) *Geometry Editability.* Whether the structural components of the model are clearly distinguishable and can be locally modified without unintended influence on other parts.
- (2) *Geometry Editing Efficiency.* Whether the structure facilitates rapid editing and ensures that geometric modifications accurately reflect the intended target shape.
- (3) *Texture Editability.* Whether texture regions are clearly defined and can be modified independently without affecting unrelated areas.
- (4) *Texture Editing Efficiency.* Whether target texture regions can be quickly identified and edited in a manner consistent with the user’s intention and alignment.
- (5) *Animation Friendliness.* Whether the model can be easily rigged with a skeleton while maintaining natural and plausible deformations during local transformations.