

Novelty Assessment Report

Paper: S2GO: Streaming Sparse Gaussian Occupancy

PDF URL: <https://openreview.net/pdf?id=z8ggdMlSco>

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-29

Abstract

Despite the efficiency and performance of sparse query-based representations for perception, state-of-the-art 3D occupancy estimation methods still rely on voxel-based or dense Gaussian-based 3D representations. However, dense representations are slow, and they lack flexibility in capturing the temporal dynamics of driving scenes. Distinct from prior work, we instead summarize the scene into a compact set of 3D queries which are propagated through time in an online, streaming fashion. These queries are then decoded into semantic Gaussians at each timestep. We couple our framework with a denoising rendering objective to guide the queries and their constituent Gaussians in effectively capturing scene geometry. Owing to its efficient, query-based representation, S2GO achieves state-of-the-art performance on the nuScenes and KITTI occupancy benchmarks, outperforming prior art (e.g., GaussianWorld) by 2.7 IoU with 4.5x faster inference.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Streaming 3D Semantic Occupancy Estimation from Multi-View Images**

A total of **9 papers** were analyzed and organized into a taxonomy with **10 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Temporal Propagation and World Modeling**
- **Embodied Progressive Perception**
- **Static Multi-View Aggregation**
- **Real-Time Egocentric and Distributed Perception**
- **Application-Specific 3D Semantic Perception**

Complete Taxonomy Tree

- Streaming 3D Semantic Occupancy Estimation from Multi-View Images Survey Taxonomy
- Temporal Propagation and World Modeling
 - Query-Based Streaming Representations ★ (1 papers)
 - [0] S2GO: Streaming Sparse Gaussian Occupancy (Anon et al., 2026) [View paper](#)
 - Dense Temporal Scene Modeling (1 papers)
 - [2] Gaussianworld: Gaussian world model for streaming 3d occupancy prediction (Sicheng Zuo, 2025) [View paper](#)
- Embodied Progressive Perception
 - Embodied Exploration with Gaussian Representations (1 papers)
 - [1] Embodiedoc: Embodied 3d occupancy prediction for vision-based online scene understanding (Wu Yuqi, 2025) [View paper](#)
 - Incremental Scene Graph Construction (1 papers)
 - [6] Incremental 3D Semantic Scene Graph Prediction from RGB Sequences (Shun Cheng Wu, 2023) [View paper](#)
- Static Multi-View Aggregation
 - View-Guided Transformer Aggregation (1 papers)
 - [4] ViewFormer: Exploring Spatiotemporal Modeling for Multi-View 3D Occupancy Perception via View-Guided Transformers (Li Jinke, 2024) [View paper](#)
 - Stereo and Multi-View Depth Integration (1 papers)
 - [3] MonSter++: Unified Stereo Matching, Multi-view Stereo, and Real-time Stereo with Monodepth Priors (Cheng, 2025) [View paper](#)
- Real-Time Egocentric and Distributed Perception
 - Egocentric Binocular Vision Pipelines (1 papers)
 - [5] Real-time 3D Semantic Scene Perception for Egocentric Robots with Binocular Vision (Nguyen, 2024) [View paper](#)
 - Distributed Smart Sensor Networks (1 papers)
 - [8] 3D Semantic Scene Perception using Distributed Smart Edge Sensors (Bultmann, 2022) [View paper](#)
- Application-Specific 3D Semantic Perception
 - Medical Robotics Collision Avoidance (1 papers)
 - [9] Multi-view Real-time 3D Occupancy Map for Machine-patient Collision Avoidance (Timothy Callemeyn, 2021) [View paper](#)
 - General Robotics Scene Understanding (1 papers)
 - [7] 3D Semantic Scene Understanding and Reconstruction (Feng, 2023) [View paper](#)

Narrative

Core task: streaming 3D semantic occupancy estimation from multi-view images. This field addresses the challenge of continuously reconstructing and semantically labeling 3D space from sequences of camera views, a problem central to autonomous navigation and

embodied AI. The taxonomy reveals several complementary research directions. Temporal Propagation and World Modeling focuses on maintaining coherent scene representations over time, often using query-based or memory-driven mechanisms to integrate new observations with past knowledge. Embodied Progressive Perception emphasizes incremental updates from an agent's perspective, building scene understanding as the viewpoint evolves. Static Multi-View Aggregation tackles the foundational problem of fusing information from multiple simultaneous camera feeds without temporal dynamics. Real-Time Egocentric and Distributed Perception targets efficiency and deployment constraints, including egocentric setups and distributed sensor networks. Application-Specific 3D Semantic Perception tailors methods to particular domains such as collision avoidance or robotic manipulation, where task-specific priors guide the representation.

Within Temporal Propagation and World Modeling, query-based streaming representations have emerged as a particularly active line of work, balancing memory efficiency with the need to propagate scene understanding across frames. S2GO[0] exemplifies this approach by maintaining structured queries that evolve with incoming observations, enabling continuous occupancy updates without reprocessing the entire history. This contrasts with methods like GaussianWorld[2], which leverages Gaussian splatting for world modeling, and EmbodiedOcc[1], which integrates embodied agent trajectories more explicitly into the representation. Meanwhile, works such as ViewFormer[4] explore transformer-based fusion for static multi-view scenarios, and Realtime Semantic Egocentric[5] prioritizes low-latency egocentric perception. The central tension across these branches lies in trading off representational richness, temporal consistency, and computational cost, with S2GO[0] positioned among methods that use lightweight, query-driven abstractions to achieve streaming performance while maintaining semantic detail over extended sequences.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

Both subtopics address temporal modeling for streaming 3D semantic occupancy estimation from multi-view images, but differ fundamentally in their scene representation paradigm. Query-Based Streaming Representations uses sparse, learnable queries that are propagated frame-to-frame for efficient online inference. Dense Temporal Scene Modeling employs dense volumetric representations (3D Gaussians or voxels) combined with world models to capture scene continuity more comprehensively.

Similarities: - Both focus on temporal propagation to leverage information across frames in streaming scenarios - Both aim to improve 3D semantic occupancy estimation through temporal consistency - Both are designed for online/streaming inference rather than batch processing

Differences: - Representation density: sparse query-based vs. dense voxel/Gaussian representations - Memory and computational footprint: queries are lightweight and selective vs. dense representations require more resources - Modeling approach: query propagation mechanisms vs. world model-based scene evolution - Spatial coverage: queries focus on salient regions vs. dense methods model the entire scene uniformly

Suggested Search Directions: - Hybrid approaches combining sparse queries with local dense refinement - Comparative studies on accuracy-efficiency trade-offs between sparse and dense temporal modeling - Adaptive methods that switch between sparse and dense representations based on scene complexity

Sibling Subtopics

- **Dense Temporal Scene Modeling** (leaves: 1, papers: 1)
- Scope: Methods using dense 3D Gaussian or voxel representations with world models to exploit scene continuity.
- Exclude: Sparse query-based streaming methods belong to Query-Based Streaming Representations.

Contributions Analysis

Overall novelty summary. The paper proposes a streaming sparse query-based framework for 3D occupancy estimation, decoding queries into semantic Gaussians at each timestep. According to the taxonomy tree, this work sits in the 'Query-Based Streaming Representations' leaf under 'Temporal Propagation and World Modeling'. Notably, this leaf contains only the original paper itself—no sibling papers are listed. This suggests the specific combination of sparse queries, temporal propagation, and Gaussian decoding for streaming occupancy is relatively unexplored in the examined literature, positioning the work in a sparse research direction within the broader temporal modeling branch.

The taxonomy reveals that the broader 'Temporal Propagation and World Modeling' branch also includes 'Dense Temporal Scene Modeling' (e.g., GaussianWorld), which uses dense Gaussian or voxel representations rather than sparse queries. Neighboring branches include 'Static Multi-View Aggregation' (transformer-based view fusion without temporal propagation) and 'Embodied Progressive Perception' (incremental scene building through agent exploration). The scope notes clarify that methods without explicit temporal modeling belong elsewhere, while dense Gaussian approaches are separated from sparse query-based streaming. This structural context suggests the paper bridges temporal modeling with query efficiency, diverging from both static aggregation and dense temporal representations.

Among 28 candidates examined, none were found to clearly refute any of the three contributions. For the streaming sparse query-based framework, 9 candidates were examined with 0 refutable; for geometry denoising pretraining, 10 candidates with 0 refutable; for the Gaussian formulation and voxel splatting, 9 candidates with 0 refutable. This limited search scope—top-K semantic matches plus citation expansion—suggests that within the examined literature, no prior work directly overlaps with the specific combination of sparse queries, temporal propagation, and Gaussian decoding. However, the absence of refutable candidates does not imply exhaustive coverage of all related work in the field.

Based on the limited search of 28 candidates, the work appears to occupy a relatively novel position, particularly in combining sparse query-based streaming with Gaussian representations for occupancy estimation. The taxonomy structure—where the paper is the sole member of its leaf—reinforces this impression within the examined scope. However, the analysis does not cover all possible prior work in dense representations, alternative query mechanisms, or related temporal modeling approaches outside the top-K semantic matches. A broader literature review might reveal additional connections or overlapping ideas not captured in this limited search.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Streaming sparse query-based framework for 3D occupancy estimation

Description: The authors propose S2GO, a streaming framework that represents driving scenes using sparse 3D queries (approximately 1k) instead of dense voxel or Gaussian representations. These queries are propagated temporally and decoded into semantic Gaussians for efficient occupancy estimation.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Gaussianflowocc: Sparse and weakly supervised occupancy estimation using gaussian splatting and temporal flow

URL: [View paper](#)

Brief Assessment

GaussianFlowOcc[10] uses a sparse Gaussian representation with a Gaussian transformer but does not employ a streaming framework with temporal query propagation as described in S2GO. GaussianFlowOcc[10] focuses on temporal flow estimation for dynamic objects rather than maintaining persistent queries across time steps.

2. OccWorld: Learning a 3D Occupancy World Model for Autonomous Driving

URL: [View paper](#)

Brief Assessment

OccWorld[11] focuses on learning a world model in 3D occupancy space for predicting scene evolution and ego-car movement using discrete scene tokens and a GPT-like transformer. It does not employ sparse 3D queries with temporal propagation as in the original paper's streaming framework.

3. Navigation-guided sparse scene representation for end-to-end autonomous driving

URL: [View paper](#)

Brief Assessment

Navigation Sparse Representation[18] focuses on end-to-end autonomous driving with navigation-guided sparse scene representation for planning tasks, not on 3D occupancy estimation with temporal propagation of queries as in the original paper.

4. SparseWorld: A Flexible, Adaptive, and Efficient 4D Occupancy World Model Powered by Sparse and Dynamic Queries

URL: [View paper](#)

Brief Assessment

SparseWorld[16] focuses on a 4D occupancy world model with range-adaptive perception and state-conditioned forecasting for planning tasks, while the original paper addresses streaming 3D occupancy estimation with temporal query propagation and Gaussian decoding for perception tasks.

5. STCOcc: Sparse Spatial-Temporal Cascade Renovation for 3D Occupancy and Scene Flow Prediction

URL: [View paper](#)

Brief Assessment

STCOcc[17] uses implicit learning-based approaches with sparse occlusion-aware attention, not a sparse query-based representation with temporal propagation as in the original paper.

6. Pointbev: A sparse approach for bev predictions

URL: [View paper](#)

Brief Assessment

PointBEV[15] operates on sparse BEV cells for 2D segmentation tasks (vehicle, pedestrian, lane), not 3D occupancy estimation with temporal query propagation and Gaussian decoding as in the original paper.

7. Trajectory prediction for autonomous driving: Progress, limitations, and future directions

URL: [View paper](#)

Brief Assessment

Trajectory Prediction Survey[19] focuses on trajectory forecasting for autonomous driving, covering prediction of future paths of traffic participants. It does not address 3D occupancy estimation, sparse query representations, or Gaussian-based scene modeling, which are the core technical contributions of the original paper.

8. Doracamom: Joint 3D Detection and Occupancy Prediction with Multi-view 4D Radars and Cameras for Omnidirectional Perception

URL: [View paper](#)

Brief Assessment

DoraCAMOM[12] focuses on multi-modal fusion (4D radar + cameras) for joint detection and occupancy tasks, not on streaming temporal propagation of sparse queries for occupancy alone.

9. DIO: Decomposable Implicit 4D Occupancy-Flow World Model

URL: [View paper](#)

Brief Assessment

DIO[13] focuses on decomposable implicit 4D occupancy-flow prediction with instance-level decomposition using source point prompting, rather than streaming sparse query propagation for occupancy estimation. The architectural approaches and core objectives differ fundamentally.

Contribution 2: Geometry denoising pretraining phase

Description: A novel pretraining stage is introduced where queries are initialized at noised LiDAR points and trained with a denoising objective combined with rendering supervision. This enables sparse queries to effectively move from empty space to occupied regions and self-organize to capture dense 3D structure.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. In-place scene labelling and understanding with implicit scene representation

URL: [View paper](#)

Brief Assessment

Inplace Scene Labelling[31] focuses on semantic scene understanding using neural radiance fields with sparse semantic annotations, not on geometry denoising pretraining for query-based 3D occupancy prediction. The candidate addresses semantic label propagation and fusion rather than query initialization and movement through denoising objectives.

2. OccludeNeRF: Geometry-aware 3D Scene Inpainting with Collaborative Score Distillation in NeRF

URL: [View paper](#)

Brief Assessment

OccludeNeRF[35] focuses on 3D scene inpainting using collaborative score distillation in NeRF, not on geometry denoising pretraining for 3D scene representation learning. The candidate addresses occlusion challenges in inpainting tasks through multi-view information sharing, which is a different technical problem from the original paper's pretraining approach for sparse query-based occupancy estimation.

3. Masked local-global representation learning for 3d point cloud domain adaptation

URL: [View paper](#)

Brief Assessment

Masked Local Global[39] focuses on domain adaptation for point clouds using masked representation learning and prototype-calibrated self-training, not on geometry denoising pretraining for 3D scene representation with sparse queries and rendering supervision.

4. Renderdiffusion: Image diffusion for 3d reconstruction, inpainting and generation

URL: [View paper](#)

Brief Assessment

RenderDiffusion[33] focuses on incorporating 3D triplane representations into image diffusion models for reconstruction and generation tasks, but does not employ a geometry denoising pretraining phase where queries are initialized at noised LiDAR points with denoising objectives as described in the original paper.

5. Point cloud denoising in outdoor real-world scenes based on measurable segmentation

URL: [View paper](#)

Brief Assessment

Outdoor Denoising Segmentation[36] focuses on denoising LiDAR point clouds in outdoor scenes through hierarchical segmentation strategies, not on pretraining neural networks with denoising objectives for 3D scene representation learning with queries and Gaussians.

6. DSplats: 3D Generation by Denoising Splats-Based Multiview Diffusion Models

URL: [View paper](#)

Brief Assessment

DSplats[38] focuses on denoising latent representations of multiview images to generate 3D Gaussians for novel view synthesis, not on pretraining queries with geometry denoising objectives for 3D occupancy estimation as in the original paper.

7. Diffuscene: Denoising diffusion models for generative indoor scene synthesis

URL: [View paper](#)

Brief Assessment

DiffuScene[30] focuses on indoor scene synthesis using diffusion models for object arrangement and retrieval, not on geometry denoising pretraining for 3D scene representation learning with sparse queries and LiDAR points as in the original paper.

8. Dmv3d: Denoising multi-view diffusion using 3d large reconstruction model

URL: [View paper](#)

Brief Assessment

DMV3D[32] focuses on denoising multi-view images for 3D reconstruction using a transformer-based model, not on pretraining sparse queries with geometry denoising objectives for occupancy estimation tasks.

9. Pre-training meets iteration: Learning for robust 3D point cloud denoising

URL: [View paper](#)

Brief Assessment

Pretraining Denoising Iteration[34] focuses on point cloud denoising with iterative learning, not on training sparse queries for 3D occupancy estimation in autonomous driving scenes.

10. S2GO: Streaming Sparse Gaussian Occupancy Prediction

URL: [View paper](#)

Brief Assessment

S2GO Streaming[37] presents a similar denoising pretraining approach but applies it to a different architectural framework (streaming sparse queries vs. the original's approach). The core innovation of using noised LiDAR initialization with denoising objectives appears in both works, but the candidate's integration with streaming query-based occupancy prediction represents a distinct application context rather than direct prior work that refutes novelty.

Contribution 3: Improved Gaussian formulation and efficient voxel splatting algorithm

Description: The authors propose opacity-weighted geometry estimation for Gaussians and develop an efficient CUDA-based Gaussian-to-voxel splatting algorithm. These improvements halve training time while improving performance by addressing unnatural Gaussian behavior and optimizing memory access patterns.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Structured 3D gaussian splatting for novel view synthesis based on single RGB-LiDAR View

URL: [View paper](#)

Brief Assessment

Structured RGB-LiDAR[25] focuses on single-view RGB-LiDAR reconstruction with virtual background fitting for scene structure optimization, not on the opacity-weighted geometry estimation or efficient CUDA-based Gaussian-to-voxel splatting algorithms proposed in the original paper for streaming occupancy prediction.

2. Volsplat: Rethinking feed-forward 3d gaussian splatting with voxel-aligned prediction

URL: [View paper](#)

Brief Assessment

VolSplat[22] focuses on voxel-aligned Gaussian prediction for feed-forward 3D reconstruction from multi-view images, not on improving Gaussian-to-voxel splatting algorithms or opacity-weighted geometry estimation for occupancy prediction tasks.

3. 4D Gaussian Splatting for Real-Time Dynamic Scene Rendering

URL: [View paper](#)

Brief Assessment

4D Gaussian Splatting[20] focuses on dynamic scene rendering with temporal deformation fields for 3D Gaussians, not on voxel-based occupancy prediction or Gaussian-to-voxel splatting algorithms as in the original paper.

4. 3DGS-ReLoc: 3D Gaussian Splatting for Map Representation and Visual ReLocalization

URL: [View paper](#)

Brief Assessment

3DGS-ReLoc[27] focuses on visual relocalization using Gaussian splatting for mapping, not on voxel-based occupancy prediction or splatting optimization. The candidate does not address opacity-weighted geometry estimation or CUDA-based Gaussian-to-voxel splatting algorithms.

5. Voxelsplat: Dynamic gaussian splatting as an effective loss for occupancy and flow prediction

URL: [View paper](#)

Brief Assessment

VoxelSplat[24] focuses on using Gaussian splatting as a training-time regularization loss for occupancy and flow prediction in autonomous driving, not on improving the Gaussian-to-voxel splatting algorithm itself or addressing unnatural Gaussian behavior through opacity-weighted geometry estimation as the original paper does.

6. STREAMINGGS: Voxel-Based Streaming 3D Gaussian Splatting with Memory Optimization and Architectural Support

URL: [View paper](#)

Brief Assessment

StreamingGS[26] focuses on memory-centric rendering for mobile devices with voxel-based streaming, while the original paper addresses query-based occupancy estimation with opacity-weighted geometry and CUDA-based Gaussian-to-voxel splatting for autonomous driving. These are fundamentally different application domains and technical approaches.

7. Dronesplat: 3d gaussian splatting for robust 3d reconstruction from in-the-wild drone imagery

URL: [View paper](#)

Brief Assessment

DroneSplat[29] focuses on drone-based 3D reconstruction with dynamic distractor elimination and limited-view constraints, not on improving Gaussian-to-voxel splatting algorithms or opacity-weighted geometry estimation for general occupancy prediction frameworks.

8. DyGASR: Dynamic Generalized Gaussian Splatting with Surface Alignment for Accelerated 3D Mesh Reconstruction

URL: [View paper](#)

Brief Assessment

DyGASR[28] focuses on dynamic 3D mesh reconstruction from multi-view images using Gaussian splatting for surface alignment, not on voxel-based occupancy prediction or the specific opacity-weighted geometry estimation and CUDA optimization techniques proposed in the original paper.

9. 3DGS-Loc: 3D Gaussian splatting for map representation and visual localization

URL: [View paper](#)

Brief Assessment

3DGS-Loc[23] focuses on visual relocalization using 3D Gaussian splatting for map representation, not on improving the Gaussian formulation or developing efficient voxel splatting algorithms for occupancy prediction tasks.

Appendix: Text Similarity Detection

Textual similarity detection checked 30 papers and found 3 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. S2GO: Streaming Sparse Gaussian Occupancy Prediction

Detected in: Contribution: contribution_2

△ **Note:** This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] S2GO: Streaming Sparse Gaussian Occupancy [View paper](#)
- [1] Embodiedocc: Embodied 3d occupancy prediction for vision-based online scene understanding [View paper](#)
- [2] Gaussianworld: Gaussian world model for streaming 3d occupancy prediction [View paper](#)
- [3] MonSter++: Unified Stereo Matching, Multi-view Stereo, and Real-time Stereo with Monodepth Priors [View paper](#)
- [4] ViewFormer: Exploring Spatiotemporal Modeling for Multi-View 3D Occupancy Perception via View-Guided Transformers [View paper](#)
- [5] Real-time 3D Semantic Scene Perception for Egocentric Robots with Binocular Vision [View paper](#)
- [6] Incremental 3D Semantic Scene Graph Prediction from RGB Sequences [View paper](#)
- [7] 3D Semantic Scene Understanding and Reconstruction [View paper](#)
- [8] 3D Semantic Scene Perception using Distributed Smart Edge Sensors [View paper](#)
- [9] Multi-view Real-time 3D Occupancy Map for Machine-patient Collision Avoidance [View paper](#)
- [10] Gaussianflowocc: Sparse and weakly supervised occupancy estimation using gaussian splatting and temporal flow [View paper](#)
- [11] OccWorld: Learning a 3D Occupancy World Model for Autonomous Driving [View paper](#)
- [12] Doracamom: Joint 3D Detection and Occupancy Prediction with Multi-view 4D Radars and Cameras for Omnidirectional Perception [View paper](#)
- [13] DIO: Decomposable Implicit 4D Occupancy-Flow World Model [View paper](#)

- [14] Bridging past and future: End-to-end autonomous driving with historical prediction and planning [View paper](#)
- [15] Pointbev: A sparse approach for bev predictions [View paper](#)
- [16] SparseWorld: A Flexible, Adaptive, and Efficient 4D Occupancy World Model Powered by Sparse and Dynamic Queries [View paper](#)
- [17] STCOcc: Sparse Spatial-Temporal Cascade Renovation for 3D Occupancy and Scene Flow Prediction [View paper](#)
- [18] Navigation-guided sparse scene representation for end-to-end autonomous driving [View paper](#)
- [19] Trajectory prediction for autonomous driving: Progress, limitations, and future directions [View paper](#)
- [20] 4D Gaussian Splatting for Real-Time Dynamic Scene Rendering [View paper](#)
- [21] 3d geometry-aware deformable gaussian splatting for dynamic view synthesis [View paper](#)
- [22] Volsplat: Rethinking feed-forward 3d gaussian splatting with voxel-aligned prediction [View paper](#)
- [23] 3DGS-Loc: 3D Gaussian splatting for map representation and visual localization [View paper](#)
- [24] Voxelsplat: Dynamic gaussian splatting as an effective loss for occupancy and flow prediction [View paper](#)
- [25] Structured 3D gaussian splatting for novel view synthesis based on single RGB-LiDAR View [View paper](#)
- [26] STREAMINGGS: Voxel-Based Streaming 3D Gaussian Splatting with Memory Optimization and Architectural Support [View paper](#)
- [27] 3DGS-ReLoc: 3D Gaussian Splatting for Map Representation and Visual ReLocalization [View paper](#)
- [28] DyGASR: Dynamic Generalized Gaussian Splatting with Surface Alignment for Accelerated 3D Mesh Reconstruction [View paper](#)
- [29] Dronesplat: 3d gaussian splatting for robust 3d reconstruction from in-the-wild drone imagery [View paper](#)
- [30] Diffuscene: Denoising diffusion models for generative indoor scene synthesis [View paper](#)
- [31] In-place scene labelling and understanding with implicit scene representation [View paper](#)
- [32] Dmv3d: Denoising multi-view diffusion using 3d large reconstruction model [View paper](#)
- [33] Renderdiffusion: Image diffusion for 3d reconstruction, inpainting and generation [View paper](#)
- [34] Pre-training meets iteration: Learning for robust 3D point cloud denoising [View paper](#)
- [35] OccludeNeRF: Geometry-aware 3D Scene Inpainting with Collaborative Score Distillation in NeRF [View paper](#)
- [36] Point cloud denoising in outdoor real-world scenes based on measurable segmentation [View paper](#)
- [37] S2GO: Streaming Sparse Gaussian Occupancy Prediction [View paper](#)
- [38] DSplats: 3D Generation by Denoising Splats-Based Multiview Diffusion Models [View paper](#)
- [39] Masked local-global representation learning for 3d point cloud domain adaptation [View paper](#)