Novelty Assessment Report

Paper: π^3: Permutation-Equivariant Visual Geometry Learning

PDF URL: https://openreview.net/pdf?id=DTQIjngDta

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-27

Abstract

We introduce π^3, a feed-forward neural network that offers a novel approach to visual geometry reconstruction, breaking the reliance on a conventional fixed reference view. Previous methods often anchor their reconstructions to a designated viewpoint, an inductive bias that can lead to instability and failures if the reference is suboptimal. In contrast, π^3 employs a fully permutation-equivariant architecture to predict affine-invariant camera poses and scale-invariant local point maps without any reference frames. This design not only makes our model inherently robust to input ordering, but also leads to higher accuracy and performance. These advantages enable our simple and bias-free approach to achieve state-of-the-art performance on a wide range of tasks, including camera pose estimation, monocular/video depth estimation, and dense point map reconstruction. Code and models will be publicly available.

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: **Permutation-Equivariant Visual Geometry Reconstruction without Reference Frames**

A total of **16 papers** were analyzed and organized into a taxonomy with **11 categories**.

Taxonomy Overview

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

- **Permutation-Equivariant Visual Geometry Reconstruction**
- **Equivariant Representations for 3D Point Clouds**
- **Geometry-Aware Attention and Positional Encoding**
- **Theoretical Foundations of Equivariant Learning**
- **Structured Reconstruction with Generative Models**
- **Visual Permutation Learning**
- **Geometry-Aware Representation Learning**
- **Equivariant Filtering for Visual-Inertial Odometry**

Complete Taxonomy Tree

- Permutation-Equivariant Visual Geometry Reconstruction without Reference Frames Survey Taxonomy
- Permutation-Equivariant Visual Geometry Reconstruction
  - Reference-Free Multi-View Reconstruction ★ (4 papers)
  - [0] π^3: Permutation-Equivariant Visual Geometry Learning (Anon et al., 2026) View paper
  - [1] : Scalable Permutation-Equivariant Visual Geometry Learning (Y Wang, 2025) View paper
  - [2] : Permutation-Equivariant Visual Geometry Learning (Y Wang, 2025) View paper
  - Deep Structure from Motion (1 papers)
  - [9] Deep permutation equivariant structure from motion (Dror Moran, 2021) View paper
- Equivariant Representations for 3D Point Clouds
  - SO(3)-Equivariant Point Cloud Registration (1 papers)
  - [5] Correspondence-Free Point Cloud Registration with SO(3)-Equivariant Implicit Shape Representations (Zhu, 2021) View paper
  - Capsule Networks for Point Clouds (2 papers)
  - [10] Canonical capsules: Self-supervised capsules in canonical pose (Sun Wei-wei, 2021) View paper
  - [11] Quaternion Equivariant Capsule Networks for 3D Point Clouds (Yongheng Zhao, 2019) View paper
  - Foundation Model Adaptation for Point Clouds (1 papers)
  - [14] Adaptive Point-Prompt Tuning: Fine-Tuning Heterogeneous Foundation Models for 3D Point Cloud Analysis (Li Meng-Ke, 2025) View paper
- Geometry-Aware Attention and Positional Encoding (1 papers)
  - [4] GTA: A Geometry-Aware Attention Mechanism for Multi-View Transformers (Miyato, 2023) View paper
- Theoretical Foundations of Equivariant Learning (1 papers)
  - [3] Unsupervised learning of group invariant and equivariant representations (Winter, 2022) View paper
- Structured Reconstruction with Generative Models (1 papers)
  - [6] PolyDiffuse: Polygonal Shape Reconstruction via Guided Set Diffusion Models (Chen Jiacheng, 2023) View paper
- Visual Permutation Learning (2 papers)
  - [13] Deeppermnet: Visual permutation learning (Cruz, 2017) View paper
  - [15] Visual permutation learning (Rodrigo Santa Cruz, 2018) View paper
- Geometry-Aware Representation Learning (2 papers)
  - [7] Graphs, Geometry, and Learning Representations: Navigating the Non-Euclidean Landscape in Computer Vision and Beyond (Skenderi, 2024) View paper

    - [8] Learning Geometry-aware Representations by Sketching (Hyundo Lee, 2023) View paper
- Equivariant Filtering for Visual-Inertial Odometry (1 papers)
    - [16] An Equivariant Filter for Visual Inertial Odometry (Pieter van Goor, 2021) View paper

Narrative

Core task: Permutation-equivariant visual geometry reconstruction without reference frames. This field addresses the challenge of reconstructing 3D geometry from multiple views while respecting the inherent symmetries—particularly permutation invariance—of the input data, without relying on a fixed coordinate system or reference frame. The taxonomy reveals several interconnected branches: the central branch of Permutation-Equivariant Visual Geometry Reconstruction focuses on methods that explicitly encode permutation symmetries in multi-view settings, often building on structure-from-motion principles adapted to deep learning (e.g., Deep Permutation SfM[9]). Adjacent branches explore Equivariant Representations for 3D Point Clouds, which develop group-theoretic architectures for point-set data (Quaternion Equivariant Capsules[11], Canonical Capsules[10]), and Geometry-Aware Attention mechanisms that incorporate spatial relationships into transformer-like models (Geometry Aware Attention[4]). Theoretical Foundations provide the mathematical underpinnings for equivariant learning, while Structured Reconstruction with Generative Models (PolyDiffuse[6]) and Visual Permutation Learning (Visual Permutation Learning[15], DeepPermNet[13]) address related symmetry-aware tasks in generation and ordering.

Recent work has intensified around scalable and reference-free reconstruction methods. Scalable Permutation Equivariant[1] and Permutation Equivariant Geometry[2] push toward handling larger view sets efficiently, while Pi Three[12] explores related permutation-invariant formulations. The original paper, Pi Cubed[0], sits squarely within the Reference-Free Multi-View Reconstruction cluster, emphasizing permutation equivariance without anchor frames—a contrast to earlier correspondence-based approaches like Correspondence Free Registration[5]. Compared to Scalable Permutation Equivariant[1], which prioritizes computational efficiency, Pi Cubed[0] appears to focus more directly on the theoretical and architectural implications of full permutation symmetry. Open questions remain around balancing expressiveness with scalability, integrating geometric priors from attention mechanisms, and extending these ideas to dynamic or non-rigid scenes.

Related Works in Same Category

The following **3 sibling papers** share the same taxonomy leaf node with the original paper:

1. : Scalable Permutation-Equivariant Visual Geometry Learning

Authors: Y Wang, J Zhou, H Zhu, W Chang, Y Zhou, et al. (6 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

â¦ network that offers a novel approach to visual geometry reconstruction, breaking the reliance on â¦ In contrast, $Ï^3$ employs a fully permutation-equivariant architecture to predict affine-â¦

⚠ Similarity Notice

This paper appears to be a variant or near-duplicate of the original paper. Both titles reference 'π^3' (pi-cubed) and 'Permutation-Equivariant Visual Geometry', and the candidate's abstract excerpt contains identical phrases about 'fully permutation-equivariant architecture to predict affine-invariant' outputs, suggesting they describe the same system and contribution.

2. : Permutation-Equivariant Visual Geometry Learning

Authors: Y Wang, J Zhou, H Zhu, W Chang, Y Zhou, et al. (6 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

â¦ the power of deep learning for reconstructing geometry from image pairs [13â¦ fully permutation-equivariant method that eliminates reference view-based biases in visual geometry learnâ¦

⚠ Similarity Notice

This paper is highly similar to the original paper; it may be a variant or near-duplicate. Please manually verify.

3. $Ï^3$: Permutation-Equivariant Visual Geometry Learning

Authors: Wang Yifan, Zhou Jianjun, Zhu, Haoyi, Zhou Yang, et al. (12 authors total) | **Year/Venue**: 2025 • arXiv (Cornell University) | **URL**: View paper

Abstract

We introduce $Ï^3$, a feed-forward neural network that offers a novel approach to visual geometry reconstruction, breaking the reliance on a conventional fixed reference view. Previous methods often anchor their reconstructions to a designated viewpoint, an inductive bias that can lead to instability and failures if the reference is suboptimal. In contrast, $Ï^3$ employs a fully permutation-equivariant architecture to predict affine-invariant camera poses and scale-invariant local point maps w...

⚠ Similarity Notice

This paper is highly similar to the original paper; it may be a variant or near-duplicate. Please manually verify.

Contributions Analysis

Overall novelty summary. The paper introduces π^3, a feed-forward network for visual geometry reconstruction that eliminates reliance on fixed reference views through a fully permutation-equivariant architecture. It resides in the 'Reference-Free Multi-View Reconstruction' leaf, which contains four papers total including the original work. This leaf sits within the broader 'Permutation-Equivariant Visual Geometry Reconstruction' branch, indicating a moderately populated but focused research direction. The taxonomy reveals this is an active area with sibling work exploring similar permutation-invariant formulations, suggesting the paper enters a space with established momentum rather than pioneering entirely uncharted territory.

The taxonomy structure shows the paper's immediate neighbors include 'Deep Structure from Motion' methods that recover camera parameters from point tracks, and broader branches addressing 'Equivariant Representations for 3D Point Clouds' with SO(3)-equivariant registration and capsule networks. The scope notes clarify boundaries: the paper's reference-free approach explicitly excludes calibrated stereo methods and differs from point cloud processing tasks. Nearby work on 'Geometry-Aware Attention and Positional Encoding' incorporates 3D structure into attention mechanisms, representing a complementary direction that could intersect with permutation-equivariant architectures. The taxonomy reveals a field balancing theoretical equivariance foundations with practical multi-view reconstruction challenges.

Among 27 candidates examined across three contributions, the 'π^3 permutation-equivariant architecture' contribution shows one refutable candidate from nine examined, indicating some prior work in permutation-equivariant designs for visual geometry. The 'fixed reference view bias identification' and 'state-of-the-art performance' contributions found zero refutable candidates among nine each, suggesting these aspects may be more novel or less directly addressed in the limited search scope. The statistics indicate a focused literature search rather than exhaustive coverage, with the single refutable match likely representing closely related architectural work within the same taxonomy leaf rather than definitive prior art.

Based on the limited search of 27 candidates, the work appears to offer meaningful contributions in eliminating reference frame dependencies, though the permutation-equivariant architecture concept has some precedent in the examined literature. The taxonomy positioning in a four-paper leaf suggests a maturing but not overcrowded research direction. The analysis covers top semantic matches and does not claim exhaustive field coverage, leaving open the possibility of additional related work beyond the examined scope.

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

Contribution 1: Identification of fixed reference view bias in visual geometry reconstruction

Description: The authors systematically identify and demonstrate that the common practice of anchoring reconstructions to a fixed reference view introduces an unnecessary inductive bias that limits model robustness and performance in visual geometry reconstruction tasks.

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. Shape anchors for data-driven multi-view reconstruction

URL: View paper

Brief Assessment

Shape Anchors[27] focuses on combining recognition and multi-view cues through patch-level depth transfer from RGB-D databases, not on analyzing or addressing fixed reference view bias in reconstruction architectures.

2. Spatial performance with perspective displays as a function of computer graphics eyepoint elevation and geometric field of view

URL: View paper

Brief Assessment

Perspective Display Performance[29] focuses on human spatial perception with computer graphics displays and eyepoint elevation effects, not on neural network architectures or reference view selection in visual geometry reconstruction systems.

3. Recon3D: High Quality 3D Reconstruction from a Single Image Using Generated Back-View Explicit Priors

URL: View paper

Brief Assessment

Recon3D[23] focuses on single-image 3D reconstruction using generated back-view priors, not on multi-view geometry reconstruction or reference view selection issues.

4. Video-Based Camera Localization Using Anchor View Detection and Recursive 3D Reconstruction

URL: View paper

Brief Assessment

Anchor View Detection[28] addresses camera localization in industrial inspection scenarios using anchor frames as spatial guides, not the systematic identification of fixed reference view bias as an inductive bias problem in neural geometry reconstruction.

5. Edit360: 2d image edits to 3d assets from any angle

URL: View paper

Brief Assessment

Edit360[24] focuses on propagating 2D image edits to 3D assets for editing tasks, not on visual geometry reconstruction or camera pose estimation. The paper does not address reference view bias in reconstruction pipelines.

6. Panoramic 3D Reconstruction of an Indoor Scene Using A Multi-view

URL: View paper

Brief Assessment

Panoramic Indoor Reconstruction[30] focuses on panoramic 3D reconstruction of indoor scenes using multi-view approaches with reference viewpoints for smoothness constraints, rather than systematically identifying or challenging the fixed reference view bias as a fundamental limitation in visual geometry reconstruction.

7. Spatial-Aware Anchor Growth for 3D Gaussian Field Reconstruction

URL: View paper

Brief Assessment

Spatial Aware Anchor[25] addresses bias in 3D Gaussian splatting anchor optimization, not the fixed reference view bias in multi-view geometry reconstruction that the original paper identifies.

8. : Scalable Permutation-Equivariant Visual Geometry Learning

URL: View paper

Brief Assessment

Scalable Permutation Equivariant[1] addresses permutation equivariance in visual geometry but does not provide evidence of prior work identifying fixed reference view bias, thus cannot refute the original paper's novelty claim on this systematic identification.

9. Multi-view Normal and Distance Guidance Gaussian Splatting for Surface Reconstruction

URL: View paper

Brief Assessment

Multi View Normal Guidance[26] addresses single-view geometric constraints in 3D Gaussian Splatting for surface reconstruction, not the fixed reference view bias in feed-forward visual geometry learning that the original paper identifies.

Contribution 2: π^3 permutation-equivariant architecture

Description: The authors introduce π^3, a fully permutation-equivariant neural network architecture that predicts affine-invariant camera poses and scale-invariant local pointmaps without requiring any reference frames, making it inherently robust to input ordering.

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. EquiPose: Exploiting Permutation Equivariance for Relative Camera Pose Estimation

URL: View paper

Brief Assessment

Cannot assess refutation as the candidate paper (EquiPose[20]) provides no full text context for comparison. The retrieval query suggests potential similarity, but without actual content, no evidence-based determination is possible.

2. Quaternion Equivariant Capsule Networks for 3D Point Clouds

URL: View paper

Brief Assessment

Quaternion Equivariant Capsules[11] focuses on SO(3) rotation equivariance for 3D point clouds using quaternion capsule networks, not affine-invariant camera pose and pointmap prediction for multi-view visual geometry reconstruction.

3. Lie group decompositions for equivariant neural networks

URL: View paper

Brief Assessment

Lie Group Decompositions[19] focuses on equivariance to affine/matrix Lie group transformations (GL+(n,R), SL(n,R)) for image classification tasks, not on permutation-equivariant architectures for camera pose and pointmap prediction in visual geometry reconstruction.

4. Deep permutation equivariant structure from motion

URL: View paper

Prior Art Analysis

Deep Permutation SfM[9] demonstrates prior work on permutation-equivariant architectures for structure from motion that predicts camera parameters and 3D point positions from point tracks. The candidate paper explicitly describes a neural network architecture designed to be equivariant to permutations of both cameras and scene points, where reordering inputs yields correspondingly reordered outputs. This directly addresses the same core concept of permutation equivariance for geometry prediction that the original paper claims as novel. Both papers use equivariant layers that respect input ordering symmetries and produce outputs that maintain consistent one-to-one mappings regardless of input permutation order.

Evidence

Evidence 1 - **Rationale**: Deep Permutation SfM[9] demonstrates a permutation-equivariant network that predicts camera poses and scene structure without requiring initialization or reference frames, showing this approach existed before the original paper's claimed novelty. - **Original**: to ensure our model's output is invariant to the arbitrary ordering of input views, we designed our networkϕto bepermutation-equivariant... for any permutationπ, letp π be an operator that permutes the order of a sequence. the networkϕ satisfies the permutation-equivariant property: ϕ(pπ(s)) =pπ(ϕ(s... - **Candidate**: our network can be applied in a single scene scenario, in which for a given scene the weights are optimized to directly minimize a reprojection loss. this minimization does not require initialization of either camera parameters or scene structure, yet it achieves accurate recovery of poses and scene...

Evidence 2 - **Rationale**: Both papers identify similar advantages of permutation-equivariant architectures: consistent mappings independent of input ordering and robustness to variations in input structure. This shows the benefits claimed by the original paper were already recognized in prior work. - **Original**: this property guarantees a consistent one-to-one correspondence between each image and its respective output (e.g., geometry or pose). this design offers several key advantages. first, reconstruction quality becomesindependent of the reference view selection, in contrast to prior methods that suffer... - **Candidate**: there are several important advantages to using our equivariant architecture compared to standard fully connected architectures. above all, our architecture encodes the structure of the task into the model, thus providing a strong inductive bias... our architecture has the crucial ability to handle ...

5. Fourier-Based Equivariant Graph Neural Networks for Camera Pose Estimation

URL: View paper

Brief Assessment

Cannot assess refutation as the candidate paper (Fourier Equivariant Pose[22]) provides no full text context for comparison.

6. Equivariant Ray Embeddings for Implicit Multi-View Depth Estimation

URL: View paper

Brief Assessment

Equivariant Ray Embeddings[17] focuses on SE(3) gauge equivariance for multi-view depth estimation with ray embeddings and spherical harmonics, not on permutation equivariance for camera pose and pointmap prediction without reference frames.

7. : Scalable Permutation-Equivariant Visual Geometry Learning

URL: View paper

Brief Assessment

The candidate paper context is too limited (only fragments visible) to establish whether it presents a permutation-equivariant architecture for affine-invariant camera poses and scale-invariant pointmaps that would refute the original paper's novelty.

8. E2PN: Efficient SE (3)-equivariant point network

URL: View paper

Brief Assessment

Efficient Equivariant Point[21] focuses on SE(3)-equivariant convolutions for 3D point clouds using quotient representations ($S^2 \times R^3$), not affine-invariant camera pose prediction or pointmap reconstruction from multi-view images as in the original paper.

9. Robust 3D Point Cloud Registration Based on Deep Learning and Optimization

URL: View paper

Brief Assessment

Robust Registration Optimization[18] focuses on 3D point cloud registration using optimization techniques. The candidate mentions permutation-invariant points in 3d space and affine-invariant methods, but does not describe a neural network architecture for predicting camera poses and pointmaps without reference frames.

Contribution 3: State-of-the-art performance across multiple benchmarks

Description: Through extensive experiments, the authors demonstrate that π^3 achieves state-of-the-art performance across multiple tasks including camera pose estimation, monocular and video depth estimation, and dense pointmap reconstruction, outperforming prior leading methods.

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. Explicit Depth-Aware Blurry Video Frame Interpolation Guided by Differential Curves

URL: View paper

Brief Assessment

Depth Aware Interpolation[33] focuses on blurry video frame interpolation with depth-aware scene flow estimation, not on camera pose estimation, monocular/video depth estimation, or dense pointmap reconstruction tasks that the original paper addresses.

2. Decotr: Enhancing depth completion with 2d and 3d attentions

URL: View paper

Brief Assessment

Decotr[40] focuses on depth completion tasks (dense depth map generation from sparse measurements), not the visual geometry reconstruction tasks (camera pose estimation, monocular/video depth estimation, pointmap reconstruction) that π^3 addresses. These are fundamentally different problem domains with different inputs and objectives.

3. Mono3r: Exploiting monocular cues for geometric 3d reconstruction

URL: View paper

Brief Assessment

Mono3r[36] focuses on integrating monocular geometric priors into multi-view reconstruction to address matching-based limitations in textureless regions. The original paper's SOTA claims span camera pose estimation, monocular/video depth estimation, and pointmap reconstruction across different benchmark suites, while Mono3r[36] evaluates primarily on multi-view pose and point cloud estimation tasks using different datasets (7scenes, NRGBD, DTU, ETH3D, Tanks & Temples). The technical approaches and evaluation protocols differ substantially.

4. E3D-Bench: A Benchmark for End-to-End 3D Geometric Foundation Models

URL: View paper

Brief Assessment

E3D Bench[37] is a benchmarking paper that evaluates existing 3D geometric foundation models rather than proposing a new method. It does not challenge π^3's novelty claims about achieving state-of-the-art performance, as it serves a different purpose: systematically comparing multiple models including methods that may be related to π^3's approach.

5. Refinement of Monocular Depth Maps via Multi-View Differentiable Rendering

URL: View paper

Brief Assessment

Depth Refinement Rendering[39] focuses on refining monocular depth maps for static scenes using multi-view differentiable rendering, not on the broader visual geometry tasks (camera pose estimation, video depth, pointmap reconstruction) that π^3 addresses.

6. Putting people in their place: Monocular regression of 3d people in depth

URL: View paper

Brief Assessment

Monocular Regression Depth[32] focuses on 3D human pose/shape estimation and depth reasoning for people in images, not the general visual geometry tasks (camera pose, depth estimation, pointmap reconstruction) addressed by the original paper.

7. Unsupervised learning of depth and ego-motion from monocular video using 3d geometric constraints

URL: View paper

Brief Assessment

Geometric Constraints Depth[31] focuses on unsupervised depth and ego-motion estimation from monocular video using 3D geometric constraints and ICP-based losses. The original paper addresses permutation-equivariant visual geometry learning across camera pose estimation, depth estimation, and pointmap reconstruction tasks with a fundamentally different architectural approach.

8. Robust 3D Human Avatar Reconstruction From Monocular Videos Using Depth Optimization and Camera Pose Estimation

URL: View paper

Brief Assessment

Robust Avatar Reconstruction[38] focuses on 3D human avatar reconstruction from monocular videos using depth optimization and camera pose estimation for deformable avatars. The original paper addresses general visual geometry reconstruction including camera poses, depth estimation, and pointmap reconstruction across diverse scenes (indoor, outdoor, static, dynamic). These are fundamentally different problem domains with different evaluation benchmarks.

9. DeepFusion: Real-time dense 3D reconstruction for monocular SLAM using single-view depth and gradient predictions

URL: View paper

Brief Assessment

DeepFusion[35] addresses monocular SLAM with dense 3D reconstruction using depth predictions, not the multi-task visual geometry learning (camera pose estimation, video depth estimation, pointmap reconstruction) that π^3 targets across diverse benchmarks.

Appendix: Text Similarity Detection

Textual similarity detection checked 29 papers and found 5 similarity segment(s) across 2 paper(s).

The following **2 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. : Permutation-Equivariant Visual Geometry Learning

Detected in: Core Task (sibling), Contribution: contribution_1, 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.

2. $Ï^3$: Permutation-Equivariant Visual Geometry Learning

Detected in: Core Task (sibling)

⚠ **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] π^3: Permutation-Equivariant Visual Geometry Learning View paper
- [1] : Scalable Permutation-Equivariant Visual Geometry Learning View paper
- [2] : Permutation-Equivariant Visual Geometry Learning View paper
- [3] Unsupervised learning of group invariant and equivariant representations View paper
- [4] GTA: A Geometry-Aware Attention Mechanism for Multi-View Transformers View paper
- [5] Correspondence-Free Point Cloud Registration with SO(3)-Equivariant Implicit Shape Representations View paper
- [6] PolyDiffuse: Polygonal Shape Reconstruction via Guided Set Diffusion Models View paper
- [7] Graphs, Geometry, and Learning Representations: Navigating the Non-Euclidean Landscape in Computer Vision and Beyond View paper
- [8] Learning Geometry-aware Representations by Sketching View paper
- [9] Deep permutation equivariant structure from motion View paper
- [10] Canonical capsules: Self-supervised capsules in canonical pose View paper
- [11] Quaternion Equivariant Capsule Networks for 3D Point Clouds View paper
- [12] $Ï^3$: Permutation-Equivariant Visual Geometry Learning View paper
- [13] Deeppermnet: Visual permutation learning View paper
- [14] Adaptive Point-Prompt Tuning: Fine-Tuning Heterogeneous Foundation Models for 3D Point Cloud Analysis View paper
- [15] Visual permutation learning View paper
- [16] An Equivariant Filter for Visual Inertial Odometry View paper
- [17] Equivariant Ray Embeddings for Implicit Multi-View Depth Estimation View paper
- [18] Robust 3D Point Cloud Registration Based on Deep Learning and Optimization View paper
- [19] Lie group decompositions for equivariant neural networks View paper
- [20] EquiPose: Exploiting Permutation Equivariance for Relative Camera Pose Estimation View paper
- [21] E2PN: Efficient SE (3)-equivariant point network View paper
- [22] Fourier-Based Equivariant Graph Neural Networks for Camera Pose Estimation View paper
- [23] Recon3D: High Quality 3D Reconstruction from a Single Image Using Generated Back-View Explicit Priors View paper
- [24] Edit360: 2d image edits to 3d assets from any angle View paper
- [25] Spatial-Aware Anchor Growth for 3D Gaussian Field Reconstruction View paper
- [26] Multi-view Normal and Distance Guidance Gaussian Splatting for Surface Reconstruction View paper
- [27] Shape anchors for data-driven multi-view reconstruction View paper
- [28] Video-Based Camera Localization Using Anchor View Detection and Recursive 3D Reconstruction View paper
- [29] Spatial performance with perspective displays as a function of computer graphics eyepoint elevation and geometric field of view View paper
- [30] Panoramic 3D Reconstruction of an Indoor Scene Using A Multi-view View paper
- [31] Unsupervised learning of depth and ego-motion from monocular video using 3d geometric constraints View paper
- [32] Putting people in their place: Monocular regression of 3d people in depth View paper
- [33] Explicit Depth-Aware Blurry Video Frame Interpolation Guided by Differential Curves View paper
- [34] Self-Supervised Monocular 4D Scene Reconstruction for Egocentric Videos View paper
- [35] DeepFusion: Real-time dense 3D reconstruction for monocular SLAM using single-view depth and gradient predictions View paper
- [36] Mono3r: Exploiting monocular cues for geometric 3d reconstruction View paper
- [37] E3D-Bench: A Benchmark for End-to-End 3D Geometric Foundation Models View paper
- [38] Robust 3D Human Avatar Reconstruction From Monocular Videos Using Depth Optimization and Camera Pose Estimation View paper
- [39] Refinement of Monocular Depth Maps via Multi-View Differentiable Rendering View paper
- [40] Decotr: Enhancing depth completion with 2d and 3d attentions View paper