

# RiemanLine: Riemannian Manifold Representation of 3D Lines for Factor Graph Optimization

Yan Li<sup>1</sup>, Ze Yang<sup>2</sup>, Keisuke Tateno<sup>3</sup>, Federico Tombari<sup>3,4</sup>, Liang Zhao<sup>5</sup>, Gim Hee Lee<sup>1</sup>

<sup>1</sup>National University of Singapore

<sup>2</sup>Peking University

<sup>3</sup>Google

<sup>4</sup>Technical University of Munich

<sup>5</sup>The University of Edinburgh

## Abstract

Minimal parametrization of 3D lines plays a critical role in camera localization and structural mapping. Existing representations in robotics and computer vision predominantly handle independent lines, overlooking structural regularities such as sets of parallel lines that are pervasive in man-made environments. This paper introduces **RiemanLine**, a unified minimal representation for 3D lines formulated on Riemannian manifolds that jointly accommodates both individual lines and parallel-line groups. Our key idea is to decouple each line landmark into global and local components: a shared vanishing direction optimized on the unit sphere  $S^2$ , and scaled normal vectors constrained on orthogonal subspaces, enabling compact encoding of structural regularities. For  $n$  parallel lines, the proposed representation reduces the parameter space from  $4n$  (orthonormal form) to  $2n + 2$ , naturally embedding parallelism without explicit constraints. We further integrate this parameterization into a factor graph framework, allowing global direction alignment and local projection optimization within a unified manifold-based bundle adjustment. Extensive experiments on ICL-NUIM, TartanAir, and synthetic benchmarks demonstrate that our method achieves significantly more accurate pose estimation and line reconstruction, while reducing parameter dimensionality and improving convergence stability.

**Code** — <https://github.com/yanyan-li/RiemanLine>

## 1 Introduction

Robustly reconstructing (Schonberger and Frahm 2016; Forster, Pizzoli, and Scaramuzza 2014) unknown three-dimensional scenes and estimating (Carlone et al. 2015; Carlone and Calafiore 2018) six degrees-of-freedom (6-DoF) camera poses from visual inputs are fundamental challenges in robotics and computer vision. However, these odometry and SLAM methods often suffer from structural inaccuracies and pose drift during the incremental camera tracking and mapping process. To mitigate these issues, techniques such as local bundle adjustment (Mur-Artal, Montiel, and Tardos 2015; Rosinol et al. 2020), sliding window optimization (Qin, Li, and Shen 2018; Engel, Koltun, and Cremers 2017), and loop closure (Labbe and Michaud 2019; Mur-Artal, Montiel, and Tardos 2015) are commonly

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

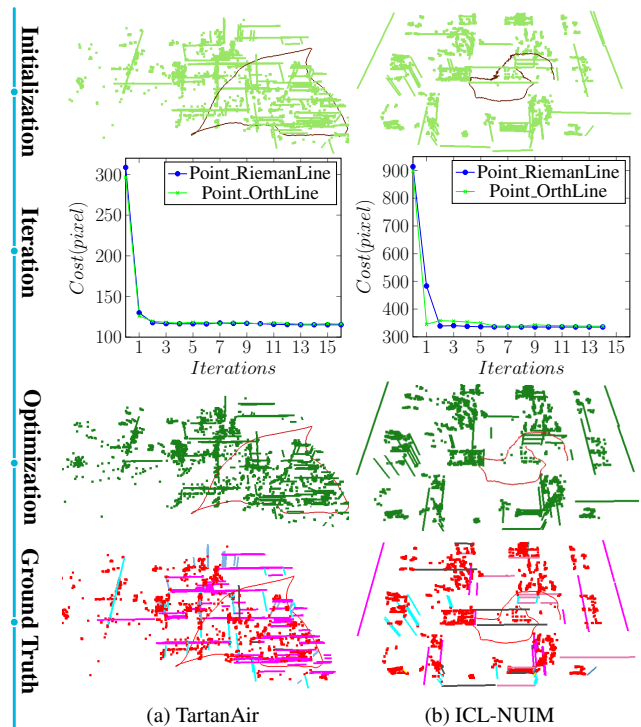


Figure 1: An illustration of co-visibility factor graph optimization based points and lines. The initial factor graphs are depicted in the first row, with landmarks and trajectories colored light green and scarlet, respectively. The convergence curves for different representations are plotted in the second row. The optimized results based on the proposed *Point\_RiemanLine* method and ground-truth graphs are presented in the last rows, respectively.

incorporated. The core of these techniques lies in the use of factor graph optimization (Carlone and Calafiore 2018), which jointly refines the structure and transformation parameters. In this paper, we present a novel solution to the line-leveraged factor graph optimization problem by proposing a compact line parameterization on a Riemannian manifold along with constraint factors that connect line landmarks to camera poses, as shown in Figure 1.

Point features have long served as the foundation for most visual pose estimation systems, as demonstrated by their widespread adoption in leading methods (Mur-Artal, Montiel, and Tardos 2015; Qin, Li, and Shen 2018; Rosinol et al. 2020). Despite their success, point features exhibit notable limitations, especially in challenging environments such as indoor scenes. To overcome the limitations of point-only factor graph optimization (Li et al. 2024), recent efforts have explored the integration of additional geometric primitives such as lines (Lu and Song 2015; Zuo et al. 2017) and planes (Zhou, Koppel, and Kaess 2021) into both tracking and optimization modules. Compared to plane detection which is typically based on depth maps (Salas-Moreno et al. 2014) or convolutional neural networks (Paigwar et al. 2020), line features can be efficiently extracted from RGB images, offering a versatile and computationally efficient means to enhance visual odometry performance. The most popular and widely adopted line parameterization method in line-based SLAM systems is the **Orthonormal algorithm** (Bartoli and Sturm 2005), which enables elegant optimization within the framework of **Lie Algebra**.

Typically, a single line segment contributes a re-projection factor (Hartley and Zisserman 2003) to the optimization module, whereas a collection of lines imparts broader structural and global regularities. Specifically, a group of parallel line segments on a 2D image plane crosses at the vanishing point, which can be projected to the camera coordinate to obtain the corresponding vanishing direction (McLean and Kotturi 1995) via a related intrinsic matrix. A group of vanishing direction vectors can model structured environments such as the Atlanta or Manhattan World assumptions (Straub et al. 2017). However, these assumptions are often too restrictive for general environments, and it is also difficult to optimize the structure as one primitive in factor graph optimization modules. Furthermore, conventional minimal parameterization (Bartoli and Sturm 2005) that is widely adopted in line-SLAM systems (Zuo et al. 2017; He et al. 2018) lacks the expressiveness to compactly parametrize a group of structure lines. Consequently, this leads to increased complexity and reduced efficiency since the optimization frameworks must rely on additional parameters and manually introduce constraints into loss functions to capture the relationships among these lines.

In this paper, we first propose a unified minimal representation for 3D lines, including individual and structure line landmarks, which has a clear geometric interpretation in representation and optimization based on Riemannian manifolds. Specifically, the line parameters are explicitly decoupled into **global** and **local** components: the global component defines the shared direction of parallel lines, and the local component that lies in the orthogonal plane encodes scaled normals that distinguish each line. Representing  $n$  parallel 3D lines with the Orthonormal representation requires  $4n$  parameters. In contrast, our method reduces this to only  $2n + 2$  parameters, significantly reducing the dimensionality while implicitly encoding parallelism without additional constraints. We integrate this representation into a joint factor graph framework with co-visibility factors, enabling both accurate pose estimation and structurally consis-

Algorithm	Parameters
Euclidean	$\mathbf{P}_s, \mathbf{P}_e$ endpoints of the 3D line
Plücker	$\mathbf{n}, \mathbf{v}$ $\mathbf{n} = [\mathbf{P}_s]_{\times} \mathbf{P}_e, \mathbf{v} = \mathbf{P}_s - \mathbf{P}_e$
Quaternion	$\bar{\mathbf{q}}, d$ $\mathcal{R}(\bar{\mathbf{q}}) = [\mathbf{n} \ \mathbf{v} \ [\mathbf{n}]_{\times} \mathbf{v}], d = \ \mathbf{n}\ /\ \mathbf{v}\ $
Closest Point	$\mathbf{j} = d\bar{\mathbf{q}}$
Spherical	$\theta, \phi, \alpha, d$ $\theta, \phi, \alpha$ are angles, $d$ is the distance

Table 1: Popular line representations and transformation relationships between them. Here  $[\cdot]_{\times}$  is the skew-s operation, and the  $\mathcal{R}$  is the transformation operation from quaternion to rotation matrix. The definitions of the remaining parameters are provided in Section 3.

tent line reconstruction. Our contributions are summarized as follows:

- We introduce a unified minimal parameterization for 3D lines on Riemannian manifolds that seamlessly extends from independent lines to sets of parallel lines by explicitly decoupling global and local components.
- A joint factor graph framework incorporating co-visibility and extensibility factors, specifically designed to leverage the proposed minimal representation;
- We conduct extensive evaluations on ICL-NUIM, TartanAir, and synthetic datasets, demonstrating that our method achieves higher accuracy with fewer parameters and improved convergence stability compared to existing representations.

## 2 Related work

As shown in Table 1, two stages of representation are typically involved in landmark reconstruction and optimization tasks (Bartoli and Sturm 2005). The first stage focuses on 3D triangulation from 2D measurements, and the second seeks a minimal parameterization for iterative refinement. The two stages can be unified into a single step when the degrees of freedom used in reconstruction meet the requirements of minimal parametrization.

Generally, Euclidean XYZ (Mur-Artal, Montiel, and Tardos 2015; Qin, Li, and Shen 2018) is used to parametrize endpoints of a finite line in the 3D space. Although this representation can be used to provide re-projection residuals of lines for camera pose optimization, it has the over-parameterization problems in landmark optimization. Similar to the *Euclidean*, a line at infinity can be represented as the *Plücker* coordinate via two three-dimensional vectors containing the direction of the line and normal based on the line and the camera coordinate frame as listed in Table 1. Furthermore, the dual quaternion approach (Kotkas and Roumeliotis 2013) represents a line in 3D space using dual quaternions, which provides a concise way to represent rotations and translations in 3D space and can be used to encode the position and orientation of lines. The advantage of this approach is that it allows for easy com-

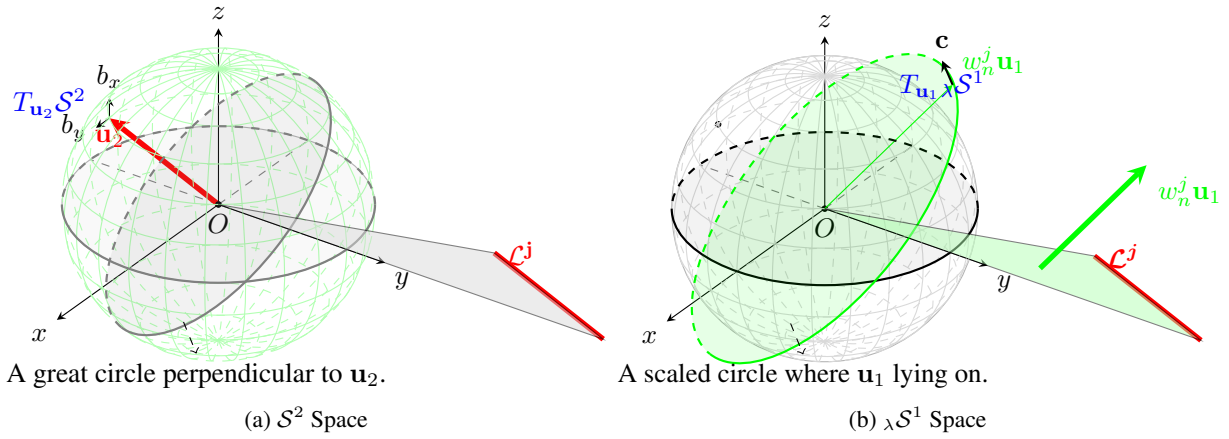


Figure 2: Illustration of the proposed parametrization for a line landmark  $\mathcal{L}^j$ . The vanishing direction vector  $\mathbf{u}_2$  ( $\|\mathbf{u}_2\| = 1$ ) and scaled normal vector  $\omega_n^j \mathbf{u}_1$  ( $\|\mathbf{u}_1\| = 1, \omega_n^j > 0$ ) are optimized on the tangent spaces  $T_{\mathbf{u}_2} \mathcal{S}^2$  of the sphere and  $T_{\mathbf{u}_1} \lambda \mathcal{S}^1$  of the scaled circle, respectively.

position of transformations, which makes it useful in applications such as robotics and animation. By multiplying the unit quaternion and the distance parameter, the *closest point* method (Yang and Huang 2019) can be considered as the ‘closest point’ for a 3D line. The transformation relationships between *Plücker*, *Quaternion* and *Closest Point* methods are listed in Table 1. The rotation matrix used in those representations can be optimized via *Lie algebras* which define the tangent spaces of related manifolds. For the *Spherical* form, a line can also be represented by three angles and a distance parameter. Instead of optimizing the orthogonal matrix through *Lie Groups* being special instances of the manifold, a more general Riemannian manifold is used to refine the vanishing and normal vectors subsequently in the proposed method.

Structure-SLAM (Li et al. 2020c) estimates the orientation using surface normals from monocular images, while Linear-SLAM (Joo et al. 2021) extracts planes from depth maps. Given a known rotation, the translation estimation becomes linear. The Manhattan World (MW) model assumes orthonormal landmarks and the Atlanta World (AW) model introduces multiple perpendicular horizontal directions. The multi-MW model (Yunus, Li, and Tombari 2021) enforces orthogonality within local regions instead of throughout the scene. However, such methods often neglect optimization within factor graphs. Recent approaches integrate structural constraints, lines and planes, directly into optimization. (Zhang et al. 2015) pioneered using parametric 3D lines as SLAM landmarks, enhancing traditional point-based methods. (Lu and Song 2015) extended this by incorporating diverse structural features within a multilayer feature graph (MFG) to improve environmental representation and pose estimation. Kimera-VIO (Rosinol et al. 2020) leverages 3D meshes to enforce coplanarity constraints within factor graphs. PLP-VIO (Li et al. 2020a) further refines this approach by incorporating line-based meshes to enhance spatial understanding. CoP (Li et al. 2020b) introduced a novel parameterization that represents points and lines using plane

parameters to preserve geometric consistency during optimization. Beyond scene-specific models, Struct-VIO (Zhou et al. 2015; Zou et al. 2019) mitigates directional errors and reducing drift by parameterizing line segments parallel to the local Manhattan world. Our method generalizes this by parameterizing all parallel lines, which uses the structural information to improve pose estimation accuracy in SLAM.

### 3 On-Manifold Representation of 3D Lines

#### Preliminary: Representation in *Orthonormal*

As shown in Table 1, a 3-dimensional finite line  $\mathbf{L}^w$  in the world coordinates can be represented by its two 3D endpoints,  $\mathbf{P}_s^w$  and  $\mathbf{P}_e^w$ , as  $\mathbf{L}^w = [\mathbf{P}_s^w \ \mathbf{P}_e^w]$ . For an infinite line, the direction and normal vectors can be used to represent the line  $\mathcal{L}^w = [\mathbf{n}^w \ \mathbf{d}^w]$  based on the *Plücker* representation, where  $\mathbf{n}^w$  and  $\mathbf{d}^w$  can be derived from the endpoints  $\mathbf{P}_s^w$  and  $\mathbf{P}_e^w$ . This can be further expressed as the multiplication of several matrices via the following function:

$$\begin{aligned} \mathcal{L}^w &= [\mathbf{n}^w \ \mathbf{d}^w] \\ &= \begin{bmatrix} \frac{\mathbf{n}^w}{\|\mathbf{n}^w\|} & \frac{\mathbf{d}^w}{\|\mathbf{d}^w\|} \end{bmatrix} \begin{bmatrix} \|\mathbf{n}^w\| & 0 \\ 0 & \|\mathbf{d}^w\| \end{bmatrix}. \end{aligned} \quad (1)$$

We then use  $\begin{bmatrix} \frac{\mathbf{n}^w}{\|\mathbf{n}^w\|} & \frac{\mathbf{d}^w}{\|\mathbf{d}^w\|} \end{bmatrix}$  and  $\begin{bmatrix} \|\mathbf{n}^w\| & 0 \\ 0 & \|\mathbf{d}^w\| \end{bmatrix}$  to establish the following matrices:

$$\begin{cases} \mathbf{U} = \begin{bmatrix} \frac{\mathbf{n}^w}{\|\mathbf{n}^w\|} & \frac{\mathbf{d}^w}{\|\mathbf{d}^w\|} & \frac{\mathbf{n}^w \times \mathbf{d}^w}{\|\mathbf{n}^w \times \mathbf{d}^w\|} \end{bmatrix}; \\ \mathbf{W} = \begin{bmatrix} \|\mathbf{n}^w\|/\lambda & -\|\mathbf{d}^w\|/\lambda \\ \|\mathbf{d}^w\|/\lambda & \|\mathbf{n}^w\|/\lambda \end{bmatrix}. \end{cases} \quad (2)$$

Here,  $\lambda$  denotes  $\sqrt{\|\mathbf{d}^w\|^2 + \|\mathbf{n}^w\|^2}$ ,  $\mathbf{U} \in SO(3)$ , and  $\mathbf{W} \in SO(2)$ .

In *quaternion-based* optimization methods (Yang and Huang 2019),  $\mathbf{U}$  is represented as a quaternion vector via  $\bar{\mathbf{q}} = \mathcal{R}^{-1}(\mathbf{U})$ , and the optimization process is implemented based on quaternion representation. With the widespread

application of differentiable manifolds in SLAM optimization problems, the Orthonormal representation (Bartoli and Sturm 2005) maps the update process of  $SO(3) \times SO(2)$  on the tangent space *Lie algebra* based on 4 degree-of-freedom  $[\rho^T \ \omega]^T$  that is the minimal line representation widely used in point-line SLAM (Zuo et al. 2017; Li et al. 2020b) and VIO (He et al. 2018) systems.

### Riemannian Representation of a Single 3D Line

We decouple the single line  $\mathcal{L}^j$  into two components: the unit direction vector  $\mathbf{u}_2$ , and the scaled normal vector  $\omega_n^j \mathbf{u}_1$ . As illustrated in Figure 2, the direction vector  $\mathbf{u}_2$  (with  $\|\mathbf{u}_2\| = 1$ ) lies on the surface of the unit sphere  $\mathcal{S}^2$ . The scaled normal vector  $\omega_n^j \mathbf{u}_1$  lies on the plane  ${}_\lambda \mathcal{S}^1$ , which is centered at the origin and orthogonal to  $\mathbf{u}_2$ .

**Optimization on  $\mathcal{S}^2$ .** As illustrated in Figure 2a, the tangent space to  $\mathcal{S}^2$  (at  $\mathbf{u}_2$ ) is given by:

$$T_{\mathbf{u}_2} \mathcal{S}^2 \doteq \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x}^T \mathbf{u}_2 = 0\}. \quad (3)$$

Here,  $\mathbf{x}^T \mathbf{u}_2 = 0$  shows that the vectors  $\mathbf{x}$  in the tangent space are perpendicular to  $\mathbf{u}_2$ . Although  $\mathcal{S}^2$  is a non-linear 3D space, its tangent space  $T_{\mathbf{u}_2} \mathcal{S}^2$  is a 2-DoF linear space. Based on the perpendicular bases  $\mathbf{b}_x$  and  $\mathbf{b}_y$ , any vector adjacent to  $\mathbf{u}_2$  on the unit sphere can be represented.

In the representation process,  $\delta\theta_1$  and  $\delta\theta_2$ ,  $[\delta\theta_1 \ \delta\theta_2]^T = \delta\theta$ , are corresponding disturbances towards  $\mathbf{b}_x$  and  $\mathbf{b}_y$ , respectively. Consequently, the perturbation  $\delta\mathbf{m}$  on the tangent space can be represented as:

$$\delta\mathbf{m} = [\mathbf{b}_x \ \mathbf{b}_y] \begin{bmatrix} \delta\theta_1 \\ \delta\theta_2 \end{bmatrix}, \quad (4)$$

and the additive operation on the unit sphere is denoted as  $\mathbf{u}_2 \boxplus \delta\mathbf{m}$ , which maps  $\delta\mathbf{m}$  back to  $\mathcal{S}^2$  via the *Riemannian exponential*  $Exp_{\mathbf{u}_2}(\delta\mathbf{m})$ , and the vector after perturbation is denoted as  $\mathbf{v}_2$  computed via:

$$\begin{aligned} \mathbf{v}_2 &= \mathbf{u}_2 \boxplus \delta\mathbf{m} \\ &= \mathbf{u}_2 \cos \|\delta\mathbf{m}\| + \frac{\delta\mathbf{m}}{\|\delta\mathbf{m}\|} \sin \|\delta\mathbf{m}\|, \end{aligned} \quad (5)$$

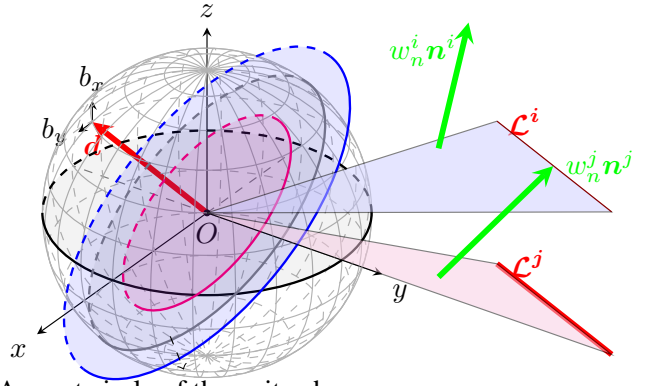
where the perturbation  $\delta\mathbf{m}$  is close to  $\mathbf{0}$ .

**Optimization on  ${}_\lambda \mathcal{S}^1$ .** Concurrently, the tangent space to  ${}_\lambda \mathcal{S}^1$  is denoted as:

$$T_{\mathbf{u}_1} {}_\lambda \mathcal{S}^1 \doteq \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x}^T \mathbf{u}_1 = \mathbf{x}^T \mathbf{u}_2 = 0\}. \quad (6)$$

Here,  $\mathbf{x}^T \mathbf{u}_1 = \mathbf{x}^T \mathbf{u}_2 = 0$  shows that the vectors on the tangent space are perpendicular to  $\mathbf{u}_1$  and  $\mathbf{u}_2$  at the same time. Although  ${}_\lambda \mathcal{S}^1$  is a non-linear 2D space, there is a base vector  $\mathbf{u}_3$  which is orthogonal to both  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . This base vector can be selected as  $\mathbf{u}_3$  since  $\mathbf{u}_3 = \mathbf{u}_2 \times \mathbf{u}_1$ . With the perturbation of  $\mathbf{u}_2$ , the tangent space of  $\mathbf{u}_1$  is also perturbed by the orthogonal relationship. Based on the updated global direction  $\mathbf{v}_2$ , we define the space  $T_{\mathbf{u}_1} {}_\lambda \mathcal{S}^1$  based on  $\mathbf{u}_3 = \mathbf{v}_2 \times \mathbf{u}_1$ .

For a small angle  $\delta\gamma$ , we rotate  $\mathbf{u}_1$  within the plane orthogonal to  $\mathbf{v}_2$  by an angle  $\delta\gamma$  to obtain the new vector  $\mathbf{v}_1$  by using the base vector of the tangent space. Based on the



A great circle of the unit sphere.

Figure 3: Illustration of the parametrization for two parallel line landmarks  $\mathcal{L}^i$  and  $\mathcal{L}^j$ . The vanishing direction vector  $\mathbf{u}_2$  and normalized normal vector  $\mathbf{u}_1$  are optimized on the tangent spaces  $T_{\mathbf{u}_2} \mathcal{S}^2$  of the sphere and  $T_{\mathbf{u}_1} \mathcal{S}^1$  of the circle, respectively.

Rodrigues rotation formula, the updated direction can be denoted as

$$\mathbf{v}_1 = \lambda \cdot (\cos(\delta\gamma) \mathbf{u}_1 + \sin(\delta\gamma) \mathbf{u}_3), \quad (7)$$

which rotates  $\mathbf{u}_1$  within the plane orthogonal to  $\mathbf{v}_2$  by angle  $\delta\gamma$ .  $\lambda \in \mathbb{R}_+$  is a scalar magnitude.

### Unified Representation for Parallel Line Sets

As illustrated in Figure 3, our method naturally extends to a set of parallel lines by building upon the representation of a single 3D line. For a group of  $k$  parallel lines denoted as  $\mathbb{S} = [\mathcal{L}_0^w, \mathcal{L}_1^w, \dots, \mathcal{L}_k^w]$ , the shared vanishing direction is represented by a single unit vector  $\mathbf{u}_2$ . Each line  $\mathcal{L}_i^w$  in the group is characterized by a local component  $\mathbf{u}_1^i$  and a scale  $\omega^i$ . This results in a compact representation of the set as:

$$[\mathbf{u}_2, \omega^0 \mathbf{u}_1^0, \omega^1 \mathbf{u}_1^1, \dots, \omega^k \mathbf{u}_1^k], \quad (8)$$

where all  $\omega^i \mathbf{u}_1^i$  lying on the circle  $\omega^i \mathcal{S}^1$  perpendicular to  $\mathbf{u}_2$ . This leads to a minimal representation of the entire line group with  $2 + 2k$  degrees of freedom: 2 DoF for the shared global direction  $\mathbf{u}_2$ , and 2 DoF per line for its local component.

As shown in Figure 3,  $\mathbf{u}_2$  lies on  $\mathcal{S}^2$  and each  $\mathbf{u}_1^i$  lies on the associated circle orthogonal to  $\mathbf{u}_2$  by the minimal representation on the unit sphere. Consequently, the minimal parameterization of the parallel line group follows the same format as the single-line case and can be written as:

$$[\delta\theta, \delta\gamma^0, \dots, \delta\gamma^k, \lambda^0, \dots, \lambda^k], \quad (9)$$

where  $\delta\theta$  encodes the global direction  $\mathbf{u}_2$ ,  $\delta\gamma^i$  represents the angular perturbation of  $\mathbf{u}_1^i$  on the circle, and  $\lambda^i$  denotes the perturbation corresponding distance scale. Since parallel lines  $\mathcal{L}^i$  and  $\mathcal{L}^j$  share the same direction vector  $\mathbf{u}_2$ , the global vector  $\mathbf{u}_2$  can thus be optimized via Equation 4 and 5. The associated local vectors  $\omega^i \mathbf{u}_1^i$  and  $\omega^j \mathbf{u}_1^j$  lie on the

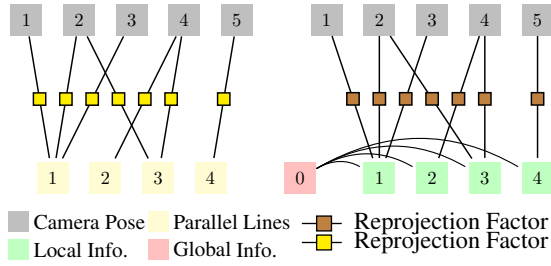


Figure 4: Factor graph representations for different line-based structures. Left: conventional line re-projection factors. Right: the proposed parallel line representation, explicitly separating global and local components with re-projection factors.

same plane defined by this tangent plane, therefore these local components can be optimized separately.

This unified formulation demonstrates the versatility of our approach that enables seamless extension from individual line landmarks to structurally consistent representations of parallel line groups within a common manifold-based optimization framework.

## 4 Optimization with Points and Lines

### Graph Construction

The vertices in the point-line factor graphs  $\mathcal{G}$  include camera poses  $\mathcal{V}_{pose}$ , point landmarks  $\mathcal{V}_p$ , line landmarks  $\mathcal{V}_l$  and parallel-line sets  $\mathcal{V}_{para}$ . Specifically, camera pose  $\mathbf{T}_{w,c_i} = \begin{bmatrix} \mathbf{R}_{w,c_i} & \mathbf{t}_{w,c_i} \\ \mathbf{0} & 1 \end{bmatrix}$ , where  $\mathbf{T}_{w,c_i} \in SE(3)$ ,  $\mathbf{R}_{w,c_i} \in SO(3)$ , and  $\mathbf{t}_{w,c_i} \in \mathbb{R}^3$ . Points used in the optimization module is parametrized as  $\mathbf{P}_w^k = [x^k \ y^k \ z^k]^T$ , and line landmarks are represented in minimal parameterization forms.

### Factors and Constraints

Co-visibility connections (Mur-Artal, Montiel, and Tardos 2015; Mur-Artal and Tardós 2017; Campos et al. 2021) are built when two images detect the same landmark, such as a point or a line on the map. In this section, we define the co-visibility factors based on the re-projection models of point and line features.

**Co-visibility factors from points.** Based on the point feature measurement model, the measurement of the  $k^{th}$  global point landmark  $\mathbf{P}_w^k$  at frame  $c_j$  is represented as  $\bar{\mathbf{p}}_k^j$  in the normalized coordinate, and the re-projection factor of a point feature is defined as  $\mathbf{r}_p(\bar{\mathbf{p}}_k^j, \mathbf{P}_w^k, \mathbf{T}_{w,c_j})$ .

**Co-visibility factors from lines.** Traditionally, the mapline  $\mathcal{L}_w^i$  in the world coordinate is transferred to the camera coordinates, and then re-projected on the image plane  $\mathbf{l}_k^j$  of viewpoint  $c_j$ . The error between the re-projected line  $\mathbf{l}_k^j$  and the two endpoints  $\bar{\mathbf{p}}_{k,s}^j$  and  $\bar{\mathbf{p}}_{k,e}^j$  of the extracted 2D line can be written as:

$$\mathbf{r}_l(\bar{\mathbf{p}}_{k,s}^j, \bar{\mathbf{p}}_{k,e}^j, \mathcal{L}_w^i, \mathbf{T}_{w,c_j}) = \begin{bmatrix} \text{dis}(\bar{\mathbf{p}}_{k,s}^j, \mathbf{l}_k^j) \\ \text{dis}(\bar{\mathbf{p}}_{k,e}^j, \mathbf{l}_k^j) \end{bmatrix}, \quad (10)$$

where  $\text{dis}(\cdot)$  gives the distance between a point and a line.

**Constraints between parallel lines.** To enforce structural consistency, we incorporate additional constraints among groups of parallel lines. Specifically, for each mapline  $\mathcal{L}_w^i$  in a set of  $N$  parallel lines  $\{\mathcal{L}_w^i\}_{i=1}^N$ , its direction should be parallel to other lines in the set. The parallelism constraint is enforced by minimizing the angular deviation between the direction of each line  $\mathbf{u}_2^j$ . The residual is defined as:

$$\mathbf{r}_{\parallel}(\mathcal{L}_w^i, \{\mathcal{L}_w^j\}_{j=1, j \neq i}^N) = \frac{1}{N-1} \sum_{j=1, j \neq i}^N (1 - \mathbf{u}_2^{i\top} \mathbf{u}_2^j). \quad (11)$$

Here, the residual  $\mathbf{r}_{\parallel}$  encourages all lines in the group to remain parallel to each other during optimization.

## 5 Experiments

We evaluate the proposed RiemannLine and StructRiemannLine parameterizations on both public benchmarks and simulated environments. Our goals are to assess: (i) pose estimation accuracy, (ii) landmark reconstruction quality, and (iii) the benefits of incorporating structural constraints within a unified Riemannian manifold representation.

**Baselines.** We compare against standard parameterizations widely used in SLAM and VO: **Euclidean XYZ** (Mur-Artal, Montiel, and Tardos 2015), **Orthonormal** (Bartoli and Sturm 2005), and **ImplicitLine** (Zhao et al. 2015). The Orthonormal form serves as the minimal representation for Closest Point (Yang and Huang 2019) and Quaternion (Kotzas and Roumeliotis 2013) methods. All baselines are fed with the same co-visibility factor graph in each sequence.

**Datasets.** We use the ICL-NUIM benchmark (Handa et al. 2014) (eight indoor RGB-D sequences) and four TartanAir (Wang et al. 2020) sequences featuring photorealistic synthetic environments. Additionally, we evaluate on three challenging simulation sequences (corridor, box, sphere) generated using Open-Structure (Li et al. 2024).

**Metrics.** We report Absolute Trajectory Error (ATE) for camera localization, and angular errors (direction and normal) for line reconstruction. All results are averaged per sequence. Computations were performed on an Intel NUC Mini PC with Core i7-8700 CPU.

### Evaluation on ICL-NUIM

Table 2 summarizes ATE RMSE and MEDIAN errors. The initial factor graphs show significant drift (e.g., RMSE 12.13 cm on *livingroom0*). For brevity, we refer to **Point\_StructRiemannLine** as **StructRiemannLine** in tables. The proposed **Point\_StructRiemannLine** achieves the best overall accuracy across all eight sequences. By explicitly encoding parallel-line constraints into the minimal Riemannian representation, it reduces errors without introducing additional parameters. For example, *livingroom2* achieves RMSE 0.75 cm (MEDIAN 0.68 cm), and *office2* maintains RMSE 0.75 cm while preserving robust convergence behavior.

Sequence	Initial		Optimization using Independent Primitives								Optimization using Structure Constraints			
	Factor Graph		Point		OrthLine		ImplicitLine		RiemanLine		OrthLine_Constr		StructRiemanLine	
	RMSE	Med.	RMSE	Med.	RMSE	Med.	RMSE	Med.	RMSE	Med.	RMSE	Med.	RMSE	Med.
livingroom0	12.13	1.95	1.14	0.69	0.81	0.38	0.96	0.40	<b>0.81</b>	0.38	<b>0.81</b>	0.38	<b>0.81</b>	<b>0.37</b>
livingroom1	2.15	1.57	1.67	1.24	4.13	4.22	0.92	0.81	1.19	0.71	3.58	3.60	<b>0.90</b>	<b>0.55</b>
livingroom2	12.95	6.95	1.61	0.80	0.79	0.71	1.18	0.90	0.78	0.71	0.79	0.72	<b>0.75</b>	<b>0.68</b>
livingroom3	18.26	13.93	11.64	11.43	6.70	6.33	6.67	6.76	6.71	6.34	6.71	6.34	<b>6.65</b>	<b>6.33</b>
office0	1.73	0.69	0.70	0.34	<b>0.38</b>	<b>0.26</b>	0.47	0.37	<b>0.38</b>	<b>0.26</b>	<b>0.38</b>	<b>0.26</b>	<b>0.38</b>	<b>0.26</b>
office1	11.93	5.48	7.68	5.22	1.81	0.89	9.78	7.81	1.89	0.87	1.81	0.87	<b>1.76</b>	<b>0.73</b>
office2	2.95	1.88	0.93	0.55	<b>0.75</b>	0.56	<b>0.75</b>	<b>0.53</b>	<b>0.75</b>	0.56	0.76	0.56	<b>0.75</b>	0.56
office3	7.56	2.47	1.22	0.66	0.78	0.49	1.86	0.77	0.79	0.51	0.85	0.53	<b>0.77</b>	<b>0.46</b>

Table 2: Comparison of translation (APE) RMSE and median errors (cm) on the ICL-NUIM (Handa et al. 2014) benchmark dataset. The best results are highlighted in **bold**.

Sequence	Carwelding		Hospital		Office		Jpn. Alley	
	Trans.	Rot.	Trans.	Rot.	Trans.	Rot.	Trans.	Rot.
Initial	14.74	10.89	33.78	26.83	57.00	47.78	10.92	5.23
OrthLine	4.47	0.16	5.32	4.83	43.25	29.23	3.75	2.80
OrthLine_Constr	4.56	0.16	5.49	4.97	50.04	39.50	3.68	2.72
RiemanLine	4.46	0.16	7.29	6.60	40.95	26.70	3.85	2.92
StructRiemanLine	<b>4.08</b>	<b>0.15</b>	<b>2.91</b>	<b>2.04</b>	<b>15.60</b>	<b>12.08</b>	<b>3.65</b>	<b>2.68</b>

Table 3: Comparison of APE RMSE (translation (cm) and rotation (degree) on the TartanAir dataset.

Method	Carweld.		Hospital		Office		Jpn. Alley	
	Dir.	Norm.	Dir.	Norm.	Dir.	Norm.	Dir.	Norm.
OrthLine	1.52	1.42	5.16	4.11	3.89	4.49	0.88	0.86
OrthLine_Constr	0.92	0.93	5.05	2.54	2.04	2.28	0.46	0.42
RiemanLine	1.55	1.43	5.38	4.68	3.89	4.48	0.88	0.73
StructRiemanLine	<b>0.82</b>	<b>0.83</b>	<b>1.60</b>	<b>0.91</b>	<b>0.04</b>	<b>0.51</b>	<b>0.07</b>	<b>0.15</b>

Table 4: Comparison of line reconstruction performance on TartanAir. Median errors (degrees) are reported for direction (Dir.) and normal (Norm.).

## Evaluation on TartanAir

Table 3 reports ATE translation and rotation errors on the TartanAir benchmark. The proposed **StructRiemanLine** achieves the most significant improvement in structurally rich indoor sequences. On *Hospital*, the translation RMSE drops from 33.78 cm (Initial) and 7.29 cm (**RiemanLine**) to only 2.91 cm, while the rotation RMSE decreases from 26.83° to 2.04°. Similarly, on *Office*, the translation error is reduced by over **40%** compared to Point\_OrthLine and Point\_RiemanLine, demonstrating that the proposed parallel-line manifold representation provides strong orientation priors that are especially beneficial in man-made environments with dominant structural regularities. These results validate that explicitly encoding shared vanishing directions within the factor graph not only improves translational accuracy but also significantly enhances rotational consistency, a critical factor for large indoor scene reconstruction.

**Line reconstruction.** Table 4 reports median angular errors. StructRiemanLine achieves the lowest direction and normal errors across all TartanAir sequences. In *Hospital*, the direction error is reduced to 1.60° and the normal error to 0.91°, validating that structural encoding improves both pose estimation and landmark quality. Figure 5 shows

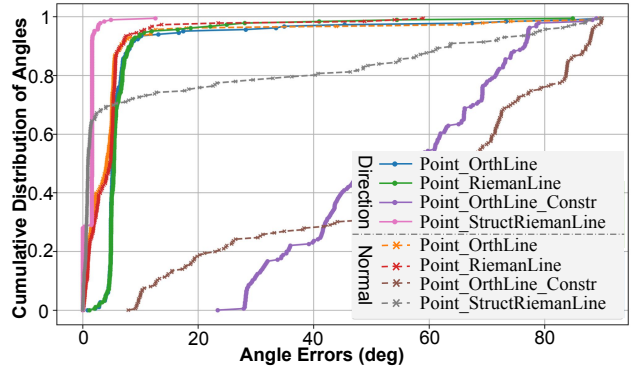


Figure 5: Line reconstruction errors of different methods in the *Hospital* sequence.

the cumulative distribution of angular errors for line direction and normal estimation on the *Hospital* sequence. The proposed StructRiemanLine achieves the steepest rise and highest saturation, indicating that over 90% of reconstructed lines fall within 2° of the ground truth for both direction and normal vectors. Compared to Point\_Orthonormal and Point\_RiemanLine, which exhibit heavier tails beyond 5°, StructRiemanLine demonstrates significantly reduced variance. This confirms that encoding parallelism within the Riemannian manifold not only lowers median error (Table 4) but also improves the consistency of line reconstruction.

## Simulations

Figure 7 illustrates the synthetic environments that are *corridor* and *sphere*. In the corridor scenario, a rectangular arrangement of green points forms a corridor-like structure with blue structural lines embedded within, while a red trajectory traces a loop along the corridor’s center. The sphere scenario depicts a roughly spherical distribution of green points surrounded by blue lines, where the red trajectory forms a dense circular pattern around the sphere’s surface.

Table 5 reports ATE RMSE. While both RiemanLine and Orthonormal perform similarly on simple structures, StructRiemanLine consistently achieves the lowest translation and rotation errors. In the *sphere* scenario, it reduces translation RMSE to 1.16 cm and maintains minimal orientation drift, confirming the benefits of enforcing parallelism con-

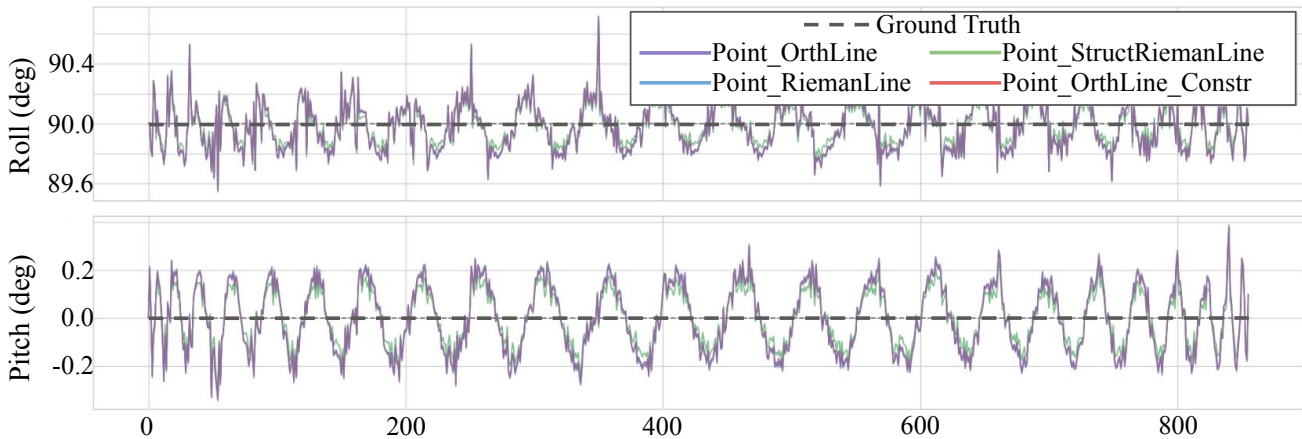


Figure 6: Pitch and roll angle errors for the sphere sequence in the simulation dataset. The plots compare the estimated orientations of different methods against the ground truth.

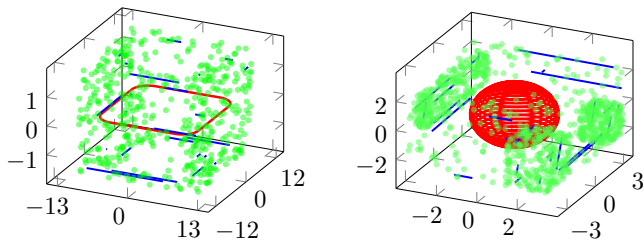


Figure 7: Visualization of the simulation scenarios: *corridor* and *sphere*. Points, lines, and trajectories are shown in green, blue, and red, respectively.

Method	corridor		box		sphere	
	Trans.	Rot.	Trans.	Rot.	Trans.	Rot.
Initial	10.57	1.54	9.67	0.85	25.99	5.49
OrthLine	4.24	0.32	2.09	<b>0.23</b>	1.19	<b>0.19</b>
OrthLine_Constr	4.29	0.32	2.09	<b>0.23</b>	1.19	<b>0.19</b>
RiemanLine	4.24	0.32	2.09	<b>0.23</b>	1.20	<b>0.19</b>
StructRiemanLine	<b>4.02</b>	<b>0.31</b>	<b>2.08</b>	<b>0.23</b>	<b>1.16</b>	<b>0.19</b>

Table 5: Comparison of translation (APE RMSE (cm)) and rotation (APE RMSE (degree)) on the simulation dataset.

straints in structurally symmetric environments.

Figure 6 further compares pitch and roll estimation errors in the *sphere* simulation. Point\_OrthoNormal and Point\_RiemanLine exhibit small but noticeable oscillations, while StructRiemanLine maintains the closest alignment with ground truth across the trajectory.

**Runtime and Complexity Analysis.** Table 6 reports the parameter complexity and solver runtime on the *box* simulation sequence, which contains 447 camera poses, 130 point landmarks, 20 line landmarks, and 2 parallel-line groups. Landmarks observed by fewer than three cameras are pruned prior to optimization to ensure a consistent and well-constrained factor graph. Starting from the *Point-only* baseline, the inclusion of 20 line landmarks in

Method	Para. Blks.	Eff. Params.	Resi. Blks.	Resi.	Time (s)
<i>Point</i>	577	3072	36010	37010	53.90
OrthLine_Constr	597	3152	41773	83452	97.62
StructRiemanLine	579	3116	41679	83358	49.62

Table 6: Comparison of parameter complexity and total solver runtime between the conventional Orthonormal representation with additional parallelism constraints and the proposed minimal parametrization on the *box* sequence

Point\_OrthLine\_Constr introduces 20 additional parameter blocks and **80** new parameters due to the conventional  $4n$  minimal parameterization. In contrast, the proposed **StructRiemanLine** requires only **44** parameters to represent the same 20 lines, as 19 of them are encoded within two parallel-line groups under the compact  $2n + 2$  formulation.

In terms of computational efficiency, the total optimization time drops from **97.62 s** to **49.62 s**, yielding an approximate 49% overall speed-up. This improvement is consistent with the reduction in parameter blocks ( $597 \rightarrow 579$ ) and state dimensionality ( $3152 \rightarrow 3116$ ), which together produce a sparser and better-conditioned Hessian.

## 6 Conclusion and Future Works

We have presented a novel minimal representation for 3D line landmarks. Unlike conventional approaches, the proposed parameterization not only encodes individual lines with minimal degrees of freedom, but also naturally accommodates structural regularities such as sets of parallel lines. Building upon this representation, we introduced a joint factor graph framework that integrates both co-visibility and structural factors, enabling more accurate and efficient optimization of camera poses and landmarks compared to traditional point-line co-visibility graphs.

Looking ahead, we envision incorporating the proposed parameterization into full SLAM and visual odometry pipelines to achieve more robust tracking and high-fidelity reconstruction in large-scale environments.

## Acknowledgments

We would like to thank Xin Li for insightful discussions and constructive feedback. This research was supported by the Tier 2 Grant (MOE-T2EP20124-0015) from the Singapore Ministry of Education.

## References

- Bartoli, A.; and Sturm, P. 2005. Structure-from-motion using lines: Representation, triangulation, and bundle adjustment. *Computer vision and image understanding*, 100(3): 416–441.
- Campos, C.; Elvira, R.; Rodríguez, J. J. G.; Montiel, J. M.; and Tardós, J. D. 2021. Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam. *IEEE Transactions on Robotics*, 37(6): 1874–1890.
- Carlone, L.; and Calafiore, G. C. 2018. Convex relaxations for pose graph optimization with outliers. *IEEE Robotics and Automation Letters*, 3(2): 1160–1167.
- Carlone, L.; Tron, R.; Daniilidis, K.; and Dellaert, F. 2015. Initialization techniques for 3D SLAM: A survey on rotation estimation and its use in pose graph optimization. In *2015 IEEE international conference on robotics and automation (ICRA)*, 4597–4604. IEEE.
- Engel, J.; Koltun, V.; and Cremers, D. 2017. Direct sparse odometry. *IEEE transactions on pattern analysis and machine intelligence*, 40(3): 611–625.
- Forster, C.; Pizzoli, M.; and Scaramuzza, D. 2014. SVO: Fast semi-direct monocular visual odometry. In *2014 IEEE international conference on robotics and automation (ICRA)*, 15–22. IEEE.
- Handa, A.; Whelan, T.; McDonald, J.; and Davison, A. J. 2014. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM. In *2014 IEEE international conference on Robotics and automation (ICRA)*, 1524–1531. IEEE.
- Hartley, R.; and Zisserman, A. 2003. *Multiple view geometry in computer vision*. Cambridge university press.
- He, Y.; Zhao, J.; Guo, Y.; He, W.; and Yuan, K. 2018. PL-VIO: Tightly-coupled monocular visual-inertial odometry using point and line features. *Sensors*, 18(4): 1159.
- Joo, K.; Kim, P.; Hebert, M.; Kweon, I. S.; and Kim, H. J. 2021. Linear RGB-D SLAM for structured environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11): 8403–8419.
- Kottas, D. G.; and Roumeliotis, S. I. 2013. Efficient and consistent vision-aided inertial navigation using line observations. In *2013 IEEE International Conference on Robotics and Automation*, 1540–1547. IEEE.
- Labbé, M.; and Michaud, F. 2019. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics*, 36(2): 416–446.
- Li, X.; He, Y.; Lin, J.; and Liu, X. 2020a. Leveraging planar regularities for point line visual-inertial odometry. In *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 5120–5127. IEEE.
- Li, X.; Li, Y.; Örnek, E. P.; Lin, J.; and Tombari, F. 2020b. Co-planar parametrization for stereo-SLAM and visual-inertial odometry. *IEEE Robotics and Automation Letters*, 5(4): 6972–6979.
- Li, Y.; Brasch, N.; Wang, Y.; Navab, N.; and Tombari, F. 2020c. Structure-slam: Low-drift monocular slam in indoor environments. *IEEE Robotics and Automation Letters*, 5(4): 6583–6590.
- Li, Y.; Guo, Z.; Yang, Z.; Sun, Y.; Zhao, L.; and Tombari, F. 2024. Open-Structure: Structural Benchmark Dataset for SLAM Algorithms. *IEEE Robotics and Automation Letters*.
- Lu, Y.; and Song, D. 2015. Visual navigation using heterogeneous landmarks and unsupervised geometric constraints. *IEEE Transactions on Robotics*, 31(3): 736–749.
- McLean, G. F.; and Kotturi, D. 1995. Vanishing point detection by line clustering. *IEEE Transactions on pattern analysis and machine intelligence*, 17(11): 1090–1095.
- Mur-Artal, R.; Montiel, J. M. M.; and Tardos, J. D. 2015. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE transactions on robotics*, 31(5): 1147–1163.
- Mur-Artal, R.; and Tardós, J. D. 2017. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE transactions on robotics*, 33(5): 1255–1262.
- Paigwar, A.; Erkent, Ö.; Sierra-Gonzalez, D.; and Laugier, C. 2020. GndNet: Fast ground plane estimation and point cloud segmentation for autonomous vehicles. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2150–2156. IEEE.
- Qin, T.; Li, P.; and Shen, S. 2018. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Transactions on Robotics*, 34(4): 1004–1020.
- Rosinol, A.; Abate, M.; Chang, Y.; and Carlone, L. 2020. Kimera: an open-source library for real-time metric-semantic localization and mapping. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 1689–1696. IEEE.
- Salas-Moreno, R. F.; Glocken, B.; Kelly, P. H.; and Davison, A. J. 2014. Dense planar SLAM. In *2014 IEEE international symposium on mixed and augmented reality (ISMAR)*, 157–164. IEEE.
- Schonberger, J. L.; and Frahm, J.-M. 2016. Structure-from-motion revisited. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4104–4113.
- Straub, J.; Freifeld, O.; Rosman, G.; Leonard, J. J.; and Fisher, J. W. 2017. The Manhattan frame model—Manhattan world inference in the space of surface normals. *IEEE transactions on pattern analysis and machine intelligence*, 40(1): 235–249.
- Wang, W.; Zhu, D.; Wang, X.; Hu, Y.; Qiu, Y.; Wang, C.; Hu, Y.; Kapoor, A.; and Scherer, S. 2020. Tartanair: A dataset to push the limits of visual slam. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4909–4916. IEEE.
- Yang, Y.; and Huang, G. 2019. Aided Inertial Navigation: Unified Feature Representations and Observability Analysis. In *2019 International Conference on Robotics and Automation (ICRA)*, 3528–3534.

- Yunus, R.; Li, Y.; and Tombari, F. 2021. Manhattanslam: Robust planar tracking and mapping leveraging mixture of manhattan frames. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 6687–6693. IEEE.
- Zhang, G.; Lee, J. H.; Lim, J.; and Suh, I. H. 2015. Building a 3-D line-based map using stereo SLAM. *IEEE Transactions on Robotics*, 31(6): 1364–1377.
- Zhao, L.; Huang, S.; Yan, L.; and Dissanayake, G. 2015. A new feature parametrization for monocular SLAM using line features. *Robotica*, 33(3): 513–536.
- Zhou, H.; Zou, D.; Pei, L.; Ying, R.; Liu, P.; and Yu, W. 2015. StructSLAM: Visual SLAM with building structure lines. *IEEE Transactions on Vehicular Technology*, 64(4): 1364–1375.
- Zhou, L.; Koppel, D.; and Kaess, M. 2021. LiDAR SLAM with plane adjustment for indoor environment. *IEEE Robotics and Automation Letters*, 6(4): 7073–7080.
- Zou, D.; Wu, Y.; Pei, L.; Ling, H.; and Yu, W. 2019. StructVIO: Visual-inertial odometry with structural regularity of man-made environments. *IEEE Transactions on Robotics*, 35(4): 999–1013.
- Zuo, X.; Xie, X.; Liu, Y.; and Huang, G. 2017. Robust visual SLAM with point and line features. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1775–1782. IEEE.