COIN-BIEVR: 3D Intensity Mapping for Robust LiDAR-Inertial Odometry

Patrick Pfreundschuh[1], Cedric Le Gentil[2], Roland Siegwart[1], and Cesar Cadena[3]

***Abstract*— Purely geometry-based LiDAR-Inertial Odometry (LIO) often fails in geometrically degenerate environments, like tunnels or flat fields. While current intensity-augmented methods mitigate this, they rely on dense intensity images for feature detection and gradient calculation. This introduces errors from approximated projection models and is incompatible with the irregular scan patterns of many modern sensors. To address this, we propose COIN-BIEVR, which integrates intensity directly into a high-resolution 3D map. Rather than performing 2D image-feature detection, we propose a map-informed sampling to identify informative intensity points. These points are used to extend a geometry-based LIO framework with a photometric optimization that calculates intensity gradients directly on the 3D representation without an intermediate projection model and thereby supports diverse scan patterns. Experiments across multiple datasets demonstrate that COIN-BIEVR significantly improves robustness in degenerate scenarios while maintaining or improving accuracy in geometrically rich environments.**

I. Introduction

LiDAR–Inertial odometry (LIO) has become a widely adopted solution for state estimation in robotics due to its strong robustness and accuracy across diverse environments. However, purely geometry-based LIO remains susceptible to failure in environments with insufficient structural constraints such as in corridors, tunnels, or planar environments, which can lead to degraded accuracy or failure. While BIEVR-LIO [1] mitigates this by leveraging fine-grained geometry through a high-resolution map representation, geometric representations alone cannot resolve scenarios that are fully degenerate, such as completely featureless tunnels. To this end, other LIO methods [2, 3] integrate LiDAR intensity as a complementary modality and project it into 2D images to extract features and construct intensity gradients. However, these projections only approximate the true LiDAR sensing model, which introduces errors during optimization and the dependency on dense images prohibits the application to sensors with irregular scanning patterns that are gaining popularity due to their lower cost.

COIN-BIEVR overcomes these limitations by integrating intensity directly into a dense 3D map representation. We augment BIEVR [1] with voxel-wise intensity images, enabling the construction of dense gradients directly in 3D space. This formulation eliminates projection errors during the photometric optimization and supports irregular LiDAR patterns. As classical image-based feature detection used in [2, 3] is not feasible for irregular scan patterns, we propose a novel map-informed sampling to identify regions with informative intensity. Our experiments demonstrate consistent improvements in robustness across multiple datasets, containing both dense and irregular scanning systems.

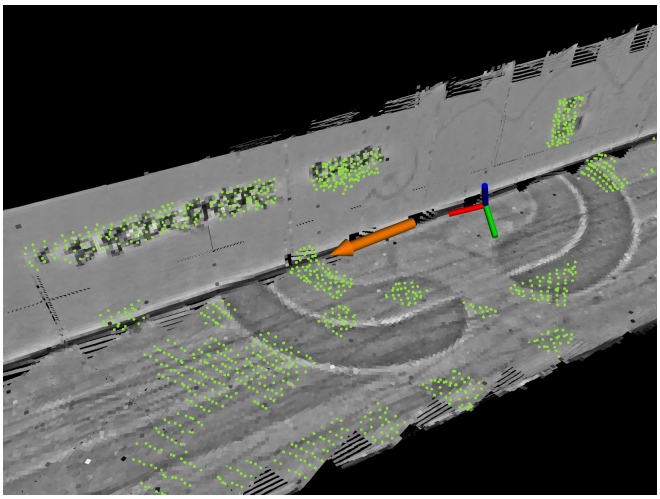


Fig. 1. COIN-BIEVR provides robust odometry in the geometrically degenereate *TunnelS* environment. The intensity map is visualized in gray colors and the estimated geometrically degenerate direction is indicated with an orange arrow. Intensity points (green) are selected in areas with informative gradients in the degenerate direction.

II. Method

COIN-BIEVR adapts intensity-based concepts from COIN-LIO into the BIEVR-LIO framework by augmenting it with voxel-wise intensity maps. This enables photometric optimization in 3D space, instead of a projected 2D image space, which supports irregular LiDAR scan patterns. To improve intensity consistency, we adapt the COIN-LIO normalization strategy to irregular scans and introduce a map-informed sampling strategy to select points that provide complementary information to existing geometric constraints.

A. Intensity Processing

A key challenge for intensity signals in photometric optimization is their dependence not only on intrinsic surface reflectivity, but also on external factors such as range, incidence angle, and sensor-specific near-range effects. When not specifically compensated for, these factors cause inconsistencies within and between scans, hindering intensity-based registration processes that assume photometric consistency.

To mitigate these effects, we adapt the image-based filter from COIN-LIO to support irregular scan patterns. To this end, we first construct an intensity image $\mathbf{I}_I$ using a spherical projection model that maps a point ${}_\mathrm{L}\mathbf{p} = [x, y, z]$ from

[1]Autonomous Systems Lab, ETH Zürich, Switzerland.
[2]Robotic Systems Lab, ETH Zürich, Switzerland.
[3]Mobile Robotics Lab, ETH Zürich, Switzerland.
This work was supported by Swiss National Science Foundation's NCCR DFab P3.

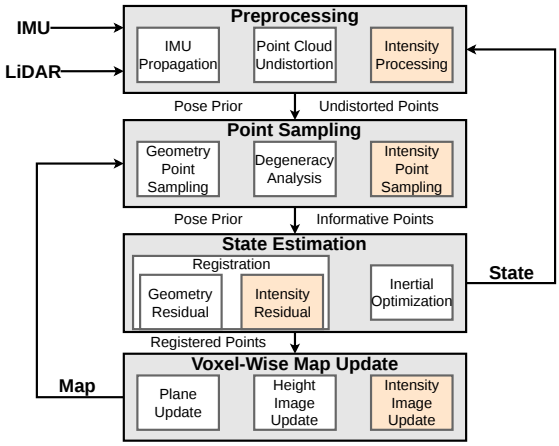


Fig. 2. Overview of COIN-BIEVR. Orange blocks indicate contributions in this work, while other blocks are adopted from BIEVR-LIO.

LiDAR frame L to image coordinates i, j:

$$\begin{bmatrix} i \\ j \end{bmatrix} = \Pi({}_{\mathrm{L}}\mathbf{p}) = \begin{bmatrix} \frac{-w}{2\pi} \operatorname{atan2}(y, x) + \frac{w}{2} \\ \frac{-h}{\Theta_{fov}} \arcsin(\frac{z}{\|{}_{\mathrm{L}}\mathbf{p}\|}) + \frac{h}{2} \end{bmatrix} \tag{1}$$

where Θ_{fov}, w and h denote the vertical field of view and horizontal and vertical resolution of the LiDAR. Importantly, the image serves only as an intermediate representation to define local neighborhoods for normalization and thus remains effective regardless of pixel sparsity or discrepancies between the projection model and the LiDAR's real geometry.

We apply a window-based normalization similar to COIN-LIO to account for smoothly varying brightness driven by global scene structure. We construct a brightness map $\mathbf{I}_B(i, j)$ by averaging intensity values within a large window, with the key modification that only non-empty pixels are included, ensuring robustness to sparsity. The final normalized intensity $\mathbf{I}_F(i, j)$ is scaled to $[0, 255]$ by a constant s:

$$\mathbf{I}_F(i, j) = s \cdot \frac{\mathbf{I}_I(i, j)}{\mathbf{I}_B(i, j) + 1} \tag{2}$$

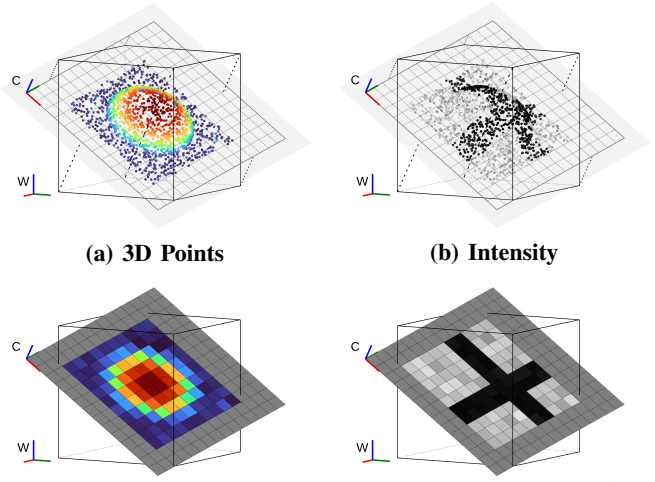
B. Voxel-wise Intensity Map

Approaches such as PG-LIO and COIN-LIO detect features in 2D intensity images and store the resulting patches as 3D points. However, doing so requires dense image representations for feature detection and is therefore not well suited for sparse or irregular scan patterns.

In contrast, COIN-BIEVR integrates intensity directly into the dense map representation of BIEVR-LIO, which models 3D surfaces using voxel-wise oriented height images $\mathbf{I}_H$. The image plane is iteratively estimated as the dominant surface in voxels of $0.5\,\mathrm{m}$ length, as illustrated in Figure 3. Building on this formulation, we augment each voxel with photometric information using an additional image layer $\mathbf{I}_P$ to capture pixel-wise intensity information. We call this layer intensity map, to avoid confusion with intensity images (II-A).

To associate incoming LiDAR points with their corresponding pixels of resolution $r = 0.05\,\mathrm{m}$, we transform them from the global frame G to the voxel image coordinate frame C and map them to pixel coordinates u,v:

$${}_{\mathrm{C}}\mathbf{p}_i = \mathbf{T}_{\mathrm{CG}}\, {}_{G}\mathbf{p}_i, u = \frac{[{}_{\mathrm{C}}\mathbf{p}_i]_x}{r}, v = \frac{[{}_{\mathrm{C}}\mathbf{p}_i]_y}{r}, z = [{}_{\mathrm{C}}\mathbf{p}_i]_z \tag{3}$$



(a) 3D Points (b) Intensity

(c) Height Image $\mathbf{I}_H$ (d) Intensity Map $\mathbf{I}_P$

Fig. 3. Voxel image update. *(a)* The dominant voxel plane is iteratively refined. Points colored by distance to the plane. *(b)* Points from (a) colored by intensity values. *(c)* The height map is generated by projecting points to image coordinates. Pixes store a weighted average height of all contributing points. *(d)* The intensity map is generated using the same projection as (b), where pixel values represent a weighted average of the point intensities.

We update the intensity map using the same weighting as the height image:

$$\mathbf{I}_P(u, v) \leftarrow \frac{\mathbf{W}(u, v) \cdot \mathbf{I}_P(u, v) + w_i \cdot \mathcal{I}_i}{\mathbf{W}(u, v) + w_i} \tag{4}$$

$$\mathbf{W}(u, v) \leftarrow \mathbf{W}(u, v) + w_i, \;\; w_i = \min(0.5, \|{}_{\mathrm{L}}\mathbf{p}_i\|^{-1}) \tag{5}$$

where $\mathbf{W}$ stores the pixel weights and $\mathcal{I}_i$ is the point's filtered intensity from Section II-A.

This map update integrates the full undistorted point cloud without any downsampling, to capture as much information as possible. Note that, differently from the intensity image (II-A), the voxel-wise intensity maps are iteratively created from multiple scans. Voxels are efficiently updated in parallel and stored based on a hash-map, which achieves $\mathcal{O}(1)$ average-case voxel lookup. We evict the least recently updated voxels if the number of stored voxels exceeds a configurable capacity to limit memory usage.

C. Map-Informed Intensity Point Sampling

To focus on informative areas and avoid noise from low-texture regions, we propose a map-informed sampling strategy. Instead of extracting features from dense images, we quantify the intensity information in each voxel to select points in salient areas. Following the geometric complementarity proposed in COIN-LIO, this sampling prioritizes intensity information that constrains the optimization in directions where geometric data is redundant or lacking. To do so, we approximate the geometric information in the registration and identify which directions require additional information. We then sample intensity points in the voxels that provide visual information in the geometrically uninformative direction.

We approximate the informative directions $\boldsymbol{\iota}$ of an intensity map by the sum of its image gradient magnitudes:

$$\boldsymbol{\iota} = 0.5 \cdot \sum_{u,v} \begin{bmatrix} \mathbf{M}_u(u, v) \cdot |\mathbf{I}_P(u+1, v) - \mathbf{I}_P(u-1, v)| \\ \mathbf{M}_v(u, v) \cdot |\mathbf{I}_P(u, v+1) - \mathbf{I}_P(u, v-1)| \end{bmatrix} \tag{6}$$

with $\mathbf{M}_u$ and $\mathbf{M}_v$ as indicator functions guaranteeing the presence of neighboring pixels.

To identify underconstrained geometric directions, we analyze the eigenvectors $\mathbf{v}_j$ and eigenvalues λ_j of the distribution of observed voxel normals $\mathbf{n}_i$. Instead of calculating them from the input point cloud, we extract the normals readily available from the BIEVR map. We find the observed voxels $\mathbf{V}$ using the registration prior.

$$\mathbf{A} = \sum_{i \in \mathbf{V}} \mathbf{n}_i \mathbf{n}_i^\top = \sum_{j=1}^{3} \lambda_j \mathbf{v}_j \mathbf{v}_j^\top \quad \text{with } \lambda_1 \leq \lambda_2 \leq \lambda_3$$

We identify the target direction $\boldsymbol{\eta}$ based on the ratio of the smallest eigenvalues to cover cases with one (e.g. tunnels) or two (e.g. flat areas) unconstrained directions:

$$\boldsymbol{\eta} = \begin{cases} \mathbf{v}_1 + \mathbf{v}_2 & \text{if } 10 \cdot \lambda_1 > \lambda_2 \\ \mathbf{v}_1 & \text{otherwise} \end{cases} \tag{7}$$

We quantify a voxel's contribution l_c, as the projection of the target direction on its informative intensity direction:

$$l_c = (\mathbf{R}_{\mathrm{CW}} * {}_{\mathrm{W}}\boldsymbol{\eta}) \cdot \boldsymbol{\iota} \tag{8}$$

We select the 100 voxels with the strongest contribution as intensity voxels. Inside these voxels, we downsample the input point cloud with a voxel size of $0.1\,\mathrm{m}$.

D. State Estimation

We define the robot state as
$\mathbf{x} \triangleq [\mathbf{R}_{\mathrm{GI}}\ {}_{\mathrm{G}}\mathbf{t}_{\mathrm{GI}}\ {}_{\mathrm{G}}\mathbf{v}_{\mathrm{I}}\ \mathbf{b}_a\ \mathbf{b}_g\ {}_{\mathrm{G}}\mathbf{g}] \in SO(3) \times \mathbb{R}^{12} \times S^2$
where $\mathbf{R}_{\mathrm{GI}}$ represents the rotation from the IMU frame I to the global frame G, ${}_{\mathrm{G}}\mathbf{t}_{\mathrm{GI}}$ and ${}_{\mathrm{G}}\mathbf{v}_{\mathrm{I}}$ are the IMU position and velocity expressed in G, $\mathbf{b}_a$ and $\mathbf{b}_g$ denote the accelerometer and gyroscope biases, and ${}_{\mathrm{G}}\mathbf{g}$ is the gravity direction in the global frame. COIN-BIEVR follows the same loosely-coupled fusion scheme as BIEVR-LIO, and computes the pose $\mathbf{T}_{\mathrm{GI}}$ solely from registration (Section II-D.2), without any inertial residuals, while the remainder of the state is optimized in a sliding-window optimization with fixed poses as constraints. We refer the reader to [1] for a more details.

1) IMU Prediction & Point Undistortion: We employ a piecewise linear continuous-time model [4] to compute pose increments $\mathbf{T}_{\mathrm{I}^k\mathrm{I}^i}$ between the start and end time t_k and t_j of a scan. These are used to undistort LiDAR points ${}_{\mathrm{L}^i}\mathbf{p}_i$ from their acquisition time t_i to the scan end-time t_k:

$${}_{\mathrm{I}^k}\mathbf{p}_i = \mathbf{T}_{\mathrm{I}^k\mathrm{I}^i} \mathbf{T}_{\mathrm{IL}}\, {}_{\mathrm{L}^i}\mathbf{p}_i, \tag{9}$$

where $\mathbf{T}_{\mathrm{IL}}$ represents the extrinsic calibration. Additionally, these increments provide the registration prior $\hat{\mathbf{T}}_{\mathrm{WI}^k}$.

2) Registration: We register the sampled point cloud directly to the map using geometric and photometric residuals.

Each input point ${}_{\mathrm{I}}\mathbf{p}_i$ is mapped to a voxel by hashing its coordinates based on the current pose estimate:

$${}_{\mathrm{C}}\mathbf{p}_i(\boldsymbol{\xi}) = \mathbf{T}_{\mathrm{CG}} \mathbf{T}_{\mathrm{GI}} \mathrm{Exp}(\boldsymbol{\xi})\, {}_{\mathrm{I}}\mathbf{p}_i \tag{10}$$

where $\mathrm{Exp}(\boldsymbol{\xi})$ maps the pose increment $\boldsymbol{\xi} \in \mathfrak{se}(3)$ to a transformation $\mathbf{T} \in SE(3)$. The respective pixel coordinates $u_i(\boldsymbol{\xi}), v_i(\boldsymbol{\xi})$ are calculated according to Equation (3). If a point projects to an unobserved pixel, it is omitted.

The geometric residuals are identical to BIEVR-LIO and minimize the difference between the height of input points and the respective pixel value in the height-map $\mathbf{I}_H$:

$$r_i^{geo}(\boldsymbol{\xi}) = {}_{\mathrm{C}}\mathbf{p}_i^z(\boldsymbol{\xi}) - \mathbf{I}_H\big(u_i(\boldsymbol{\xi}), v_i(\boldsymbol{\xi})\big), \tag{11}$$

where ${}_{\mathrm{C}}\mathbf{p}_i^z$ is the point height above the image plane.

Similarly, the photometric residuals minimize the difference between the filtered intensity value $\mathcal{I}_i$ of a point $\mathbf{p}_i$ and the pixel value in the intensity map $\mathbf{I}_P$:

$$r_i^{pho}(\boldsymbol{\xi}) = \mathcal{I}_i - \mathbf{I}_P\big(u_i(\boldsymbol{\xi}), v_i(\boldsymbol{\xi})\big) \tag{12}$$

The combined non-linear least-squares problem is solved using Levenberg-Marquardt minimization:

$$\boldsymbol{\xi}^* = \arg\min_{\boldsymbol{\xi}} \sum_{i \in \{G,P\}} \rho\Big(\big\|r_i^{geo}(\boldsymbol{\xi})\big\|\Big) + \sum_{j \in P} \rho\Big(\big\|\lambda \cdot r_i^{pho}(\boldsymbol{\xi})\big\|\Big) \tag{13}$$

where $\rho(\cdot)$ denotes the Huber loss and λ is a constant to compensate the different magnitude between photometric and geometric values. G is the set of points selected from informed dual-resolution sampling in BIEVR-LIO, and P contains intensity points sampled in Section II-C.

III. Experiments

We evaluate COIN-BIEVR across structure-rich and geometrically degenerate environments, comparing it against state-of-the-art geometry-only [5, 6] and intensity-augmented [2, 3] LIO frameworks. The test sequences contain both dense 3D and irregular pattern LiDAR systems. We evaluate the baseline methods using sensor-specific configurations where publicly available and otherwise use default parameters. For our approach we use identical parameters across all experiments, aside from different intensity preprocessing parameters required from different LiDAR sensors (Section II-A). For Ouster LiDARs, we use the factory-provided point-to-pixel lookup table to create intensity images and remove line artifacts [2]. We report the Absolute Trajectory Error (ATE) in Table I, considering runs with relative errors exceeding 20% as failures (×).

Experiments were conducted on an Intel i7-11800H CPU and the algorithm ran at more than 25Hz in all sequences.

TABLE I
Absolute Trajectory Error RMSE [m]
Geometry-only and intensity-augmented methods.

L	D	Sequence	FAST-LIO2	DLIO	BIEVR-LIO	COIN-LIO	PG-LIO	COIN-BIEVR
OS0-128	NCD	QuadHard	0.049	0.14	0.051	**0.046**	0.054	0.049
		Cloister	0.078	0.162	0.053	0.078	0.097	**0.051**
		Stairs	×	0.122	0.056	0.102	0.077	**0.051**
		Park	0.31	0.335	0.798	**0.287**	0.331	0.645
	ENWIDE	TunnelS	×	×	×	0.743	0.635	**0.432**
		TunnelD	×	×	×	0.487	0.425	**0.369**
		IntersectionS	12.473	0.177	0.231	0.466	0.279	**0.166**
		IntersectionD	23.8	3.394	0.404	1.912	0.473	**0.303**
		RunwayS	×	×	0.44	1.033	0.491	**0.312**
		RunwayD	×	×	4.35	2.437	3.273	**2.245**
		FieldS	0.163	0.17	0.159	0.232	0.274	**0.153**
		FieldD	9.209	0.617	**0.174**	0.581	0.206	0.175
		KatzenseeS	1.122	0.188	0.194	0.412	0.26	**0.186**
		KatzenseeD	1.02	0.328	0.243	0.592	0.355	**0.22**
Avia	GEODE	Shield 1	×	×	0.256	–	–	**0.22**
		Shield 4	×	×	0.275	–	–	**0.245**
		Shield 5	×	×	**0.146**	–	–	0.219
		FlatSurfacesS	×	×	×	–	–	**0.064**
M360	GT	ETH-1	0.073	0.153	0.086	–	–	**0.065**
		ETH-1 <5	×	×	6.53	–	–	**0.97**

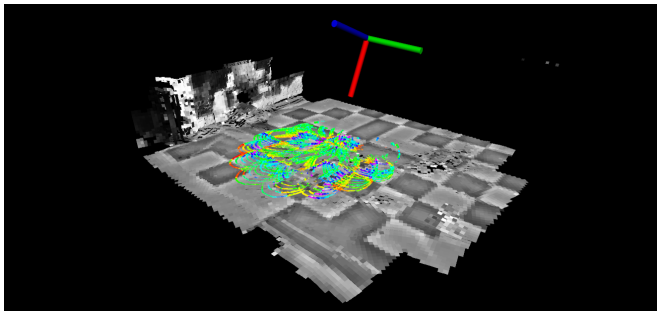


Fig. 4. Visualization of the *FlatSurfacesS* sequence. The intensity map is visualized in grayscale, LiDAR points in color. Despite the irregular scan pattern, the floor pattern is visible in the intensity map and constrains the registration in the presence of entirely flat, degenerate geometry.

1) Newer College Dataset: We evaluate the Newer College Dataset [7] that contains sequences from a handheld Ouster OS0-128 with an integrated IMU. While *Cloister*, *Park*, and *Quad-Hard* provide rich structural features, *Stairs* presents a narrow, constrained geometry. COIN-BIEVR achieves the highest accuracy in two of the four sequences and consistently outperforms the BIEVR-LIO baseline. This demonstrates that LIO performance can also benefit from integrating intensity in geometrically well-constrained settings.

2) ENWIDE Dataset: The ENWIDE dataset [2] that also uses a handheld Ouster OS0-128 LiDAR was specifically recorded to evaluate intensity-based methods in environments with little informative structure. COIN-BIEVR clearly outperforms the baseline approaches and notably achieves the most accurate odometry in the challenging Tunnel environment, visualized in Figure 1, in which the geometry is fully degenerate. This showcases the robustness added through the use of voxel-based intensity information.

3) GEODE Dataset: To evaluate performance under irregular scan patterns, we use sequences from the GEODE dataset [8] captured with a Livox Avia. Results for COIN-LIO and PG-LIO are unavailable here as they require dense intensity images. COIN-BIEVR improves accuracy over BIEVR-LIO in the *Shield 1* and *Shield 4* sequences. While near-range intensity artifacts ($< 1.5\,\mathrm{m}$) lead to a slightly higher ATE in *Shield 5*, COIN-BIEVR is the only method that avoids divergence in *Flat Surfaces S*. In this sequence, the LiDAR is moved around for multiple seconds while only observing flat ground, as shown in Figure 4, which causes geometry-only methods to diverge. Our approach leverages patterns on the floor tiles to maintain accurate odometry.

4) GrandTour Dataset: We assess the performance of our method on the Livox Mid-360 LiDAR that combines a $360°$ horizontal field of view with an irregular scan pattern using the *ETH-1* sequence [9]. We did not find any publicly available datasets using this LiDAR with ground truth trajectories in which BIEVR-LIO drifts significantly. To evaluate the effectiveness of COIN-BIEVR, we artificially induce degeneracy by capping the LiDAR range at $5\,\mathrm{m}$ (*ETH-1* $<$*5m*). The results demonstrate that COIN-BIEVR effectively reduces drift in both the original geometrically constrained environment and the artificially degenerated scenario, which underlines the utility of our method to provide relevant intensity constraints across varying sensor types.

IV. Conclusion

In this work, we presented COIN-BIEVR, a novel LIO approach that augments a high-resolution map representation with intensity information to enable direct 3D photometric optimization during registration. As our approach does not rely on dense projected intensity images, we provide a robust solution for sensors with irregular scanning patterns that are incompatible with traditional image-based feature detection and our map-informed sampling strategy provides intensity constraints that complement uninformative geometric directions. Experimental results across varied hardware and environments confirm that our method effectively leverages intensity to provide accurate odometry in multiple settings where geometry-only approaches fail.

References

[1] P. Pfreundschuh, T. Tuna, C. L. Gentil, R. Siegwart, C. Cadena, and H. Oleynikova, "Bievr-lio: Robust lidar-inertial odometry through bump-image-enhanced voxel maps," *arXiv preprint arXiv:2604.14421*, 2026. [Online]. Available: https://arxiv.org/abs/2604.14421.

[2] P. Pfreundschuh, H. Oleynikova, C. Cadena, R. Siegwart, and O. Andersson, "Coin-lio: Complementary intensity-augmented lidar inertial odometry," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2024, pp. 1730–1737.

[3] N. Khedekar and K. Alexis, "Pg-lio: Photometric-geometric fusion for robust lidar-inertial odometry," *arXiv preprint arXiv:2506.18583*, 2025.

[4] C. Le Gentil and T. Vidal-Calleja, "Continuous latent state preintegration for inertial-aided systems," *The International Journal of Robotics Research*, vol. 42, no. 10, pp. 874–900, 2023.

[5] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lio2: Fast direct lidar-inertial odometry," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2053–2073, 2022.

[6] K. Chen, R. Nemiroff, and B. T. Lopez, "Direct lidar-inertial odometry: Lightweight lio with continuous-time motion correction," *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3983–3989, 2023. DOI: 10.1109/ICRA48891.2023.10160508.

[7] L. Zhang, M. Camurri, D. Wisth, and M. Fallon, "Multi-camera lidar inertial extension to the newer college dataset," *arXiv preprint arXiv:2112.08854*, 2021.

[8] Z. Chen et al., "Heterogeneous lidar dataset for benchmarking robust localization in diverse degenerate scenarios," *The International Journal of Robotics Research*, p. 02 783 649 251 344 967, 2024.

[9] J. Frey et al., "Grandtour: A legged robotics dataset in the wild for multi-modal perception and state estimation," *arXiv preprint arXiv:2602.18164*, 2026.