



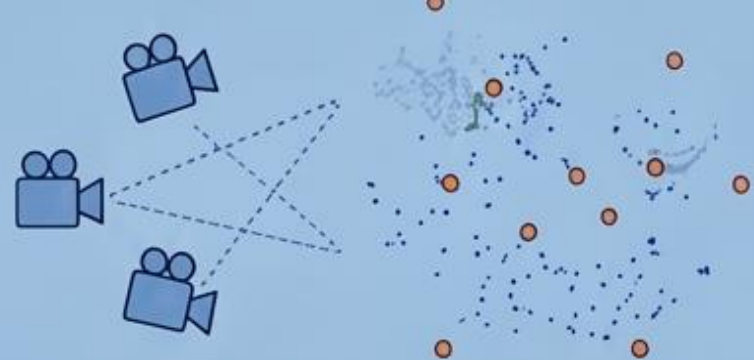
PRISM-SLAM: Probabilistic Ray-Grounded Inference for Scale-aware Metric SLAM

Eunsoo Im, Gyeongwan Lee, and Junghun Suh

Architecture

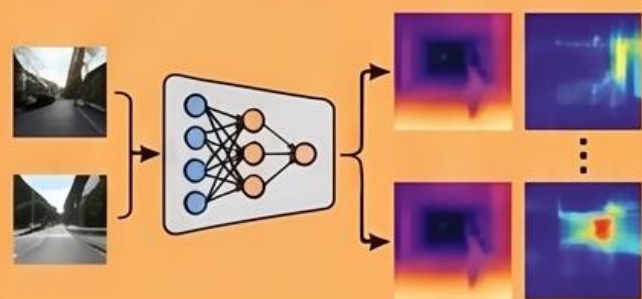
Process 1: Frontend Tracking (~30 Hz)

- Real-time local PnP-RANSAC pose estimation.
- Sparse feature extraction from RGB inputs.



Process 2: VFM(vision Foundation Model) Prior Extraction

- Asynchronous GPU worker for DA3 foundation model.
- Dense metric depth & spatial confidence prediction.



Process 3: Log-Domain Scale Filter

- WLS fusion of sparse points & dense VFM priors.
- Additive scale drift tracking in log-space.
- Strictly positive, continuous metric scale feedback

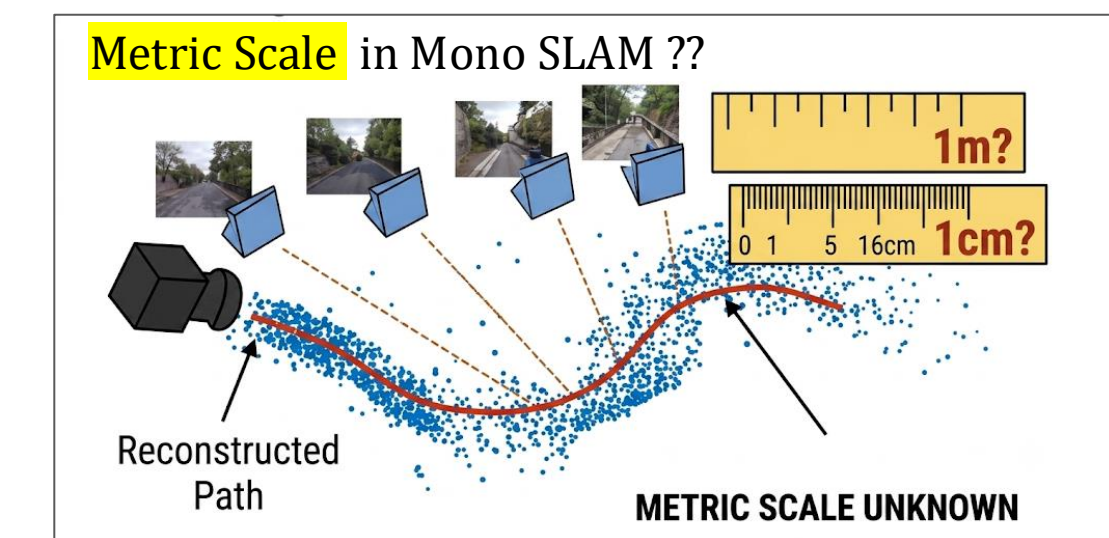
Process 4: Metric Graph Optimization

- Bayesian factor graph window, global Bundle Adjustment.
- Metric scale anchoring via Plücker ray-distance. Soft-gating of dynamic structures via DSUG(Dynamic Scene Uncertainty Gating).

Motivation

Existing monocular SLAM methods are inherently bound to a Sim(3) paradigm, where absolute scale remains unobservable, inevitably requiring post-hoc oracle alignment to ground truth to achieve metric accuracy.

Method	Depth Prior	Alignment	FPS
ORB-SLAM3 (Campos et al., 2021)	—	Sim(3)	30
DROID-SLAM (Teed & Deng, 2021)	—	Sim(3)	5
DPV-SLAM++ (Lipson et al., 2024)	—	Sim(3)	50
GO-SLAM (Zhang et al., 2023)	—	Sim(3)	3
MonoGS (Matsuki et al., 2024)	—	Sim(3)	3
Splat-SLAM (Sandström et al., 2025)	OmniData (Eftekhar et al., 2021)	Sim(3)	1.2
MASt3R-SLAM (Murai et al., 2025)	MASt3R (Leroy et al., 2024)	Sim(3)	15
VGGT-SLAM (Goldman et al., 2025)	VGGT (Wang et al., 2025)	Sim(3)	20
EC3R-SLAM (Hu et al., 2025)	VGGT (Wang et al., 2025)	Sim(3)	36
WildGS-SLAM (Zheng et al., 2025)	Metric3D v2 (Hu et al., 2024)	Sim(3)	0.5
GigaSLAM (Deng et al., 2025)	UniDepth (Piccinelli et al., 2024)	n.s.	—
PRISM (Ours)	DA3 (Yang et al., 2025)	SE(3)	30

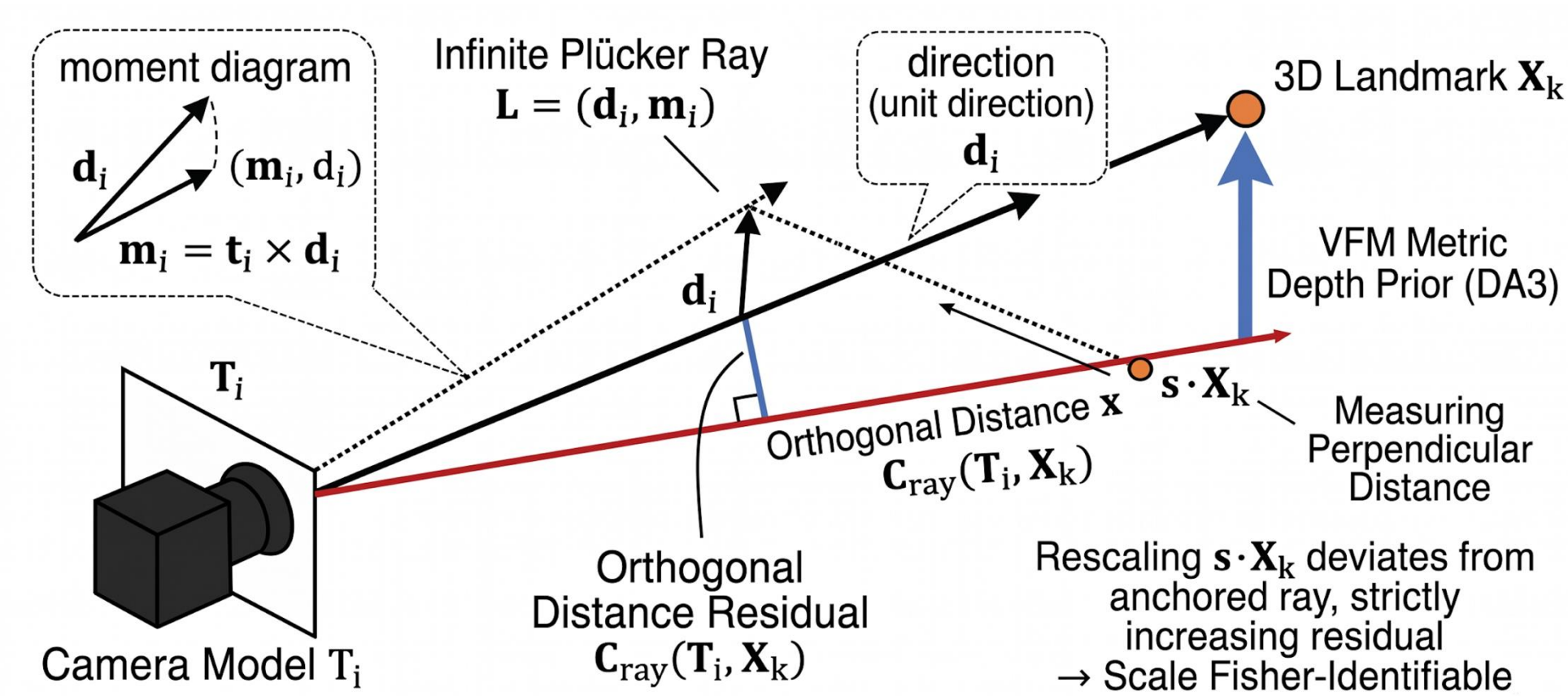


Unlike prior work, we deliver full SE(3) metric scale natively

RAY Optimization

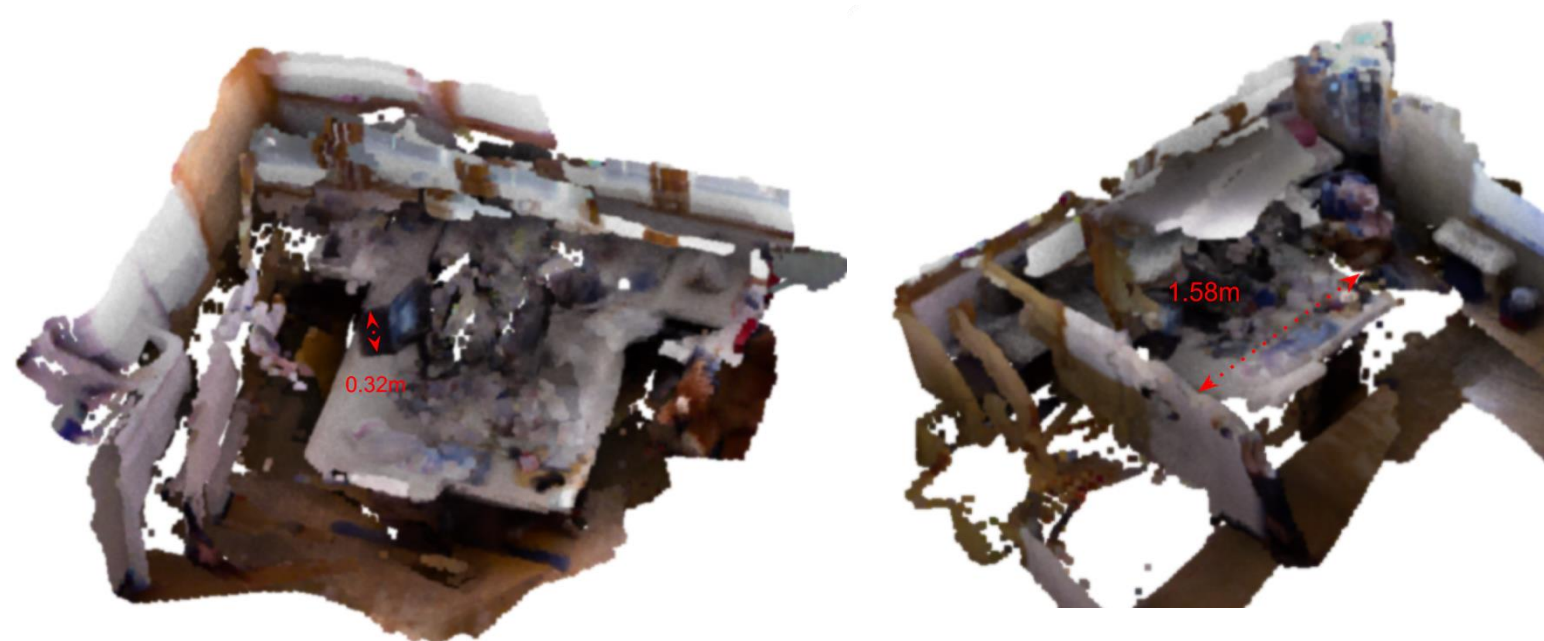
Plücker Ray-Distance Factor:

Introduces a 3D ray-distance formulation to anchor monocular observations in a globally consistent metric coordinate system.



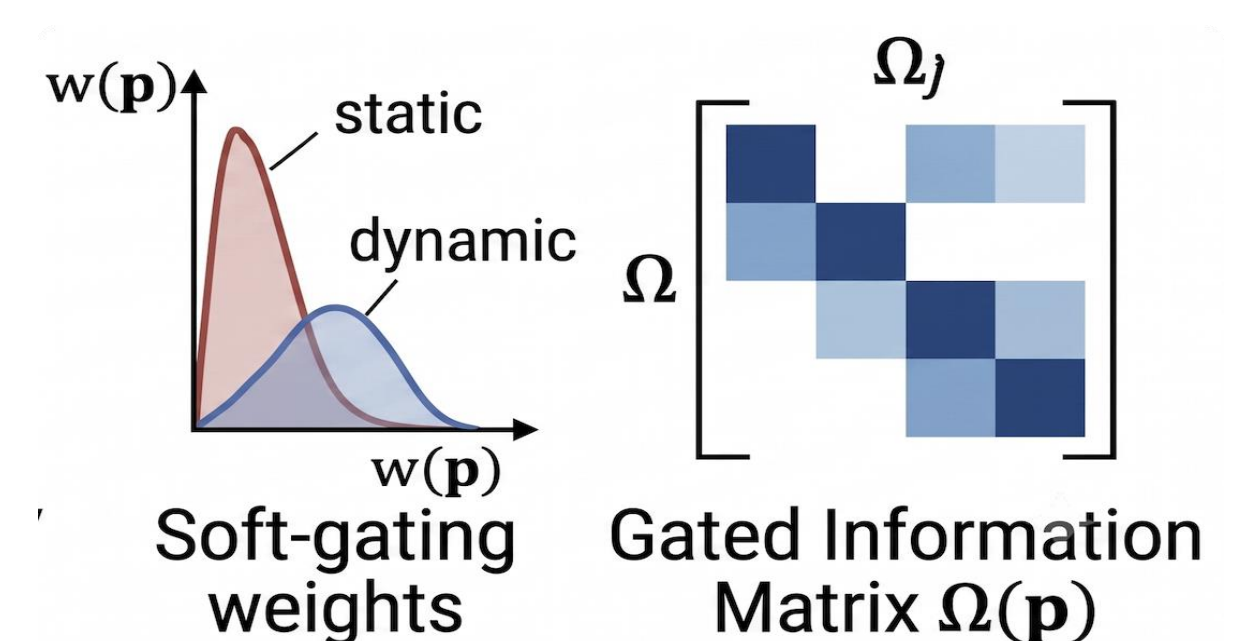
MAP

Decoupled 3D Mapping: Separates tracking from heavy mapping to ensure backend efficiency.



DSUG(Dynamic Scene Uncertainty Gating)

Probabilistic Soft-Gating: Dynamically downweights dynamic distractors and unreliable depth regions without incurring the computational overhead of hard semantic masks



Experiments

<ATE RMSE(cm) on BONN Dynamic

	Method	Metric	Dynamic Sequences(Sim(3) / SE(3))			
			balloon	Balloon2	pers trk	balloon trk
RGB-D	ORB-SLAM3	o	5.8	17.7	70.7	-
	DynaSLAM	o	3.0	2.9	6.1	-
	ReFusion	o	17.5	25.4	28.9	-
	RoDyn-SLAM	o	7.9	11.5	14.5	-
Uncalibrated RGB	PRISM(Ours)	o	9.8 / 18.1	14.0 / 17.7	36.7 / 39.5	7.8 / 9.1

<ATE RMSE(cm) on TUM RGB-D)

	Method	Metric	Static(Sim(3) / SE(3))		Dynamic(Sim(3) / SE(3))	
			sit	walk	Sit-xyz	Walk-xyz
Calibrated RGB	ORB-SLAM3	x	0.7	0.9	25.3	48.8
	DROID-SLAM	x	0.4	0.3	17.6	21.4
	MonoGS	x	1.1	3.6	16.2	18.9
	WildGS-SLAM	x	0.5	0.6	4.1	5.2
UnCalibrated RGB	PRISM(Ours)	o	1.6 / 1.8	1.9 / 2.7	12.7 / 20.0	23.2 / 26.8