

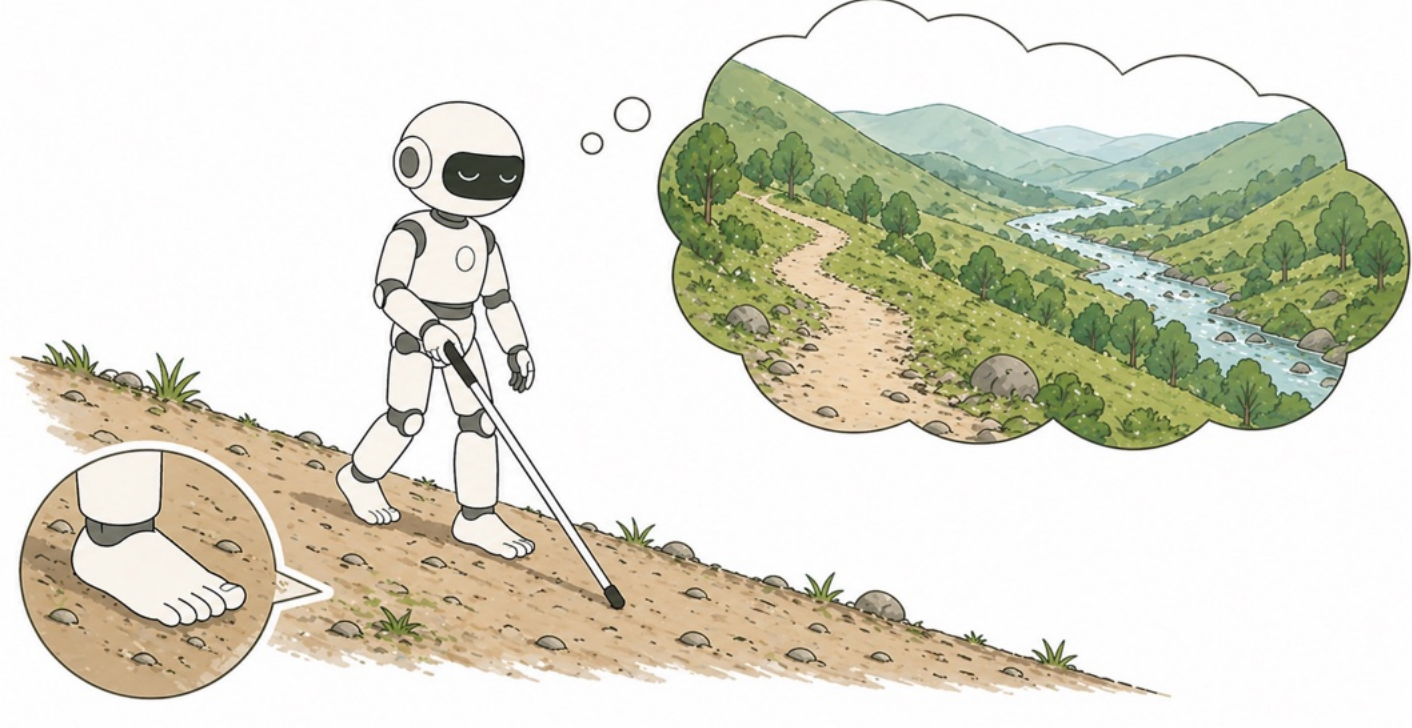
FootRecon: Quadrupedal Terrain Reconstruction from Sparse Foot Contacts with Geometric Prior

Yujin Park, Seungoh Han, Taebok Lee, Chanwoo Lee, Hyungyu Park, Kyungdon Joo

3D Vision & Robotics Lab., UNIST, South Korea

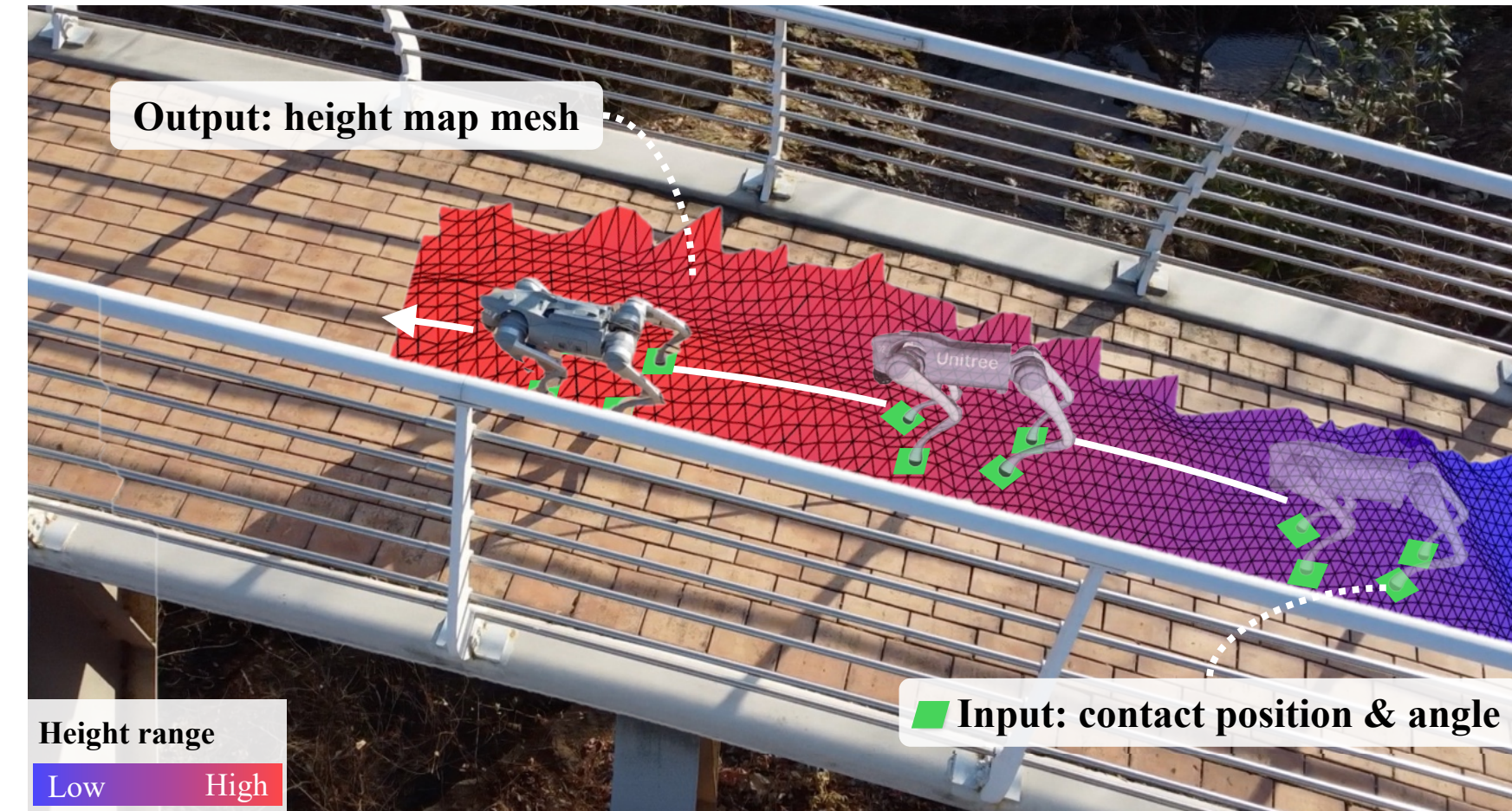


Motivation



- **Reliable Geometry is Essential:** Safe locomotion fundamentally requires accurate terrain understanding.
- **Exteroceptive Fragility:** Vision and LiDAR suffer from severe degradation in real-world conditions.
- **Contact as Physical Signal:** Physical contacts between the robot and terrain offer a complementary sensing modality that remains reliable.

Problem statement



Ill-posed Problem: Sparse contacts do not uniquely determine terrain geometry, resulting in multiple feasible sets.

=> **Reconstructing a consistent local terrain representation with learned priors conditioned on contact observations**

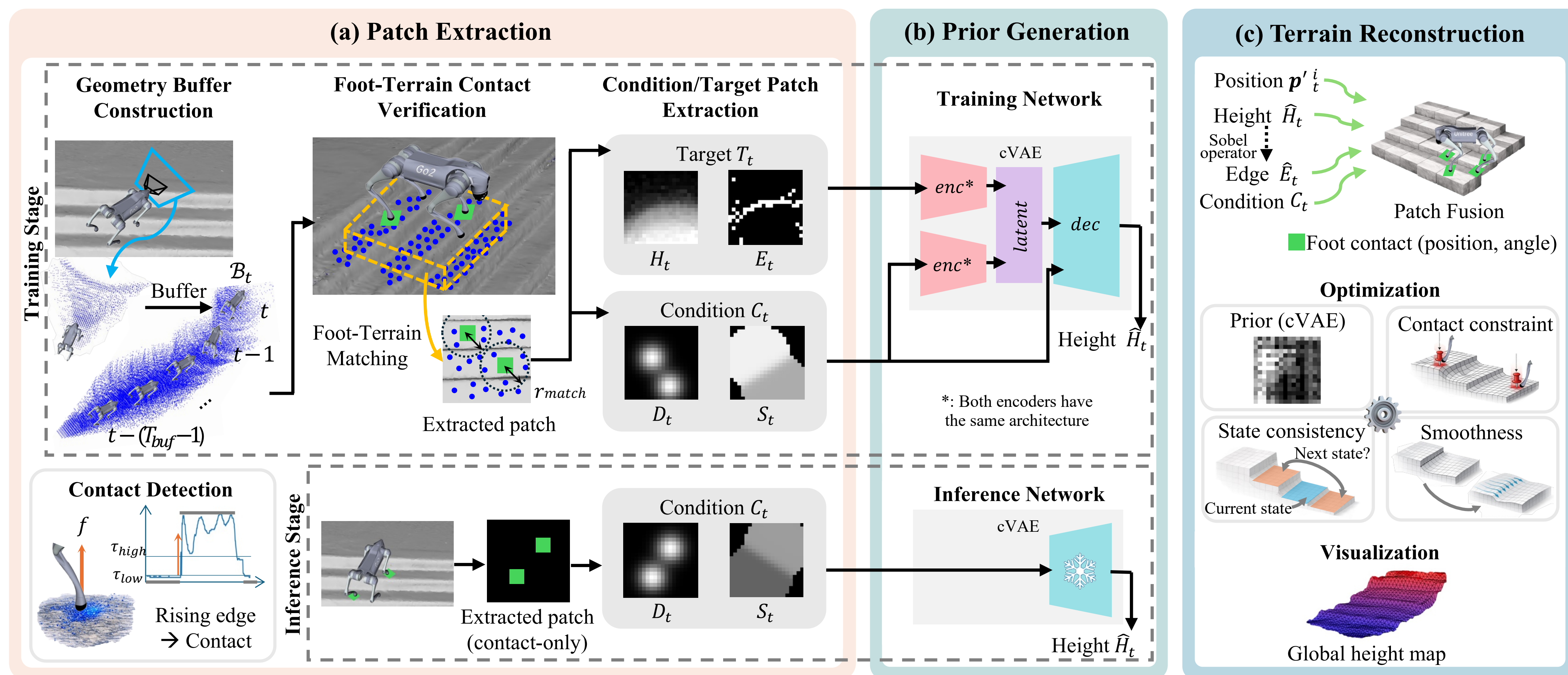
Contributions

👉 **Key Insight: Learning priors to infer plausible terrain beyond sparse contacts**

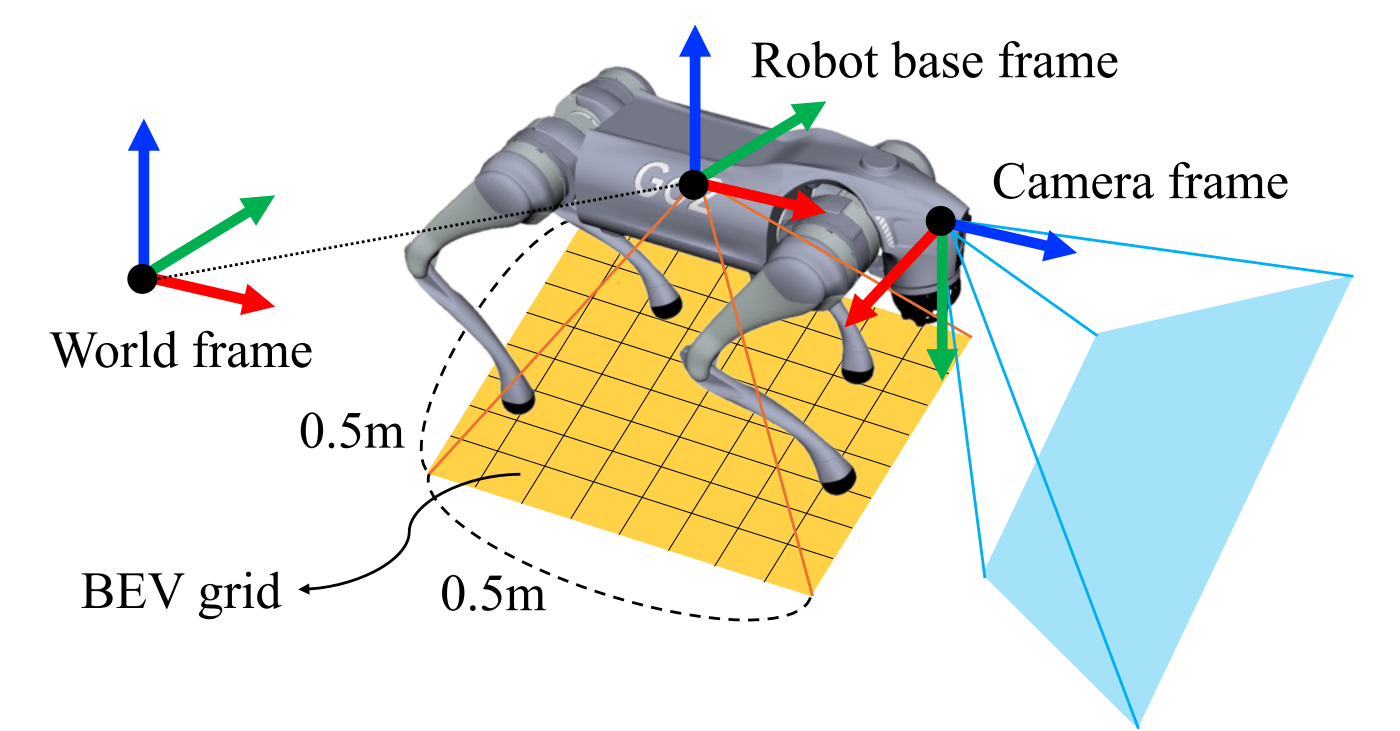
- We formulate terrain reconstruction as a contact-conditioned inference problem under sparse and ill-posed observations.
- We propose a generative model that learns priors over terrain geometry to infer plausible reconstructions from limited contact signals.
- We demonstrate that our approach enables robust terrain estimation even in scenarios where exteroceptive sensing fails.

Method

Overall pipeline

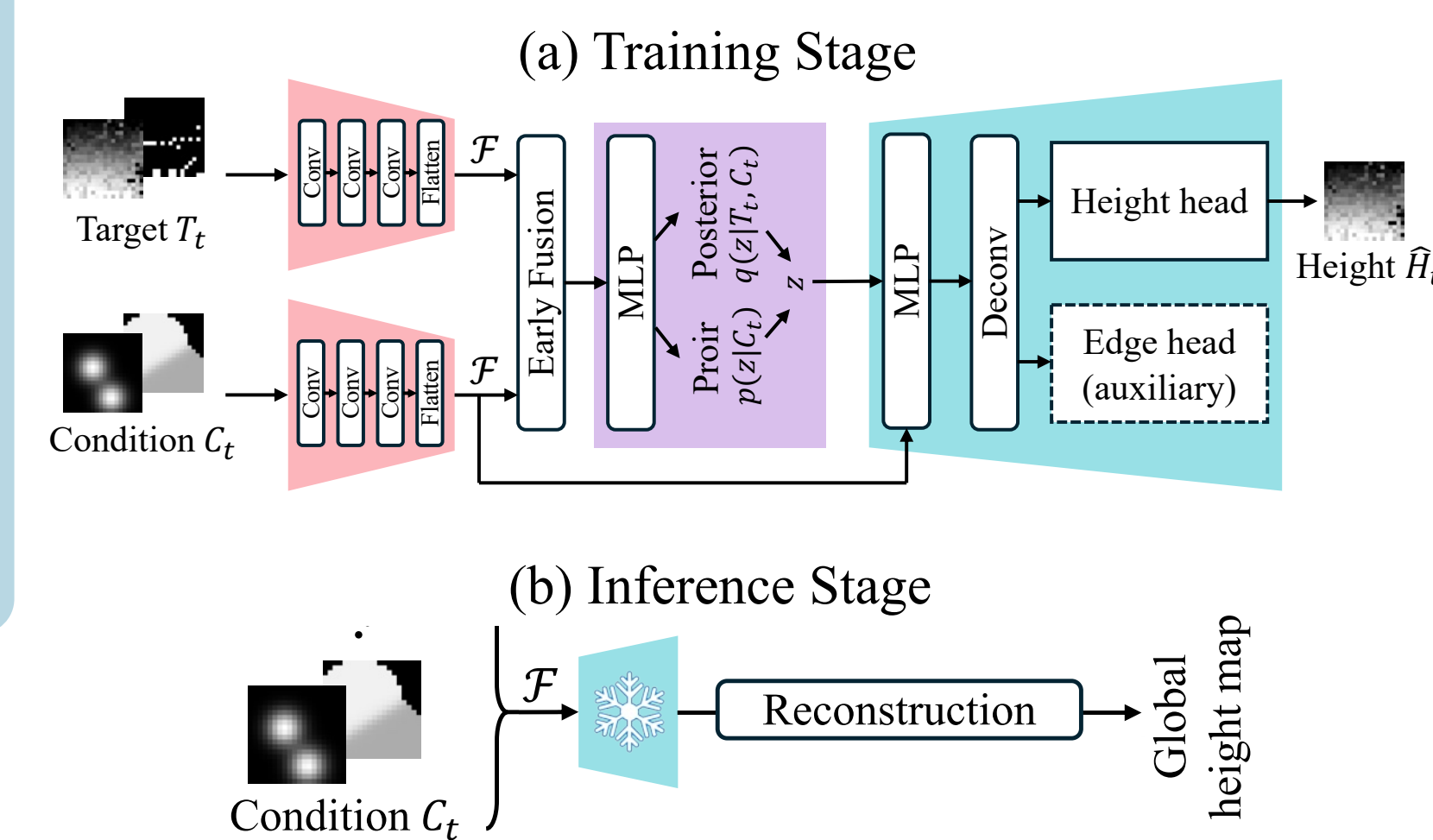


Coordinate systems



- Verified contacts are used to extract paired 2D patches, which are both represented on a robot-centered BEV grid.

Network architecture



- **Patch extraction:** Verified foot-terrain contacts are spatially aligned with geometry and converted into paired target (height/edge) and condition (contact-derived cues) patches for local terrain representation
- **Prior generation:** A cVAE models the distribution of local terrain geometry and generates a plausible height patch from sparse contact-derived condition inputs by sampling a latent prior learned during training
- **Terrain reconstruction:** The predicted height patch is refined via geometry-aware optimization that enforces contact consistency, smoothness, and prior regularization, and is then fused into a globally consistent height map

Experiment

Runtime

Stage	Time (ms)
Training.	
Geometry Buffer Construction	17.34
Verification + Patch Extraction	190.69
Inference.	
Condition Patch Extraction	0.31
Prior Generation	1.44
Refinement	133.09
Total inference time per patch	134.84

Training env.



Quantitative results

Terrain	Method	$RMSE_{global} \downarrow$	$MAE_{patch} \downarrow$	$RMSE_{patch} \downarrow$
Wide Staircase	Kernel Reg.	0.1428	0.5503	0.5579
	MLS Plane Fit.	0.2200	0.5903	0.6096
	TSDF	0.2696	0.4642	0.4819
	Ours	0.0917	0.1228	0.1384
Coir Mat	Kernel Reg.	0.1009	0.7669	0.7690
	MLS Plane Fit.	0.1155	0.7510	0.7619
	TSDF	0.2018	0.4764	0.4857
	Ours	0.0772	0.1080	0.1309
Bamboo Forest	Kernel Reg.	0.1362	1.1365	1.1397
	MLS Plane Fit.	0.1797	1.1177	1.1234
	TSDF	0.2637	1.0213	1.0264
	Ours	0.0560	0.1078	0.1269
Steep Slope	Kernel Reg.	0.1504	2.3107	2.3214
	MLS Plane Fit.	0.2118	2.1983	2.2126
	TSDF	0.3373	2.0107	2.0164
	Ours	0.0892	0.2640	0.3030
Cave	Kernel Reg.	0.2677	0.0883	0.0946
	MLS Plane Fit.	0.2555	0.0853	0.1008
	TSDF	0.3039	0.0541	0.0641
	Ours	<u>0.2657</u>	<u>0.0748</u>	<u>0.0819</u>

Bold and underline denote the best and second-best results, respectively.

Qualitative results

