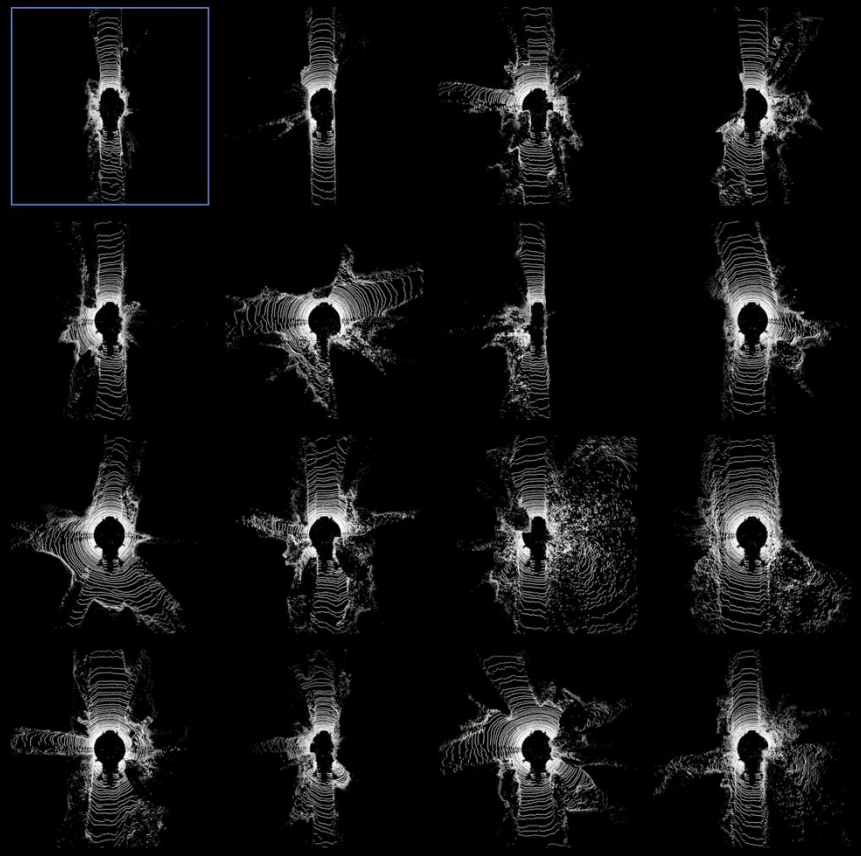
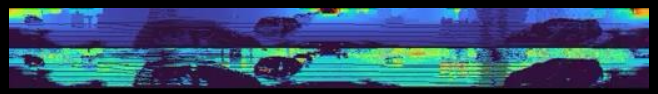
ICRA 2025

Fast LiDAR Data Generation with Rectified Flows

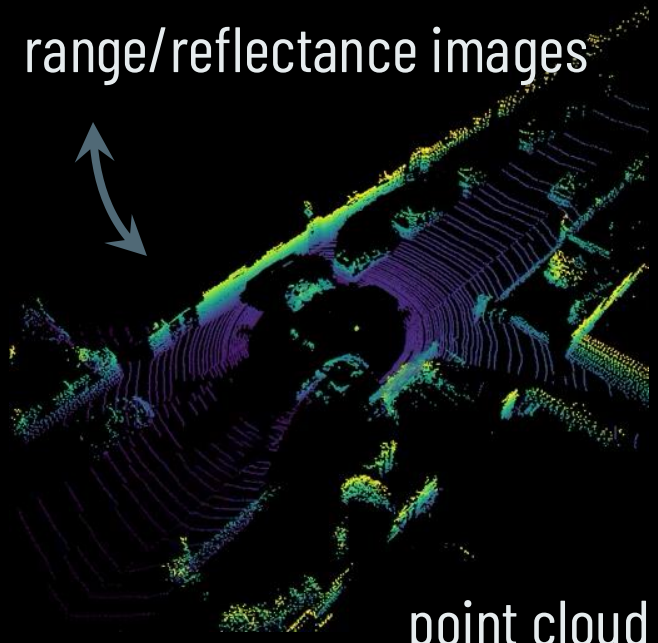
Kazuto Nakashima	Kyushu University
Xiaowen Liu	Kyushu University
Tomoya Miyawaki	Kyushu University
Yumi Iwashita	Jet Propulsion Laboratory
Ryo Kurazume	Kyushu University



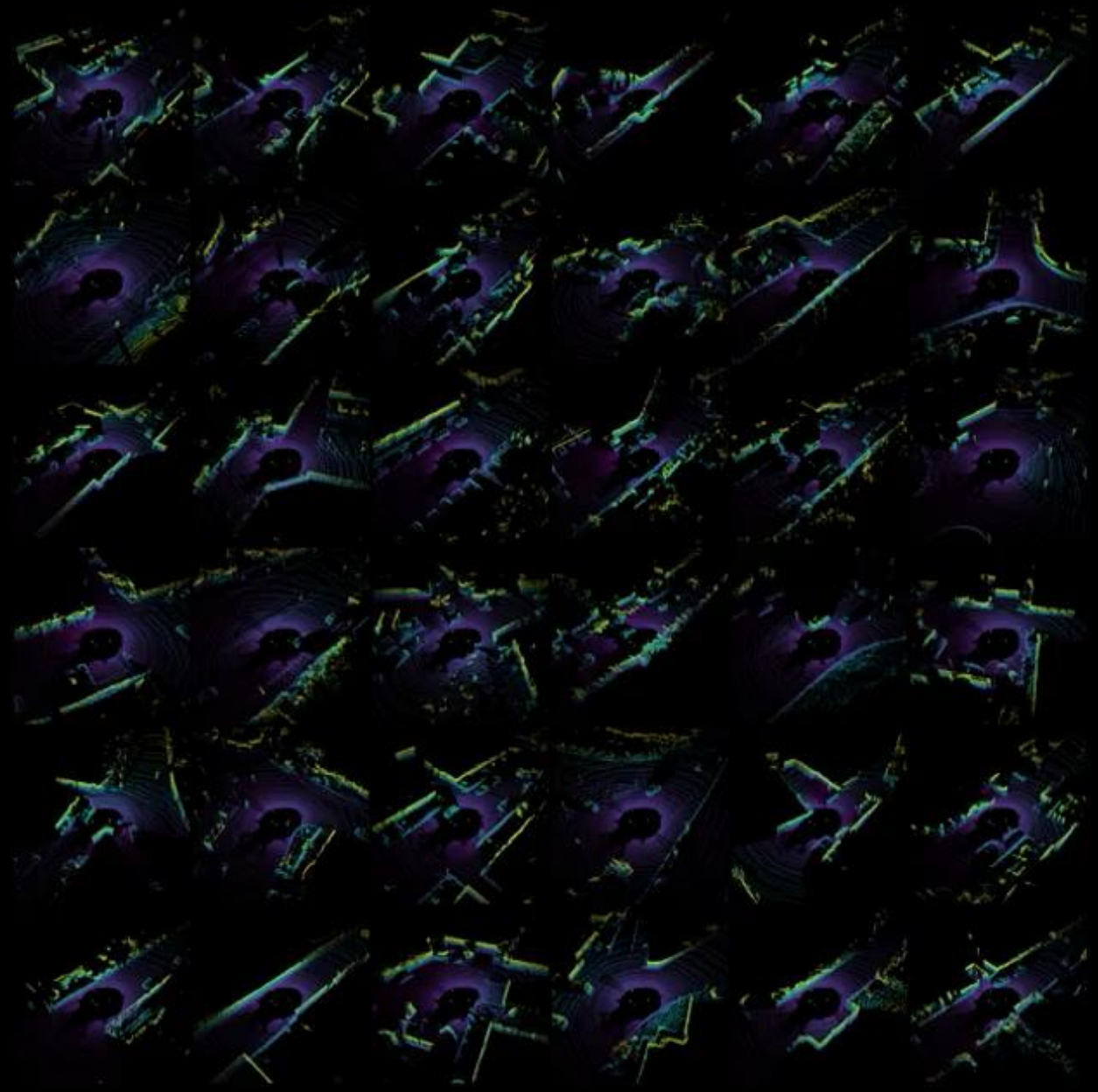
LiDAR data



range/reflectance images



point cloud

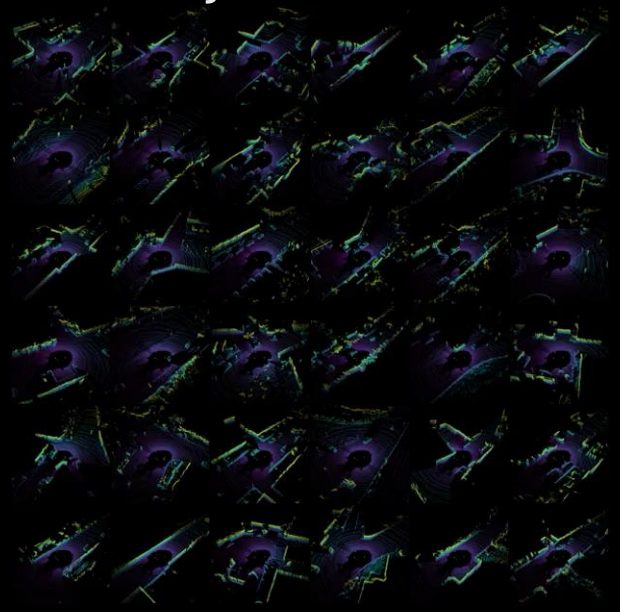


Motivation

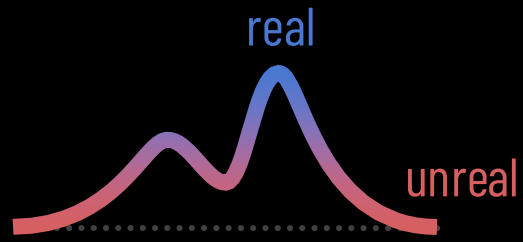
LiDAR Generative Models

Representing the probability distribution of LiDAR data (range image/point cloud)

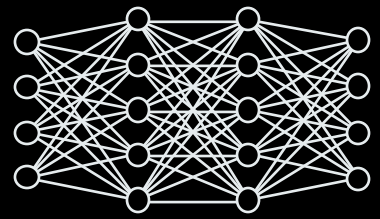
A large-scale dataset



Generative modeling



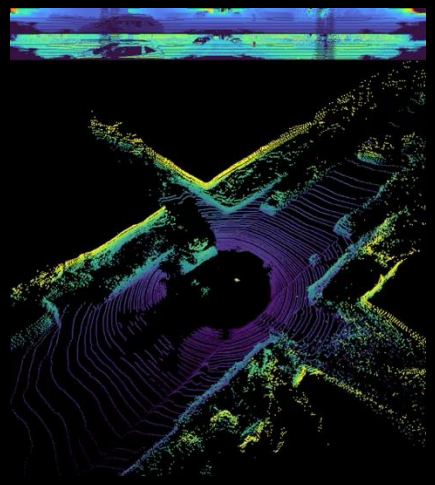
Probability distribution



modeled by neural networks (NNs)

Sampling

Novel data



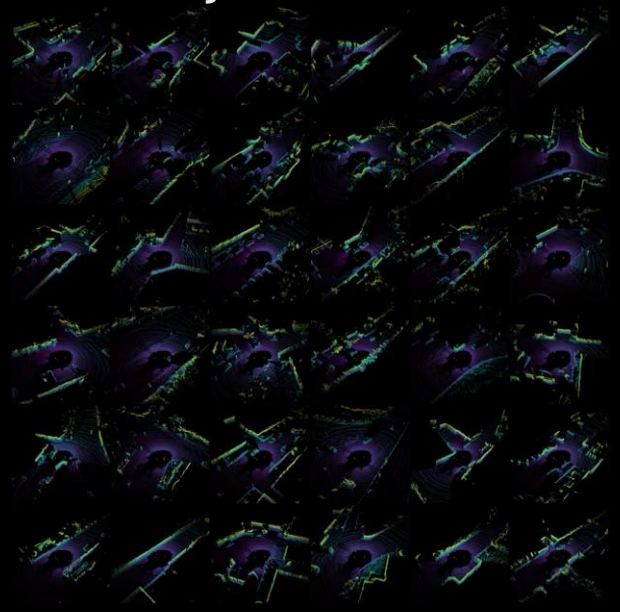
Motivation

LiDAR Generative Models

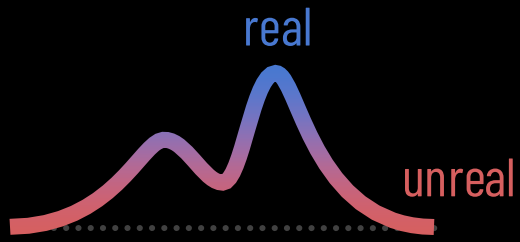
A powerful data prior for downstream tasks (e.g. completion and sim-to-real)

[Zyrianov+ ECCV'22] [Nakashima+ ICRA'24] [Ran+ CVPR'24] [Nakashima+ WACV'23] ...

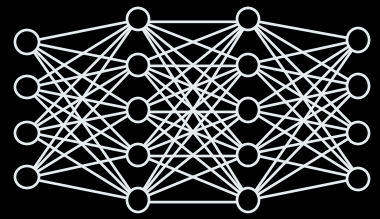
A large-scale dataset



Generative modeling



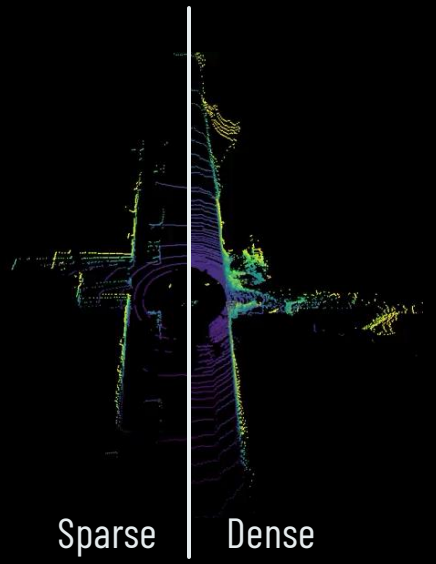
Probability distribution



modeled by neural networks (NNs)

Inference

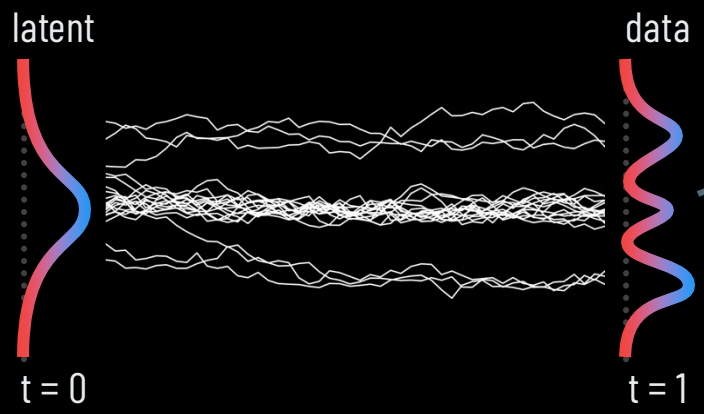
e.g. Completion



Related Work

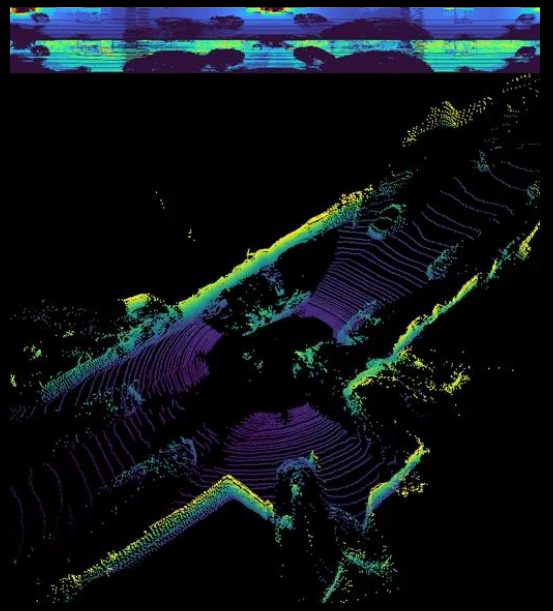
Diffusion Models of Range Images [Zyrianov+ ECCV'22] [Nakashima+ ICRA'24] [Ran+ CVPR'24]

Pros: High-quality samples, stable training, post-hoc conditioning w/o re-training



Stochastic denoising process
to generate data from latent noise

target



range/reflectance images

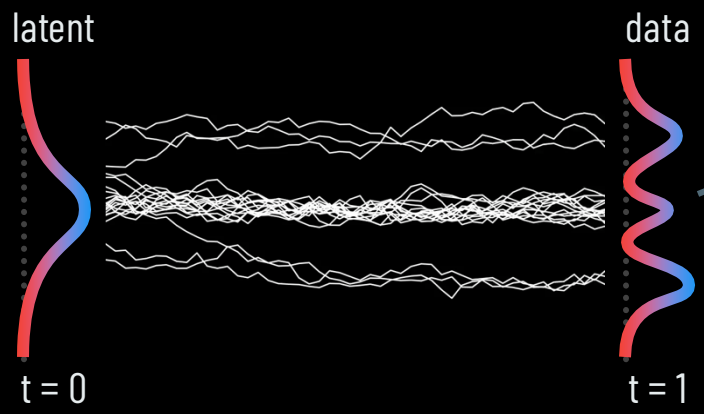
point cloud

Related Work

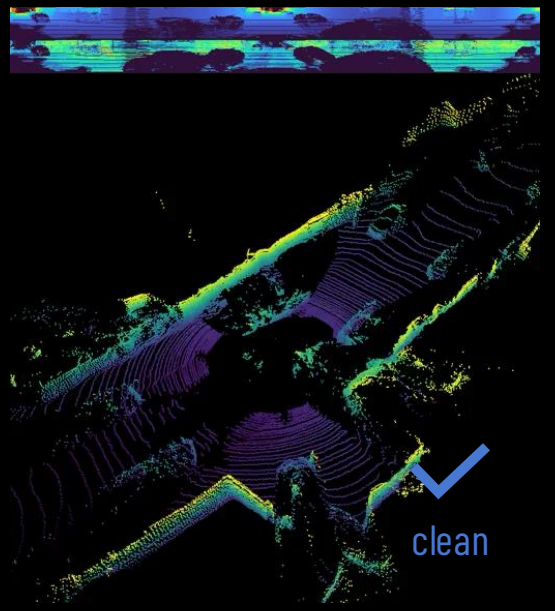
Diffusion Models of Range Images [Zyrianov+ ECCV'22] [Nakashima+ ICRA'24] [Ran+ CVPR'24]

Pros: High-quality samples, stable training, post-hoc conditioning w/o re-training

Cons: A large number of steps (NN evaluation) are required for *high-quality* sampling

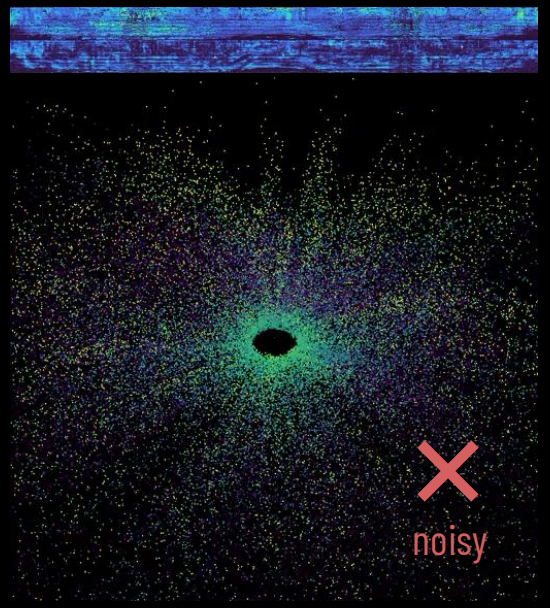


The stochastic & curved trajectories
prone to discretization errors



256 steps
(3.7 sec)

slow



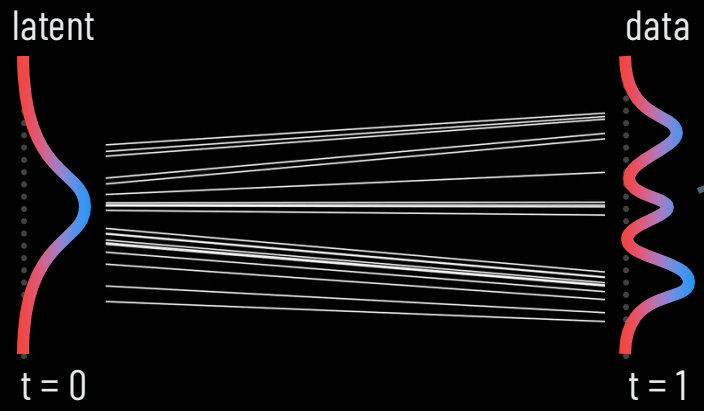
2 steps
(56 msec)

fast

Our Approach

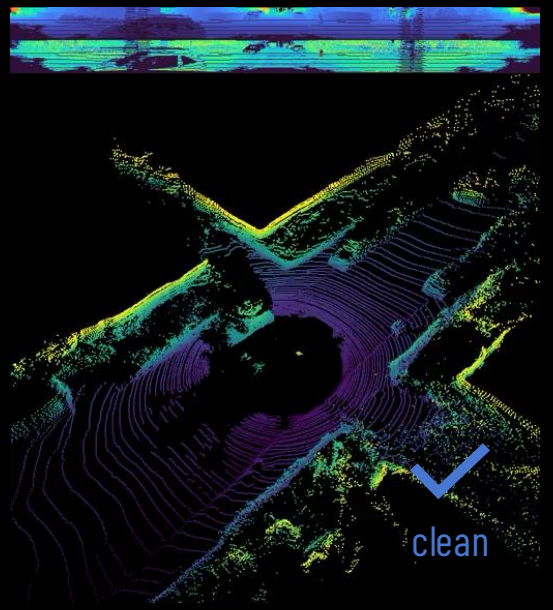
R2Flow (Range–Reflectance Flow)

Modeling LiDAR images w/ easy-to-approximate straight trajectories + efficient NN architecture

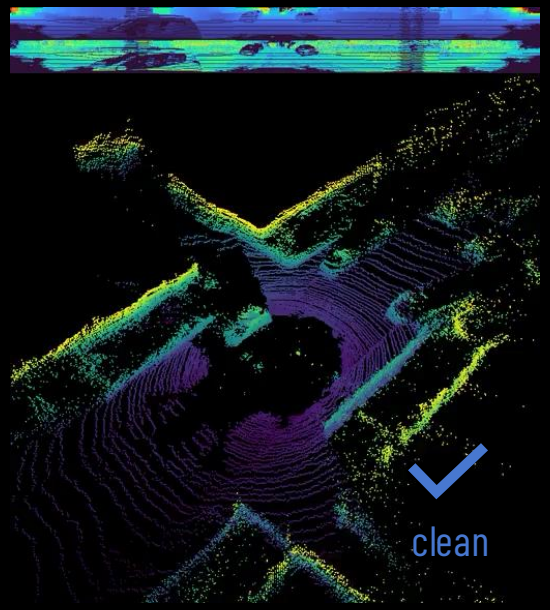


The straight trajectory-based modeling is robust to the step size in sampling

Our results



256 steps (7.6 sec) ✗ slow



2 steps (62 msec) ✓ fast

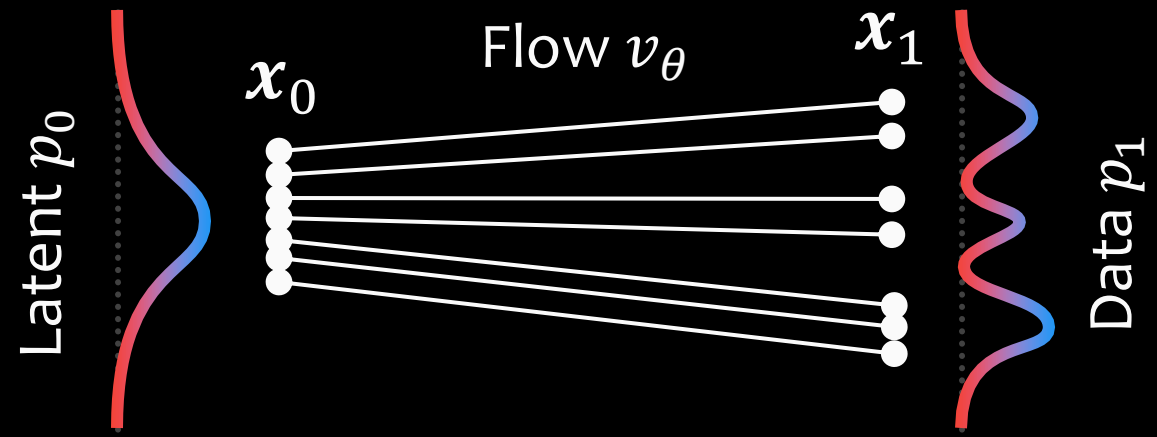
Our Approach

Formulation using **Rectified Flow** [Liu+ ICLR'23] [Lee+ arXiv'24]

A neural ODE that yields the straight trajectories $\{\boldsymbol{x}_t | 0 \leq t \leq 1\}$

$$d\boldsymbol{x}_t = v_\theta(\boldsymbol{x}_t, t)\, dt$$ from $\boldsymbol{x}_0$ (latent) to $\boldsymbol{x}_1$ (data point)

Flow field v_θ is a neural network trained to follow a uniform velocity $\boldsymbol{x}_1 - \boldsymbol{x}_0$



Our Approach

Formulation using **Rectified Flow** [Liu+ ICLR'23] [Lee+ arXiv'24]

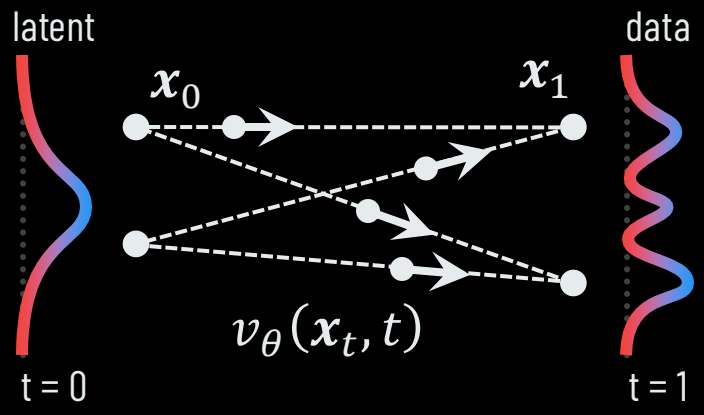
A neural ODE that yields the straight trajectories $\{\boldsymbol{x}_t | 0 \leq t \leq 1\}$

$$d\boldsymbol{x}_t = v_\theta(\boldsymbol{x}_t, t)\, dt \qquad \text{from } \boldsymbol{x}_0 \text{ (latent) to } \boldsymbol{x}_1 \text{ (data point)}$$

Flow field v_θ is a neural network trained to follow a uniform velocity $\boldsymbol{x}_1 - \boldsymbol{x}_0$

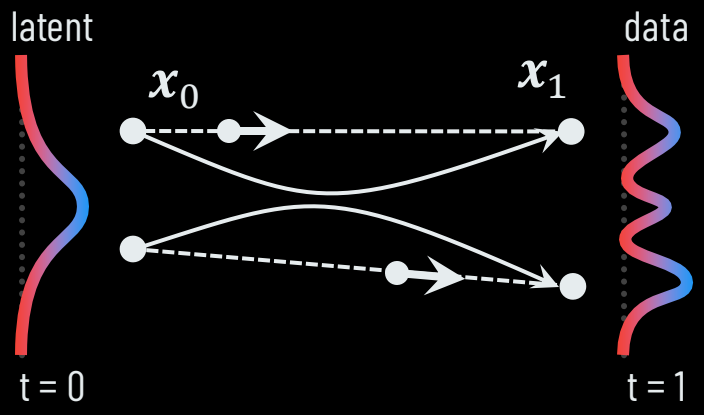
Step 1: Training initial flows

Train v_θ w/ random pairs



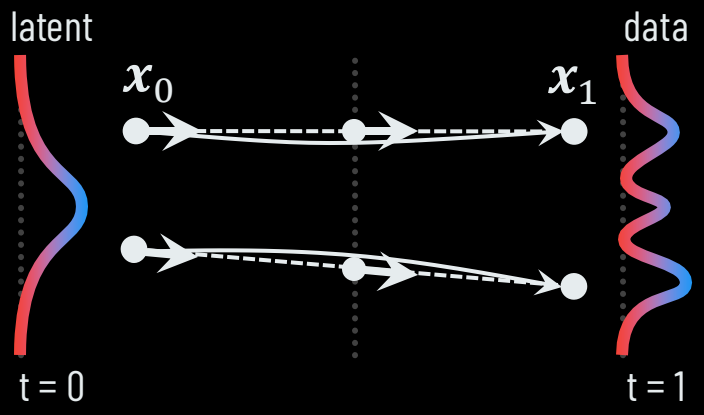
Step 2: Straightening flows

Re-train v_θ w/ generated pairs



Step 3: Time-step distillation

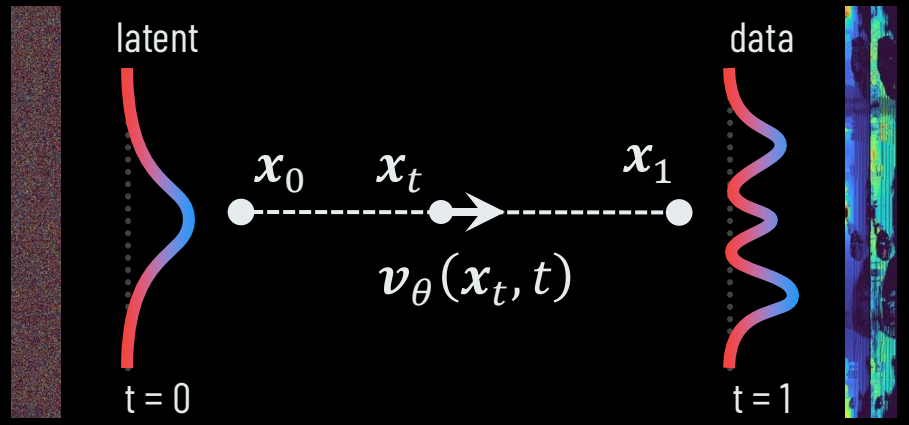
Distill v_θ at specific timesteps



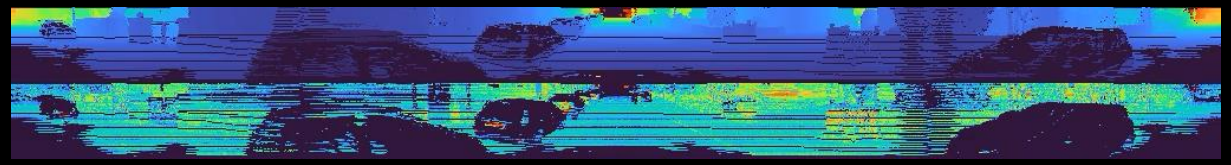
Our Approach

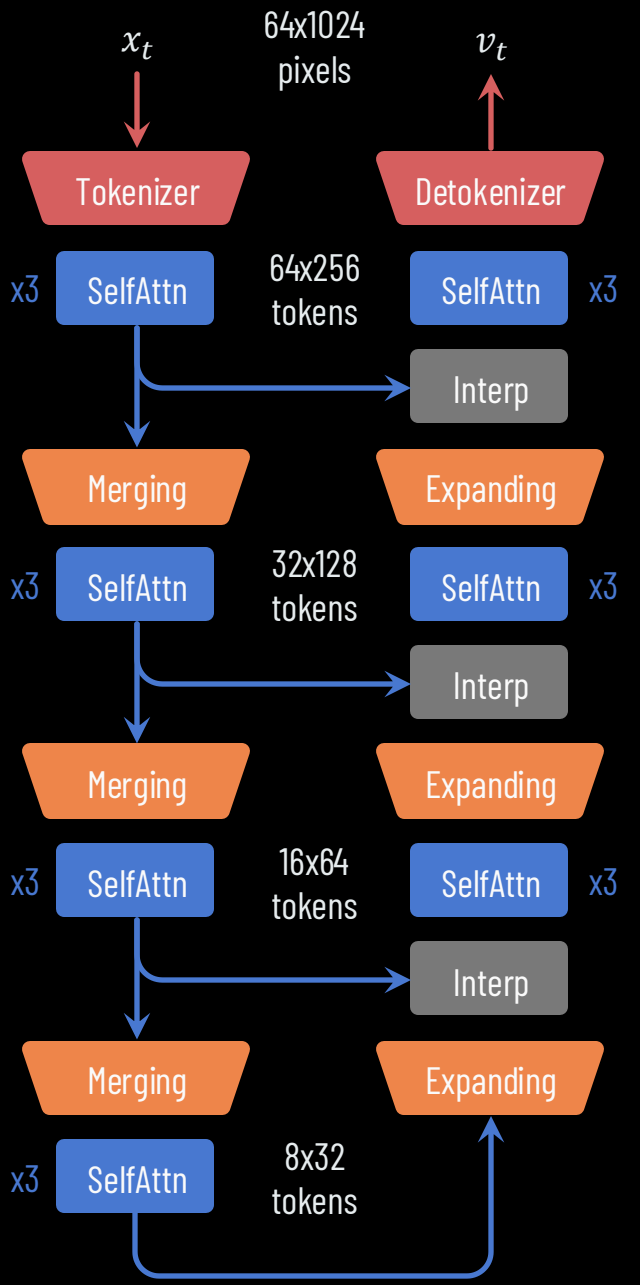
Architecture of Estimator v_θ

We learn the estimator v_θ in the pixel space for precision



We modify **HDiT** (efficient ViT w/ local attention) [Crowson+ ICML'24] to process the *panoramic* and *spatially-aligned* LiDAR structure

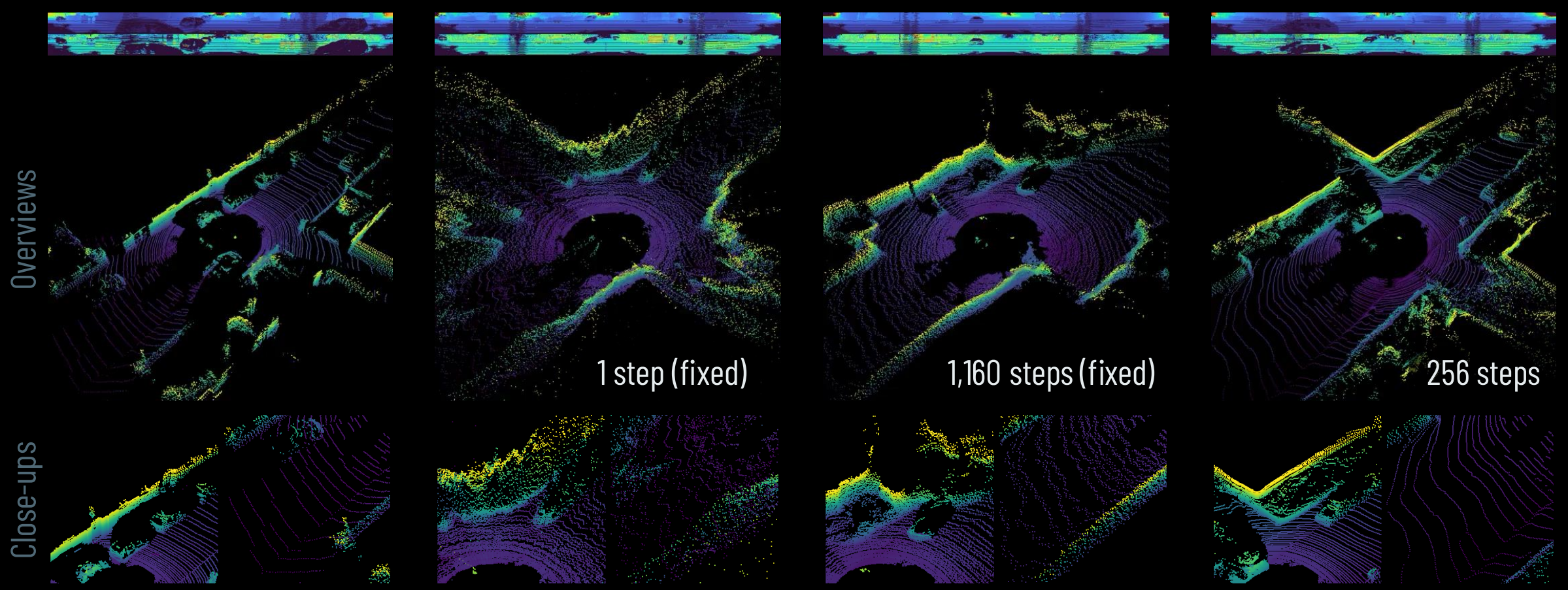




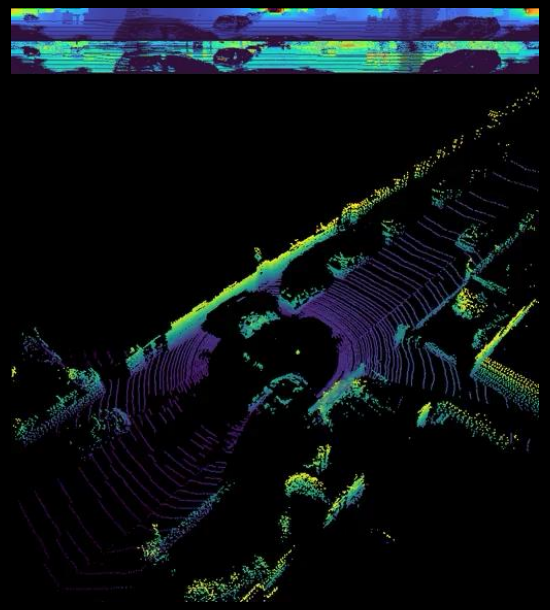
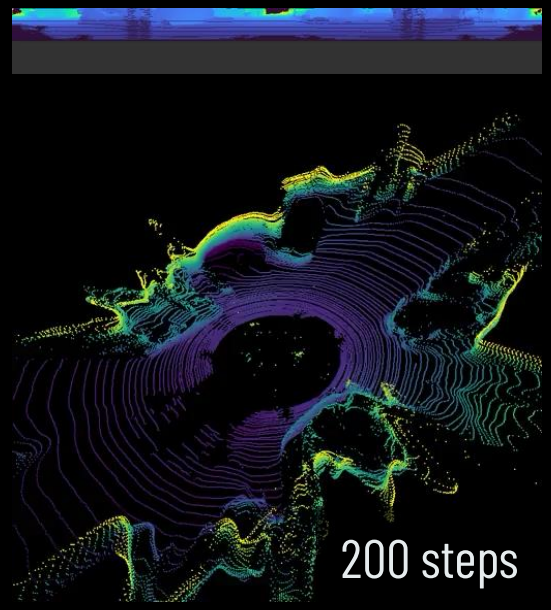
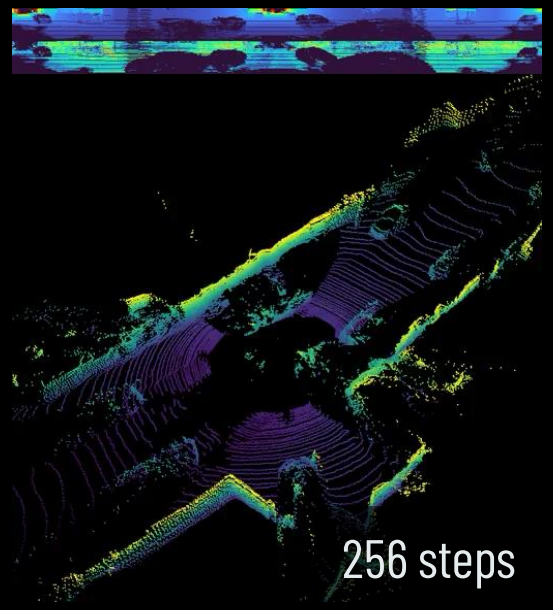
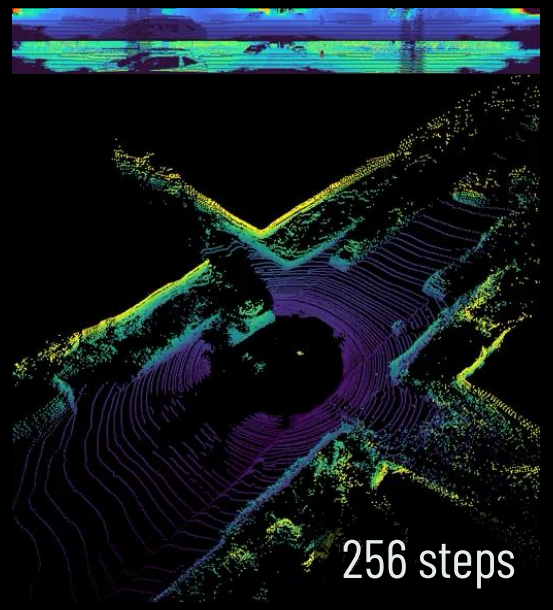
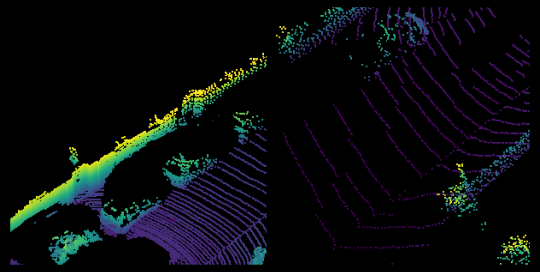
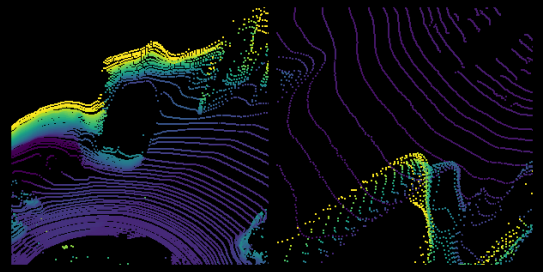
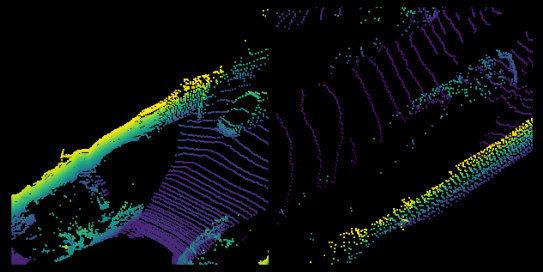
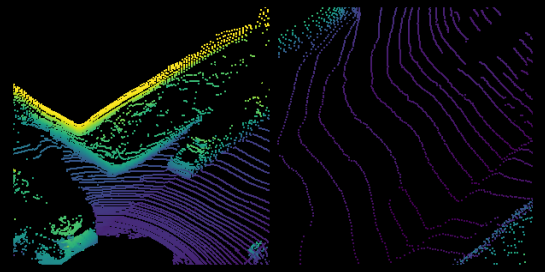
Experiments

Qualitative Comparison w/ Baselines

Unconditional Generation



Unconditional Generation

Training data KITTI-360 [Liao+ TPAMI'22]	LiDM (diffusion) [Ran+ CVPR'24]	R2DM (diffusion) [Nakashima+ ICRA'24]	***R2Flow*** (Ours)
			
	200 steps	256 steps	256 steps
			
	Clean but wavy structures	Sharp lines & structures	Sharp lines & structures

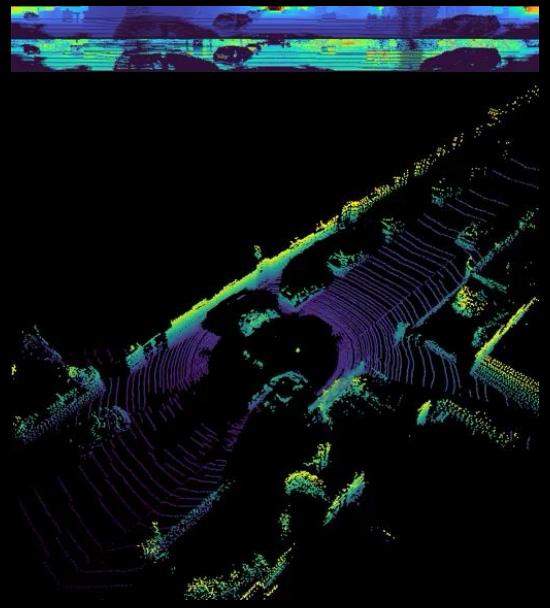
Overviews

Close-ups

Training data

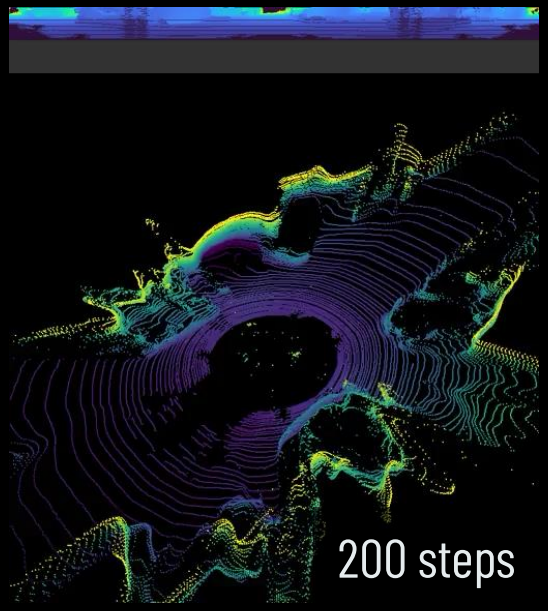
KITTI-360 [Liao+ TPAMI'22]

Overviews

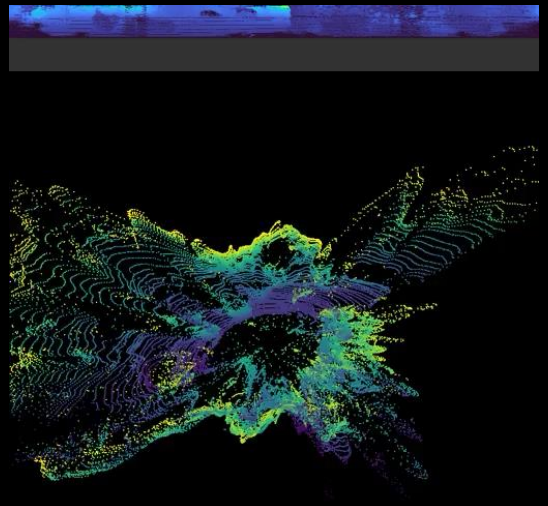


LiDM (diffusion)

[Ran+ CVPR'24]



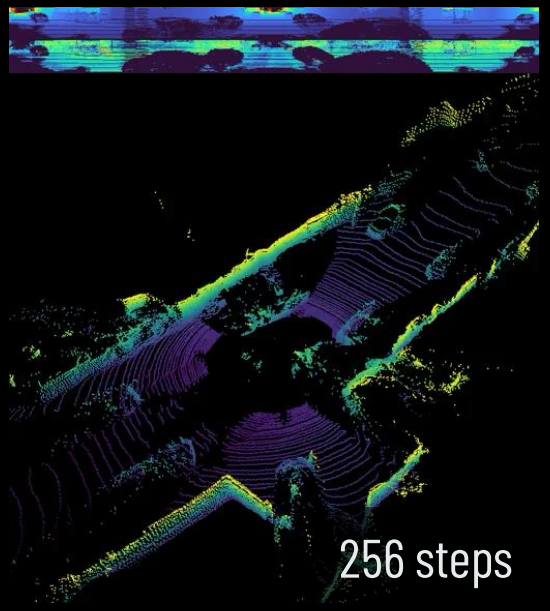
200 steps



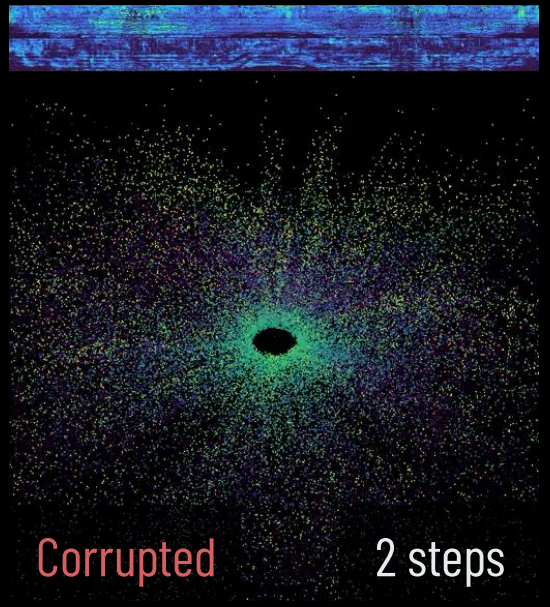
Corrupted 2 steps

R2DM (diffusion)

[Nakashima+ ICRA'24]



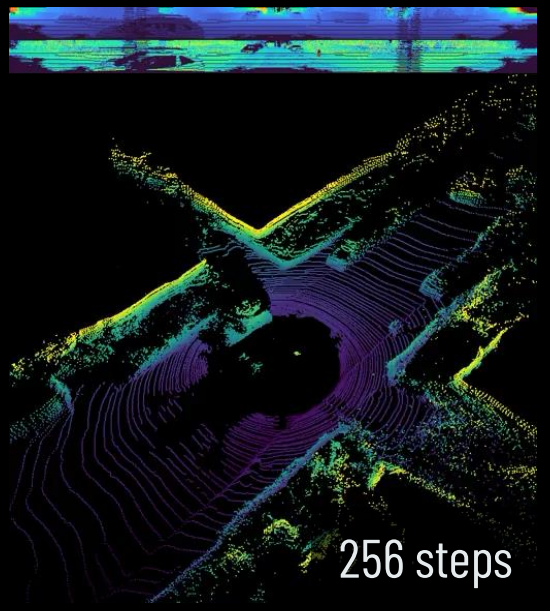
256 steps



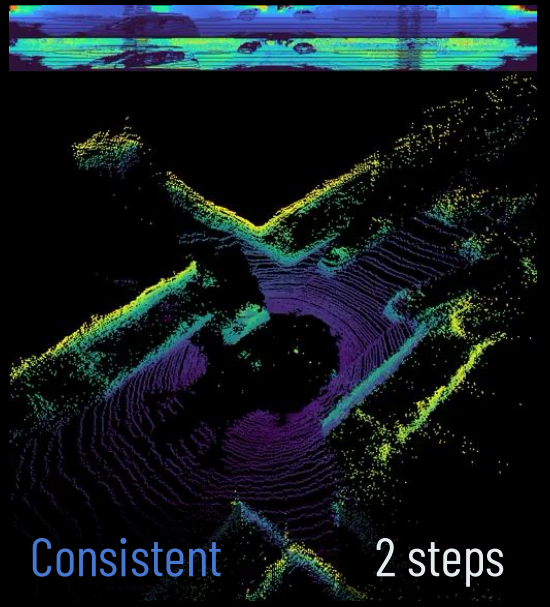
Corrupted 2 steps

R2Flow

(Ours)



256 steps



Consistent 2 steps