Taming Transformers for Realistic Lidar Point Cloud Generation

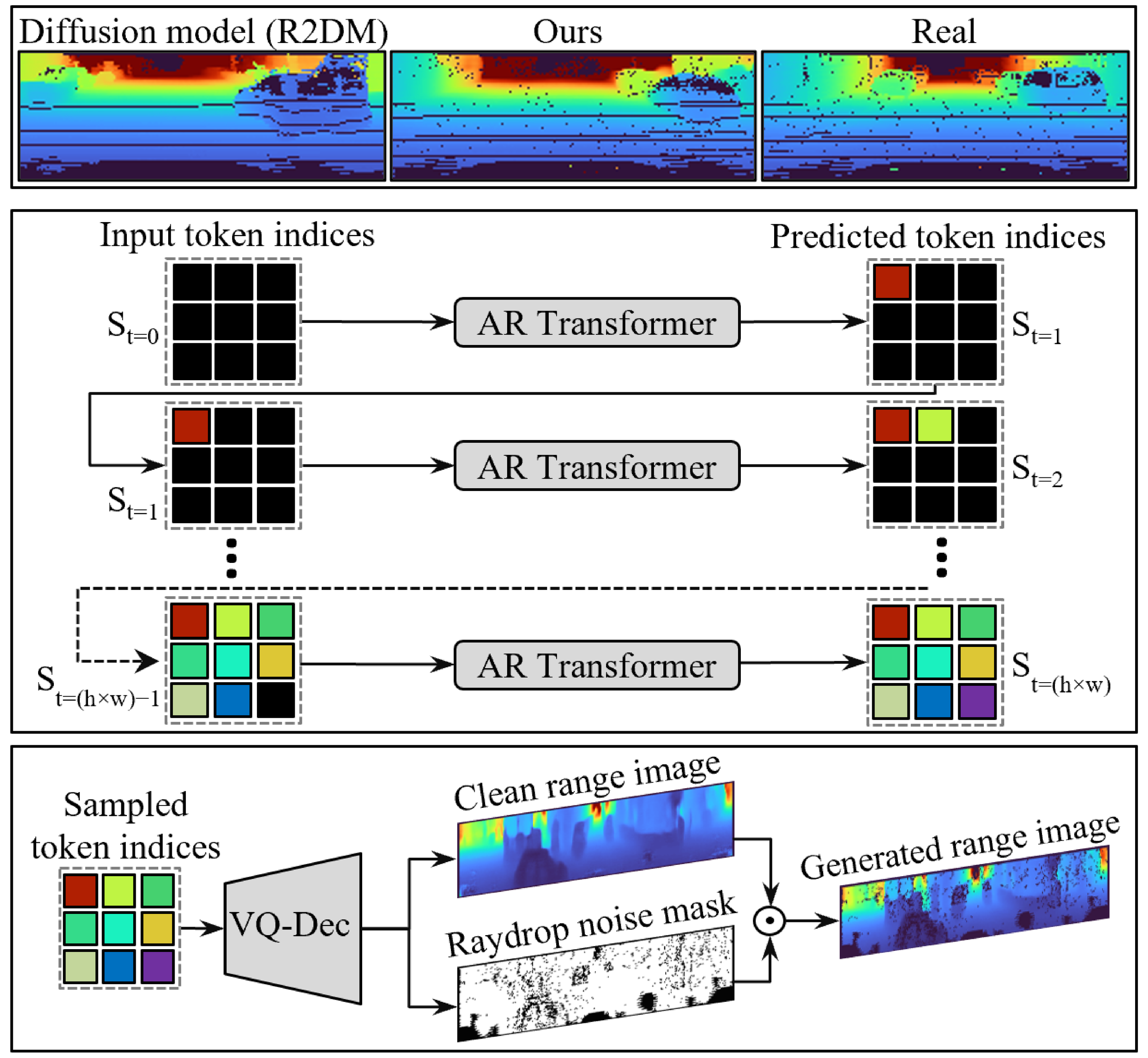
Hamed Haghighi, Amir Samadi, Mehrdad Dianati, Valentina Donzella, Kurt Debattista

WMG, THE UNIVERSITY OF WARWICK · QUEEN'S UNIVERSITY BELFAST · Hi-Drive Designing Automation

Paper · Contact

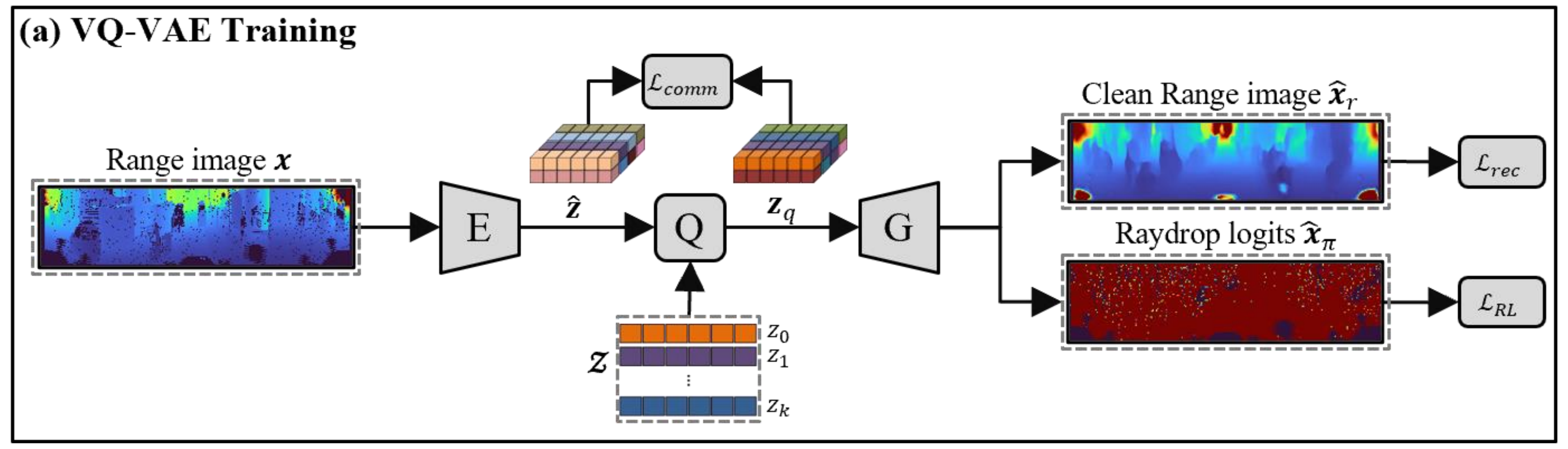
Introduction

- Motivation & Problem Statement:
  - Lidar sensors are crucial for autonomous vehicles, providing precise 3D environmental scans. However, realistic simulation of Lidar is challenging.
  - The traditional physics-based simulation algorithms require hefty computations; therefore, recent data-driven models have emerged.
  - Diffusion Models (DMs) have achieved SOTA in Lidar point cloud generation, however, they struggle to realistically model Lidar raydrop noise due to their denoising nature.
- Solution:
  - Retaining the strengths of DMs in iterative sampling and stable training by using auto-regressive transformers, while mitigating their deficiencies by decomposing range image and raydrop synthesis via an adapted VQ-VAE.

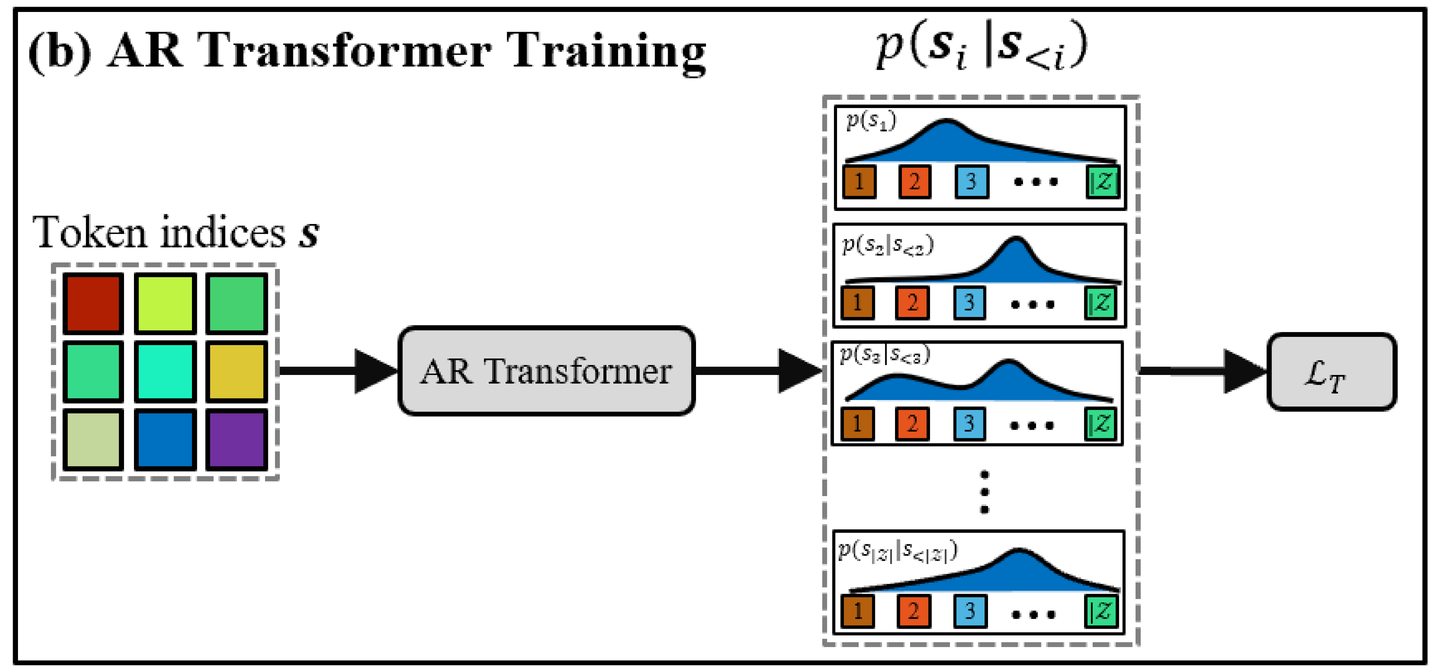


Methodology

- Representation of Lidar Point Cloud:
  - Spherical projection for KITTI-360: Convert (x, y, z) to (r, θ, φ).
  - Scan unfolding for KITTI odometry: Partition sequences into H sub-sequences.
- Adapting VQ-VAE:
  - Decomposing range image synthesis and raydrop estimation in VQ-VAE decoder
  - Incorporating geometric perseverance during training to improve the generalisability of the VQ-VAE..



- Auto-Regressive Transformer:
  - Training Objective: Model token interactions to predict token indices sequentially.
  - Output: Iterative token index sampling followed by VQ-VAE decoding to generate realistic range images and raydrop masks.



Results

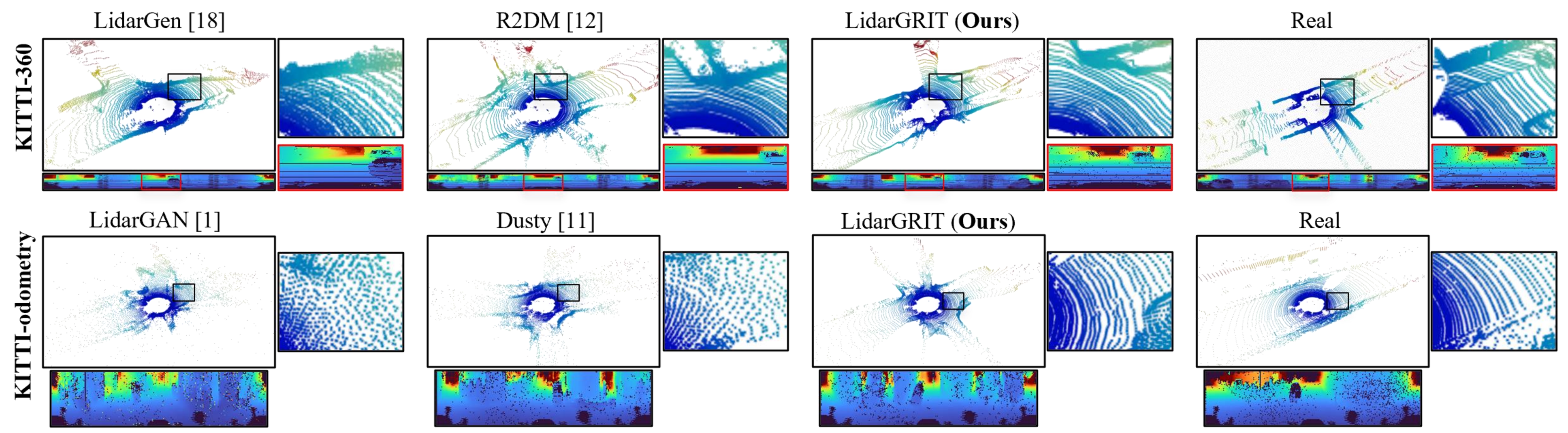
- Datasets: KITTI-360, KITTI-Odometry
- Quantitative Comparison:
  - KITTI-360:

	Image	BEV		Point cloud
Method	SWD×10² ↓	MMD×10⁴ ↓	JSD×10² ↓	FPD ↓
LidarGen [14]	33.93	2.19	5.70	43.27
UltraLiDAR [13]	N/A	2.23	10.52	N/A
R2DM [9]	20.82	4.00	4.55	**10.84**
LidarGRIT (**Ours**)	**10.29**	**2.16**	**3.93**	12.54

  - KITTI-odometry:

	Image	Point cloud		
Method	SWD×10² ↓	MD ×10³ ↓	JSD×10² ↓	FPD ↓
LidarGAN [1]	82.29	6.70	15.98	700
Dusty [8]	52.81	2.07	5.10	389
LidarGRIT (**Ours**)	**15.15**	**1.65**	**2.06**	**116**

- Qualitative Comparison:



Summary

- Introduced a novel generative model that integrated auto-regressive transformers with VQ-VAE for realistic Lidar point cloud generation.
- Incorporated raydrop estimation loss for accurate noise synthesis and geometric perseverance to improve VQ-VAE generalisability.
- Outperformed state-of-the-art models on KITTI-360 and KITTI odometry datasets, showing significant improvements across various metrics.