NFL 🏈: Neural LiDAR Fields for Novel View Synthesis

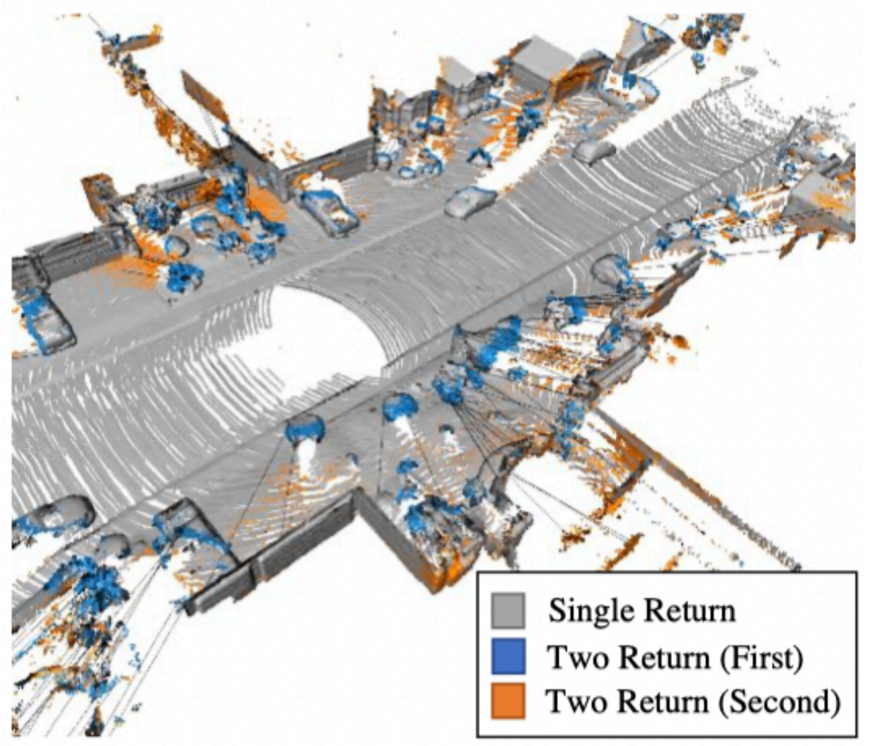
1 ETH zürich 2 NVIDIA 3 UNIVERSITY OF TORONTO 4 VECTOR INSTITUTE

Shengyu Huang[1,2] , Zan Gojcic[2] , Zian Wang[2,3,4], Francis Williams[2], Yoni Kasten[2], Sanja Fidler[2,3,4], Konrad Schindler[1], Or Litany[2]

ICCV23 PARIS

Project page

Problem statement



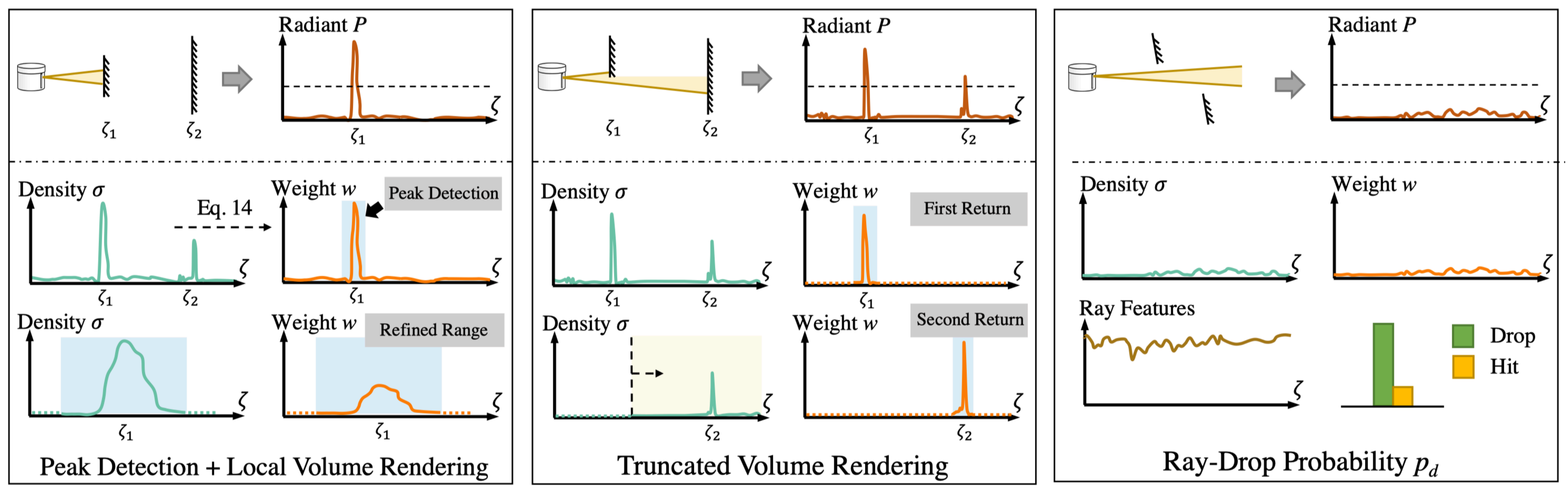
- **Input:** LiDAR scans $\mathbf{X} = \{\mathbf{X}_v\}_{v=1}^{n_v}$, where each scan $\mathbf{X}_v$ is associated with a sensor pose $\mathbf{T}_v \in \mathrm{SE}(3)$ and consists of n_r rays. Every ray $\mathbf{r}(\mathbf{o}, \mathbf{d})$ records measurements $(\zeta_1, e_1, p_d, p_s, \zeta_2, e_2)$ as:
  - ζ_* : range of the first/second return
  - e_* : intensity of the first /second return
  - $p_d \in \{0,1\}$: ray drop mask
  - $p_s \in \{0,1\}$: two return mask
- **Goal:** Render virtual LiDAR scans $\mathbf{X}_{tgt}$ from novel sensor poses $\mathbf{T}_{tgt}$.

Volume rendering for active sensor

- Range-dependent radiant power $P(\zeta)$ at range ζ:
$$P(\zeta) = \int_0^{2\zeta/c} P_e(t) H(\zeta - \frac{ct}{2})\, dt\,, \quad H(\zeta) = H_T(\zeta) H_C(\zeta)$$
$$H_T(\zeta) = \frac{\rho}{\pi} \cos(\theta) \delta(\zeta - \zeta_0)\,, \quad H_C(\zeta) = T_\zeta^2 \frac{A_e}{\zeta^2}\,.$$
- Probabilistic radiant power P_ζ at range ζ:
$$P_\zeta = C \frac{T_\zeta^2 \cdot \sigma_\zeta \rho_\zeta}{\zeta^2} \cos(\theta)\,.$$
- Volume rendering of the probabilistic radiant:
$$P = \sum_{j=1}^{N} \int_{\zeta_j}^{\zeta_{j+1}} C \frac{T_\zeta^2 \cdot \sigma_\zeta \rho_\zeta}{\zeta^2} \cos(\theta_j)\, d\zeta = \sum_{j=1}^{N} w_j \rho'_{\zeta_j}\,.$$
- Active sensing versus passive sensing:
$$\boxed{w_j = 2\alpha_{\zeta_j} \cdot \prod_{k=1}^{j-1} (1 - 2\alpha_{\zeta_k})}, \quad \boxed{w_j = \alpha_{\zeta_j} \cdot \prod_{k=1}^{j-1} (1 - \alpha_{\zeta_k})}.$$

Neural Field for LiDAR NVS

- Neural LiDAR field: $F : (\mathbf{x}, \mathbf{d}) \mapsto (\sigma, \rho, p_d)$.
- LiDAR beam rendering:

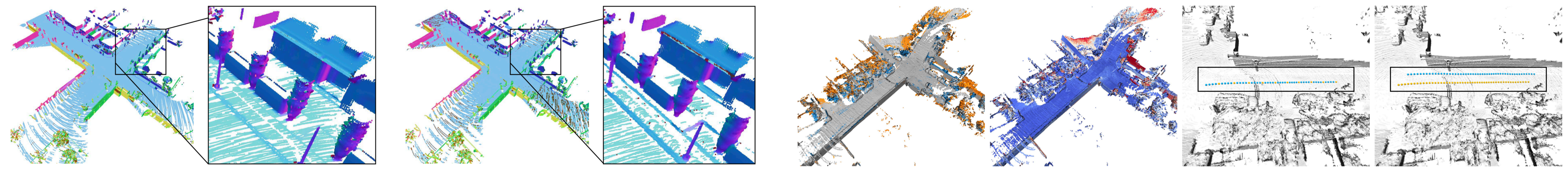


Peak Detection + Local Volume Rendering | Truncated Volume Rendering | Ray-Drop Probability p_d

- Loss terms: $L = L_{\text{range}} + \lambda_e L_e + \lambda_d L_d + \lambda_s L_s$.

Datasets and evaluation protocol

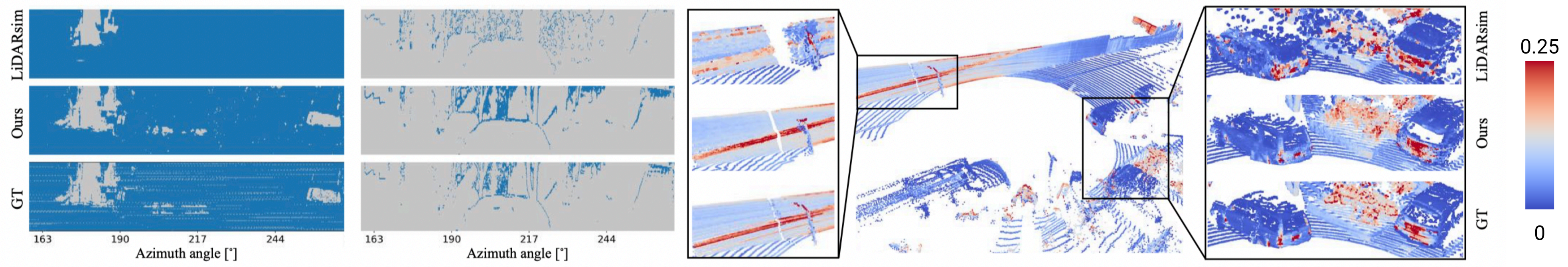
- *TownClean* and *TownReal* datasets that are simulated using *idealised* and *diverging beam* profile, respectively;
- *Waymo Interp:* four static scenes from Waymo Open datasets, 40/10 scans for training/testing;
- *Waymo NVS*: same scenes, train on entire 50 scans → test from different sensor trajectory → re-train on the novels views → test on the original 50 scans.



(a) TownClean (b) TownReal (c) Waymo Interp. (d) Waymo NVS

Experimental results

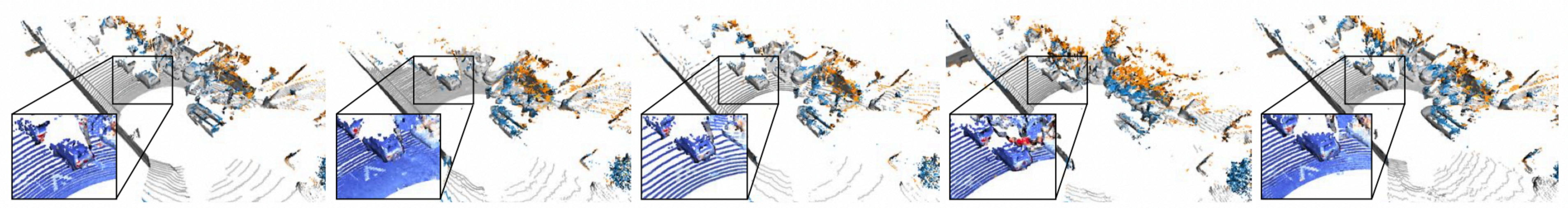
- Comprehensive ray measurement evaluation of LiDAR NVS on *Waymo Interp.* dataset.



(a) First return mask (b) Second return mask (c) LiDAR scans coloured by intensity values

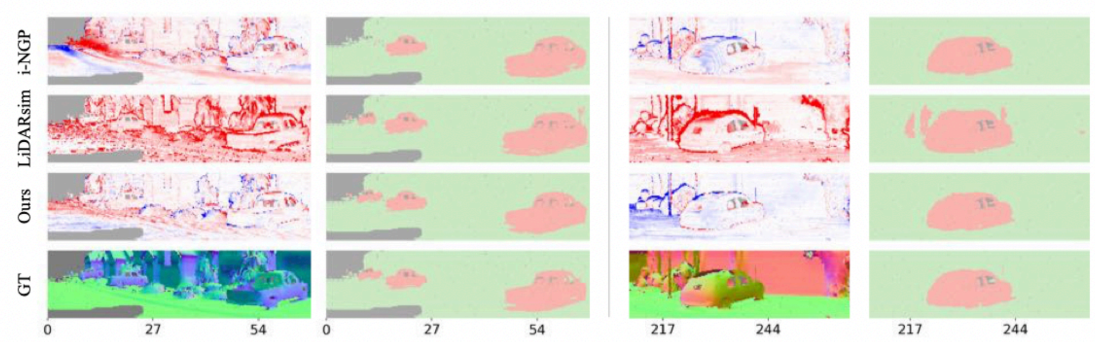
Method		First range			Second range					Intensity		Ray drop		
		Recall@50↑	MAE↓	MedAE↓	Seg. recall↑	Seg. precision↑	Recall@50↑	MAE↓	MedAE↓	MSE^{1st}↓	MSE^{2nd}↓	Recall↑	Precision↑	IoU↑
LiDARsim		74.1	105.4	18.5	3.5	11.5	1.0	2258.0	1898.2	0.013	0.018	32.5	**85.5**	30.5
Ours	Central ray	92.8	32.8	5.6	79.8	62.9	61.1	589.1	21.8	0.004	0.009	64.3	81.7	**57.1**
	Diverged beam	92.3	36.1	5.7	82.1	55.6	67.4	505.1	13.4	**0.004**	**0.008**	**65.1**	78.0	56.1
	GT mask	**93.2**	**29.7**	**5.6**	**100.0**	**100.0**	**79.8**	**116.0**	**8.1**	0.004	0.011	65.1	78.0	56.1

- LiDAR novel view synthesis by changing the sensor elevation angle θ, pose (x, y, z), and number of beams.



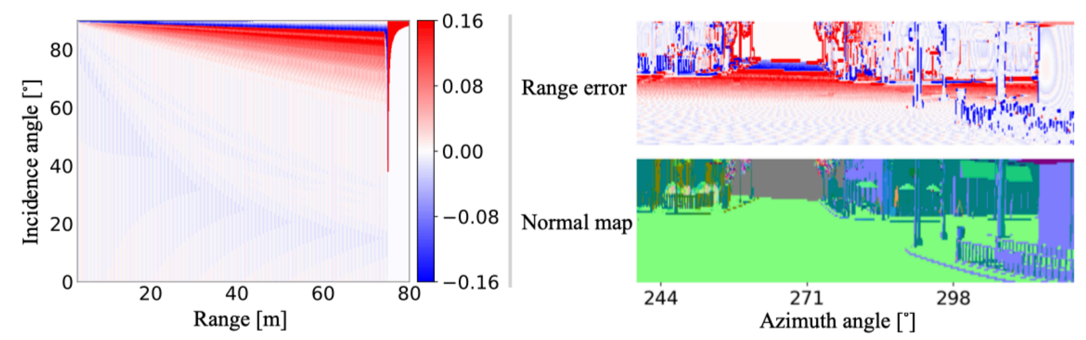
Ref. frame $(\theta_0,\ x_0,\ y_0,\ z_0)$ (a) $\theta = \theta_0 - 5$ (b) 64 →32 beam; $\theta = \theta_0 - 2$ (c) $z = z_0 + 1$ (d) $x = x_0 + 1;\ \ y = y_0 + 1$

- Evaluation of synthesised novel views via proxy task. In this example, the geometry in-accuracy (-100 100 cm) leads to erroneous semantic segmentation.



Method	Vehicle			Background		
	Recall↑	Precision↑	IoU↑	Recall↑	Precision↑	IoU↑
i-NGP	93.2	85.9	80.9	98.3	99.2	97.6
DS-NeRF	90.7	**87.1**	80.2	**98.5**	98.9	97.4
URF	87.8	81.7	73.7	98.0	98.4	96.5
Lidarsim	90.5	70.5	65.9	94.9	99.0	94.0
Ours	**95.9**	87.0	**83.9**	98.3	**99.5**	**97.8**

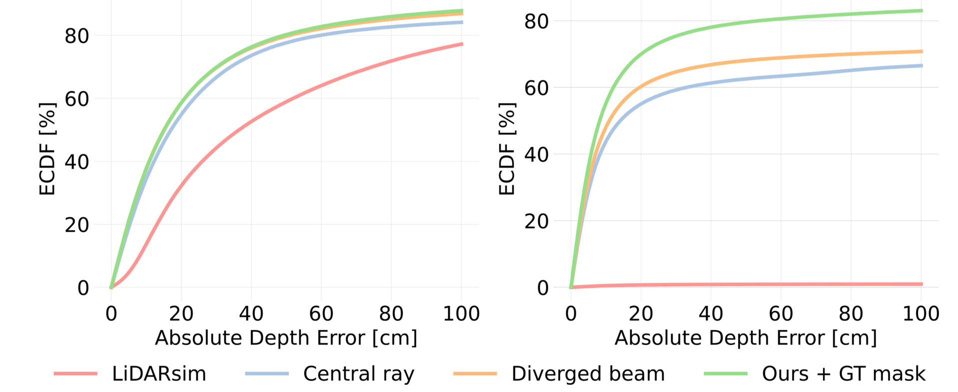
- Effects of waveform discretisation and beam divergence.



- Effectiveness of volume rendering for active sensing.

Method	TownClean			Waymo Interp.		
	MAE↓	MedAE↓	CD↓	MAE↓	MedAE↓	CD↓
i-NGP	41.0 (-1.2)	4.1 (0.0)	17.6 (0.2)	25.3 (-1.1)	4.5 (-1.0)	10.5 (-1.1)
DS-NeRF	37.4 (-4.2)	3.0 (-0.9)	14.4 (-2.2)	27.4 (-0.8)	5.4 (-1.0)	13.6 (-0.9)
URF	46.4 (3.0)	4.5 (0.3)	18.4 (1.6)	28.3 (0.1)	5.3 (-0.1)	13.1 (0.2)
Ours	32.0 (-2.1)	2.3 (-2.5)	9.0 (-3.9)	30.8 (-2.1)	5.1 (-2.0)	12.1 (-2.3)

- Benefits of *diverging beam* profile.



- Results on *Waymo NVS* dataset by varying the displacement. Numbers are reported as *MedAE/CD* [cm].

	i-NGP	DS-NeRF	URF	LiDARsim	Ours
(0.5, 0.5, 0.5)	7.0 / 14.4	7.0 / 16.0	9.0 / 19.6	16.1 / 33.1	**5.4 / 13.0**
(1.5, 1.5, 1.0)	8.4 / 17.6	7.8 / 18.5	11.0 / 27.5	16.5 / 37.9	**5.8 / 14.3**
(2.5, 2.5, 1.5)	11.6 / 28.0	9.3 / 22.8	13.9 / 35.5	17.2 / 46.3	**6.4 / 18.4**

Our contribution

- Devised the volume rendering scheme for active sensors;
- Proposed a neural field model grounded on the physical LiDAR sensing process;
- Proposed a closed-loop evaluation protocol to evaluate NVS performance;
- Developed a LiDAR simulator that supports *diverging beam* profile.

Future work

- Extend to handle dynamic scenes;
- Additionally handle scattering and attenuation effects in adverse weather;
- Recover “true” geometry from biased measurements.