NLOS-NeuS: None-line-of-sight Neural Implicit Surface

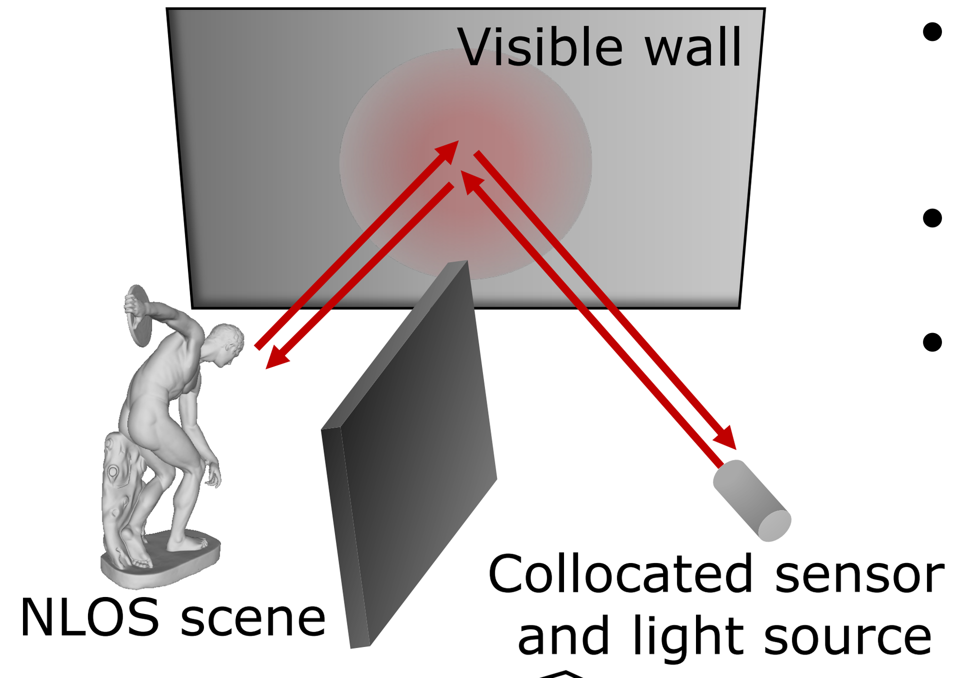
Yuki Fujimura Takahiro Kushida Takuya Funatomi Yasuhiro Mukaigawa

NAIST Nara Institute of Science and Technology

ICCV23 PARIS

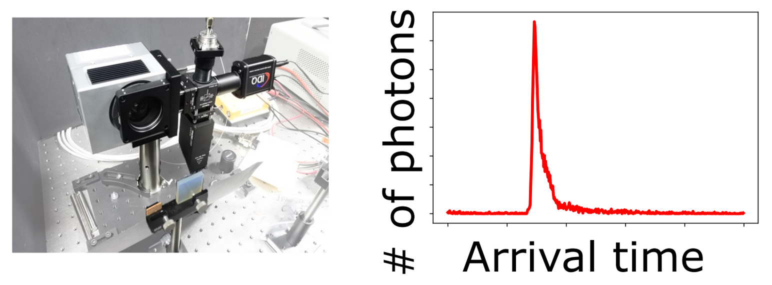
Project page

Non-line-of-sight (NLOS) imaging



- NLOS imaging is to reconstruct the invisible scene from the sensor and light source
- In the NLOS setup, the sensor only see the visible wall
- The typical input of the NLOS imaging is indirect light captured by a time-resolved sensor such as a single photon avalanche diode (SPAD)

Transient histogram captured by a SPAD



- Photon counts at each arrival time
- The time resolution is from [ns] to [ps]

Related work

- DLCT [CVPR2020] is proposed for NLOS surface reconstruction with discretized voxel grid representation
- NeTF [ICCP2021] uses neural field similar to NeRF for NLOS imaging
- We propose a neural field approach for NLOS surface reconstruction with continuous implicit surface (signed distance function (SDF))

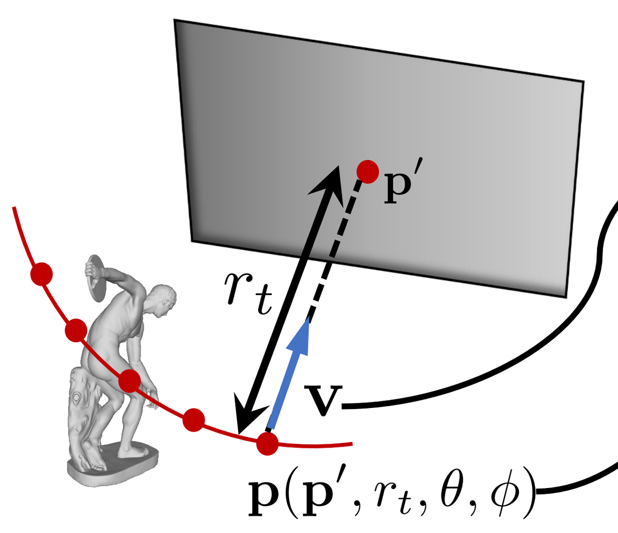
Method	Scene representation	Output geometry
LCT [Nature2018]	Voxel grid	Volumetric density
DLCT [CVPR2020]	Voxel grid	Volumetric density + surface normals
NeTF [ICCP2021]	Neural field	Volumetric density
NLOS-NeuS (Ours)	**Neural field**	**Implicit surface (SDF)**

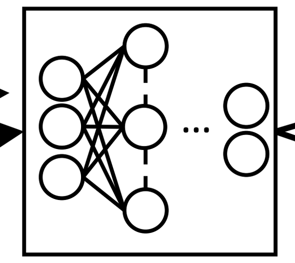
Overview: Volume rendering for transient histogram

(1) We sample points on a sphere centered at $\mathbf{p}'$ with radius $r = ct/2$ (c: light speed), which corresponds to time bin t

(2) MLP takes each sample point for estimating reflectance and signed distance, which is then converted to density

(3) We apply volume rendering to reconstruct transient histogram, which is compared to the observed histogram to compute a training loss





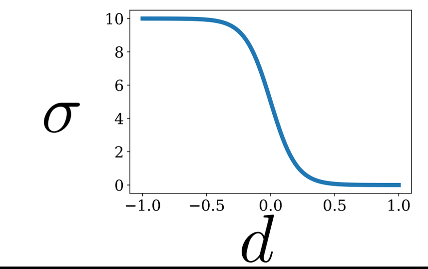
Reflectance $\rho(\mathbf{p}, \mathbf{v})$

Signed distance $d(\mathbf{p})$ → Density $\sigma(\mathbf{p})$

The sum on the sphere

$$\tau(\mathbf{p}', t) = \sum_{\theta,\phi} A(r_t, \theta) w(\mathbf{p}, \mathbf{v}) \rho(\mathbf{p}, \mathbf{v}) \Delta\theta\Delta\phi$$

$$\sigma(\mathbf{p}) = \frac{1}{\alpha} Sigmoid\left(-\frac{d(\mathbf{p})}{\alpha}\right)$$

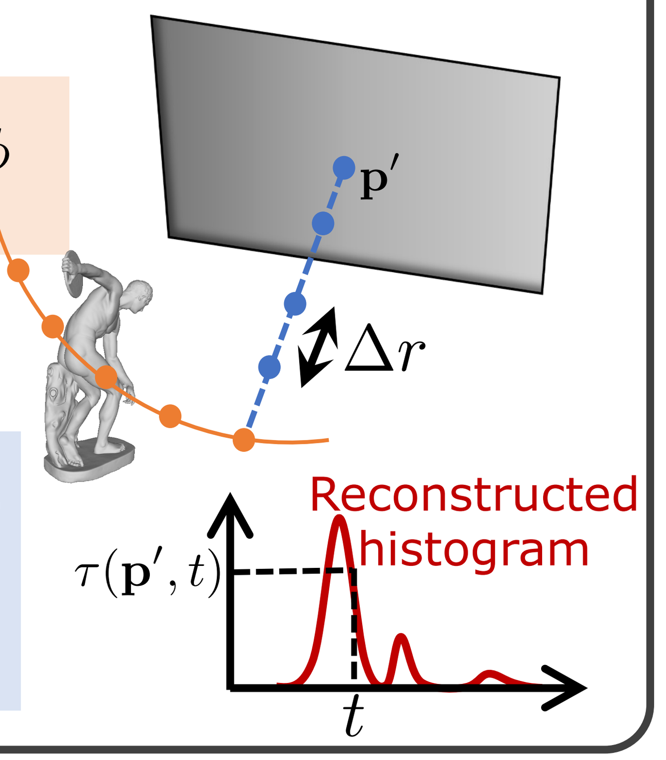


- We use the conversion proposed in StyleSDF [CVPR2022] because of its computational efficiency
- Positive distance (outside object) returns small density
- α controls tightness, i.e., $\alpha \to 0$ represents perfectly sharp surface

Volume rendering on a ray from the wall

$$w(\mathbf{p}, \mathbf{v}) = \sum_{s=t_{min}}^{t} T_s \left(1 - \exp\{-\sigma(\mathbf{p}(\mathbf{p}', r_s, \theta, \phi))\Delta r_s\}\right)$$

$$T_s = \exp\left\{-\sum_{u=t_{min}}^{s-1} \sigma(\mathbf{p}(\mathbf{p}', r_u, \theta, \phi))\Delta r_u\right\}$$



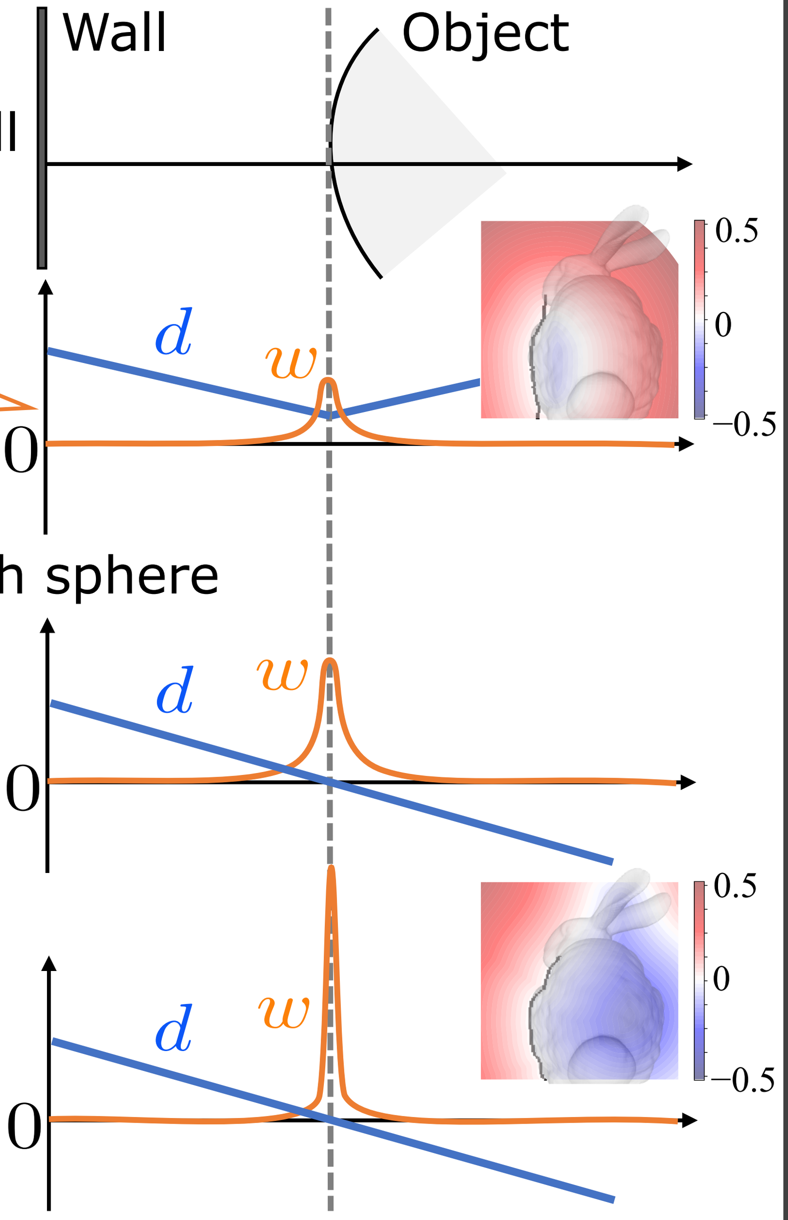
Key: Constraints for learning SDF in the NLOS setup

In the NLOS setup,

- The object is not observed directly
- Only one side of the object is observed from the wall

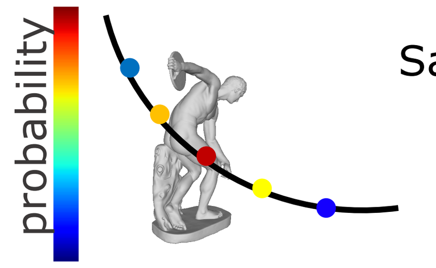
This leads to incorrect SDF

Common failure case:
Volume rendering weight is the highest at a point with non-zero signed distance



(1) <u>Self-supervised zero level-set learning</u>

During training, we compute PDF based on $w\rho$ on each sphere
The signed distances at sampled points with the PDF are forced to be 0



Sample points

$$\mathcal{L}_z = \sum_i d(\mathbf{p}_i)$$

(2) <u>Constraint on volume rendering weights</u>

By reducing α, we generate sharp w for suppressing effects from non-zero signed distance points.

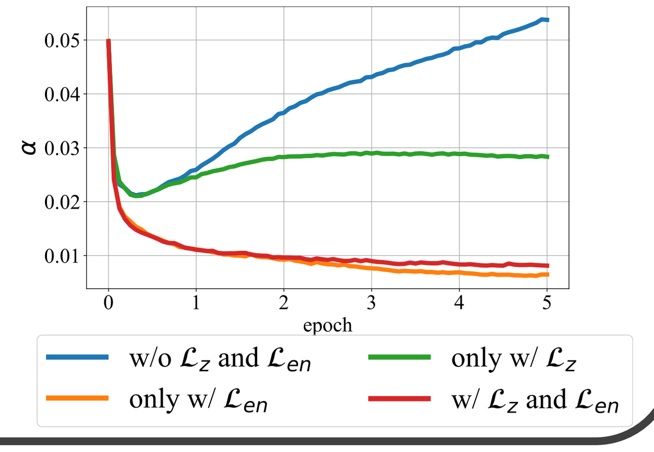
Specifically, we use

$$\mathcal{L}_{en} = \sum_{\mathbf{p}',\theta,\phi} -\hat{o}\log_2\hat{o} - (1-\hat{o})\log_2(1-\hat{o})$$

where $\hat{o} = \sum_{t=t_{min}}^{t_{max}} w(\mathbf{p}, \mathbf{v})$

(Intuitively, all densities outside the object should be 0, mathematical discussion is in the supplementary)

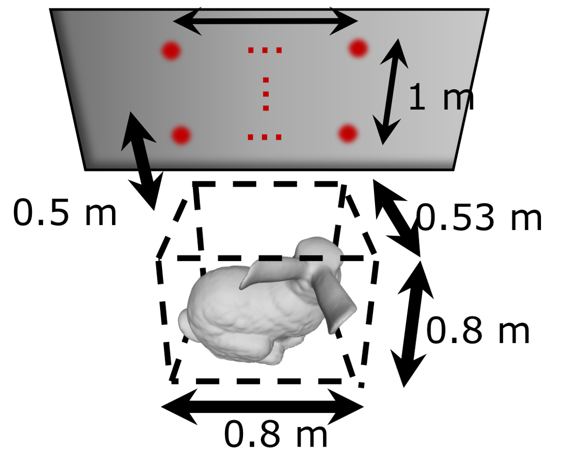
α plots during training w/ and w/o each loss

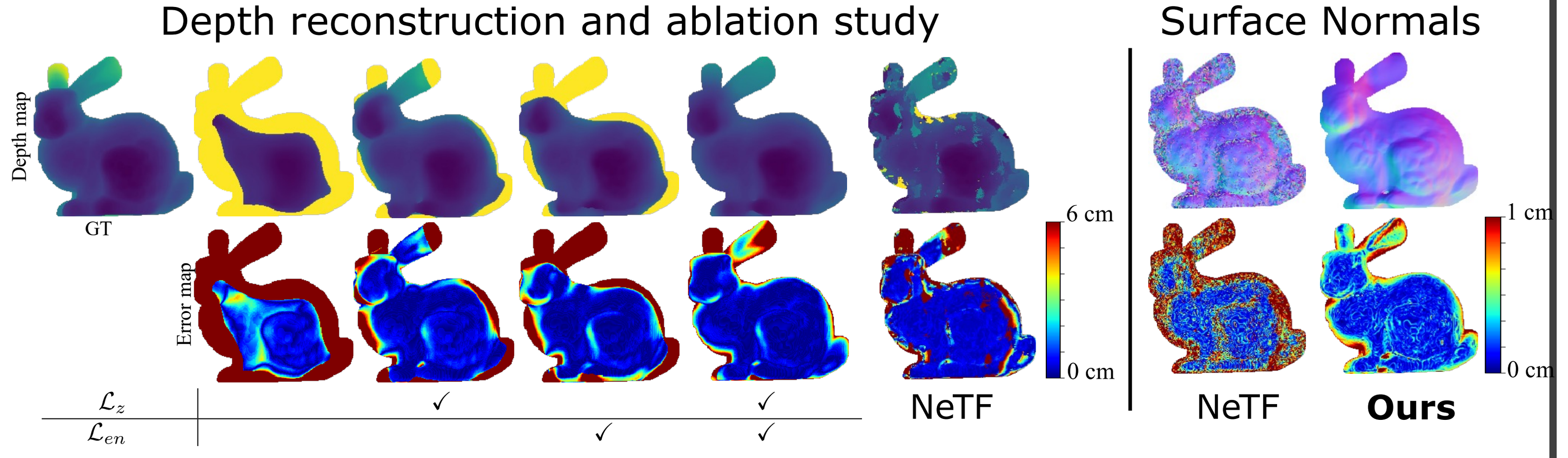


Experiments

(1) Synthetic dataset (ZNLOS dataset [ICCP2019])

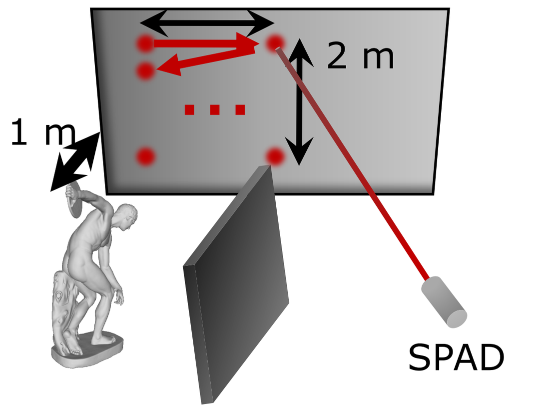
- Scan region is 1m x 1m
- 256 x 256 observed points
- # of histogram time bins is 200

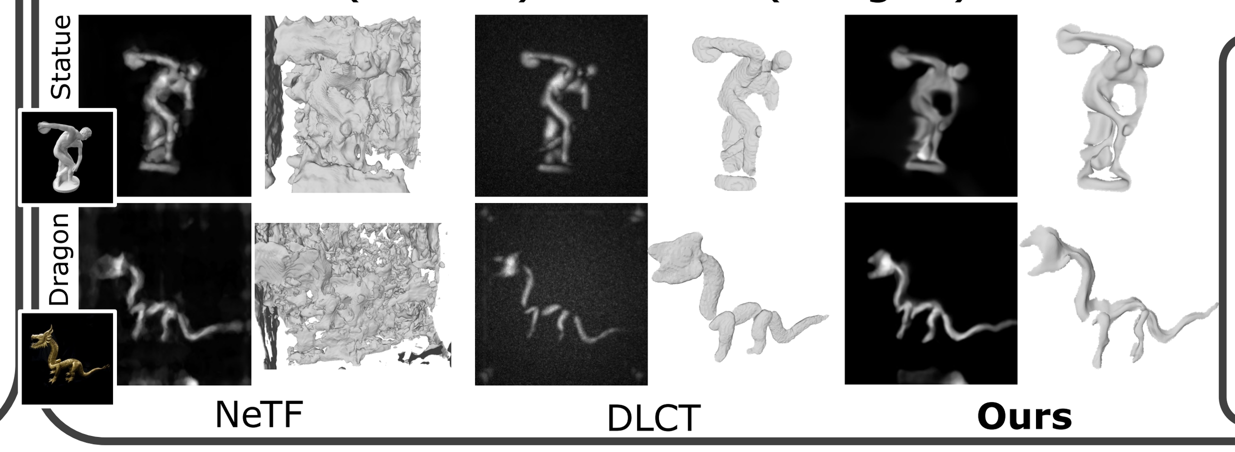




(2) Real dataset (f-k dataset [SIGGRAPH2019])

- Captured by SPAD
- Scan region is 2m x 2m
- 256 x 256 observed points
- # of histogram time bins is 160 (Statue) and 120 (Dragon)





Challenging scene (limitation)

Thin structures are difficult for SDF representation