Onsite Non-line-of-sight Imaging via Online Calibration

I. A derivation for NLOS imaging models

We provide a detailed derivation for the NLOS imaging model in the main text, based on the physics of light transport. When light is emitted from a laser beam $\mathbf{o}$ toward a relay wall W, the irradiance at an illumination point $\mathbf{l}$ is expressed as:

$$E_{\mathbf{o}\to\mathbf{l}} = \frac{\Phi_{\mathbf{o}}}{A_{\mathbf{l}}} \tag{S1}$$

where $A_{\mathbf{l}}$ represents the size of the laser speckle on the relay wall, and $\Phi_{\mathbf{o}}$ the radiant flux (i.e., photon energy emitted per unit time). Assuming that the relay wall is Lambertian, and its bidirectional reflectance distribution function (BRDF) $f_W = \rho/\pi$ (where ρ is the albedo), the radiance from $\mathbf{l}$ toward a point $\mathbf{p}$ is defined as:

$$L_{\mathbf{l}\to\mathbf{p}} = f_W E_{\mathbf{l}\leftarrow\mathbf{o}} = \frac{\rho\Phi_{\mathbf{o}}}{\pi A_{\mathbf{l}}} \tag{S2}$$

and the irradiance received at $\mathbf{p}$ from $\mathbf{l}$ is:

$$E_{\mathbf{p}\leftarrow\mathbf{l}} = \int_{A_{\mathbf{l}}} L_{\mathbf{l}\to\mathbf{p}} \cos(\theta_{\mathbf{p},\mathbf{l}}) \mathrm{d}\Omega_{\mathbf{p},\mathbf{l}} \tag{S3}$$

where $\theta_{\mathbf{p},\mathbf{l}}$ is the angle between the incident ray, from $\mathbf{l}$ to $\mathbf{p}$, and the normal $\mathbf{n}_{\mathbf{p}}$ at the point $\mathbf{p}$. $\mathrm{d}\Omega_{\mathbf{p},\mathbf{l}}$ denotes the solid angle centered at the point $\mathbf{p}$ subtended by area $A_{\mathbf{l}}$. When the speckle is sufficiently small such that $\sqrt{A_{\mathbf{l}}} \ll |\mathbf{l}-\mathbf{p}|$, $\cos(\theta_{\mathbf{p},\mathbf{l}})$ can be considered as a constant, as well as $L_{\mathbf{l}\to\mathbf{p}}$. The integral in (S3) becomes:

$$E_{\mathbf{p}\leftarrow\mathbf{l}} = L_{\mathbf{l}\to\mathbf{p}} \cos(\theta_{\mathbf{p},\mathbf{l}}) \Omega_{\mathbf{p},\mathbf{l}} = \frac{\rho\Phi_{\mathbf{o}}}{\pi A_{\mathbf{l}}} (\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}}) \frac{(\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}}) A_{\mathbf{l}}}{|\mathbf{l}-\mathbf{p}|^2} \tag{S4}$$

For simplicity, we denote the unit vector from $\mathbf{a}$ to $\mathbf{b}$ as $\boldsymbol{\omega}_{\mathbf{a}\to\mathbf{b}} = \frac{\mathbf{b}-\mathbf{a}}{|\mathbf{b}-\mathbf{a}|}$. After reflection from the hidden object P, the radiance at the detection point $\mathbf{s}$ on the relay wall is defined as:

$$\begin{aligned} L_{\mathbf{p}\to\mathbf{s}} &= f_P(\mathbf{p}; \boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}}, \boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}}) E_{\mathbf{p}\leftarrow\mathbf{l}} \\ &= f_P(\mathbf{p}; \boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}}, \boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}}) \frac{\rho\Phi_{\mathbf{o}}}{\pi} (\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}}) (\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}}) \frac{1}{|\mathbf{l}-\mathbf{p}|^2} \end{aligned} \tag{S5}$$

and the irradiance received at the detection point $\mathbf{s}$ from the hidden object P is expressed as:

$$\begin{aligned} E_{\mathbf{s}\leftarrow P} &= \int_P L_{\mathbf{p}\to\mathbf{s}} \cos(\theta_{\mathbf{s},\mathbf{p}}) \mathrm{d}\Omega_{\mathbf{s},\mathbf{p}} \\ &= \int_P f_P(\mathbf{p}; \boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}}, \boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}}) \frac{\rho\Phi_{\mathbf{o}}}{\pi} (\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}}) (\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}}) \frac{1}{|\mathbf{l}-\mathbf{p}|^2} \\ &\quad \cdot (\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}}) (\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{p}}) \frac{1}{|\mathbf{p}-\mathbf{s}|^2} \mathrm{d}A_{\mathbf{p}} \end{aligned} \tag{S6}$$

The radiance that arrives at the detector $\mathbf{d}$, after scattered off the relay wall W, is thus formulated as:

$$\begin{aligned} L_{\mathbf{s}\to\mathbf{d}} &= f_W E_{\mathbf{s}\leftarrow P} \\ &= \Phi_{\mathbf{o}} \left(\frac{\rho}{\pi}\right)^2 \int_P f_P(\mathbf{p}; \boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}}, \boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}}) \\ &\quad \cdot \frac{(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}}) (\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}}) (\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}}) (\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{p}})}{|\mathbf{l}-\mathbf{p}|^2 |\mathbf{p}-\mathbf{s}|^2} \mathrm{d}A_{\mathbf{p}} \end{aligned} \tag{S7}$$

Similarly, assuming that the sensing area of the detector is sufficiently small ($\sqrt{A_{\mathbf{s}}} \ll |\mathbf{s}-\mathbf{d}|$) on the relay wall, the irradiance at the detector becomes:

$$E_{\mathbf{d}\leftarrow\mathbf{s}} = L_{\mathbf{s}\to\mathbf{d}} \cos(\theta_{\mathbf{d},\mathbf{s}}) \Omega_{\mathbf{d},\mathbf{s}} \tag{S8}$$

In general, $\theta_{\mathbf{d},\mathbf{s}}$ is close to 0 (i.e., the ray from $\mathbf{s}$ to $\mathbf{d}$ is orthogonal to the detector), so:

$$\cos(\theta_{\mathbf{d},\mathbf{s}}) = (\boldsymbol{\omega}_{\mathbf{d}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{d}}) = 1 \tag{S9}$$

and $\Omega_{\mathbf{d},\mathbf{s}}$ is the solid angle centered at $\mathbf{d}$ subtended by area $A_{\mathbf{s}}$:

$$\Omega_{\mathbf{d},\mathbf{s}} = \frac{(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{d}} \cdot \mathbf{n}_{\mathbf{s}})A_{\mathbf{s}}}{|\mathbf{s}-\mathbf{d}|^2} \tag{S10}$$

The radiant flux onto the detector becomes:

$$\begin{aligned}\Phi_{\mathbf{d}} &= E_{\mathbf{d}\leftarrow\mathbf{s}}A_{\mathbf{d}} \\ &= \Phi_{\mathbf{o}}\left(\frac{\rho}{\pi}\right)^2(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{d}} \cdot \mathbf{n}_{\mathbf{s}})\frac{A_{\mathbf{d}}A_{\mathbf{s}}}{|\mathbf{s}-\mathbf{d}|^2}\int_P f_P(\mathbf{p};\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}},\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}}) \\ &\cdot \frac{(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}})(\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}})(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}})(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{p}})}{|\mathbf{l}-\mathbf{p}|^2|\mathbf{p}-\mathbf{s}|^2}\mathrm{d}A_{\mathbf{p}}\end{aligned} \tag{S11}$$

The aforementioned equations are derived for a continuous wave laser. For a pulsed laser, we further account for the Dirac delta function:

$$\begin{aligned}\Phi_{\mathbf{d}}(t) = N_{\mathbf{o}}h\nu\left(\frac{\rho}{\pi}\right)^2(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{d}} \cdot \mathbf{n}_{\mathbf{s}})\frac{A_{\mathbf{d}}A_{\mathbf{s}}}{|\mathbf{s}-\mathbf{d}|^2}\int_P \delta(t-(|\mathbf{l}-\mathbf{o}|+|\mathbf{p}-\mathbf{l}|+|\mathbf{s}-\mathbf{p}|+|\mathbf{d}-\mathbf{s}|)/c) \\ \cdot f_P(\mathbf{p};\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}},\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}})\frac{(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}})(\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}})(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}})(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{p}})}{|\mathbf{l}-\mathbf{p}|^2|\mathbf{p}-\mathbf{s}|^2}\mathrm{d}A_{\mathbf{p}}\end{aligned} \tag{S12}$$

Moreover, we consider occlusions in the scenario, υ, with respect to $\mathbf{p}$, $\mathbf{l}$, and $\mathbf{s}$. We denote f as the BRDF of a hidden object P, instead of f_P. The transient (i.e., photons detected per unit time) can be expressed by dividing the photon energy $h\nu$, where ν is frequency:

$$\begin{aligned}\tau(t;\mathbf{o},\mathbf{l},\mathbf{s},\mathbf{d}) = N_{\mathbf{o}}\left(\frac{\rho}{\pi}\right)^2(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{d}} \cdot \mathbf{n}_{\mathbf{s}})\frac{A_{\mathbf{d}}A_{\mathbf{s}}}{|\mathbf{s}-\mathbf{d}|^2}\int_P \delta(t-(|\mathbf{l}-\mathbf{o}|+|\mathbf{p}-\mathbf{l}|+|\mathbf{s}-\mathbf{p}|+|\mathbf{d}-\mathbf{s}|)/c) \\ \cdot f(\mathbf{p};\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}},\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}})\frac{(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}})(\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}})(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}})(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{p}})}{|\mathbf{l}-\mathbf{p}|^2|\mathbf{p}-\mathbf{s}|^2}\upsilon(\mathbf{p};\mathbf{l},\mathbf{s})\mathrm{d}A_{\mathbf{p}}\end{aligned} \tag{S13}$$

After that, we make some denotations to simplify (S13): $t_{\mathbf{a}\to\mathbf{b}} = \frac{|\mathbf{b}-\mathbf{a}|}{c}$ denotes the travel time from point $\mathbf{a}$ to point $\mathbf{b}$. Further, we define the *Gamma* map coefficients $\Gamma(\mathbf{s})$ and attenuation factor as

$$\Gamma(\mathbf{s}) = N_{\mathbf{o}}\left(\frac{\rho}{\pi}\right)^2(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{d}} \cdot \mathbf{n}_{\mathbf{s}})\frac{A_{\mathbf{d}}A_{\mathbf{s}}}{|\mathbf{s}-\mathbf{d}|^2} \tag{S14}$$

$$g(\mathbf{p};\mathbf{l},\mathbf{s}) = \frac{(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}})(\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}})(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}})(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{p}})}{|\mathbf{l}-\mathbf{p}|^2|\mathbf{p}-\mathbf{s}|^2}\upsilon(\mathbf{p};\mathbf{l},\mathbf{s}) \tag{S15}$$

We present the meaning and effect of *Gamma* map in the manuscript. And the attenuation factor include attenuation due to distance $|\mathbf{l}-\mathbf{p}|^{-2}|\mathbf{p}-\mathbf{s}|^{-2}$, due to shading effects $(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{l}} \cdot \mathbf{n}_{\mathbf{p}})(\boldsymbol{\omega}_{\mathbf{l}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{l}})(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}})(\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}} \cdot \mathbf{n}_{\mathbf{p}})$, and the visibility due to occlusions $\upsilon(\mathbf{p};\mathbf{l},\mathbf{s})$.

The maximal depth of the bounding box

Since the distance between hidden objects and the relay surface plays a significant role in the attenuation of photons, we consider it to be the maximal depth $z_{\max}$ of the bounding box. Here we assume that the hidden object is a perfectly diffuse white sphere, i.e., its BRDF $f(\mathbf{p}) = \frac{1}{\pi}$, and that the signal is larger than a bias b, as:

$$\Gamma(\mathbf{s})\int_P \frac{(\boldsymbol{\omega}_{\mathbf{s}\to\mathbf{p}} \cdot \mathbf{n}_{\mathbf{s}})^2}{\pi|\mathbf{p}-\mathbf{s}|^4}\mathrm{d}A_{\mathbf{p}} = \Gamma(\mathbf{s})\int_0^{2\pi}\int_0^{\pi/2}\frac{\cos^2\theta}{\pi z_{\max}^4}\mathrm{d}\theta\sin\theta\mathrm{d}\phi \geq b \tag{S16}$$

when $\Gamma(\mathbf{s})$ is minimal. The bias b includes the dark counts of the capture system and the ambient photon flux. $z_{\max}$ is then computed:

$$z_{\max} = \min\left(\frac{2}{3b}\Gamma(\mathbf{s})\right)^{1/4} \tag{S17}$$

Temporal jitter details

In the main paper, we formulate the transient recorded with a SPAD coupled with a TCSPC as

$$\tau^{\text{SPAD}} = Pois(\tau * j + b) \tag{S18}$$

where j represents the temporal jitter of the entire system. The temporal jitter typically yields a curve having two parts: a Gaussian peak and an exponential tail,

$$j(t;\mu,\sigma,\kappa_0,\kappa_1,\gamma) = Gaus(t;\mu,\sigma) + \gamma Exp(t;\mu,\kappa_0,\kappa_1) \tag{S19}$$

Here the Gaussian peak $Gaus$ and exponential tail Exp expressions are

$$Gaus(t;\mu,\sigma) = \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \tag{S20}$$

$$Exp(t;\mu,\kappa_0,\kappa_1) = \frac{1}{\sqrt{t}}\exp\left(-\frac{(t-\mu)^2}{\kappa_0 t}\right)\left(1+\frac{t-\mu}{\kappa_1 t}\right) \tag{S21}$$

where μ, σ, κ_0, and κ_1 are the coefficients of the temporal jitter, and γ is the weight of the exponential term.

II. Baseline optimization method

Many algorithms have been proposed to reconstruct a hidden scene from its transients. Volume-based algorithms, e.g., LCT [1], FK [2], and PF [3], assume that the hidden scene is represented as a volume of 3D voxels, and solve the inverse problem using FFT for reconstruction. These algorithms require a regular form of transients as input. Alternatively, optimization-based and learning-based algorithms (e.g., NeTF [4]) support arbitrary forms of input transients, while the latter require tens of hours for training.

We implement an optimization algorithm based on the confocal imaging model. In common with most existing methods, the optimization algorithm assumes that the hidden object is perfectly diffuse, and that there are no occlusions between the object and the relay surface, and no inter-reflection between the points on the hidden object. This implies that in Equation (4), the attenuation $g(\mathbf{p};\mathbf{s})$ is considered to be $\frac{1}{|\mathbf{p}-\mathbf{s}|^4}$, and the BRDF $f(\mathbf{p};\boldsymbol{\omega}_{\mathbf{p}\to\mathbf{s}})$ becomes view-independent $\mathbf{f}$. Equation (4)is thus simplified as a linear formulation:

$$\boldsymbol{\tau} = \Psi\mathbf{f} \tag{S22}$$

where $\boldsymbol{\tau}$ is the discretized measurements and Ψ a linearized measurement matrix with $g(\mathbf{p};\mathbf{s})$. The albedo of the hidden object, $\mathbf{f}$, can be optimized by minimizing the differences between the predicted $\Psi\mathbf{f}$ and the measured $\boldsymbol{\tau}$. We adopt the Poisson likelihood function to evaluate the similarity of the following observation function with total variation (TV) as a regularization term [5], [6]:

$$\mathcal{L}_{\text{NLOS}}(\mathbf{f}) = -\ln(\prod_i \frac{(\mathbf{e}_i^\intercal\Psi\mathbf{f})^{\tau_i}}{\tau_i!}\exp(-\mathbf{e}_i^\intercal\Psi\mathbf{f})) + \lambda||\mathbf{f}||_{\text{TV}} \tag{S23}$$

where $\mathbf{e}_i$ is the ith standard unit vector and τ_i the ith element of vector $\boldsymbol{\tau}$, while λ is the weight of the TV. We adopt a gradient descent method to implement the optimization algorithm.

III. Simulation algorithm

We develop an algorithm to synthesize transients based on the confocal imaging model in Equation (4). The illumination spot collocates with the detection point $\mathbf{s}$ on the relay surface, and the light scatters in a spherical wavefront toward a hidden object. We assume that at the detection point, an intensity map $L(u,v;\mathbf{s})$, and a depth map $D(u,v;\mathbf{s})$ are captured, where u,v represents a corresponding pixel of the pair of maps. We then extract the depth and the intensity at each pixel, and compute the transient $\tau(t;\mathbf{s})$ at the time instant t. This process is repeated for all detection points, as outlined in Algorithm 1.
We render transients $\tau(t;\mathbf{s})$ of a bunny (as shown in Fig. S1) of $0.6\,\text{m} \times 0.6\,\text{m}$, with $4\,\text{ps}$ temporal resolution and 64 $\times$ 64 spatial resolution. By considering the temporal jitter, the bias, and the Poisson distribution in Equation (6), the transients with noise are also synthesized.

Algorithm 1 Transient simulation

Input: depth maps $D(u,v;\mathbf{s})$, intensity maps $L(u,v;\mathbf{s})$, speed of light c
Output: transients $\tau(t;\mathbf{s})$

```
initialize τ(t;s)
for all sensing points s do
    for all image pixels (u,v) do
        t = round(2 * D(u,v)/c)
        τ(t;s) = τ(t;s) + L(u,v)
    end for
end for
```

Fig. S1: Simulated bunny

IV. More results

In the main paper, we use NeTF to reconstruct results with three different sampling patterns: a regular grid, concentric circles, and one non-uniform pattern. Here we use a baseline optimization method to compare them again.

The reconstruction results are a little blurrier than NeTF, but we can figure out that the reconstruction results in the concentric-circle pattern and in the non-uniform pattern shows strong intensity at the center (or the torso).

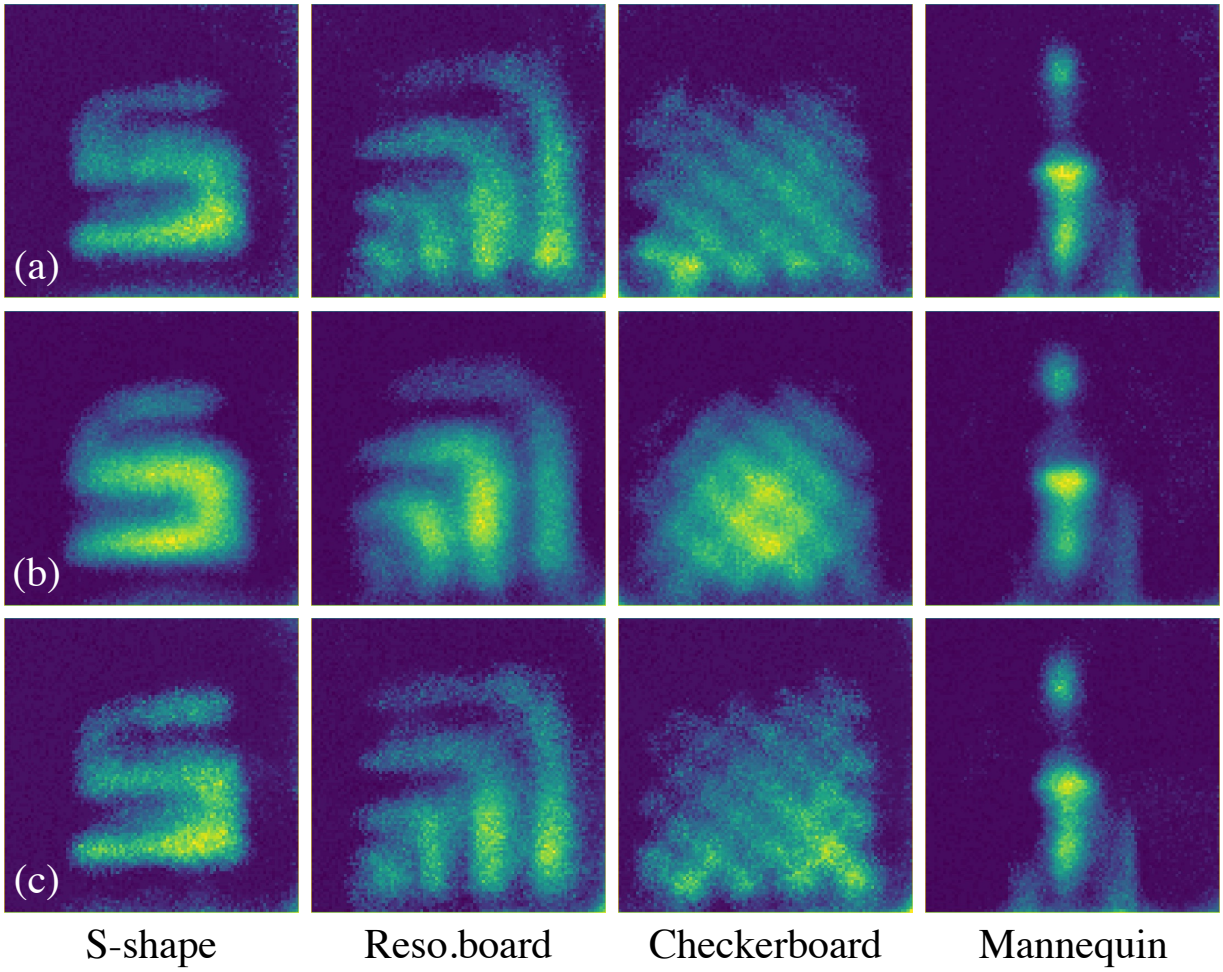

Fig. S2: Reconstruction evaluation using transients measured from our calibrated NLOS imaging system with baseline optimization method. From left to right: S-shape, Reso.board, Checkerboard, and Mannequin. From top to bottom: Sampling patters are a regular grid (a), concentric circles (b), and a non-uniform pattern (c).

Here we also provide additional results of our ablation study, which shows show how calibration errors of the relay wall may affect the reconstruction result of hidden objects. As shown in Fig. S3, we can find that no matter whether the data has noise or not, the calibration error will lead to distortion of the reconstruction results, and with the increase of the calibration error angle of the relay wall, the reconstruction results will become worse. In addition, different deviation directions will bring different distortion effects.

References

[1] M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," *Nature*, vol. 555, no. 7696, pp. 338–341, 2018.

[2] D. B. Lindell, G. Wetzstein, and M. O'Toole, "Wave-based non-line-of-sight imaging using fast fk migration," *ACM Transactions on Graphics*, vol. 38, no. 4, pp. 1–13, 2019.

[3] X. Liu, I. Guillén, M. La Manna, J. H. Nam, S. A. Reza, T. H. Le, A. Jarabo, D. Gutierrez, and A. Velten, "Non-line-of-sight imaging using phasor-field virtual wave optics," *Nature*, vol. 572, no. 7771, pp. 620–623, 2019.

[4] S. Shen, Z. Wang, P. Liu, Z. Pan, R. Li, T. Gao, S. Li, and J. Yu, "Non-line-of-sight imaging via neural transient fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[5] J.-T. Ye, X. Huang, Z.-P. Li, and F. Xu, "Compressed sensing for active non-line-of-sight imaging," *Optics Express*, vol. 29, no. 2, pp. 1749–1763, 2021.

[6] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is spiral-tap: Sparse poisson intensity reconstruction algorithms—theory and practice," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1084–1096, 2011.

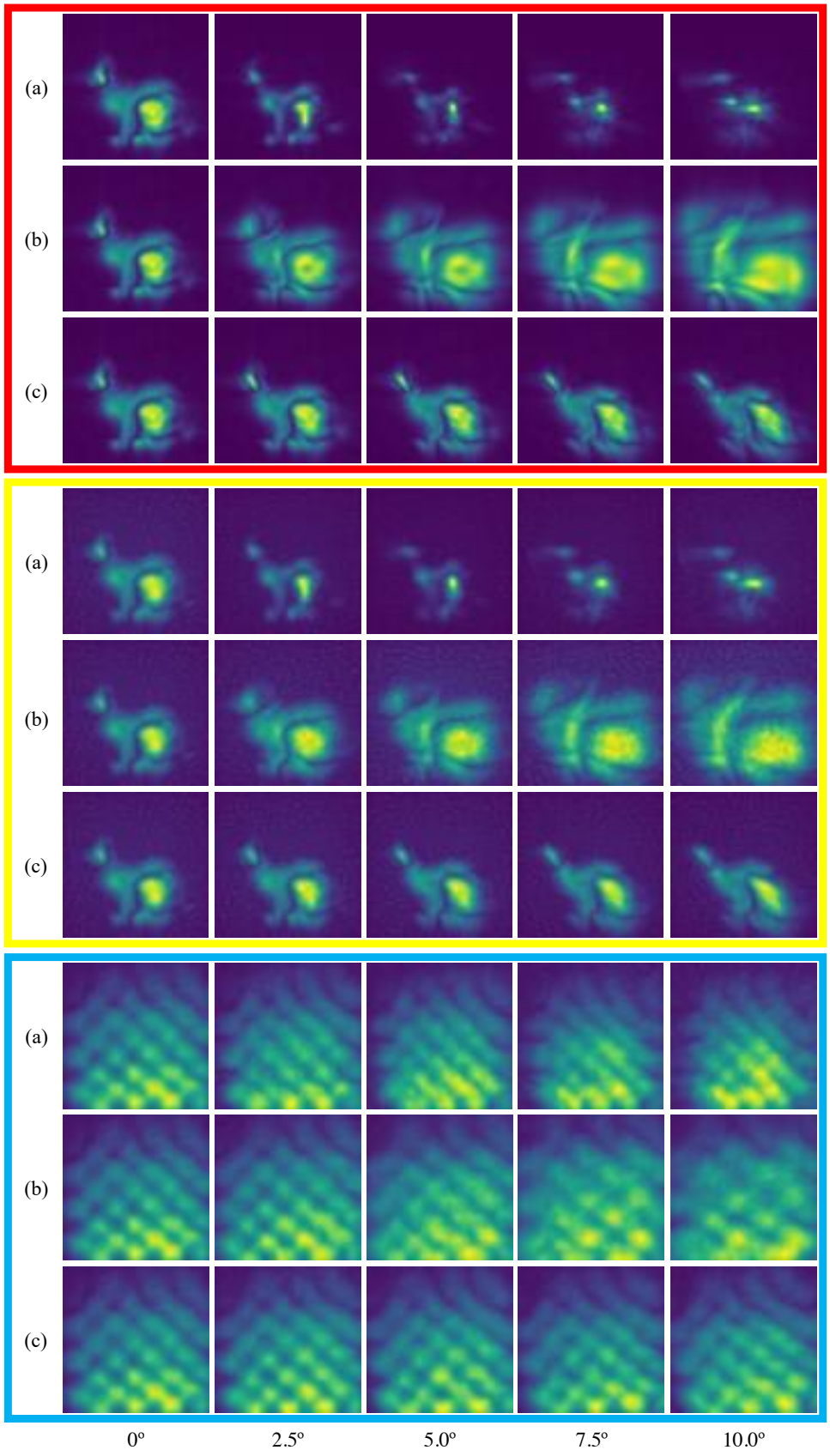

Fig. S3: Additional ablation studies for calibration errors of the relay wall. The NLOS reconstruction results of a simulated Bunny (in red frame), NLOS reconstruction results of a simulated Bunny with noise (in yellow frame), and NLOS reconstruction results of measured Checkerboard (in blue frame). For each case, the mis-calibrated wall is rotating clockwise (a), or conterclockwise (b), or forward (c) in comparison with the accurate position. From left to right: Results with the relay wall at $0.0°$ (accurately calibrated), and at $2.5°$ to $10.0°$ at an interval of $2.5°$, where the relay wall is mis-calibrated.