Supplementary materials

Paper title: Analysis of Feature Visibility in Non-Line-of-Sight Measurements

Local MTF pattern with/without intensity drop off

As mentioned in **main text section 4.1**, the generalized NLOS forward model without the intensity drop off term $\Upsilon(d_i, d_d)$ can be described as a Radon integral. All previous works approximate the intensity drop off by the squared traveling distance:

$$g(\boldsymbol{p}_i, \boldsymbol{p}_d, t) = \int_{\mathcal{D}} \Upsilon(d_i, d_d) \cdot \delta(d_i + d_d - t \cdot c) \cdot f(\boldsymbol{p})\, d\boldsymbol{p} \tag{1}$$

$$\approx \int_{\mathcal{D}} \frac{1}{d_i^2} \cdot \frac{1}{d_d^2} \cdot \delta(d_i + d_d - t \cdot c) \cdot f(\boldsymbol{p})\, d\boldsymbol{p}\,. \tag{2}$$

In the main text, we provide the MTF pattern without the intensity drop off. Here we provide the MTF pattern at the same five positions but considering the intensity drop off effect. As it is shown in Fig. 1, this intensity drop off is negligible for the spatial frequency analysis.

An intuition behind Equation (1) is that the modeling process starting from the unknown reflectance $f(\boldsymbol{p})$ instead of the field amplitude should be observed. Furthermore, the kernel $\delta(d_i + d_d - t \cdot c)$ governs the integral geometry, whereas the intensity drop-off $\Upsilon(d_i, d_d)$ reduces the value by a specific weight. This effect is similar to the amplitude error made in the diffraction integral approximation.

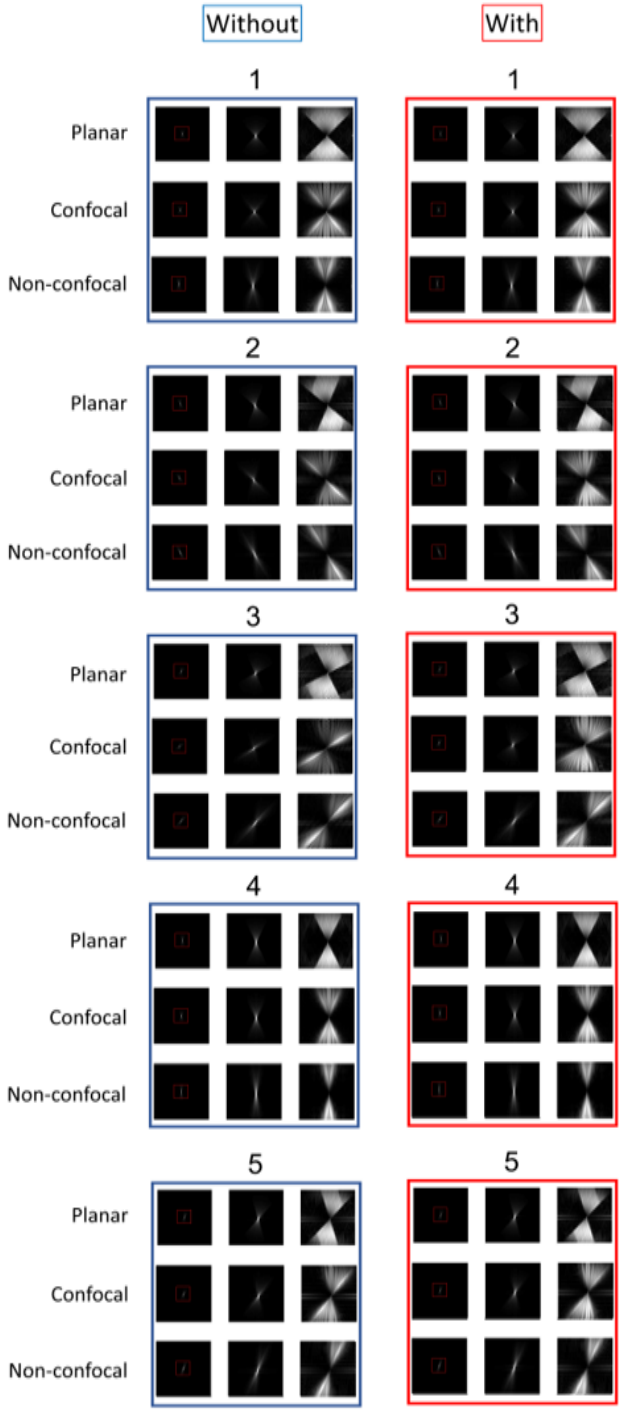

Figure 1: **MTF pattern without/with intensity drop off term:** This shows the MTF pattern differences without and with distance drop-off in the forward model. The first column represents no distance drop off term. The second one incorporates the distance drop-off. Each box contains the planar, confocal and non-confocal MTF patterns. Number 1-5 represent the different local positions which are the same as in **main text Figure 3**.

Non-confocal illumination extended MTF cone

There are two main types of the NLOS forward sampling method: confocal and non-confocal setup. As it can be seen in **main text Figure 2**, confocal illumination and detection positions are co-located whereas for the non-confocal setup, both positions are different. To allow for easier comparison, we consider the same detection grid sampling at a finite relay wall. Then for the non-confocal measurement, we analyze the different Fourier cone behaviors along with varying illumination source position. This is shown in Figure 2. We consider a fixed local window (symmetry at the center, $0.5\,\mathrm{m}$ in depth) and three illumination positions within the limited sampling area. Different illumination positions also provide rotation in the Fourier cone for the non-confocal measurement. Even though the Fourier cone varies, as long as the illumination positions only cover the limited relay area, the cone stays within the boundary estimated from the planar model.

We show the generalized extended MTF cone by changing the illumination position at the limited sampling area. Also, we show the consistency with our proposed model.

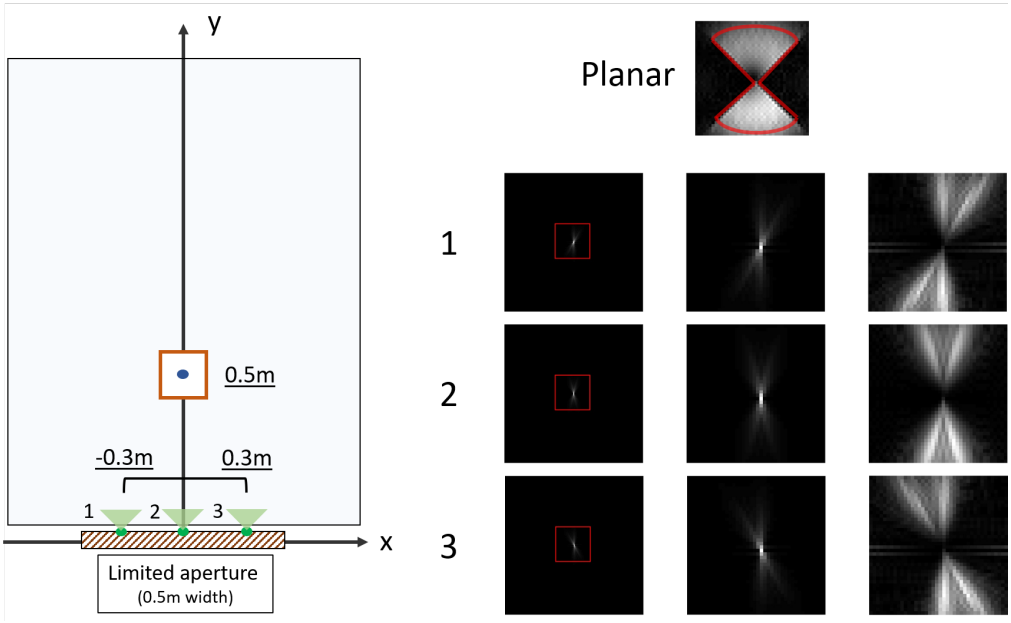

Figure 2: **Non-confocal illumination variation:** The non-confocal acquisition setup has two degrees of freedom for focused detection and illumination positions. In the figure, we fixed the array of detection positions at the limited aperture (width $0.5\,\mathrm{m}$). We pick three illumination positions (1-3) within the limited aperture. We consider the same local volume position, and each local MTF pattern is shown on the right. Even though each non-confocal Fourier cone is narrower than the planar cone on the top, by using the multiple illumination source positions, the same Fourier cone coverage can be achieved.

An intuition for why the non-confocal measurement cone varies depending on the illumination position is provided in the following. We know that the non-confocal forward model can be described as the integration along a thin ellipsoid surface in space having two foci (one is the illumination, the other the detection position). For a fixed position window, the ellipsoid curvature is changed by moving one of the foci. In this case, the tangent line with its normal vector changes also. Thus each local cone also rotates when moving the illumination source along the hyperplane.

Here we list a summary considering two types of NLOS measurement in both mathematical and practical sense. Mathematically speaking, the non-confocal NLOS measurement is a more generalized version of confocal measurement. However, it can be seen in **main text Figure 3** that using the same limited size relay wall, the confocal setup covers a slightly wider area in the Fourier domain compared to one single illumination for the non-confocal measurement. As we show in this section, this downside of the non-confocal measurement can be reduced in practice by simply adding some extra illumination source positions.

NLOS Radon integral modeling error

Some modeling error still constrains the model used for the NLOS problem. Notice that this *modeling error* is the error in the mathematical sense. Researchers avoid conflicting this error by adjusting the scene targets to fulfill the modeling gap. Here we provide a specific description for the modeling error may cause the *inverse crime* (Inverse crime means the forward model is not accurate enough causing the corresponding inverse solutions to fail in the real-world measurements).

Basic notation: We use the symbol f to represent the unknown function we would like to recover, A for a linear operator as the forward model, g for one measurement, G for a finite set of measurements, $\boldsymbol{p}$ represent the space variables (x, y, z). $\boldsymbol{p}_i$ and $\boldsymbol{p}_d$ stands for illumination point and detector point on the visible relay wall.

1. Lambertian approximation This is the approximation for the unknown function. By assuming the hidden scene reflects light isotropically, the unknown function f can be described as $f(\boldsymbol{p}) = f(x, y, z)$ meaning that the reflectance is a constant value in space. However, in reality, surface reflectance should at least based on incident and observation which leads to a higher dimension. This is well known in the computer graphic society.

2. Direct bounce modeling approximation With Approximation 1, one can model direct bounce in the integral equation as described in Eq. (3). However, this approximation ignores the indirect bounces (multi-bounces) signal within the invisible area. To make a distinction, we use B to represent the direct bounce model:

$$g(\boldsymbol{p}_i, \boldsymbol{p}_d, t) \approx Bf(\boldsymbol{p}) = \int \Upsilon(|\boldsymbol{p} - \boldsymbol{p}_i|, |\boldsymbol{p} - \boldsymbol{p}_d|) \cdot f(\boldsymbol{p}) \cdot \delta(|\boldsymbol{p} - \boldsymbol{p}_i| + |\boldsymbol{p} - \boldsymbol{p}_d| - t \cdot c)\, d\boldsymbol{p}\,. \tag{3}$$

In this equation, $g(\boldsymbol{p}_i, \boldsymbol{p}_d, t)$ stands for a single time response captured from given illumination and detector positions. $\Upsilon(|\boldsymbol{p} - \boldsymbol{p}_i|, |\boldsymbol{p} - \boldsymbol{p}_d|)$ stands for the intensity drop-off term. The kernel function $\delta(|\boldsymbol{p} - \boldsymbol{p}_i| + |\boldsymbol{p} - \boldsymbol{p}_d| - t \cdot c)$ mainly describes the geometry of the integral. Based on whether illumination position $\boldsymbol{p}_i$ and detector position $\boldsymbol{p}_d$ are co-located or not, the geometry looks either spherically or ellipsoidally.

However, the actual forward model A is not equal to B only if the multibounce light is negligible. In fact, we can model the time response measurement $g(\boldsymbol{p})$ into two parts $g^{'}(\boldsymbol{p})_{\text{direct}}$ and $g^{'}(f(\boldsymbol{p}))_{\text{multi}}$.

The actual physically accurate model results to

$$g(\boldsymbol{p}_i, \boldsymbol{p}_d, t) = Af(\boldsymbol{p}) = g^{'}(\boldsymbol{p}_i, \boldsymbol{p}_d, t)_{direct} + g^{'}(f(\boldsymbol{p}))_{multi} = \underbrace{Bf(\boldsymbol{p})}_{\textbf{Direct}} + \underbrace{g^{'}(f(\boldsymbol{p}))}_{\textbf{MPI}} \tag{4}$$

It is obvious to see that the multibounce light depends on the unknown function f. Solving the entire inverse problem accounting for the multibounce signal is a bi-convex problem since A and f are both unknown. Most of the methods presented so far are trying to approximate $B^{-1}(g(\boldsymbol{p}_i, p_d, t))$ by numerical optimization or a linear combination of backprojection method. Both of them fail mathematically with the term $g^{'}(f(p))$ in Eq. (4) causing some multibounce signal artifacts in the reconstruction volume.

3. Delta kernel model error Even though the model used before gives a well-approximated result (Eq. (3)), the direct bounce modeling still has an internal error in reality.

The kernel function $\delta(|\boldsymbol{p}-\boldsymbol{p}_i|+|\boldsymbol{p}-\boldsymbol{p}_d|-t\cdot c)$ in the integral (3) refers to infinite temporal resolution. Since inverse methods trying to solve the model refers to the infinite temporal resolution, based on our observation, this inaccurate modeling may harm the inverse solution more and more when the temporal resolution from the measurement becomes lower. This means the *inverse crime* happens.

In reality, the actual light pulse coming from the physical setup may have a specific temporal degradation or shape distortion resulting from the complex relay wall surface. Previous methods try to ignore the first returning pulse reflected from the visible surface. We argue that it would be useful to use it to correct this delta modeling error.

The good thing for the approximation is that it makes the inverse process tractable by only adding extra physical constraints, such as looking at a simple patch or reducing the scene complexity. On the other hand, modeling a proper NLOS formation may result in the inverse becoming mathematically intractable. We want to address those difficulties and gaps for the readers for future research.

Three-dimensional Fourier slice sampling

In our **main text Section 4.1**, we mention the planar model for the NLOS forward measurement function. Here we provide a detailed model to represent the 3D Fourier slice sampling (**main text Section 5.1 refers to 2D case**).

Basic notation: $I(x, y, z)$ represents the unknown volume function in the cartesian coordinates x, y, z. $F(u, v, h)$ refers to its Fourier transform. $t(x, y, z)$ refers to the measurement (integral) data.

Cartesian to spherical coordinates: If we regard the center of the 3D unknown volume as the origin, ρ can be regarded as the distance of each plane integral and θ and ϕ directly refer to the direction of each projection integral.

For the measurement data:

$$g(\rho, \theta, \phi) = t(x, y, z)\,. \tag{5}$$

For the unknown volume function, $g(\rho, \theta, \phi)$ refers to the measurement data in spherical coordinates. To agree with the coordinate, we translate the coordinate for $F(u, v, h)$ to $G(\omega, \theta, \phi)$ as well.

3D Plane Radon integral: Based on the 3D Radon transform, the transformation between the original data $I(x, y, z)$ and measurement data in spherical coordinates $g(\rho, \theta, \phi)$ can be interpreted as follows:

$$g(\rho, \theta, \phi) = \iiint_{-\infty}^{\infty} I(x, y, z)\delta(\sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z - \rho)\,dx\,dy\,dz\,. \tag{6}$$

Then, we calculate the 1D Fourier transform of the measurement $g(\rho, \theta, \phi)$ along ϕ as follows:

$$G(\omega, \theta, \phi) = \int_{-\infty}^{\infty} g(\rho, \theta, \phi)e^{-j2\pi\omega\rho}\,d\rho\,. \tag{7}$$

Replacing $g(\rho, \theta, \phi)$ with equation in 6 yields

$$G(\omega, \theta, \phi) = \int_{-\infty}^{\infty} [\iiint_{-\infty}^{\infty} I(x, y, z)\delta(\sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z - \rho)\,dx\,dy\,dz]e^{-j2\pi\omega\rho}\,d\rho\,. \tag{8}$$

Then the to make delta kernel $\delta(\sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z - \rho)$ inside the equation become one, we can get rid of the ρ variable by using $\rho = \sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z$. Then we simplify the equation:

$$G(\omega, \theta, \phi) = \iiint_{-\infty}^{\infty} I(x, y, z)e^{-j2\pi(\omega\sin\theta\cos\phi x + \omega\sin\theta\sin\phi y + \omega\cos\theta z)}\,dx\,dy\,dz\,, \tag{9}$$

whereas the Fourier transform of the unknown volume function $f(x, y, z)$ in Cartesian coordinate results as follows:

$$F(u, v, h) = \iiint_{-\infty}^{\infty} I(x, y, z)e^{-j2\pi(ux+vy+hz)}\,dx\,dy\,dz\,. \tag{10}$$

It is easy to see comparing Eqs. (9) and (10) that the 3D Fourier slice can be achieved as follows:

$$G(\omega, \theta, \phi) = F(\omega\sin\theta\cos\phi, \omega\sin\theta\sin\phi, \omega\cos\theta)\,. \tag{11}$$

From Eq. (11) we can see that, with a given projection direction, the Fourier transform of the spatial-temporal measurement $G(\omega, \theta, \phi)$ is a slice of the Fourier transform of the original data $F(u, v, h)$.

Planar model filtered backprojection method

The previous section mentioned the 3D Fourier slice form of the 3D planar Radon integral. Here we list the filter backprojection method for this type of integration. We keep the same notation as in the previous section.

To recover $I(x, y, z)$ from its Fourier spectrum $F(u, v, h)$, we can simply apply the inverse Fourier transform as follows:

$$I(x, y, z) = \iiint_{-\infty}^{\infty} F(u, v, h)e^{j2\pi(ux+vy+hz)}\,du\,dv\,dh\,. \tag{12}$$

Rewriting Equation (12) by replacing the Cartesian coordinate with spherical coordinates yields

$$I(x, y, z) = \iiint_{-\infty}^{\infty} F(\omega\sin\theta\cos\phi, \omega\sin\theta\sin\phi, \omega\cos\theta)e^{j2\pi(\omega\sin\theta\cos\phi x + \omega\sin\theta\sin\phi y + \omega\cos\theta z)}\omega^2\sin\theta\,d\omega\,d\theta\,d\phi\,. \tag{13}$$

As we know from Equation (11), we could replace the term $F(\omega \sin\theta \cos\phi, \omega \sin\theta \sin\phi, \omega \cos\theta)$ in Equation (13) as follows:

$$I(x,y,z) = \iiint_{-\infty}^{\infty} G(\omega,\theta,\phi) e^{j2\pi(\omega \sin\theta \cos\phi x + \omega \sin\theta \sin\phi y + \omega \cos\theta z)} \omega^2 \sin\theta \, d\omega \, d\theta \, d\phi \,. \tag{14}$$

We reuse the defined $\rho = \sin\theta \cos\phi x + \sin\theta \sin\phi y + \cos\theta z$ again to simplify the equation and rewrite Equation (13) as:

$$\begin{aligned} I(x,y,z) &= \iiint_{-\infty}^{\infty} G(\omega,\theta,\phi) e^{j2\pi\omega\rho} \omega^2 \sin\theta \, d\omega \, d\theta \, d\phi \\ &= \int_0^{2\pi} \int_0^{\pi} [\int_{-\infty}^{\infty} \omega^2 \cdot G(\omega,\theta,\phi) e^{j2\pi\omega\rho} \, d\rho] \sin\theta \, d\theta \, d\phi \\ &= \int_0^{2\pi} \int_0^{\pi} [kernel * g(\rho,\theta,\phi)] \sin\theta \, d\theta \, d\phi \end{aligned} \tag{15}$$

Notice that in this equation, the final line $*$ represents the convolution operator which its frequency response is ω^2. Based on the definition, it is Laplacian filter.