On F-K Migration and Non-Line-of-Sight Imaging

Dorian Yao Chan
Carnegie Mellon University
dychan@andrew.cmu.edu

Abstract

The introduction of F-K migration, a Fourier-domain tool derived from scalar theory in seismology, into non-line-of-sight (NLOS) imaging has provided an intriguing alternative to existing NLOS approaches. In this paper, we investigate the relationship between F-K migration and traditional NLOS approaches in a few different ways. First, we observe that the Light Cone Transform (LCT) approach can be essentially formulated as prefiltered backprojection. We connect that finding to Kirchhoff migration, another seismology-inspired technique derived from scalar theory, that can be expressed in a similar fashion and is theoretically equivalent to F-K migration but yields different results in practice. From Kirchhoff migration, we derive an operator that properly models the forward propagation between the light source and the hidden scene that very closely resembles the LCT. We then demonstrate that a geometric model for NLOS imaging can be derived in a very similar fashion as F-K migration from scalar theory. Finally, we show a derivation of F-K migration from purely geometric principles, inspired by the radar literature. In short, at the end of the day, F-K migration is not so different from other NLOS techniques.

1. Introduction

Non-line-of-sight (NLOS) imaging has a number of important applications across a number of fields, such as in defense, medical imaging, autonomous driving, and search and rescue. In each of these cases, a fast and robust NLOS imaging setup could potentially revolutionize the way we approach these domains and save many lives.

However, NLOS imaging is fundamentally difficult. In general, large transient measurements need to be captured by a single photon avalanche diode (SPAD) or other ultrafast sensor in order to capture enough information about the hidden scene. As a result, a naive approach to processing this data can take many hours to compute. In addition, improperly dealing with reflectance in NLOS imaging can result in significant artifacting in output reconstructions.

In order to meet these challenges, a number of techniques have been proposed over the years that attempt to target each of these different aspects. Notably, work based on applying wave theory to NLOS imaging [18, 20, 19, 7] has shown significant success in terms of both computational efficiency as well as robustness to BRDF. In particular, a method from seismology called F-K migration [18] has shown state-of-the-art results across a wide variety of scenes and materials. Furthermore, this migration technique can be implemented extremely efficiently through just a resampling and filtering in the Fourier domain.

However, it is somewhat unclear what the exact relationship between F-K migration and other NLOS techniques are, as F-K does not use an explicit imaging model like most existing approaches. In this paper, we investigate this space in a two different ways. First, we show that other state-of-the-art approaches, like the LCT [23] and Phasor Fields [20], can be effectively viewed as "prefiltered" backprojections. We connect that to another technique known as Kirchhoff Migration, which is derived from scalar theory and therefore should theoretically provide equivalent results to F-K migration. We derive a second-order Kirchhoff migration that properly models the propagation between the light source and hidden scene, which yields an equation very similar to the LCT. We test out both of these new approaches on real data, and empirically show that they yield similar results to F-K migration and the LCT. Second, we connect the volumetric albedo model to F-K migration in two ways. We derive the volumetric albedo model using a slightly modified version of the derivation behind F-K migration. Next, we derive a variant of F-K migration from a volumetric albedo model. From both of these approaches, we find that the main difference between a volumetric albedo approach and F-K migration is just an extra high pass filter.

2. Background

NLOS imaging has received significant attention in recent years across multiple modalities and methodologies, including transient imaging [1, 3, 4, 7, 8, 9, 10, 12, 13, 15, 18, 20, 22, 23, 24, 27, 28, 29, 30], conventional cameras

[5, 14, 2, 25, 26], Wi-Fi[16], and sound [17]. In this paper, we focus on recent work in transient ultrafast imaging, which has demonstrated high quality measurements and reconstructions at fast speeds. In these techniques, a pulsed laser is pointed at a certain location on the visible wall, and a transient sensor pointed at another location records the time it takes for light to reflect back from the hidden scene. The measured signal is known as a transient measurement.

Because transient measurements tend to be quite large, many recent techniques have attempted to tackle the computational tractability of inverting the NLOS transport matrix to find a volumetric representation of the hidden scene. Notably, [23] proposed taking the confocal subset of a full transient measurement, where the pulsed laser and transient sensor are pointed at the same location. In this confocal case, the authors showed that a simple model for NLOS transport could be efficiently inverted using a deconvolution operation following an appropriate resampling - we elaborate on this approach, colloquially known as the LCT, in the next section. [30] extended the LCT to perform a variant of photometric stereo on hidden scenes. [12] proposed taking a circular subset of a confocal measurement, reducing the NLOS problem to a computed tomography problem and resulting in even faster reconstructions as well as measurement times. [1] demonstrated that a modified non-confocal reconstruction problem could be approximated using a convolutional operator.

Another class of techniques attempt to directly find a surface representation of the NLOS scene. [29] observed that BRDF-independent Fermat paths result in quantifiable discontinuities in recovered measurements, and that these paths can be used to recover fine-grain normals and positions of a NLOS surface. [27] used a differentiable rendering approach to recover surface models and BRDFs of hidden scenes. [11] represent the NLOS surface using a series of Gaussian blobs, and develop a sophisticated optimization scheme to minimize reconstruction error.

More related to our work, a number of contemporary papers have also attempted to apply wave theories of light to the NLOS problem, that demonstrate seemingly BRDF-robust reconstructions. [20] proposed using a Rayleigh-Sommerfeld-like propagation integral to simulate common imaging systems on NLOS scenes, like a thin lens, a confocal time-gated system, or a transient camera. [19] developed a faster version for non-confocal scenes. [7] showed that these "phasor field" methods are empirically robust to BRDF.

[18] observed that researchers in seismology attempt to solve similar problems to NLOS, and applied a technique known as F-K migration to the NLOS imaging problem. Derived from scalar theory, F-K models the scalar field generated by a NLOS scene using the following equation:

$$\Psi(x,y,z,t) = \iiint \Phi(k_x,k_y,k_z) e^{2\pi i(k_x x + k_y y + k_z z - ft)} \,\mathrm{d}k_x \,\mathrm{d}k_y \,\mathrm{d}k_z \tag{1}$$

where $f = c\sqrt{k_x^2 + k_y^2 + k_z^2}$ and c is the speed of light. When $t = 0$, corresponding to the geometry of the hidden scene, the above equation turns into a perfect Fourier Transform. Alternatively, by applying a change of variables from k_z to f, the above integral can be reformulated as:

$$\Psi(x,y,z,t) = \iiint \bar{\Phi}(k_x,k_y,f) e^{2\pi i(k_x x + k_y y + k_z z - ft)} \,\mathrm{d}k_x \,\mathrm{d}k_y \,\mathrm{d}f \tag{2}$$

where

$$\Phi(k_x,k_y,k_z) = \frac{c\|k_z\|}{\sqrt{k_x^2 + k_y^2 + k_z^2}} \bar{\Phi}\left(k_x, k_y, c\sqrt{k_x^2 + k_y^2 + k_z^2}\right) \tag{3}$$

Here, when $z = 0$, corresponding to the visible wall, Equation 2 again turns into a perfect Fourier Transform. Equations 1 and 2 suggest a possible reconstruction operator. We can take the Fourier Transform of the measured signal at the visible wall to compute $\bar{\Phi}(k_x, k_y, f)$. We can then apply the interpolation and filtering given by Equation 3 to recover $\Phi(k_x, k_y, k_z)$. Finally, we take the inverse Fourier Transform to get an approximation of the scene geometry. In practice, F-K migration shows high quality results that can be computed very efficiently. However, it is unclear the exact relationship between the interpolation given in Equation 3 and the other aforementioned techniques. In this paper, we attempt to investigate the differences.

3. The LCT is a Prefiltered Backprojection

Most past techniques have relied on some variant of the volumetric albedo model, given in the confocal case by the following imaging equation:

$$\tau(x',y',t) = \iiint_\Omega \mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z \frac{\rho(x,y,z)}{r^4} \cdot \delta(2\sqrt{(x-x')^2 + (y-y')^2 + z^2} - tc) \tag{4}$$

where ρ is a three-dimensional albedo volume within Ω, $\tau(x', y', t)$ denotes the transient measurement captured at wall location x', y' at time t, and c is the speed of light. Many older techniques heuristically invert this equation using filtered backprojection [28], which involves computing the adjoint of Equation 4 followed by a Laplacian filter. [23] noted that this equation could be rewritten as a

three-dimensional convolution through an appropriate resampling, called the light cone transform (LCT):

$$v^{3/2}\tau(x', y', 2\sqrt{v}/c) = \iiint_{\Omega} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \frac{\rho(x, y, \sqrt{u})}{2\sqrt{u}} \cdot \delta((x - x')^2 + (y - y')^2 + u - v) \tag{5}$$

with $z = \sqrt{u}$ and $v = (tc/2)^2$. In this form, a least-squares inverse can be efficiently computed in the Fourier domain using a Wiener filter with a selected signal-to-noise ratio.

However, what does this inverse kernel actually look like in the spatial domain? We note that the forward kernel $h = \delta((x-x')^2+(y-y')^2+u-v)$ can be written in the Fourier domain as:

$$\mathcal{F}\{h\}(k_x, k_y, k_v) = H(k_x, k_y, k_v) = \frac{1}{i2k_v} e^{i\pi(\frac{k_x^2+k_y^2}{k_v})} \tag{6}$$

The backprojection kernel will be given by the forward kernel's conjugate in the Fourier domain:

$$H^*(k_x, k_y, k_v) = -\frac{1}{i2k_v} e^{-i\pi\left(\frac{k_x^2+k_y^2}{k_v}\right)} \tag{7}$$

The inverse kernel is given by the reciprocal of the forward kernel in the Fourier Domain:

$$\begin{aligned} K(k_x, k_y, k_v) &= 1/H(k_x, k_y, k_v) = i2k_v e^{-i\pi\left(\frac{k_x^2+k_y^2}{k_v}\right)} \\ &= -(i2k_v)^2 \frac{-1}{i2k_v} e^{-i\pi\left(\frac{k_x^2+k_y^2}{k_v}\right)} \\ &= -(i2k_v)^2 H^*(k_x, k_y, k_v) \end{aligned} \tag{8}$$

Note that during the reconstruction process, $K(k_x, k_y, k_v)$ will be multiplied by the Fourier Transform of the resampled transient $\hat{\tau}$. Then, we can write the reconstructed geometry $\hat{\rho}$ as the following:

$$\begin{aligned} \mathcal{F}\{\hat{\rho}\}(k_x, k_y, k_u) &= \mathcal{F}\{\hat{\tau}\}(k_x, k_y, k_u) K(k_x, k_y, k_u) \\ &= -(i2k_u)^2 \mathcal{F}\{\hat{\tau}\}(k_x, k_y, k_u) H^*(k_x, k_y, k_u) \end{aligned} \tag{9}$$

In the spatial domain, this expression corresponds to:

$$\hat{\rho}(x, y, u) = \left(\frac{\partial^2 \hat{\tau}}{\partial v^2} * h^*\right)(x, y, u) \tag{10}$$

where $*$ is the convolution operator, h^* is the spatial kernel of the backprojection, and $\hat{\rho}$ is the resampled output geometry. Therefore, the LCT reconstruction is simply just the backprojection of the second time/depth derivative of the resampled transient. We can rewrite this in terms of the original transient (hopefully I got all of these constants right):

$$\begin{aligned} \frac{\partial^2 \hat{\tau}}{\partial v^2}(x, y, v) = &\frac{0.75\tau\left(x, y, \frac{2\sqrt{v}}{c}\right)}{\sqrt{v}} \\ &+ \frac{2.5}{c}\frac{\partial \tau}{\partial t}\left(x, y, \frac{2\sqrt{v}}{c}\right) + \frac{\sqrt{v}}{c^2}\frac{\partial^2 \tau}{\partial t^2}\left(x, y, \frac{2\sqrt{v}}{c}\right) \end{aligned} \tag{11}$$

The constant factors vary depending on the chosen falloff factor.

What does this derivation tell us about inverting the volumetric albedo model? In short, inverting Equation 4 requires backprojecting a linear combination of the original transient, its first temporal derivative, and its second temporal derivative. This corresponds in effect to "prefiltering" the input transient, and then backprojecting the modified transient back into the hidden scene to recover albedo.

We also note that the state-of-the-art method Phasor Fields in its base form also utilizes a similar approach [20]. The paper proposes prefiltering the input transient by a modulated Gaussian before backprojection:

$$\tau_{phasor}(x', y', t) = \left(e^{i\omega t} e^{-\frac{(t'-t_0)^2}{2\sigma^2}}\right) * \tau(x', y', t) \tag{12}$$

where ω, t_0, and σ are user-selected parameters to model an illumination pulse.

An interesting thing to note is that in practice, the second order term $\frac{\sqrt{v}}{c^2}\frac{\partial^2 \tau}{\partial t^2}\left(x, y, \frac{2\sqrt{v}}{c}\right)$ dominates the computation of Equation 11 (the author hasn't robustly checked this but it's been usually true). If we drop the other terms and ignore the constant coefficient and falloff, in the case of a visible wall with infinite extent, Equation 11 reduces to convolving the naively backprojected geometry with a 3D Laplacian filter. If falloff is considered, the contributed "Laplacian" will become lower magnitude as the spatial location of the projected measurement moves farther away, explaining why the "dips" of the inverse kernels visualized in Figure 3 of [1] become smaller along the x and y axes.

4. Kirchhoff migration

Given our observation from the previous section that state-of-the-art NLOS techniques are variants of prefiltered backprojection, one might ask whether there is some way F-K migration can be expressed in a similar fashion. To investigate this question, we introduce Kirchhoff migration, a technique originally described in seismology that can be implemented using backprojection. Notably, work in seismology [21] has argued that the uniqueness theorems of partial differential equations guarantee that Kirchhoff and F-K migration are equivalent. If this is true, then the results of F-K migration can be replicated using a more traditional

backprojection-based approach like Kirchoff migration. In this section, we show a short derivation of Kirchhoff migration, and then demonstrate this new technique applied to existing NLOS data.

4.1. Derivation

Kirchhoff migration can be derived in a very similar fashion to Kirchhoff or Rayleigh Sommerfeld diffraction theory. In this section, we show an abridged version of a typical derivation. We refer the reader to the seismology literature [21] for a more detailed description.

Consider the Helmholtz equation:

$$(\nabla^2 + k^2)\hat{\psi}(\mathbf{x}) = 0 \tag{13}$$

where $k^2 = 4\pi^2 f^2/c^2$, c is the speed of light, and $\hat{\psi}$ is the time-independent scalar field coefficient at $\mathbf{x}$. The Green's functions for this equation can be written as a linear combination of the two functions:

$$G^{\pm} = \frac{e^{\pm ikr}}{r} \tag{14}$$

G^+ can be viewed as a spherical wave expanding about $r = 0$, while G^- is a spherical wave traveling inward to $r = 0$. Because of this factor, G^+ is known as the causal Green's function and is used for forward propagation, while G^- is known as the anticausal Green's function and is instead used for migration.

Applying Green's theorem with the Green's function $G^- = e^{-ikr}/r$ and simplifying, we get the following integral:

$$\hat{\psi}(\mathbf{x}) = \oint_{\partial V} \left(\frac{e^{-ikr}}{r} \frac{\partial \hat{\psi}(\mathbf{x}_s)}{\partial n} - \hat{\psi}(\mathbf{x}_s) \frac{\partial}{\partial n} \frac{e^{-ikr}}{r} \right) \mathrm{d}s \tag{15}$$

where ∂V denotes the chosen surface of integration, and $\mathbf{x}_s$ denotes different locations along this surface. Expanding the normal derivative of G^-, we derive the expression:

$$\hat{\psi}(\mathbf{x}) = \oint_{\partial V} \left(\frac{e^{-ikr}}{r} \frac{\partial \hat{\psi}(\mathbf{x}_s)}{\partial n} + ik\hat{\psi}(\mathbf{x}_s) \frac{e^{-ikr}}{r} \frac{\partial r}{\partial n} + \hat{\psi}(\mathbf{x}_s) \frac{e^{-ikr}}{r^2} \frac{\partial r}{\partial n} \right) \mathrm{d}s \tag{16}$$

We can restore time dependence by multiplying by $e^{-2\pi i f t}$, yielding:

$$\psi(\mathbf{x}, t) = \oint_{\partial V} \left(\frac{e^{-2\pi i f(t+r/c}}{r} \frac{\partial \hat{\psi}(\mathbf{x}_s)}{\partial n} + i2\pi f \hat{\psi}(\mathbf{x}_s) \frac{e^{-2\pi i f(t+r/c)}}{cr} \frac{\partial r}{\partial n} + \hat{\psi}(\mathbf{x}_s) \frac{e^{-2\pi i f(t+r/c)}}{r^2} \frac{\partial r}{\partial n} \right) \mathrm{d}s \tag{17}$$

If we denote $\hat{\psi}(\mathbf{x}) e^{-2\pi i f(t+r/c)} = \psi(\mathbf{x}, t+r/c)$ as $[v]_{t+r/c}$, we can rewrite the above expression in simpler form:

$$\psi(\mathbf{x}, t) = \oint_{\partial V} \left(\frac{1}{r} \left[\frac{\partial}{\partial n} \right]_{t+r/c} - \frac{1}{cr} \frac{\partial r}{\partial n} \left[\frac{\partial \psi}{\partial t} \right]_{t+r/c} + \frac{1}{r^2} \frac{\partial r}{\partial n} [\psi]_{t+r/c} \right) \mathrm{d}s \tag{18}$$

We now need to relate this expression to the geometry of our NLOS scene. We set the top portion of the surface of integration dV to contain the visible wall S_0 where we have captured measurements, and set the bottom portion S_z directly beneath the reflectors in the NLOS scene. We assume S_z contributes negligibly to the total integral. We join these surfaces at infinity using cylindrical "walls" perpendicular to the visible wall S_∞, which also have little contribution. With these ideas in mind, our expression becomes:

$$\psi(\mathbf{x}, t) = \oint_{S_0} \left(-\frac{1}{r} \left[\frac{\partial \psi}{\partial z} \right]_{t+r/c} + \frac{1}{cr} \frac{\partial r}{\partial z} \left[\frac{\partial \psi}{\partial t} \right]_{t+r/c} - \frac{1}{r^2} \frac{\partial r}{\partial z} [\psi]_{t+r/c} \right) \mathrm{d}s \tag{19}$$

Sign changes occur because $\overrightarrow{n}$ is the outgoing normal (into the wall) and z increases downward, resulting in $\partial n = -\partial z$.

If we assume that $\psi(\mathbf{x}, t)$ can be approximately modeled as the wavefield received from a point source placed at the location of the light source, then we can write the following expression:

$$\frac{\partial \psi}{\partial z} = -\frac{1}{c} \frac{\partial r}{\partial z} \frac{\partial \psi}{\partial t} + \frac{\partial r}{\partial z} \frac{\psi}{r} \tag{20}$$

Plugging into Equation 19, we simplify to:

$$\psi(\mathbf{x}, t) = \oint_{S_0} \left(\frac{2\cos\theta}{cr} \left[\frac{\partial \psi}{\partial t} \right]_{t+r/c} - \frac{2\cos\theta}{r^2} [\psi]_{t+r/c} \right) \mathrm{d}s \tag{21}$$

with θ equal to the angle between the wall normal and the vector connecting $\mathbf{x}_s$ to $\mathbf{x}$. If we assume a confocal setup for NLOS imaging, we can replace c with $c/2$ and set $t = 0$:

$$\psi(\mathbf{x}, 0) = \oint_{S_0} \left(\frac{4\cos\theta}{cr} \left[\frac{\partial \psi}{\partial t} \right]_{2r/c} - \frac{2\cos\theta}{r^2} [\psi]_{2r/c} \right) \mathrm{d}s \tag{22}$$

In general, the near-field term $\frac{2\cos\theta}{r^2} \frac{\partial r}{\partial z} [\psi]_{2r/c}$ is dropped in seismology, with the assumption that the wavelength is much larger than r. In the case of NLOS imaging, because

we work with a wide range of frequencies in practice, we keep the near-field term (However, I empirically found that the near-field term has very little impact on the final reconstructions. I left it in for theoretical purposes but you can probably just drop it in practice). Our final migration equation is therefore given by:

$$\rho(\mathbf{x}) \approx \psi(\mathbf{x}, 0) = \oint_{S_0} \left(\frac{4\cos\theta}{cr} \frac{\partial \psi}{\partial t}(\mathbf{x}_s, 2r/c) - \frac{2\cos\theta}{r^2}\psi(\mathbf{x}_s, 2r/c) \right) \mathrm{d}s \tag{23}$$

where $\psi(\mathbf{x}_s, t)$ corresponds to the scalar field measured at the visible wall. Note that the same expression can be derived by using a slightly different Green's function as in Rayleigh-Sommerfeld Diffraction.

Equation 23 can be easily expressed as a prefiltered backprojection of two signals. To illustrate, consider the far-field term of the integral:

$$\rho(\mathbf{x}) = \oint_{S_0} \frac{4\cos\theta}{cr} \frac{\partial \psi}{\partial t}(\mathbf{x}_s, 2r/c)\, \mathrm{d}s \tag{24}$$

We can expand the surface integral, introduce a delta function over the time dimension, and expand $\cos\theta = z/r$ to rewrite the above equation as a familiar prefiltered backprojection:

$$\frac{\rho(x,y,z)}{z} = \iiint \frac{4}{cr^2} \frac{\partial \psi}{\partial t}(x_s, y_s, t)\delta\left(t - \frac{2r}{c}\right) \mathrm{d}x_s\, \mathrm{d}y_s\, \mathrm{d}t \tag{25}$$

A similar process can be applied for the near-field term.

4.2. Second-order Kirchhoff Migration

The migration equation from the previous section implies the usage of what is known as an exploding-reflector model (ERM) in seismology, where every point in the scene is effectively viewed as an independent point emitter. The same assumption is used as part of F-K migration [21]. While this model can provide good results in practice, it only accounts for one way propagation between the measurement wall and the hidden scene.

To also model the forward propagation between the light source and the hidden scene, we propose basically applying Kirchhoff migration twice. As we showed in the previous section, we can write the scalar field at some location in the hidden scene using the following expression:

$$\psi(\mathbf{x}, t) = \oint_{S_0} \left(\frac{2\cos\theta_s}{cr_s} \frac{\partial \psi}{\partial t}(\mathbf{x}_s, t + r_s/c) - \frac{2\cos\theta_s}{r_s^2}\psi(\mathbf{x}_s, t + r_s/c) \right) \mathrm{d}s \tag{26}$$

Here, the variables θ_s and r_s denote the same quantities as without the subscript in the previous section, between the visible wall and the hidden scene.

Now, what if we wanted to migrate this field again back to a location on the visible wall, like for example the position of the virtual light source x_l that generated this particular measurement? We can write the contribution of spatial location $\mathbf{x}$ using Kirchhoff migration:

$$f(\mathbf{x}, t) = \frac{2\cos\theta_l}{cr_l} \frac{\partial \psi}{\partial t}(\mathbf{x}, t + r_l/c) - \frac{2\cos\theta_l}{r_l^2}\psi(\mathbf{x}, t + r_l/c) \tag{27}$$

where θ_l is the angle between the normal of the visible wall and the vector connecting $\mathbf{x}_l$ to $\mathbf{x}$, and r_l is the distance between $\mathbf{x}_l$ and $\mathbf{x}$. We can expand this expression, and simplify to the following equation:

$$f(\mathbf{x}, t) = 4 \oint_{S_0} \frac{\cos\theta_l \cos\theta_s}{r_l r_s} \left(\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2}(\mathbf{x}_s, t + r_l/c + r_s/c) - \frac{1}{c}\left(\frac{1}{r_l} + \frac{1}{r_s}\right) \frac{\partial \psi}{\partial t}(\mathbf{x}_s, t + r_l/c + r_s/c) + \frac{1}{r_l r_s}\psi(\mathbf{x}_s, t + r_l/c + r_s/c) \right) \mathrm{d}s \tag{28}$$

In the case of a confocal measurement, we only measure S_0 at a single location for a particular light source. As a result, we drop the surface integral over the visible wall. However, we sum $f(\mathbf{x}, t)$ over all confocal measurement locations and set $t = 0$, resulting in the following simplified expression:

$$f(\mathbf{x}) = 4 \oint_{S_0} \frac{\cos^2\theta}{r^2} \left(\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2}(\mathbf{x}_s, 2r/c) - \frac{2}{cr} \frac{\partial \psi}{\partial t}(\mathbf{x}_s, 2r/c) + \frac{1}{r^2}\psi(\mathbf{x}_s, 2r/c) \right) \mathrm{d}s \tag{29}$$

The author views $f(\mathbf{x})$ as representing a variant of the albedo at $\mathbf{x}$, but he is not entirely sure and needs to think about it more. The author also thinks that you can set $t = 0$ because the sum of $f(\mathbf{x}, t)$ over all measurements should correspond to the input waveform at the virtual light source x_l, if we can accomplish a perfect migration. If the source pulse is a perfect Dirac delta, then the only part of $f(\mathbf{x}, t)$ that we are interested in corresponds to $t = 0$. The author also needs to think about this more.

This approach can once again be easily implemented as a prefiltered backprojection. We note that Equation 29 nearly matches the expression given in the radar literature by [31] - however, their version is derived in a much less concrete

fashion and uses Kirchhoff diffraction instead of migration, resulting in different signs.

Equation 29 also resembles the expression we derived in Section 3 for the LCT. Both techniques utilize a linear combination of the transient and its first and second derivatives, once again implying that there may be some equivalence between volumetric albedo techniques and wave-based techniques.

4.3. Implementation and Results

To actually implement Kirchhoff migration in practice, we approximate the time derivative by first applying a small Gaussian blur to the input transient, and then computing second order central differences. We use a similar approach for the second derivative for our second-order Kirchhoff migration. We compute the backprojections using the resampling process from the LCT [23]. To approximate the scalar field received at the wall from intensity measurements, we follow the naive approach of F-K migration [18] by taking a square root. Like F-K migration, we multiply the input transient by t in a preprocessing step.

We show comparisons of Kirchhoff migration and our second-order Kirchhoff migration with F-K migration [18] and the LCT [23] in Figure 1 on three scenes from the confocal dataset captured by [18]. In general, it seems that all four methods reconstruct scenes at roughly similar qualities. The baseline Kirchhoff migration seems to more accurately recover the shape of books on the shelf in the bottom-most scene that other techniques distort. However, at the same time, baseline Kirchhoff migration seems to result in some blurring around certain features, like the resolution chart in the bottom-most scene or around the body of the statue. These effects could potentially be due to the estimation of the temporal derivative for the far-field term of Equation 23 - if the data was higher temporal resolution, this far-field term could probably be estimated much more accurately and robustly. In contrast, our second-order migration produces the sharpest results of all tested methods at the cost of noise.

To test out the impact of reflectance, we compare the performance of the above methods on a specular discoball scene in Figure 2. The LCT clearly shows intensity streaking, while the other wave-based methods are for the most part robust. Our second-order Kirchhoff migration again shows the sharpest result by far, but noise is once again amplified.

We mentioned earlier in this section that Kirchhoff and F-K migration should theoretically yield equivalent results according to the seismology literature, but in practice there are a number of differences between the reconstructions. This could be due to a number of practical reasons beyond issues with temporal derivative estimation. One possible reason is that our implementation of Kirchhoff migration leverages a modified version of the LCT, which requires explicit discretization of a light cone kernel as well as significant resampling. Similarly, an implementation of F-K migration requires interpolation in the Fourier domain. Administering any of these steps in practice could introduce a number of artifacts.

5. F-K migration and Volumetric Albedo

In the previous section, we showed that the results of F-K migration can be approximated using a backprojection approach. If this is correct, then one might ask why F-K migration is particularly robust to BRDF, when the vast majority of previous techniques rely on variants of backprojection?

One potential answer is that Kirchoff migration, which is theoretically equivalent to F-K migration, propagates the temporal derivatives of the measured transient. [29] argued that discontinuities, corresponding to points in a transient with very large temporal derivatives, inherently arise thanks to the geometry of the hidden scene, irrespective of its BRDF. As a result, the reconstructions of Kirchoff and by extension F-K migration will be dominated by these geometry-related discontinuities, mitigating the impact of varying reflectance.

What if we put aside Kirchoff migration for now, and consider F-K migration in isolation? To investigate this question, we attempt to connect the F-K model and more traditional volumetric albedo models in the rest of this section. First, we show that a derivation from scalar theory akin to F-K migration can result in the volumetric albedo model. Second, inspired by the radar literature, we show that an expression very similar to F-K migration can be derived from a purely geometric standpoint. Our work demonstrates that the F-K model and a volumetric albedo approach are separated by just a high pass filter.

5.1. Volumetric Albedo from Scalar Theory

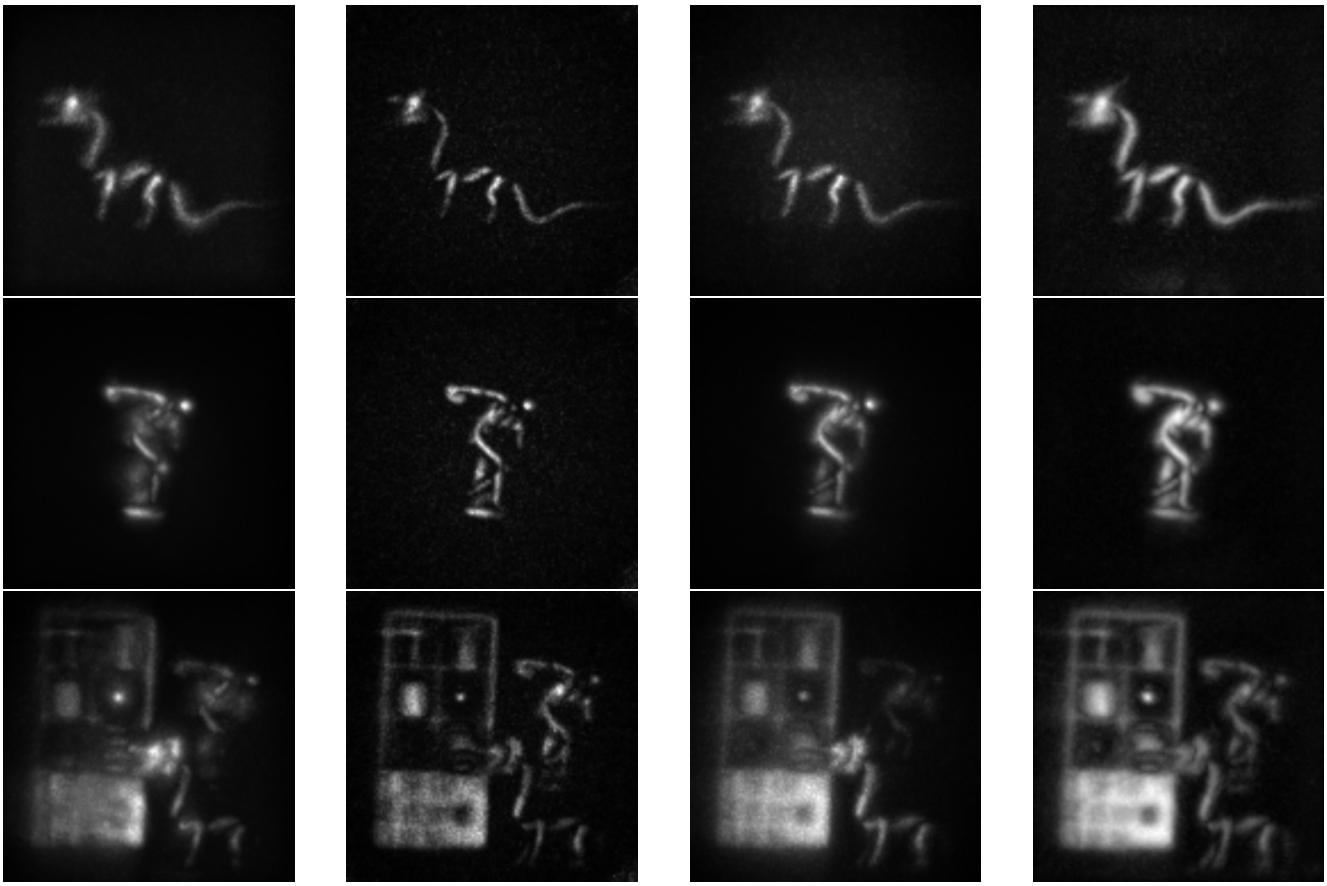
Consider the scalar wave equation:

$$(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2})\psi = 0 \tag{30}$$

where c is again the speed of light, ∇^2 is the Laplacian defined over the spatial dimensions, and ψ is the scalar field. Alternatively, we can define ψ in terms of its spectrum using a Fourier Transform operator over the spatial dimensions:

$$\psi(x, y, z, t) = \iiint \phi(k_x, k_y, k_z, t) e^{2\pi i(k_x x + k_y y + k_z z)} \, \mathrm{d}k_x \, \mathrm{d}k_y \, \mathrm{d}k_z \tag{31}$$

We can substitute this expression back into the scalar wave



(a) Kirchhoff migration (b) 2nd-order migration (c) F-K migration [18] (d) LCT [23]

Figure 1: Comparison of Kirchhoff migration and our modified second-order Kirchhoff migration with other state-of-the-art NLOS reconstruction approaches.

equation, and get the following relation for the spectrum:

$$4\pi^2 f^2 \phi(k_x, k_y, k_z, t) + \frac{\partial^2 \phi(k_x, k_y, k_z, t)}{\partial t^2} = 0 \quad (32)$$

where $f^2 = c^2(k_x^2 + k_y^2 + k_z^2)$. Solutions to Equation 32 take the form of:

$$\phi(k_x,k_y,k_z,t) = \phi_+(k_x, k_y, k_z)e^{2\pi i f t} + \phi_-(k_x, k_y, k_z)e^{-2\pi i f t} \quad (33)$$

If ϕ_- is set to zero, then the equations of F-K migration arise when plugged back into Equation 31 [18]. Alternatively, consider the case of $\phi_+(k_x, k_y, k_z) = -\phi_-(k_x, k_y, k_z) = \frac{\hat{\phi}(k_x,k_y,k_z)}{2i}$. Substituting into Equation 31:

$$\psi(x, y, z, t) = \iiint \frac{\hat{\phi}(k_x, k_y, k_z)}{2i}(e^{2\pi i f t} - e^{-2\pi i f t}) \, e^{2\pi i(k_x x + k_y y + k_z z)} \, \mathrm{d}k_x \, \mathrm{d}k_y \, \mathrm{d}k_z \quad (34)$$

Factoring out a f and simplifying the difference of complex exponentials, we get a simplified expression:

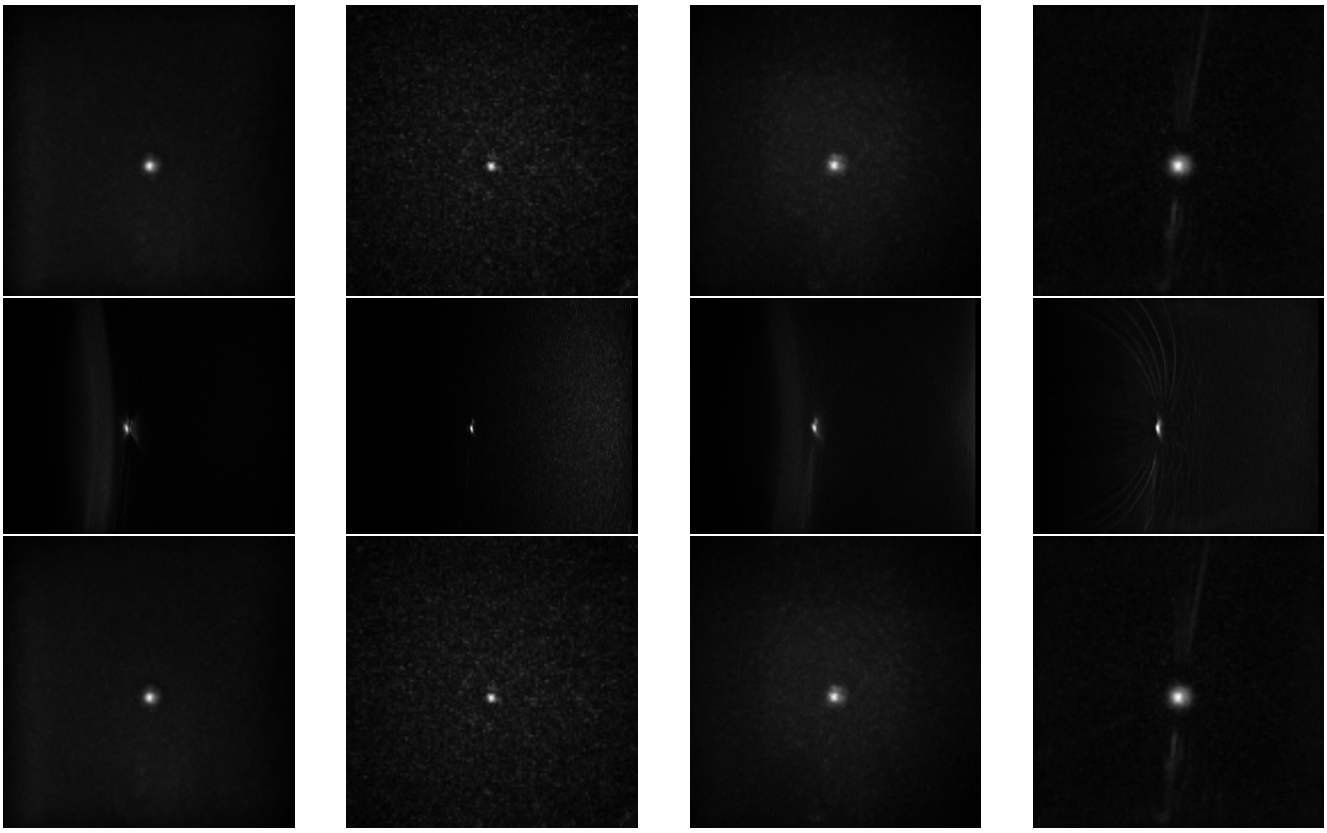
$$\psi(x, y, z, t) = \iiint f\hat{\phi}(k_x, k_y, k_z)\frac{\sin(2\pi f t)}{f} \, e^{2\pi i(k_x x + k_y y + k_z z)} \, \mathrm{d}k_x \, \mathrm{d}k_y \, \mathrm{d}k_z \quad (35)$$

This expression is easier to interpret in the spatial domain. Using the Hankel Transform or Spherical Bessel Transform, the following Fourier Transform pair can be proven:

$$\frac{\sin(2\pi f t)}{f} \rightarrow \frac{\delta(r - ct)}{2r} \quad (36)$$

where r is the distance from the origin. Then, by the Fourier Convolution Theorem, Equation 35 can be written in the spatial domain as:

$$\psi(x, y, z, t) = \iiint \frac{1}{2r}\rho(x', y', z')\delta(r - ct) \, \mathrm{d}x' \, \mathrm{d}y' \, \mathrm{d}z' \quad (37)$$



(a) Kirchhoff migration (b) 2nd-order migration (c) F-K migration [18] (d) LCT [23]

Figure 2: Comparison of Kirchhoff migration and our modified second-order Kirchhoff migration with other state-of-the-art NLOS reconstruction approaches on a specular discoball scene.

where ρ is the inverse Fourier Transform of $f\hat{\phi}$, and $r = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$.

Ignoring scaling factors, Equation 37 can be seen to exactly be the confocal volumetric albedo model with isotropic scattering, where every location in the NLOS scene can be effectively viewed as a point light source. In total, by just selecting a slightly different solution to Equation 32, we can derive the volumetric albedo model used in NLOS imaging from the theory behind F-K migration. This suggests that methods based on volumetric albedo, like the LCT [23], should be able to successfully reconstruct the same scenes as F-K. Furthermore, if we assume that the spectrum recovered through F-K $\phi_{fk}(k_x, k_y, k_z) \approx \hat{\phi}(k_x, k_y, k_z)/2i$, we can say that a volumetric albedo approach recovers the same geometry as F-K migration, high-pass filtered by f as shown in Equation 37. We show this idea more concretely in the next section.

5.2. F-K Migration from Volumetric Albedo

In this section, we follow the methodology previously outlined in the radar literature by [31] to derive an expression very similar to F-K migration from a volumetric albedo model.

Consider a simple scene with just a single scatterer, located at x', y', z'. Under a confocal volumetric albedo model with isotropic scattering, assuming $1/r$ falloff and just one way propagation, the measurement captured at the visible wall is given by:

$$s(x, y, z = 0, t) = \frac{1}{r}\rho(x', y', z')\delta(t - r/c) \tag{38}$$

We can compute a Fourier Transform over the temporal dimension, yielding:

$$s_k(x, y, f) = \frac{1}{r}\rho(x', y', z')e^{-i2\pi fr/c} \tag{39}$$

Computing more Fourier Transforms over the spatial domains yields:

$$s_f(k_x, k_y, f) = \rho(x', y', z')E(k_x, k_y, f) \tag{40}$$

where $E(k_x, k_y, f)$ is the spatial Fourier transform of $\frac{1}{r}e^{-i2\pi fr/c}$. In general, $E(k_x, k_y, f)$ is not well-formed:

however, it can be approximated using the Method of Stationary Phase:

$$E(k_x, k_y, f) = \frac{i}{k_z} e^{-2\pi i(x'k_x + y'k_y + z'k_z)} \tag{41}$$

with dispersion relation:

$$k_z = \sqrt{(f/c)^2 - k_x^2 - k_y^2} \tag{42}$$

Plugging back in:

$$s_f(k_x, k_y, f) = \frac{i}{k_z} \rho(x', y', z') e^{-2\pi i(x'k_x + y'k_y + z'k_z)} \tag{43}$$

To model a full NLOS scene, we integrate over the responses of individual scatterers, with τ_f equal to the spatiotemporal Fourier Transform of the measured transient:

$$\tau_f(k_x, k_y, f) = \iiint \frac{i}{k_z} \rho(x', y', z') e^{-i(x'k_x + y'k_y + z'k_z)} \, \mathrm{d}x' \, \mathrm{d}y' \, \mathrm{d}z' \tag{44}$$

Equation 44 shows that the transient measurement is the input albedo map, resampled and filtered in the Fourier domain. To reconstruct ρ from a measured transient τ, ρ can be written as the inverse Fourier Transform of a modified spectrum of τ:

$$\rho(x', y', z') = \iiint \bar{\tau_f}(k_x, k_y, k_z) e^{i(x'k_x + y'k_y + z'k_z)} \, \mathrm{d}k_x \, \mathrm{d}k_y \, \mathrm{d}k_z \tag{45}$$

where $\bar{\tau_f}$ is a resampled and filtered version of the Fourier Transform of the input transient:

$$\bar{\tau_f}(k_x, k_y, k_z) = \frac{k_z}{i} \tau_f(k_x, k_y, c\sqrt{k_x^2 + k_y^2 + k_z^2}) \tag{46}$$

Equations 45 and 46 very nearly match F-K migration - ignoring constant factors, the only difference is an extra filter $1/\sqrt{k_x^2 + k_y^2 + k_z^2}$ in the Stolt interpolation, repeated here for convenience:

$$\Phi(k_x, k_y, k_z) = \frac{ck_z}{\sqrt{k_x^2 + k_y^2 + k_z^2}} \bar{\Phi}(k_x, k_y, c\sqrt{k_x^2 + k_y^2 + k_z^2}) \tag{47}$$

where Φ is the spectrum at $t = 0$ corresponding to $\bar{\tau_f}$, and $\bar{\Phi}$ is the spectrum at the wall $z = 0$ corresponding τ_f. In other words, just as we found in the previous section, an approach based on volumetric albedo should recover the same geometry as F-K migration, high-pass filtered by f.

We note that a non-confocal version of F-K migration can be derived from the above approach, as shown in [31]. However, a practical implementation requires a rather ad-hoc interpolation between a 5D transient to a 3D transient in the Fourier Domain, which can potentially introduce a large number of artifacts. We show a sample result on a NLOS scene in Figure 3.

(a) Non-confocal F-K

(b) F-K on normal moveout corrected measurements [18]

Figure 3: Comparison of a non-confocal F-K implementation with F-K migration on normal moveout corrected measurements [18]. Non-confocal scene from [6].

6. Conclusion and Discussion

In this paper, we have shown that F-K migration and other NLOS approaches are intimately related. We showed that a method equivalent to F-K migration called Kirchhoff migration can be written as a prefiltered backprojection, an approach that other state-of-the-art techniques reduce to. We derived a second-order Kirchhoff migration that closely resembles the LCT. We showed that a volumetric albedo model can be used to derive a variant of F-K, and a variant of the F-K derivation can be used to derive the volumetric albedo model - in fact, the main difference is just a high pass filter f.

While this paper has derived a number of theoretical contributions, we have not significantly investigated the effects of practical issues. Perhaps the most important to consider is the impact of finite LOS wall size - we refer the reader to [1] for some of the potential impacts. Similarly, we have focused purely on a continuous case - discretizations of each of the approaches mentioned in this paper may yield different artifacts that are hard to quantify (and may contribute to BRDF effects as well). Regardless, the author hopes that this paper highlights the fundamental differences between state-of-the-art approaches in NLOS imaging.

References

[1] Byeongjoo Ahn, Akshat Dave, Ashok Veeraraghavan, Ioannis Gkioulekas, and Aswin C Sankaranarayanan. Convolutional approximations to the general non-line-of-sight imaging operator. In *IEEE International Conference on Computer Vision (ICCV)*, pages 7888–7898, Oct 2019.

[2] Katherine L. Bouman, Vickie Ye, Adam B. Yedidia, Frédo Durand, Gregory W. Wornell, Antonio Torralba, and William T. Freeman. Turning corners into cameras: Principles and methods. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2289–2297, Oct 2017.

[3] Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics Express*, 23(16):20997–21011, Aug 2015.

[4] Susan Chan, Ryan E. Warburton, Genevieve Gariepy, Jonathan Leach, and Daniele Faccio. Non-line-of-sight tracking of people at long range. *Optics Express*, 25(9):10109–10117, 2017.

[5] Sreenithy Chandran and Suren Jayasuriya. Adaptive lighting for data-driven non-line-of-sight 3d localization and object identification. 2019.

[6] Miguel Galindo, Julio Marco, Matthew O'Toole, Gordon Wetzstein, Diego Gutierrez, and Adrian Jarabo. A dataset for benchmarking time-resolved non-line-of-sight imaging, 2019.

[7] I. Guillén, X. Liu, A. Velten, D. Gutierrez, and A. Jarabo. On the effect of reflectance on phasor field non-line-of-sight imaging. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 9269–9273, 2020.

[8] Otkrist Gupta, Thomas Willwacher, Andreas Velten, Ashok Veeraraghavan, and Ramesh Raskar. Reconstruction of hidden 3D shapes using diffuse reflections. *Optics Express*, (20(17):19096–19108), 2012.

[9] Felix Heide, Wolfgang Heidrich, and Matthias B Hullin. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sens. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3222–3229, 2014.

[10] Felix Heide, Matthew O'Toole, Kai Zang, David B. Lindell, Steven Diamond, and Gordon Wetzstein. Non-line-of-sight imaging with partial occluders and surface normals. *ACM Trans. Graph.*, 2019.

[11] Julian Iseringhausen and Matthias B Hullin. Non-line-of-sight reconstruction using efficient transient rendering. *ACM Transactions on Graphics (TOG)*, 39(1):1–14, 2020.

[12] Mariko Isogawa, Dorian Chan, Ye Yuan, Kris Kitani, and Matthew O'Toole. Efficient non-line-of-sight imaging from transient sinograms. In *European Conference on Computer Vision*, pages 193–208. Springer, 2020.

[13] Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using transient imaging. In *IEEE International Conference on Computer Vision (ICCV)*, pages 159–166, 2009.

[14] Jonathan Klein, Christoph Peters, Martin Laurenzis, and Matthias Hullin. Tracking objects outside the line of sight using 2d intensity images. *Scientific Reports*, 6(32491):32491:1–32491:9, 2016.

[15] M. La Manna, F. Kine, E. Breitbach, J. Jackson, T. Sultan, and A. Velten. Error backprojection algorithms for non-line-of-sight imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(7):1615–1626, July 2019.

[16] Tianhong Li, Lijie Fan, Mingmin Zhao, Yingcheng Liu, and Dina Katabi. Making the invisible visible: Action recognition through walls and occlusions. In *arXiv preprint*, 2019.

[17] David B. Lindell, Gordon Wetzstein, and Vladlen Koltun. Acoustic non-line-of-sight imaging. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6780–6789, 2019.

[18] David B Lindell, Gordon Wetzstein, and Matthew O'Toole. Wave-based non-line-of-sight imaging using fast fk migration. *ACM Transactions on Graphics (TOG)*, 38(4):1–13, 2019.

[19] Xiaochun Liu, Sebastian Bauer, and Andreas Velten. Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems. *Nature communications*, 11(1):1–13, 2020.

[20] Xiaochun Liu, Ibón Guillén, Marco La Manna, Ji Hyun Nam, Syed Azer Reza, Toan Huu Le, Adrian Jarabo, Diego Gutierrez, and Andreas Velten. Non-line-of-sight imaging using phasor-field virtual wave optics. *Nature*, 572(7771):620–623, 2019.

[21] Gary F. Margrave and Michael P. Lamoureux. *Numerical Methods of Exploration Seismology: With Algorithms in MATLAB®*. Cambridge University Press, 2019.

[22] Christopher A. Metzler, David B. Lindell, and Gordon Wetzstein. Keyhole imaging: Non-line-of-sight imaging and

tracking of moving objects along a single optical path at long standoff distances, 2019.

[23] Matthew O'Toole, David B Lindell, and Gordon Wetzstein. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature*, 555(7696):338–341, 2018.

[24] Adithya Pediredla, Akshat Dave, , and Ashok Veeraraghavan. Snlos: Non-line-of-sight scanning through temporal focusing. In *2019 IEEE International Conference on Computational Photography (ICCP)*, pages 1–13, 2019.

[25] Charles Saunders, John Murray-Bruce, and Vivek K. Goyal. Computational periscopy with an ordinary digital camera. *Nature*, 565(7740):472–475, 2019.

[26] Matthew Tancik, Guy Satat, and Ramesh Raskar. Flash photography for data-driven hidden scene recovery. In *arXiv preprint*, 2018.

[27] C. Tsai, A. C. Sankaranarayanan, and I. Gkioulekas. Beyond volumetric albedo — a surface optimization framework for non-line-of-sight imaging. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1545–1555, June 2019.

[28] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature communications*, 3(1):1–8, 2012.

[29] Shumian Xin, Sotiris Nousias, Kiriakos N Kutulakos, Aswin C Sankaranarayanan, Srinivasa G Narasimhan, and Ioannis Gkioulekas. A theory of fermat paths for non-line-of-sight shape reconstruction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6800–6809, 2019.

[30] Sean I. Young, David B. Lindell, Bernd Girod, David Taubman, and Gordon Wetzstein. Non-line-of-sight surface reconstruction using the directional light-cone transform. In *Proc. CVPR*, 2020.

[31] Xiaodong Zhuge and Alexander G Yarovoy. Three-dimensional near-field mimo array imaging using range migration techniques. *IEEE Transactions on Image Processing*, 21(6):3026–3033, 2012.