Occluded Imaging with Time of Flight Sensors

ACHUTA KADAMBI[1], HANG ZHAO[1], BOXIN SHI[1,2] and RAMESH RASKAR[1]
[1]MIT Media Lab, Cambridge MA,[2]SUTD, Singapore

We explore the question of whether phase based Time of Flight (ToF) range cameras can be used for looking around corners and through scattering diffusers. By connecting time of flight measurements with theory from array signal processing we conclude that performance depends on two primary factors: camera modulation frequency and the width of the specular lobe ("shininess") of the wall. For purely Lambertian walls, commodity time of flight sensors achieve resolution on the order of meters between targets. For seemingly diffuse walls, such as posterboard, the resolution is drastically reduced, to the order of ten centimeters. In particular, we find that the relationship between reflectance and resolution is nonlinear—a slight amount of shininess can lead to a dramatic improvement in resolution. Since many realistic scenes exhibit a slight amount of shininess we believe that off-the-shelf ToF cameras can look around corners.

Categories and Subject Descriptors: I.2.10 [**Artificial Intelligence**]: Vision and Scene Understanding—*3D/Stereo scene analysis*

Additional Key Words and Phrases: Time of Flight, Scattering, Computational Photography

1. INTRODUCTION

With the emergence of low cost, yet powerful Time of Flight (ToF) range cameras (such as the new Microsoft Kinect) a question that is raised is whether such off the shelf devices can be used to look around corners and through diffusers.

The looking around corners problem (Figure 6a) is an important area of study. It is a theoretically challenging problem in image formation where the image sensor receives photons that scatter from line-of-sight (LOS) objects, non-line-of-sight (NLOS) objects, and background illumination. Separating these reflections has been a substantial challenge. The current state of the art is the solution from [Velten et al. 2012] where they demonstrate looking around corners in a controlled laboratory setup. Although this result does not use scene priors, it is dependent on ultrafast optical hardware, which comes at a high price tag and significant limitations. Recent work has demonstrated a solution that may drastically lower the cost of such a system [Heide et al. 2014]. However, this solution exploits scene-dependent priors in the reconstruction and uses customized hardware to acquire "transient images".

Following the problem's introduction in [Velten et al. 2012], several fundamental questions remain unanswered. These include:

—*How high does the camera modulation frequency need to be?* We show that the camera modulation frequency has an approximately linear relationship with desired resolution.

—*How "diffuse" or "shiny" does the wall have to be so we can look around the corner?* We show that the width of the specular lobe has a nonlinear relationship to recovery.

—*Do we need only amplitude or both phase and amplitude?* In practice, amplitude-only localization is susceptible to noise; we compare the two strategies.

—*Under what conditions is recovery possible?* It depends on the physical constraints and computational choices we make.

It turns out that all of these questions can be addressed through a unified forward model that we propose in the paper. We call this model a Virtual Sensor Array (VSA) as it connects ToF range measurements with array signal processing. We recognize that poblems such as looking around corners (also denoted as "corners" problem in the following text, for simplicity) are specialized. However, the VSA model generalizes to handle imaging through diffusers as well.

In summary, our key contribution is:

—A unifying theoretical framework for occluded imaging with time of flight cameras; we use array signal processing to formulate limits on recoverability and add specularity to the formulation.

Secondary technical contributions:

—Demonstration of occluded imaging in the context of the corners problem.

—Generalizing the theory to handle imaging through diffusers, and a practical demonstration of this case.

Benefits As compared to prior art [Velten et al. 2012; Heide et al. 2014], our model is rooted in array signal processing theory. In addition, while comparison work exploits customized hardware for "transient imaging", our hardware uses standard data from a ToF camera (i.e. only phase/amplitude at each pixel). A key benefit—beyond reproducibility and low cost—is real-time acquisition. To our knowledge this paper is the first to provide detailed bounds on recovery for looking around corners. This may be used as a blueprint for future camera designs.

Limitations In this paper, we propose a theoretical framework to understand the problem. Due to the extensive customization of transient imaging hardware, we validate our theory on off-the-shelf devices. This has the benefit of reproducibility, low cost, and real-time potential but a drawback in the perceived quality of results. Specific engineering challenges include a lack of customized illumination and the use of a single modulation frequency that restrict our demonstrations to relatively simple scenes. Nevertheless, to validate the theory, our scenes convey important information, such as the resolution between two objects.

2. RELATED WORK

Holography literature contains a closely related work that overlaps with our theoretical model. Specifically, [Rivenson et al. 2013], provide reconstruction guarantees for compressive holography using discrete spatial sampling to map a sparse set of 3D points to intensity and phase measurements of a 2D surface. Reconstruction guarantees are provided using the Gram matrix of a Fresnel sensing operator. In contrast, our paper is entirely in the realm of ray-based optics and maps intensity and phase of TOF measurements to a surface.

Time profile imaging represents an increasingly popular research area where captured photons are parameterized by both space and time. The "Femtophotography" technique by [Velten et al. 2013] uses a laboratory grade optical setup to capture visualizations of light in flight, however the technique is expensive (half million dollars). Recently, [Heide et al. 2013] and [Kadambi et al. 2013] have repurposed low cost 3-D, time of flight sensors to achieve some of the capabilities of Velten's system. However, beyond customized hardware, both of these techniques require the acquisition of a time-frequency shift matrix [Heide et al. 2013; Lin et al. 2014] or a time shift vector [Kadambi et al. 2013], which mitigates the real-time advantage that time of flight sensors usually enjoy. Time profile imaging has several applications, and in particular, [Velten et al. 2012] and [Heide et al. 2014] use such data to address the corners problem. Multipath interference correction is closely related to this paper. While this paper exploits information in scattered light, related techniques in light transport (cf. [O'Toole et al. 2014; Gupta et al. 2014], or time-frequency analysis [Bhandari et al. 2014] correct for such interference. A light transport analysis of transient imaging can be found in [Velten et al. 2012; Wu et al. 2014], and a comprehensive review of transient technique was collected by [Masia 2014].

Non-line-of-sight target localization is a classic inverse problem that has been studied in a variety of domains. A prominent example is multi-path radar system. For example in [Sen and Nehorai 2011], a Doppler radar system is equipped with spatial diversity (i.e. detectors at different spatial locations) to allow the system to obtain multiple "looks" of a target and resolve NLOS objects in motion. In [Sume et al. 2011], they demonstrate a radar system designed to track sources around a corner. In [Adib et al. 2014], radio waves are used to track humans: however, this technique works for NLOS only when the medium is transparent to radio waves, and makes the limiting assumption that the target is in motion. Optically localizing a source is challenging as diffuse scattering needs to be taken into account. Within computational imaging, various techniques have used indirect reflections to infer various scene properties [Reshetouski and Ihrke 2013; Reshetouski et al. 2011; Naik et al. 2011].

Phased array source localization can be described as follows: given phase and amplitude measurements from multiple sensors located near a source signal, how can we localize the source? The computational technique of choice may depend on whether the source is in the near or far field. The classic technique is time shifted beamforming which was introduced in [Carter 1981]. Holography methods have also been used for decades to achieve near-field acoustic source localization [Maynard et al. 1985]. In discrete approaches, such as [Malioutov et al. 2005] and [Cevher et al. 2008], a coordinate grid system of N voxels is drawn in the search space. Assuming that K targets are located on grid, and since $K \ll N$, by coupling grid source localization with sparse priors, it is possible to resolve very closely spaced sources. More recent work in signal processing often leverages model-based algorithms to enhance recovery [Boufounos et al. 2011; Hegde et al. 2014].

Sparse approximation refers to the problem of estimating sparse vector that satisfies a linear system of equations. Concretely, given a measurement vector $\mathbf{y}$ and a dictionary matrix $\mathbf{D}$, the goal is to solve for $\mathbf{x}$ where $\mathbf{y} = \mathbf{D}\mathbf{x}$ and $\mathbf{x}$ is known to be sparse. To solve the linear system and enforce sparsity on $\mathbf{x}$, popular solutions include iterative approaches that use an ℓ_1 regularization penalty or greedy techniques that require multiple projections. Examples of the former include LASSO [Tibshirani 1996] and Basis Pursuit [Chen et al. 2001], while examples of the latter include the matching pursuits originally introduced in [Mallat and Zhang 1993]. Another key aspect is how to select the dictionary matrix $\mathbf{D}$, such that the problem is well posed. In [Elad 2010], sparse recovery can be guaranteed when columns of $\mathbf{D}$ have a sufficiently small inner product, or more concretely, a low *mutual coherence*.

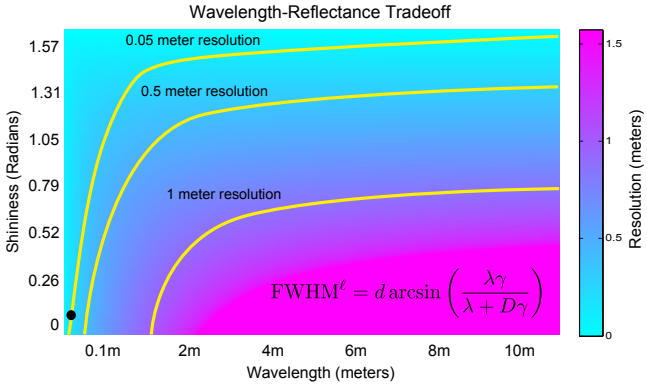


Fig. 1: The corners problem is increasing in popularity as the solution becomes within reach. The solution by Velten et al. is "impulse" based, using a very small wavelength to achieve high resolution. We explore the space of amplitude modulated, continuous wave cameras which achieve a resolution that depends on the modulation frequency and shininess (γ in radians) of the wall. In this paper we derive a parametric bound for target resolution.

3. TIME OF FLIGHT LOCALIZATION

We begin by recasting time of flight 3-D imaging into the realm of array signal processing.

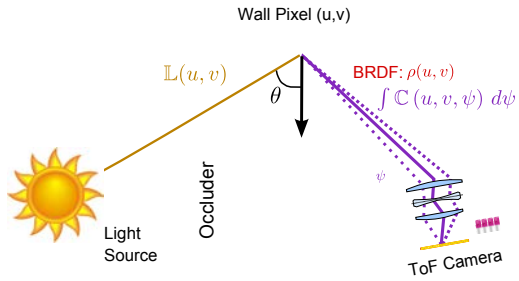
Time of Flight imaging We use the term *time of flight* (ToF) to refer to the time it takes for photons to travel through a medium. Although there are several devices to measure ToF, we restrict subsequent technical discussion to amplitude modulated continuous wave (AMCW) time of flight cameras.

To obtain depth, a light source strobes in a periodic pattern and photons are captured with a lock-in CMOS sensor. The carrier signal is the optical signal and the modulation envelope is the strobing pattern with modulation frequency $f_{\mathcal{M}}$. The phase difference between the received and emitted modulation codes, $\varphi^{\mathcal{M}}$, encodes the propagation distance via the following linear relation:

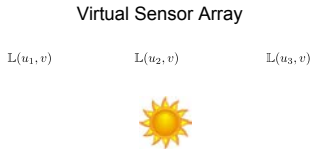
$$z = \frac{c\varphi^{\mathcal{M}}}{2\pi f_{\mathcal{M}}}, \quad c \approx 3 \times 10^8 \text{ m/s}. \tag{1}$$

Here, z is the propagation (in meters) of the optical path. A common value of $f_{\mathcal{M}}$ is 30 MHz, which corresponds to a λ of 10 meters. The camera also measures the amplitude of the reflected light, denoted as A. In summary, a ToF camera is unique in that the pair of phase and amplitude is measured at each pixel.

Source localization The 2-D source localization problem is as follows. Consider a set of M sensors spaced evenly on a horizontal axis, u. There are K transmitting sources located on the 2-D space parametrized by u and w axis. Denote the signal time delay from k-th source to m-th sensor as $\tau_{k,m}$. Then, in frequency domain, the m-th sensor receives $Y_m(2\pi f_{\mathcal{M}}) =$



(a) Lensless imaging.



(b) Triangulation with lensless measurements.

Fig. 2: Recasting time of flight cameras into a (virtual) sensor array. (a) Mapping out the transport paths from a light source to coordinates (u, v) on the wall. The wall serves as a virtual sensor array (VSA). By taking a picture of the virtual sensor we obtain the measurement model. The goal of Section 4.1 is to extract $\mathbb{L}(u, v)$ from the ToF camera measurements. (b) Given such measurements, a reasonable next step is to use backprojection to localize the source. This second step connects to lensless transient imaging, as introduced in [Wu et al. 2014].

$\sum_{k=1}^{K} A_k \exp\left(-j2\pi f_{\mathcal{M}} \tau_{k,m}\right)$. Substitution using $z = c\tau_{k,m}$ and Equation 1 yields: $Y_m(2\pi f_{\mathcal{M}}) = \sum_{k=1}^{K} A_k \exp\left(-j\varphi_{k,m}^{\mathcal{M}}\right)$. The superscript $\mathcal{M}$ on φ emphasizes that this is the phase associated with modulation frequency $f_{\mathcal{M}}$, i.e., for a fixed z, $\varphi^{\mathcal{I}} \neq \varphi^{\mathcal{J}}$ for $\mathcal{I} \neq \mathcal{J}$. As most of the analysis is concerned with narrowband scenarios, we drop the superscript. Therefore, the observation model is written as:

$$\begin{aligned}\vec{y} &\triangleq \left[Y_1\left(2\pi f_{\mathcal{M}}\right), \ldots, Y_M\left(2\pi f_{\mathcal{M}}\right)\right]^T \\ &= \left[\sum_{k=1}^{K} A_k \exp\left(-j\varphi_{k,1}\right), \ldots, \sum_{k=1}^{K} A_k \exp\left(-j\varphi_{k,M}\right)\right]^T,\end{aligned} \tag{2}$$

where $\vec{y}$ is an M-dimensional measurement vector defined over the complex field. We summarize the key intuition: each entry of $\vec{y}$ represents the measured amplitude and phase at a single sensor.

4. VIRTUAL SENSOR ARRAY FOR RECONSTRUCTION

Recall that a sensor array is an array of M sensors each measuring phase and amplitude. A *virtual* sensor array (VSA) probes the idea of turning ordinary surfaces into a sensor array. Consider the toy problem in Figure 2a: a point source emitter is hidden around the corner and the goal is to recover its location and amplitude from VSA measurements. Taken together, Sections 4.1 and 4.2 introduce the virtual sensor array in the context of this toy problem, following which Sections 4.3 and 4.4 generalize the model to broader scenes.

4.1 Virtual Sensor Array

In a time of flight camera, intensities are parameterized both spatially and temporally as

$$c(u, v, t) = A\left(u, v\right) \sin\left(2\pi f_{\mathcal{M}} t + \varphi\left(u, v\right)\right) + \zeta\left(u, v\right), \tag{3}$$

where $c(u, v, t)$ is the correlation waveform with amplitude $A(u, v)$ and phase $\varphi(u, v)$. The quantity $\zeta(u, v)$ is an offset term that represents ambient lighting. Note that $\zeta(u, v)$, $\varphi(u, v)$, and $A(u, v)$ are not parameterized in time — we assume these quantities are constant over a short integration time. Then, the amplitude and phase can be expressed as a measurement phasor

$$\mathbb{M}(u, v) \triangleq A_{\mathbb{M}}(u, v) e^{j\varphi_{\mathbb{M}}(u,v)}, \tag{4}$$

where the addition of subscript $\mathbb{M}$ to amplitude and phase links the two quantities with the phasor $\mathbb{M}$. Note that the DC offset from Equation 3 is not captured in the phasor notation of Equation 4. This is perfectly fine, as the offset is not useful (it is uncontrolled, ambient light).

Consider the case in Figure 2a, where a single omni-directional point source emits rays of light onto a wall. Localizing the source is trivial when the wall is mirrored, which allows the point source to be observed directly by the camera. This section is concerned with the more general scenario of localizing the point source when the wall is modeled as a Lambertian surface. The key insight is to represent the wall itself as a virtual, lensless imaging sensor in the (u, v) plane. We are interested in obtaining this "lensless image" formed on the virtual sensor.

Using ray optics we begin by analyzing the complex domain light transport of a unit amplitude strobing signal. As illustrated in Figure 2a, the transport phasor from source to wall is represented as $\mathbb{L}(u, v, \theta)$, where u and v are the coordinates of the wall that the ray strikes at an angle of θ to the normal. Therefore, the phasor that models transport from light source to the diffuse wall is written as

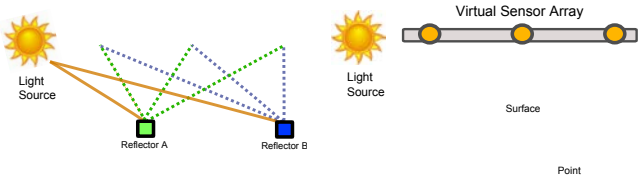
$$\mathbb{L}(u, v) = \underbrace{\frac{\cos\theta}{\varphi_{\mathbb{L}}(u, v)^2}}_{\text{Amplitude Decay}} e^{j\varphi_{\mathbb{L}}(u,v)}. \tag{5}$$

Here, note that $\mathbb{L}$ is parametrized by only u and v because, assuming the geometry in Figure 2a, the angle θ is a function of (u, v). The amplitude of the transport phasor is designed to represent an amplitude decay term. A similar transport phasor can be formulated from rays emitted from the wall to the camera. As illustrated in Figure 2a, an outgoing ray makes an angle ψ with the normal vector of the wall. The corresponding phase is $\varphi_{\mathbb{C}}(u, v)$. Therefore, the transport phasor is

$$\mathbb{C}(u, v, \psi) = \underbrace{\rho\left(u, v\right) \frac{\cos\psi}{\varphi_{\mathbb{C}}(u, v)^2}}_{\text{Amplitude Decay}} e^{j\varphi_{\mathbb{C}}(u,v)}, \tag{6}$$

where now $\rho(u, v)$ represents the Lambertian albedo of the wall at coordinates (u, v). Using the two transport phasors as well as the original amplitude of the strobing signal, A_0, the combined phasor transport from source to camera is a phasor multiplication:

$$\mathbb{M}\left(u, v, \psi\right) = A_0 \mathbb{L}\left(u, v\right) \mathbb{C}\left(u, v, \psi\right). \tag{7}$$



(a) Point reflectors are emitters (b) Surfaces are point reflectors

Fig. 3: Generalizing the toy problem from Figure 2a; (a) point reflectors can be thought of as emitters; and (b) over the voxel grid, the blue surfaces can be thought of as multiple point reflectors (green).

Since the camera is perfectly focused on the wall, Equation 7 can be written as

$$\mathbb{M}(u,v) = A_0 \mathbb{L}(u,v) \int \mathbb{C}(u,v,\psi)\, d\psi, \tag{8}$$

where the emitting angle has been integrated out. Embedded within Equation 8 is the phasor $\mathbb{L}(u,v)$, which is the projection of the source onto the wall or the virtual imaging plane. If this phasor could be isolated, then these measurements could be used to formulate a phased array, source localization problem. Because the forward problem is phasor multiplication, it is simple to isolate $\mathbb{L}(u,v)$ as

$$\mathbb{L}(u,v) = \left(\frac{1}{A_0}\right)\left(\frac{\mathbb{M}(u,v)}{\mathbb{C}(u,v)}\right). \tag{9}$$

Although $\mathbb{C}(u,v)$ is an unknown, simply computing a depth map of the wall provides the phase $\varphi_{\mathbb{C}}$. Using just $\varphi_{\mathbb{C}}$ and unit amplitude as a proxy for $\mathbb{C}(u,v)$ we obtain

$$\begin{aligned}\widehat{\mathbb{L}}(u,v) &= \left(\frac{1}{A_0}\right)\left(\frac{\mathbb{M}(u,v)}{e^{j\varphi_{\mathbb{C}}(u,v)}}\right) \\ &= \left(\frac{\cos\theta}{\varphi_{\mathbb{L}}(u,v)^2}\int \rho(u,v)\,\frac{\cos\psi}{\varphi_{\mathbb{C}}(u,v)^2}d\psi\right) e^{j\varphi_{\mathbb{L}}(u,v)} \\ &= \widehat{A}_{\mathbb{L}}(u,v)\, e^{j\widehat{\varphi}_{\mathbb{L}}(u,v)}.\end{aligned} \tag{10}$$

Here, $\widehat{\mathbb{L}}(u,v)$ is an estimate of $\mathbb{L}(u,v)$ with the correct phase and different amplitude. In order to treat the amplitude as a uniform difference in scaling (across all array elements), we assume that the reflectance profile is uniform, i.e., $\rho(u,v)$ is the same for all u and v. In summary, the key measurements are $\widehat{\varphi}_{\mathbb{L}}(u,v)$ and $\widehat{A}_{\mathbb{L}}(u,v)$. These represent the projection of phases and amplitudes of the light source onto the virtual sensor which can be used in the context of source localization.

4.2 Reconstruction

The problem has now been abstracted to 3-D source localization with a 2-D array of sensors, parameterized by (u,v). Each virtual sensor element gives phase and amplitude measurement $\widehat{\varphi}_{\mathbb{L}}(u,v)$ and $\widehat{A}_{\mathbb{L}}(u,v)$. The target is a point source whose real-world location is parameterized in 3-D spatial coordinates (X,Y,Z). Our goal is to find the points $(\widehat{u},\widehat{v},\widehat{w})$ which correspond to the target coordinates with respect to the wall and camera sensor.

Without loss of generality we will consider 2-D source localization using a 1-D slice of measurements, i.e., to obtain $(\widehat{u},\widehat{w})$ by using sensor measurements only along the horizontal u-axis. The measurement vector is of the form

$$\vec{y} = \left[\, \widehat{\mathbb{L}}(u_1)\ \ \widehat{\mathbb{L}}(u_2)\ \ \cdots\ \ \widehat{\mathbb{L}}(u_M)\,\right]^\top. \tag{11}$$

Following a parallel technique in holography [Rivenson et al. 2013], the search space is discretized to a grid of voxels (see Figure 3b). This allows standard source localization algorithms to be used [Cevher et al. 2008], whereby the set

$$G = \{t_i \mid i = 1,\ldots,N;\ t_i = [u_i, w_i]\} \tag{12}$$

denotes the set of N possible grid points. There are R possible locations on the u-axis and Q possible locations on the w-axis, such that $N = RQ$. The set

$$T = \{t_i \mid i = 1,\ldots,K;\ t_i = [u_i, w_i]\} \tag{13}$$

describes a set of K targets that are located on the voxel grid. We assume for now that the targets are located on-grid and that $K \ll N$. The task is to find out which voxels contain a target. Concretely, we define

$$\vec{x} = \left[\, x_1\ \ x_2\ \ \cdots\ \ x_N\,\right]^\top \tag{14}$$

where each entry of $\vec{x} \in \mathcal{C}^N$ is the weight of confidence that the target is at that voxel. The connection to sparsity is apparent: targets will lie on K grid locations and the vector $\vec{x}$ is therefore sparse. To complete the model we must formulate the dictionary matrix $\mathbf{D}$. Define the operator

$$\mathcal{S}(u,w) : (u,w) \to \mathcal{C}^M, \tag{15}$$

which takes as input a potential target location and generates the expected measurements along the sensor array. Define the dictionary as

$$\mathbf{D} = \left[\, \mathcal{S}(u_1,w_1)\ \ \mathcal{S}(u_2,w_1)\ \ \cdots\ \ \mathcal{S}(u_R,w_Q)\,\right], \tag{16}$$

where the columns of $\mathbf{D}$ have unit norm. In the *backprojection* problem it is expected that the observed measurement $\vec{y}$ can be represented by K columns of $\mathbf{D}$, such that

$$\vec{y} = \mathbf{D}\vec{x} \quad \text{s.t.} \quad \|\vec{x}\|_0 = K. \tag{17}$$

This is the key objective function from which we want to recover $\vec{x}$. To attack this objective function directly, sparse solvers are used. In cases where sparse solvers are not appropriate, e.g., if a good estimate for K is not available, then the standard option in signal processing is to use *beamforming*:

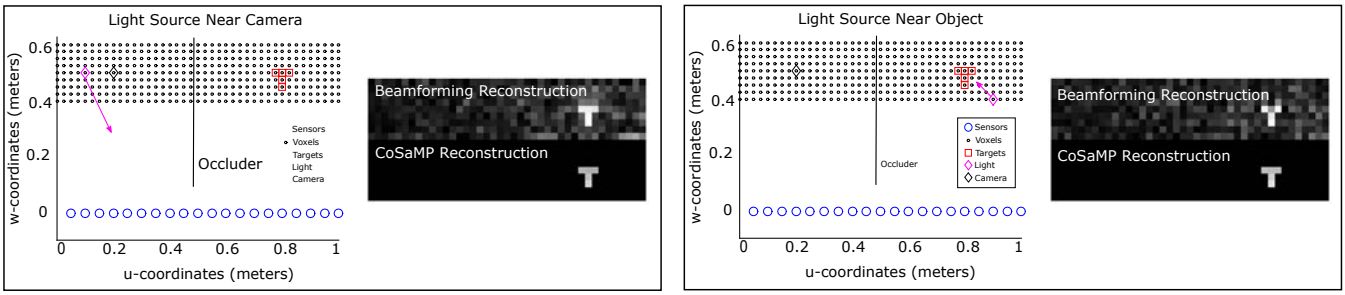
$$\widehat{\vec{x}} = \mathbf{D}^H \vec{y}. \tag{18}$$

Intuitively, the beamforming equation provides a crude estimate of source location by computing an inner product of the measured signal with each dictionary atom. Therefore, entries of $\widehat{\vec{x}}$ with a large magnitude indicate possible source locations under the beamforming model and K does not need to be known. These are the two variants that a practitioner needs to consider for the VSA problem; a list of standard solvers and objective functions for this problem can be found in Table I.

4.3 Generalizing the Model

While the toy problem allows clean derivations, it does not accurately address the goal of *imaging* around the corner. In particular, the toy problem includes the following simplifications:

—The occluded scene consists of only emitters.
—Targets are on grid.
—The occluded scene is composed of discrete points.
—The light source is not on the same side as the camera.
—The reflectance of the wall is Lambertian.



(a) Light on Camera Side

(b) Light on Object Side

Fig. 4: If sufficient light is measured, comparable quality of reconstruction is observed whether the light source is on the (a) camera side or (b) object side (magenta arrow shows direction of light propagation). For both conditions, reconstructions are shown with beamforming and a sparse solver, CoSaMP. Recovery is governed by two primary noise effects: (1) photon shot noise and (2) the inherent noise introduced from the reconstruction. When a system is reconstruction-noise limited, reconstruction quality is governed almost exclusively by mutual coherence. (not-shown) If the source is not very bright, then signal-attenuating effects, such as the extra indirect bounces in (a) wrt. (b), start to limit noise more than the reconstruction-noise. Refer to the supplement for details on the noise model used for this simulation.

Emitters to reflectors Recall that the goal of the toy problem in Figure 2a was to localize an active light source around the corner. Of course this is not a realistic scenario: it is unlikely that the hidden object is an active light source. Figure 3a illustrates a more common scenario of localizing point reflector(s) are around the corner. To solve this problem the position of the light source must be known *a priori*. Then the path to each possible voxel location is known and a valid dictionary for the space can be constructed. We must mention that the reflectors do not have to be Lambertian. More precisely, the directionality of the object is equivalent to having a directional array, which actually facilitates recovery (see Section 5).

On-grid to off-grid To this point we have assumed that point targets, e.g., a point reflector or point source, lies on a search voxel. In realistic scenarios points are not guaranteed to lie on-grid and "off-grid" localization must be performed. Fortunately, the source localization community has developed powerful tools to address this very scenario. Approaches are in the style of iterative multiresolution methods which upsample the dictionary [Malioutov et al. 2005] or Continuous Basis Pursuit which interpolates the dictionary [Ekanadham et al. 2011].

Points to surfaces As illustrated in Figure 3b, the object of interest is usually a continuous surface. In this case, the object can be modelled as many closely spaced point emitters. In this case, the recovered surface would be a convolution of the surface with the beampattern of the single point source. This would allow us to recover a blurred version of the occluded surface, where the degree of blur depends on the width of the beampattern of a single point source.

Illumination position The goal is to build a camera that can look around corners without any gadgets in the line-of-sight. Therefore it is desirable to have the source on the same side as the camera. Achieving this turns out to be an engineering challenge as opposed to a theoretical one. Following from Figure 3a, as long as the position of the light source is known—whether it is next to the camera or not—a dictionary can be constructed for the space. The engineering challenges are:

—Saturation from an area source. ToF cameras are designed to illuminate an area, and thus if an area source is aimed at a wall the direct reflections will saturate the sensor.

—Very little light comes back: light has to bounce off the wall twice and the object once before returning to the camera.

A solution to the first problem is to use a collimated beam, either from a laser or by blocking the area source. A solution to the latter problem is to use a more powerful light source than what is stock on time of flight cameras. The simulation shown in Figure 4 verifies that the VSA model holds when the illumination is on the same side as the occluder. Refer to the figure caption for details.

Shininess of the wall and reflectors An interesting link between computer graphics and array signal processing exists between the bidirectional reflectance distribution function (BRDF) and antenna directionality. In the signal processing community, when sensors show directional preference (as opposed to being omnidirectional), resolving targets within the aperture becomes much easier. Indeed, in the corners problem the directionality of the reflectance is critical. Consider two opposing cases: (i) the wall is a mirror and thus the BRDF has strong directional preference, and (ii) the wall is purely diffuse corresponding to a constant BRDF. This confirms the physical intuition where we expect high resolvability of targets with specular BRDFs (e.g. mirrors) and low resolvability with diffuse BRDFs. We must also mention that the directionality of the VSA (determined by the BRDF of the wall) is dual to the directionality of the reflectors. Practitioners should note that to generate $\mathbf{D}$, an estimate of the BRDF should be obtained for optimal results. Sections 5 and 6 probe further into the reflectivity of the wall (i.e. shininess).

4.4 Imaging through Diffusers

To this point we have proposed the VSA model and shown its utility for looking around corners. We now show that it can also be used to image through diffuse media. Specifically, we consider a *transmissive* toy problem: localization of a source through a diffuser, shown in Figure 5. Here, the key idea is that the scattered paths are deflected and thus have a fraction longer time duration to some of the sensors.

Note that the problem is almost identical to the looking around corners problem in Figure 2a. In the corners problem, the virtual

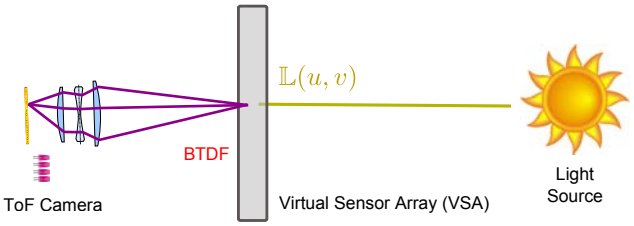


Fig. 5: The virtual sensor array model generalizes to handle scattering diffusers as well. In this context the virtual sensors are located on the visible surface of the diffuser.

sensor array was the wall itself. In the diffuser problem the virtual sensor array is the visible surface of the diffuser. Similar parallels exist for light transport analysis. Since the corners problem is using reflected measurements, the BRDF controls the directionality of sensors. In contrast, since the diffuser problem uses transmission measurements, sensor directionality is determined by the bidirectional transmittance distribution function (BTDF). Recall that the trivial case for the corners problem was a mirror, which has a peaked BRDF. The analogous trivial case for the diffuser problems occurs for a clear object, which has a peaked BTDF. We can therefore address the corresponding questions for the transmissive system, such as "how opaque or clear does the diffuser have to be"?

In crux, Equations 1 to 18 all apply to the case of imaging through diffusers.

5. ANALYSIS OF RECOVERABILITY

In this section we provide numerical guarantees to quantify when the VSA model can recover the occluded image and when it can't. To provide guarantees, parallel work in holography [Rivenson et al. 2013] proposes the use of mutual coherence as a key metric. Using our notation, we write the mutual coherence $\mu(\mathbf{D})$ as:

$$\mu(\mathbf{D}) = \max \mathbf{G}_{ij}, \quad \mathbf{G} = \left|\mathbf{D}^H \mathbf{D}\right|, \quad i \neq j, \tag{19}$$

where $\|\mathbf{D_n}\|_2 = 1$ for $n = 1, \ldots, N$. Intuitively, the mutual coherence computes the similarity between the columns of $\mathbf{D}$. For robust recovery, it is important to reduce the coherence, which is achieved through the choice of physical parameters. In the following paragraph we show that the specularity of the wall has an inverse relationship to the FWHM, and since all our functions are Gaussian, the relationship holds for mutual coherence, providing us a bound on target resolution. For more details on using mutual coherence to provide reconstruction guarantees we refer to the reader to [Rivenson et al. 2013] and our supplementary material.

Recovery guarantees for specular surfaces: When the wall is non-Lambertian, the virtual sensors are no longer omnidirectional. To model the directionality of virtual sensors, we first define the *beampattern* as a row of the Gramian matrix $\mathbf{G}$.[1] Like in optics, the FWHM of the beampattern specifies how far apart two targets must be to resolve both of them. As one can imagine, this has been explored in array signal processing. For instance, in [Van Trees 2004] a derivation is provided for the FWHM of omnidirectional sensors, which takes the form of the familiar Rayleigh limit:

$$\mathrm{FWHM}^{\angle} = \arcsin\left(\lambda / D\right), \tag{20}$$

where D is the diameter of the sensor array (in meters) and $\mathrm{FWHM}^{\angle}$ is the angular resolution (in radians). This equation provides the resolution to which targets can be resolved. A 300 MHz camera has a λ of approximately 1 meter and typically our virtual sensor array is about $D = 1$ meter wide. Then, for omnidirectional sensors, the resolution to which one can distinguish targets is approximately 1 meter, which is poor. Fortunately, if the virtual sensors were directional (e.g. if the wall is shiny), then the resolution limit improves.

We now derive the FWHM for a directional sensor system (such as a specular wall). Let $\gamma^{\angle}$ denote the FWHM of the directional response function of an individual virtual sensor, with units in radians. In the corners problem, the directional response function is the BRDF of the wall. Therefore, we use simplifications from computer graphics [Ramamoorthi and Hanrahan 2001; Han et al. 2007] to approximate the specular lobe of the BRDF by a Gaussian. From Chapter 3, of [Van Trees 2004], the FWHM for a directional sensor system is a composite of the omnidirectional FWHM (Equation 20) with the FWHM of the individual sensor response ($\gamma^{\angle}$). Following this recipe, we obtain the FWHM for the directional system:

$$\mathrm{FWHM}^{\angle} = \arcsin\left(\left(\lambda\gamma^{\angle}\right) / \left(\lambda + D\gamma^{\angle}\right)\right). \tag{21}$$

As a sanity check, it can be verified that if the sensor is omnidirectional, then Equation 21 simplifies to Equation 20.[2]. One can also verify that a low value of $\gamma^{\angle}$, i.e., a specular BRDF, corresponds to a narrower FWHM for the system. We arrived at Equation 21 using angular quantities for FWHM, but in this paper, we are also interested in spatial resolution of two targets (i.e. how many meters apart do they have to be). The relation between angular resolution and spatial resolution is written as

$$\mathrm{FWHM}^{\ell} = d\,\mathrm{FWHM}^{\angle}, \tag{22}$$

where d is the depth of the object from the array (in meters). We use the superscripts to denote the units of a scalar variable: ℓ for length (meters), $\angle$ for angular quantities in radians, and $^{\circ}$ to denote angular quantities in degrees.

General recovery guarantees based on rank and span constraints: General guarantees can be obtained using rank and span constraints. For example, $\mathbf{D}$ needs to have sufficient linear independence to uniquely recover sources. This can be expressed as a *rank-constraint*, where $\mathrm{rank}(\mathbf{D}) - 2$ encodes an upper bound on the dimensionality of the convex hull of targets (cf. pages 81-97 of [Gower 1985]). A complementary, *span constraint* characterizes appropriate arrangements of virtual sensors (e.g., the wall geometry) that avoid degenerate solutions. Using the general frameworks of rank, span, and mutual coherence, the supplemental material explores how other parameters—beyond specularity—influence reconstruction. This includes gridding, modulation frequency, aperture size, non-planar walls, and choice of reconstruction algorithm.

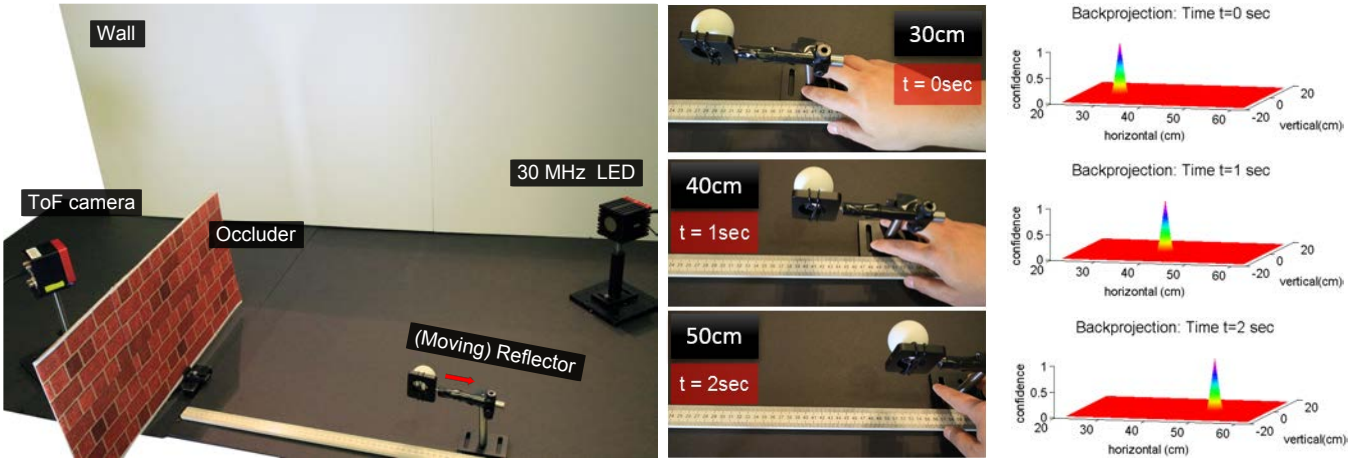
6. RESULTS

For all experiments, the time of flight camera used is the Mesa Swissranger SR4050 lock-in module. It can be purchased directly from MESA Imaging [3] in Zurich, Switzerland. This time of flight camera has a decoupled light source and operates at a modulation frequency of 30 MHz.

[1] Equivalently, beampattern is a column of $\mathbf{G}$ since $\mathbf{G}$ is Hermitian.

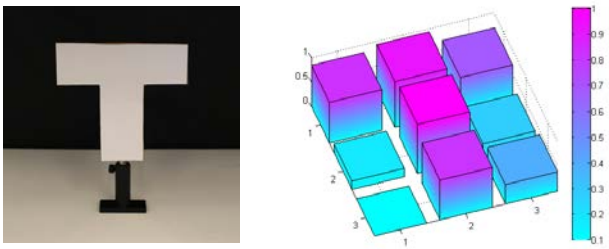
[2] Specifically, $\lim_{\gamma^{\angle} \to \infty} \left(\lambda\gamma^{\angle}\right) / \left(\lambda + D\gamma^{\angle}\right) = \lambda / D$

[3] `http://www.mesa-imaging.ch/`



(a) Moving object around the corner (b) Video of object (c) Object tracking

Fig. 6: Localizing a moving target in real-time. (a) The ping pong ball is outside the line of sight of the camera and moves horizontally. The camera measures backscattered reflections from the wall. (b) A conventional video of the object in motion. Here, three frames of the video are shown at times 0, 1 and 2 seconds where the ball position is 30, 40, and 50 cm. (c) Our technique is able to find the location of the target in real time. Please see supplementary video.



Occluded Scene Reconstruction

Fig. 7: Imaging around the corner. (left) We use the same scene from Figure 6a but replace the moving ping pong ball with a "T" shaped object. (right) Using pseudoinverse beamforming, we are able to recover the hidden image. The height represents confidence in a given voxel.

For the corners setup, white ping pong balls were chosen as reflected point targets as they are close to pure Lambertian objects. The ping pong balls are placed at a depth of 1 meter away from the wall; this allows us to easily convert from angular coordinates (degrees) to spatial coordinates (centimeters). In this section, diffuse posterboard material is used for the wall. The wall dimensions are approximately 2m by 1m and a patch of pixels that span 50cm by 50cm is used to form the observation vector. The main setup is shown in Figure 6a. Because the light source is an area source it cannot be placed on the same side as the camera; this would lead to saturation.

6.1 Qualitative Results

Real-time occluded imaging As illustrated in Figure 6, a single point reflector is occluded from the camera's line-of-sight. The reflector is placed in motion and can be localized in real-time using the backprojection algorithm. The supplement includes a video of this demonstration.

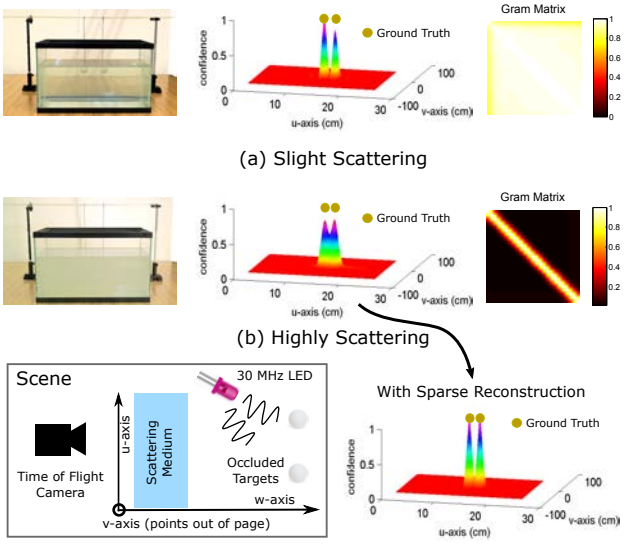


Fig. 8: Extending the problem to handle localization through diffusers. The scene is shown at the bottom left, where the ping pong ball and camera are on opposite sides of a bulk diffuser, i.e., milky water in a tank. In (a) only a few drops of milk are added. In (b) the water is very milky. Note that the balls are more sharply localized in the highly scattering scenario when sparse priors are used.

Imaging around the corner In Figure 7 we replace the moving ping pong ball from the real-time result with a small, "T shaped object" that is outside the line of sight.[4] By using pseudoinverse beamforming we are able to recover the image shown on the right. The wall is a posterboard, which has a BRDF similar to that of

[4]The size of the object is 20 by 20 centimeters

paper. To keep the framework scene-independent, priors are not placed on the reconstruction—implementing total variation or edge constraints would improve the reconstruction for some scenes.

Imaging through scattering media In Section 4.4, the VSA model has been analytically shown to generalize to the diffuser problem. Here, an experimental result is demonstrated in Figure 8 for imaging through scattering media. We place a tank of milky water in between the camera and two ping pong balls. For this problem the BTDF is the dual of BRDF, which allows us to use the same computation that was used in the corner experiments.

First we add relatively little milk and show that it is possible to localize the ball (Figure 8a). Then we add a much greater quantity of milk and show it is still possible to localize the ball (Figure 8b). Of course, in the latter case the coherence is much greater, and therefore the two peaks are not distinctly separated. To distinctly separate the peaks, a scene-dependent prior, such as sparsity can be used (Figure 8b). Imaging through scattering media is such a well established area that we must emphasize that our results are very preliminary results that need to be compared against other methods (e.g. structured light, phase conjugation, etc.). However, these experiments are sufficient to show the generality of the VSA model.

6.2 Quantitative Assessment

In this section we will perform real and simulated experiments to quantify the conditions for successful recovery.

6.2.1 *Quantitative Physical Experiments.*

Directionality of the virtual sensor array Creating a camera that can look around corners requires an understanding of the material properties of the virtual sensor array. To form the wall for the corners problem we collect four materials in increasing order of specularity: (i) posterboard; (ii) photo paper; (iii) metal; and (iv) a mirror. Figure 11 illustrates the measured directionality of the first three materials (we assume the mirror has a delta function for directionality).[5] In Table II, we list quantitative reflectance parameters, where ρ_s and α measure the specular intensity and surface roughness, as defined in the Ward BRDF model [Ward 1992].[6] We draw specific attention to γ^ℓ and the FWHM^ℓ in centimeters, which factors into the bounds we derived in Equation 22. We must also mention that the BRDF of the wall needs to be known to generate the forward operator $\mathbf{D}$.

Resolving multiple point sources Our end goal is to image around the corner; therefore, a critical performance metric is how close two point reflectors can be localized (without relying on sparsity assumptions). As this is a material dependent property, Figure 12 illustrates localization of two point sources for different materials. For the posterboard (Figure 12a) using pseudoinverse backprojection we are able to localize the two point sources that are 10 centimeters apart. We then plot the beampattern, which is one row of the Gram matrix $\mathbf{G}$. Note how, because the beampattern is very wide, the Gram matrix is very coherent. Figures 12b, 12c and 12d show results for the photo paper, metal, and mirrored objects. Observe that the beampattern narrows as the material changes from the posterboard to the mirror. In particular, for the mirror, the beampattern is a Dirac and the mutual coherence reaches the minimum value of 0.

Define the minimum resolvable distance as the minimum separation between two ping pong balls that can be detected. In our experimental results we found that for the posterboard it was 10 cm, for the photo paper 3 cm, and for the metal 2 cm.

A practical implication of this result is that even when using the relatively diffuse posterboard it is possible to obtain an image of the object around the corner (if the objects are large enough). For a fixed reflectance, using our theory, it is clear that increasing the modulation frequency scales linearly with expected resolution.

What is angular response like in the wild? If the virtual sensor array was omnidirectional, recovery of $\vec{x}$ is challenging and—in the context of phased array processing—not possible with today's time of flight cameras. If the Lambertian assumption can be relaxed then the technique would be more readily applicable today. We use the Mitsubishi Electric Research Labs (MERL) BRDF database [Matusik et al. 2003] to evaluate the directionality of various real materials. We use the fitted Ward parameters in [Ngan et al. 2005] to calculate FWHM° and FWHM^ℓ in the same manner as Table II. The results of all 100 BRDFs are shown in Figure 13. From this we can conclude that:

—Highly specular materials (like metals) have a very small FWHM^ℓ and thus high localization (about 2cm).

—Materials with a dominant specular lobe (relative to the diffuse lobe) have a reasonably small FWHM^ℓ and acceptable localization (about 5cm).

It is interesting to note that 38 of the 100 materials in the database could produce a beamforming resolution of smaller than 10cm and 29 materials can make it smaller than 5cm. We can therefore expect beamforming to be feasible for many real-world materials, and, from our empirical study, the resolution could improve by an order of magnitude using more sophisticated solvers.

6.2.2 *Quantitative Simulated Experiments.*

All simulated results are performed on the array shown in Figure 10. Sensors are spaced along the u-axis and two targets exist at $(0.2, 0.5)$ and $(0.8, 0.5)$ meters. The width of the array, i.e., aperture size, is fixed to $D = 1$ meter.

Coherence and physical parameters Earlier we were able to analytically connect mutual coherence to physical parameters of the array system. We focus our results on three such claims:

—γ vs coherence;

—$f_{\mathcal{M}}$ vs coherence; and

—grid spacing vs coherence.

Simply looking at the form of Equation 21 it is evident that γ has a nonlinear relationship to coherence. Our simulated results are consistent with the derived form (Figure 10b). In particular, note the critical region of the curve where $\gamma^\angle$ is small and Equation 21 approximately reduces to a linear form.

Superresolution via sparsity A great deal of research in signal processing centers around techniques for solving linear inverse problems. Here, we probe this idea by using a simulated array where the expected FWHM of the beampattern is approximately 2 meters. Since simple beamforming cannot resolve targets that are spaced closer than 2 meters apart, the interesting question that follows, is whether using more sophisticated solvers will allow for better resolution.

[5]The Ward lobe is rendered with the BRDF Explorer (http://www.disneyanimation.com/technology/brdf.html).

[6]The directionality of the sensor is determined by the BRDF of the surface. Details on measuring and fiting the directionality are in the supplement.

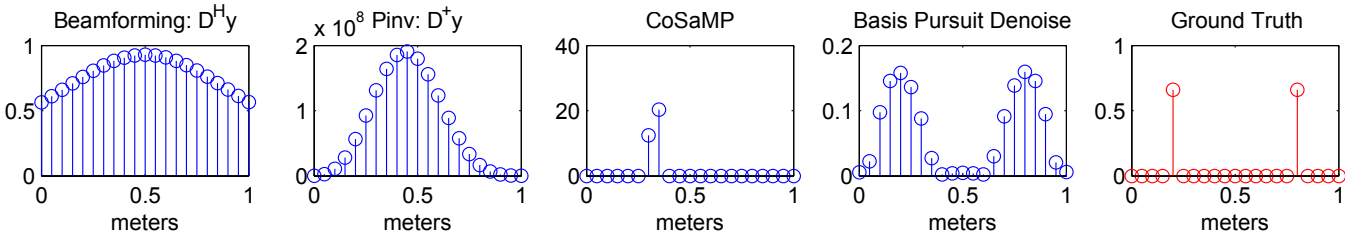


Fig. 9: Sparse priors allow superresolution of reflectors, and hence, potentially higher resolution images of the occluded scene. Ground truth reflectors are at 20cm and 80cm, which is closer than the Rayleigh limit for this scene (Equation 21). Here, Basis Pursuit Denoising (the convex relaxation) is the only technique that resolves target positions even though they are within the Rayleigh limit. Note the varying orders of magnitude of the y-axis due to sensitivity of some of these techniques to additive noise (e.g., pseudoinverse) or coherence (e.g. CoSaMP).

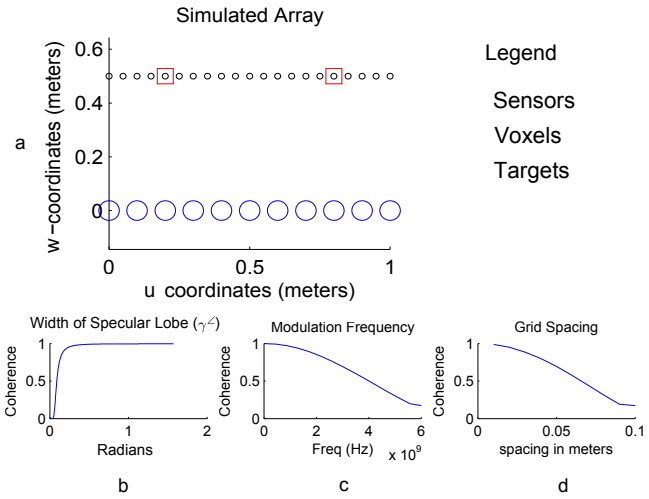


Fig. 10: (a) The simulated array used. (b-d) Impact of physical parameters on coherence.

Table I. : Different solvers for reconstruction and their objective functions.

Backprojection Solvers	Objective Formulation
Classic Beamforming	$\hat{\vec{x}} = \mathbf{D}^H \vec{y}$
Pseudoinverse	$\hat{\vec{x}} = \arg\min_{\vec{x}} \|\mathbf{D}\vec{x} - \vec{y}\|_2^2$
CoSaMP	$\hat{\vec{x}} = \arg\min_{\vec{x}} \|x\|_0 \; s.t. \; \vec{y} = \mathbf{D}\vec{x}$
Basis Pursuit Denoising	$\hat{\vec{x}} = \arg\min_{\vec{x}} \|x\|_1 \; s.t. \; \|\mathbf{D}\vec{x} - \vec{y}\|_2^2 \leq \epsilon$

Figure 9 compares the four techniques shown in Table I. Both CoSaMP and Basis Pursuit Denoising are solvers that enforce sparsity on $\vec{x}$. However, the former is a greedy algorithm while the latter is a convex relaxation. For a well defined optimization program, a convex relaxation is guaranteed to find the correct solution. No such guarantees exist for greedy methods, and therefore the latter are considered superior for recovery.[7] Our results are consistent with this intuition. As illustrated in Figure 9 we observe that both beamforming and the pseudoinverse are unable to resolve the targets, CoSaMP converges to a poor solution, and Basis Pursuit Denoising is able to superresolve the targets.

Table II. : Quantitative parameters extracted from Figure 11 and Equation 21. ρ_s and α represent specular intensity and surface roughness from the Ward BRDF model. γ° and γ^ℓ represent the FWHM of the angular response of virtual sensor, in radians and cm. Finally, FWHM$^\circ$ and FWHM$^\ell$, correspond to the minimum width of the beampattern in radians and cm.

Material	ρ_s	α	γ°	γ^ℓ	FWHM$^\circ$	FWHM$^\ell$
Posterboard	238	0.17	32.87	0.59	35.0	0.61
Photo Paper	2	0.01	2.02	0.03	1.72	0.03
Metal	1	0.002	0.93	0.01	0.57	0.01

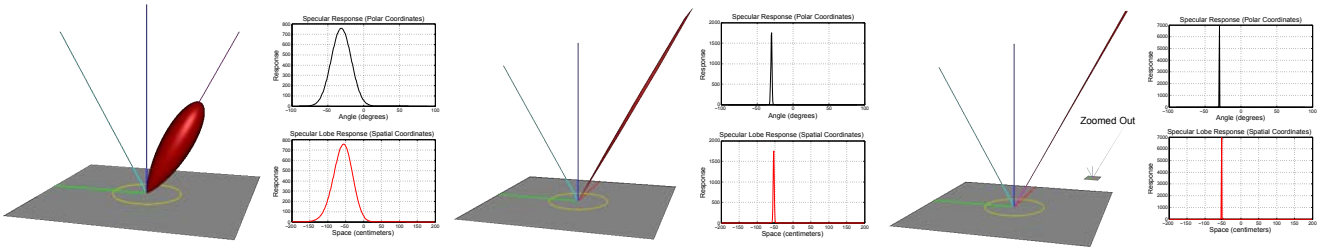
7. DISCUSSION

Do we approach the bounds? A natural question is whether the bounds we have proposed are meaningful in practice. We will restrict ourselves to the bound provided in Equation 21 as this is the most general bound: it is invariant to any model assumptions on $\mathbf{D}$ or $\vec{x}$. Another way to describe Equation 21 is that it provides a lower bound for the width of the beampattern. In the last column of Table II we list the computed bound based upon the acquired directionality of the materials. The width of the experimental beampatterns collected for different materials (Figure 12) approaches, but does not violate the bound. Any discrepancy from the bound is due to experimental error either in the measurements of $\gamma^\angle$ used to calculate the bound or in obtaining the beampattern.[8] In practice, the bound is perhaps useful when comparing materials with distinct properties. For example, the slack in the bound for the photo paper is about 2 cm, but the difference in beampattern width between the photo paper and posterboard is about 60 cm.

We must also mention that the bound provided in Equation 21 guarantees success when using the most basic solver (i.e. $\hat{\vec{x}} = \mathbf{D}^H \vec{y}$). Empirically we observe that we can often obtain a resolution close to an order of magnitude better using a stronger solver, such as the pseudoinverse. For example, in Figure 12a although the beampattern is approximately 1 meter wide, we are able to resolve targets 10 cm apart. There are many ways to derive sharper guar-

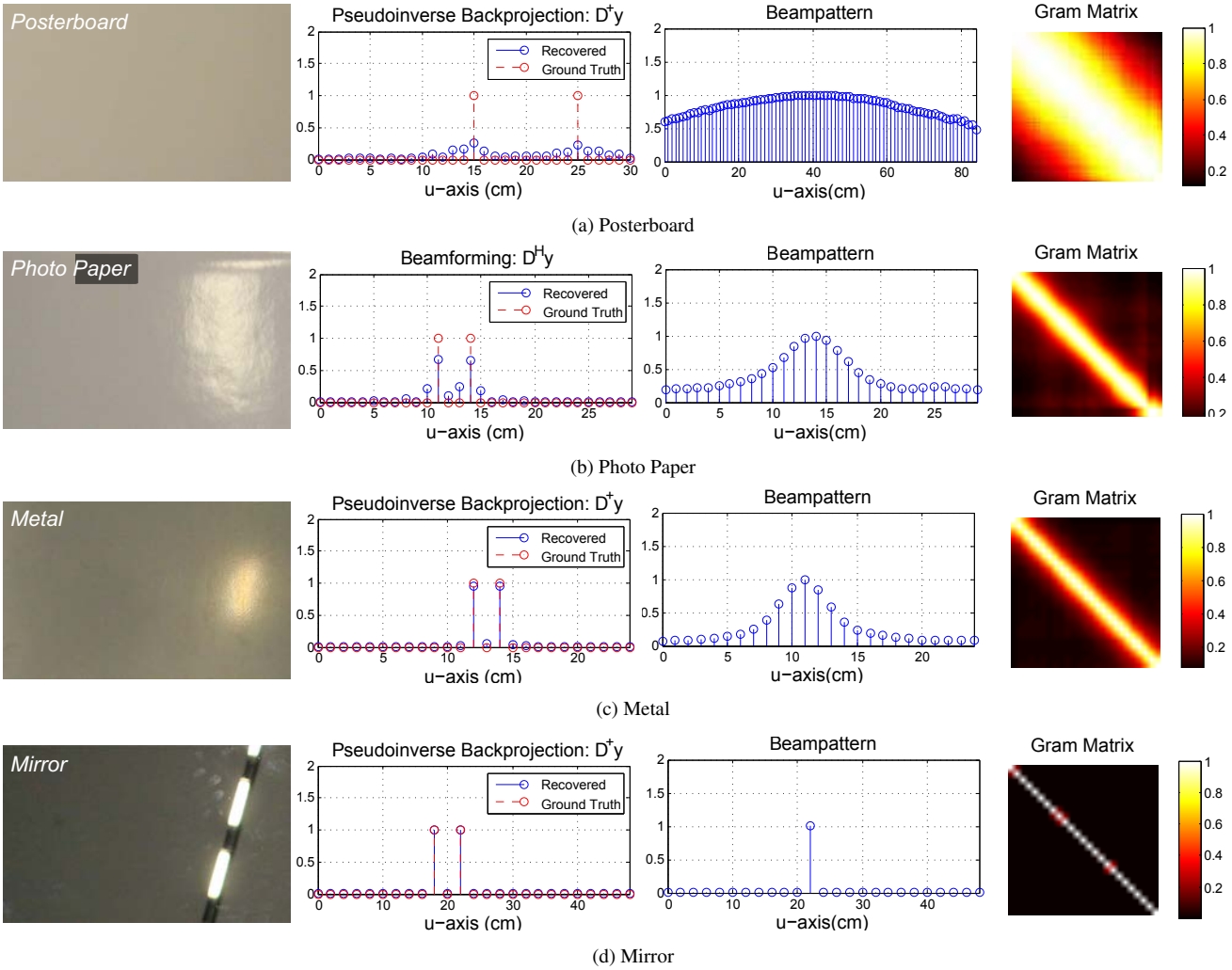
[7] Greedy methods are not without merit; they are simpler and well suited for model-based approaches.

[8] Errors inherent to ToF 3-D depth estimation can disrupt calibration between the VSA and camera.



(a) Posterboard (b) Photo Paper (c) Metal

Fig. 11: Rendered directionality for three different walls from Figure 12 along with 1-D plots of the directional response in both degrees (black) and centimeters (red). Only the specular lobe is shown (the diffuse lobe is omitted). Quantitative parameters derived from these plots can be found in Table II.



(a) Posterboard

(b) Photo Paper

(c) Metal

(d) Mirror

Fig. 12: How close can we resolve two point sources around the corner without using any prior assumptions? Consistent with the text, it depends on the material properties of the wall. Across each row the images represent: (i) a photograph of the wall, (ii) psuedoinverse backprojection ($\mathbf{D}^{\dagger}\vec{y}$), (iii) the beampattern, and (iv) the matrix $\mathbf{G}$. Note that as expected, for specular objects, $\mathbf{G}$ is sharply diagonal and algebraically incoherent, while for more diffuse objects the Gram matrix is more coherent.

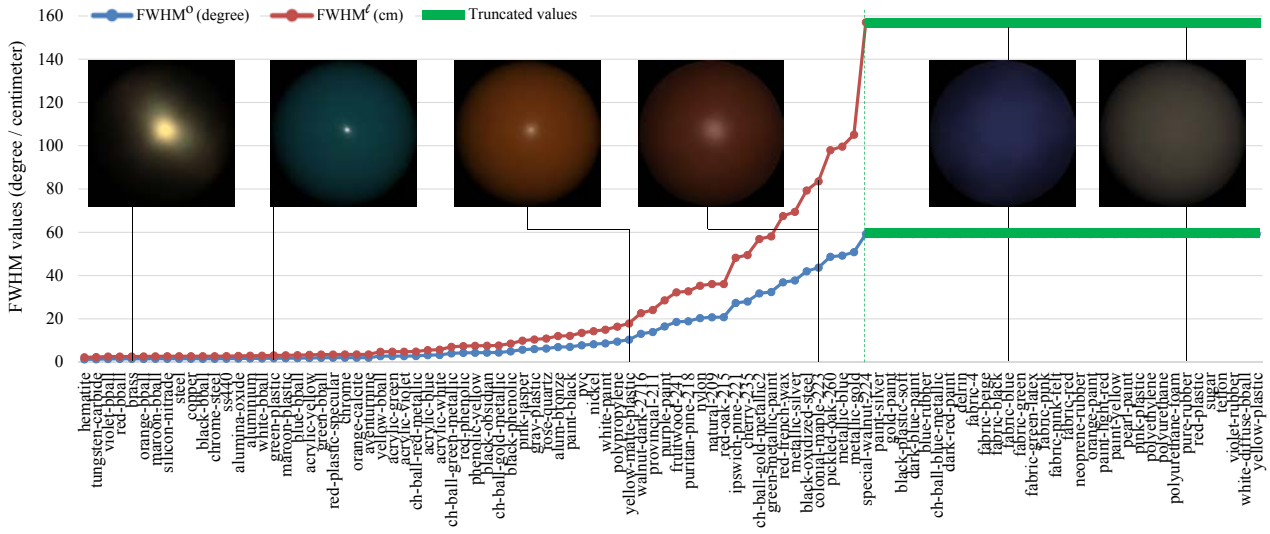


Fig. 13: Relating shininess with resolution. FWHM^{ℓ} and FWHM° calculated from all 100 materials in the MERL BRDF database. The material names are sorted according to the FWHM values in ascending order. Rendered spheres of selected materials are also displayed. Green lines indicate the regime where resolution becomes ill-defined due to the physical nature of wave equations. Equivalently, it is the region for which Equation 21 yields complex values.

antees on recovery based on sparsity, priors, rank constraints, and more; however, such guarantees require specific model assumptions and in-depth calculations. Moreover, the current guarantees, for example, in model based sparse approximation are still not fully understood within the signal processing community.

What are current limitations of time of flight technology? Assessing the limits of current time of flight cameras has been of significant interest in the last few years. Today's time of flight cameras can be counted on to have a modulation frequency of at least 30 MHz, or $\lambda = 10$ meters. First we will consider a scenario where the wall is omnidirectional. The bound in Equation 20 is ill-defined since $\lambda/D > 1$.[9]

We now consider a second case where we assume the wall is a posterboard, i.e., $\gamma = 0.59$ meters (see Table II). Then, from Equation 21 the lower bound on resolution between two targets is 55 cm if simple beamforming was used as the reconstruction technique. Fortunately the modulation frequency continues to increase with each product cycle. Fusing computation with newer technology will facilitate much more capable systems for occluded imaging.

Comparisons to existing solutions for looking-around-corners: As described in Sections 1 and 2, there are two solutions for looking around corners that have been proposed [Velten et al. 2012; Heide et al. 2014]. Both of these solutions utilize insights from "'transient imaging"', which in the former is optically complicated, and in the latter computationally involved. Then, additional computation is performed on the transient imaging data. This means that both techniques cannot capture scenes in real-time. In this paper, we see two comparative benefits of our simpler framework: (i) we are able to show a real-time scene capture and (ii) we are able to provide bounds on recovery.

Future work This paper has proposed a unifying framework for occluded imaging via off the shelf, unmodified time of flight cameras. The next step is to apply the theory to a heavily customized ToF camera.

Although we have made the connection between concepts such as time of flight and scattering, this paper only shows that the virtual sensor array model is valid for localization through scattering. Future work would go into depth into what is a classic problem in optics and would also provide comparisons to known techniques, such as structured light and phase conjugation.

Conclusion In summary, we reveal the link between looking around corners and phased array processing. We have proposed the first model for occluded imaging with standard time of flight cameras. Such cameras are increasing in popularity, both for their applications to 3-D imaging as well as novelty in computational photography applications.

Looking around the corner is a complex problem with many variables. We conclude that today's time of flight cameras are able to image around the corner with low spatial resolution by exploiting the property that standard walls are not purely Lambertian. We hope that this paper is a step toward having commodity cameras that can look around corners.

Acknowledgments

We thank the anonymous reviewers for their detailed comments, and for pointing out the mathematical connection to compressive holography. We thank Ayush Bhandari for technical comments. Achuta Kadambi was funded by the Charles S. Draper Doctoral Fellowship and Boxin Shi was partially supported the Singapore MOE Academic Research Fund MOE2013-T2-1-159 and the

[9] It is ill-defined numerically because the argument for the $arcsin$ is greater than one. It is also physically ill defined due to the wavelength being greater than the aperture size, resulting in an isotropic beampattern.

SUTD Digital Manufacturing and Design (DManD) Centre which is supported by the Singapore National Research Foundation.

REFERENCES

ADIB, F., KABELAC, Z., KATABI, D., AND MILLER, R. C. 2014. 3d tracking via body radio reflections. *Usenix NSDI 2014*.

BHANDARI, A., KADAMBI, A., WHYTE, R., BARSI, C., FEIGIN, M., DORRINGTON, A., AND RASKAR, R. 2014. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. *Optics letters 39,* 6, 1705–1708.

BOUFOUNOS, P. T., SMARAGDIS, P., AND RAJ, B. 2011. Joint sparsity models for wideband array processing. In *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics.

CARTER, G. 1981. Time delay estimation for passive sonar signal processing. *Acoustics, Speech and Signal Processing, IEEE Transactions on 29,* 3, 463–470.

CEVHER, V., DUARTE, M., BARANIUK, R. G., ET AL. 2008. Distributed target localization via spatial sparsity. In *European Signal Processing Conference (EUSIPCO)*.

CHEN, S. S., DONOHO, D. L., AND SAUNDERS, M. A. 2001. Atomic decomposition by basis pursuit. *SIAM review 43,* 1, 129–159.

EKANADHAM, C., TRANCHINA, D., AND SIMONCELLI, E. P. 2011. Recovery of sparse translation-invariant signals with continuous basis pursuit. *Signal Processing, IEEE Transactions on 59,* 10, 4735–4744.

ELAD, M. 2010. *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer.

GOWER, J. C. 1985. Properties of euclidean and non-euclidean distance matrices. *Linear Algebra and its Applications 67*, 81–97.

GUPTA, M., NAYAR, S. K., HULLIN, M. B., AND MARTIN, J. 2014. Phasor imaging: A generalization of correlation-based time-of-flight imaging.

HAN, C., SUN, B., RAMAMOORTHI, R., AND GRINSPUN, E. 2007. Frequency domain normal map filtering. In *ACM Transactions on Graphics (TOG)*. Vol. 26. ACM, 28.

HEGDE, C., INDYK, P., AND SCHMIDT, L. 2014. Approximation-tolerant model-based compressive sensing. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 1544–1561.

HEIDE, F., HULLIN, M. B., GREGSON, J., AND HEIDRICH, W. 2013. Low-budget transient imaging using photonic mixer devices. *ACM SIGGRAPH 2013*.

HEIDE, F., XIAO, L., HEIDRICH, W., AND HULLIN, M. B. 2014. Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

KADAMBI, A., WHYTE, R., BHANDARI, A., STREETER, L., BARSI, C., DORRINGTON, A., AND RASKAR, R. 2013. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics (TOG) 32,* 6, 167.

LIN, J., LIU, Y., HULLIN, M. B., AND DAI, Q. 2014. Fourier analysis on transient imaging with a multifrequency time-of-flight camera. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 3230–3237.

MALIOUTOV, D., ÇETIN, M., AND WILLSKY, A. S. 2005. A sparse signal reconstruction perspective for source localization with sensor arrays. *Signal Processing, IEEE Transactions on 53,* 8, 3010–3022.

MALLAT, S. G. AND ZHANG, Z. 1993. Matching pursuits with time-frequency dictionaries. *Signal Processing, IEEE Transactions on 41,* 12, 3397–3415.

MASIA, B. 2014. Computational imaging: Combining optics, computation and perception. *ACM Trans. on Graphics 32*, 6.

MATUSIK, W., PFISTER, H., BRAND, M., AND MCMILLAN, L. 2003. A data-driven reflectance model. *ACM Transactions on Graphics 22,* 3 (July), 759–769.

MAYNARD, J. D., WILLIAMS, E. G., AND LEE, Y. 1985. Nearfield acoustic holography: I. theory of generalized holography and the development of nah. *The Journal of the Acoustical Society of America 78*, 1395.

NAIK, N., ZHAO, S., VELTEN, A., RASKAR, R., AND BALA, K. 2011. Single view reflectance capture using multiplexed scattering and time-of-flight imaging. In *ACM Transactions on Graphics (TOG)*. Vol. 30. ACM, 171.

NGAN, A., DURAND, F., AND MATUSIK, W. 2005. Experimental analysis of brdf models. In *Proceedings of the Sixteenth Eurographics conference on Rendering Techniques*. Eurographics Association, 117–126.

O'TOOLE, M., HEIDE, F., XIAO, L., HULLIN, M. B., HEIDRICH, W., AND KUTULAKOS, K. N. 2014. Temporal frequency probing for 5d transient analysis of global light transport. *ACM Transactions on Graphics (TOG) 33,* 4, 87.

RAMAMOORTHI, R. AND HANRAHAN, P. 2001. A signal-processing framework for inverse rendering. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, 117–128.

RESHETOUSKI, I. AND IHRKE, I. 2013. *Mirrors in Computer Graphics, Computer Vision and Time-of-Flight Imaging*. Springer.

RESHETOUSKI, I., MANAKOV, A., SEIDEL, H.-P., AND IHRKE, I. 2011. Three-dimensional kaleidoscopic imaging. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 353–360.

RIVENSON, Y., STERN, A., AND ROSEN, J. 2013. Reconstruction guarantees for compressive tomographic holography. *Optics letters 38,* 14, 2509–2511.

SEN, S. AND NEHORAI, A. 2011. Adaptive ofdm radar for target detection in multipath scenarios. *Signal Processing, IEEE Transactions on 59,* 1, 78–90.

SUME, A., GUSTAFSSON, M., HERBERTHSON, M., JANIS, A., NILSSON, S., RAHM, J., AND ORBOM, A. 2011. Radar detection of moving targets behind corners. *Geoscience and Remote Sensing, IEEE Transactions on 49,* 6, 2259–2267.

TIBSHIRANI, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

VAN TREES, H. L. 2004. *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. John Wiley & Sons.

VELTEN, A., WILLWACHER, T., GUPTA, O., VEERARAGHAVAN, A., BAWENDI, M. G., AND RASKAR, R. 2012. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications 3*, 745.

VELTEN, A., WU, D., JARABO, A., MASIA, B., BARSI, C., JOSHI, C., LAWSON, E., BAWENDI, M., GUTIERREZ, D., AND RASKAR, R. 2013. Femto-photography: Capturing and visualizing the propagation of light. *ACM Transactions on Graphics (TOG) 32,* 4, 44.

VELTEN, A., WU, D., JARABO, A., MASIA, B., BARSI, C., LAWSON, E., JOSHI, C., GUTIERREZ, D., BAWENDI, M. G., AND RASKAR, R. 2012. Relativistic ultrafast rendering using time-of-flight imaging. In *ACM SIGGRAPH 2012 Talks*. ACM, 41.

WARD, G. 1992. Measuring and modeling anisotropic reflection. *Computer Graphics 26,* 2, 265–272.

WU, D., WETZSTEIN, G., BARSI, C., WILLWACHER, T., DAI, Q., AND RASKAR, R. 2014. Ultra-fast lensless computational imaging through 5d frequency analysis of time-resolved light transport. *International Journal of Computer Vision 110,* 2, 128–140.