Phasor Imaging: A Generalization of Correlation-Based Time-of-Flight Imaging

MOHIT GUPTA and SHREE K. NAYAR
Columbia University
and
MATTHIAS B. HULLIN and JAIME MARTIN
University of Bonn

In correlation-based time-of-flight (C-ToF) imaging systems, light sources with temporally varying intensities illuminate the scene. Due to global illumination, the temporally varying radiance received at the sensor is a combination of light received along multiple paths. Recovering scene properties (e.g., scene depths) from the received radiance requires separating these contributions, which is challenging due to the complexity of global illumination and the additional temporal dimension of the radiance.

We propose phasor imaging, a framework for performing fast inverse light transport analysis using C-ToF sensors. Phasor imaging is based on the idea that by representing light transport quantities as phasors and light transport events as phasor transformations, light transport analysis can be simplified in the temporal frequency domain. We study the effect of temporal illumination frequencies on light transport, and show that for a broad range of scenes, global radiance (multi-path interference) vanishes for frequencies higher than a scene-dependent threshold. We use this observation for developing two novel scene recovery techniques. First, we present Micro ToF imaging, a ToF based shape recovery technique that is robust to errors due to global illumination. Second, we present a technique for separating the direct and global components of radiance. Both techniques require capturing as few as $3-4$ images and minimal computations. We demonstrate the validity of the presented techniques via simulations and experiments performed with our hardware prototype.

Categories and Subject Descriptors: I.4.8 [**Image Processing and Computer Vision**]: Scene Analysis—*Shape*

Additional Key Words and Phrases: Computational photography, time-of-flight imaging, light transport, interreflections, scattering, correlation imaging, transient imaging, phasors, multi-path interference, global illumination, radiosity

1. INTRODUCTION

Correlation-based time-of-flight (C-ToF) imaging systems consist of temporally modulated light sources and sensors with temporally modulated exposures. The brightness measured by the sensor is the correlation between the temporally varying radiance incident on the sensor and the exposure function. This is illustrated in Figure 1. Because of their ability to measure scene depths with high precision and speed, these systems are fast becoming the method of choice for depth sensing in a wide range of applications. Several low cost and compact C-ToF systems are available as commodity devices, including the Microsoft Kinect and the Intel SoftKinectic sensors.

Global light transport in C-ToF imaging: Conventional C-ToF imaging systems assume that sensor pixels receive light only due to direct illumination of scene points from the source. However, due to global illumination, the sensor receives radiance along several paths, after multiple reflection/scattering events. Recovering scene properties (e.g., scene depths) from the received radiance requires separation of contributions from different paths. This is a difficult task due to the complexity of global illumination, and is made even more challenging because of the additional temporal dimension of the radiance.

Phasor representation of radiance: Our goal is to develop a compact model for generalized C-ToF imaging, i.e., a model of C-ToF imaging that accounts for full global illumination. To this end, we make the following observations. If the scene is illuminated with sinusoids of a given temporal frequency, the radiance at any point and direction is always a sinusoid of the same frequency, irrespective of the scene. Since all the sinusoids are of the same frequency, the frequency can be factored out and the radiance at any point and direction can be represented by a *single complex number, or phasor*. With phasor representation, light transport at each temporal frequency can be analyzed separately, thus significantly reducing the complexity. Also, since phasor radiance corresponds to a particular modulation frequency, it can be captured by a C-ToF sensor operating at that frequency with only two measurements.

Phasor imaging: Based on these observations, we propose phasor imaging, a framework for analyzing light transport in C-ToF imaging, using phasor representations of radiance and light transport events. In particular, we analyze the effect of temporal frequency on light transport and show that for a broad range of scenes, global radiance decreases with increasing frequency, eventually vanishing beyond a threshold frequency. Using this property, we develop two scene analysis techniques:

—Transport-robust shape recovery,

—Fast separation of direct and global radiance.

Transport-robust shape recovery: An important problem faced by C-ToF based depth recovery systems is the errors caused by global illumination (multi path interference). These errors are systematic and scene dependent, and can be orders of magnitude larger than the random errors occurring due to system noise. This problem has received a lot of attention recently, with a variety of techniques having been proposed to mitigate the errors [Godbaz et al. 2008; Dorrington et al. 2011; Kirmani et al. 2013]. These approaches assume global illumination to be a discrete sum of contributions along a small number $(2-3)$ of light paths. For general scenes, pixels may receive light along several, potentially infinite, light paths. Consequently, these approaches are limited to scenes with only high frequency light transport (e.g., specular interreflections).

We present Micro ToF imaging, a technique for recovering shape that is robust to errors due to global illumination, and is applicable to scenes with a broad range of light transport effects. It is based on using high temporal frequencies at which global illumination vanishes, and hence does not introduce errors in the phase of the received radiance. Although using high frequencies achieves robustness to global illumination, the unambiguous depth range is small due to phase ambiguities. Micro ToF uses two (or more) high frequencies and standard phase unwrapping techniques to disambiguate the high frequency phases, thus achieving robustness to global illumination as well as a large depth range, with as few as *four measurements*.

Fast separation of direct and global radiance: We present a technique for separation of direct and global radiance components. One way to separate the two components (using temporal light modulation) is to measure the full transient image of the scene [Heide et al. 2013; Velten et al. 2013]. These approaches, although theoretically valid, require prohibitively large acquisition times. We show that it is possible to perform the separation by capturing only *three measurements at a single high temporal frequency*. The proposed technique can be thought of as the temporal counterpart to the technique presented by Nayar *et al.* [2006] which performed separation using high spatial frequency illumination.

Limitations and implications: We have demonstrated our scene-analysis techniques by building a hardware prototype based on a low cost C-ToF sensor. Currently, these sensors have a limited range of modulation frequencies, which restricts the application of our techniques to relatively large scale scenes. However, this is not a theoretical limitation. As device frequencies increase [Akbulut et al. 2001; Wu et al. 2010; Buxbaum et al. 2002; Schwarte 2004; Busck and Heiselberg 2004], it will be possible to apply our techniques on smaller scale scenes. Due to their generality, near real time acquisition and computation times, we believe that the proposed techniques will be readily integrated into future C-ToF imaging systems for performing a variety of scene analysis tasks.

2. RELATED WORK

Impulse Time-of-Flight Imaging: Impulse ToF imaging techniques measure the *temporal impulse* response of the scene by illuminating it with very short (pico/nanosecond) laser pulses and recording the reflected light at high temporal resolution. Impulse ToF imaging was the basis of one of the first ToF range imaging systems [Koechner 1968]. While earlier systems assumed only a single direct reflection of light from the scene, recent techniques (called transient imaging) have used the impulse ToF principle to measure and analyze both direct and indirect light transport for capturing images around a corner [Kirmani et al. 2009], measuring 3D shape [Velten et al. 2012] and motion of objects [Pandharkar et al. 2011] around the corner, performing separation of light transport components [Wu et al. 2012], measuring BRDF [Naik et al. 2011], capturing images with a lens-less sensor [Wu et al. 2012], and capturing the propagation of light [Velten et al. 2013].

Correlation-Based Time-of-Flight Imaging: These techniques were introduced as a low cost alternative to impulse ToF imaging. The scene is illuminated with continuous temporally modulated light (e.g., with sinusoids), and the sensor measures the temporal correlation of the incident light with a reference function [Schwarte et al. 1997; Lange and Seitz 2001]. Scene depths are computed by measuring the relative phase-shift between the incident light and the emitted light. While there has been research on optimizing the modulation waveform [Payne et al. 2010; Ferriere et al. 2008; Ai et al. 2011] for achieving high precision and for handling interference among multiple ToF cameras [Buttgen et al. 2007], it is mostly assumed that sensor receives only direct reflection from the scene. Our work seeks to generalize correlation-based ToF imaging to include a variety of indirect (global) light transport effects.

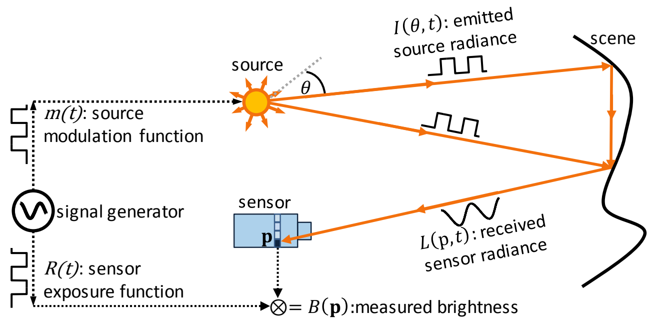


Fig. 1. **Correlation-based ToF image formation model.** The scene is illuminated by a temporally modulated light source, with radiant intensity $I(\theta, t)$ along direction θ. The sensor's exposure is also temporally modulated during the integration time according to the function $R(t)$. The brightness $B(\mathbf{p})$ measured at a sensor pixel $\mathbf{p}$ is the correlation of the incoming radiance $L(\mathbf{p}, t)$ and the exposure function $R(t)$.

Multi Path Interference in Time-of-Flight Imaging: Recently, there has been a lot of research towards mitigating the effect of global illumination (multi-path) in ToF cameras. In general, this is a difficult problem because global illumination depends on scene structure, which is unknown at time of capture. There have been several attempts at solving the problem for special cases, such as piecewise planar Lambertian scenes [Fuchs 2010; Fuchs et al. 2013; Jimenez et al. 2012] or temporally sparse signals [Godbaz et al. 2008; 2009; 2012; Dorrington et al. 2011; Jimenez et al. 2012; Kadambi et al. 2013; Kirmani et al. 2013]. These approaches do not generalize to all forms of light transport. Moreover, they often require capturing a large number of images and/or computationally intensive optimization-based reconstruction algorithms. The approach of Freedman *et al.* [2014] considers compressible signals (instead of sparse signals), and can handle limited amount of diffuse interreflections. However, since the signal is assumed to be compressible, it is limited to scenes where the dominant amount of global illumination is due to only a small number of light paths. The approach presented in this paper requires taking as few as four measurements and only a few linear operations, and is applicable to scenes with a wide range of light transport effects.

Light Transport Analysis Using Spatial Light Modulation: In the last few years, several techniques performing light transport analysis using spatially modulated light have been presented. This includes methods for inverting light transport [Steven M. Seitz 2005], performing global-transport-robust shape recovery [Gupta et al. 2009; Gupta and Nayar 2012; Gupta et al. 2013; Couture et al. 2014] and separating or selectively enhancing light transport components [Nayar et al. 2006; Reddy et al. 2012; OToole et al. 2012].

Recently, O'Toole *et al.* [2014] have used a combination of spatial and temporal light modulation for performing a variety of scene analysis tasks. While their techniques rely on *high-spatial-frequency* light modulation, our focus is on studying the behavior of light transport as a function of temporal frequencies. We develop techniques that use only *high-temporal-frequency* light modulation, and achieve near real-time capture rates.

3. BACKGROUND AND IMAGING MODEL

A C-ToF imaging system consists of a temporally modulated light source, and a sensor whose exposure can be temporally modulated during integration time. This is illustrated in Figure 1. Let the source be modulated with a periodic function $m(t)$ (normalized to be between 0 and 1). Then, the radiant intensity $I(\theta, t)$ of the source in direction θ is given as:

$$I(\theta, t) = i(\theta)m(t). \tag{1}$$

The sensor exposure is temporally modulated according to the exposure function $R(t)$, which can be realized either by on-chip gain modulation (e.g., photonic mixer devices [Schwarte et al. 1997]) or by external optical shutters [Carnegie et al. 2011].

Let the radiance incident at a sensor pixel $\mathbf{p}$ be $L(\mathbf{p}, t)$. The brightness $B(\mathbf{p})$ measured at pixel $\mathbf{p}$ is given by the correlation between the incoming radiance and the exposure function:

$$B(\mathbf{p}) = \int_0^{\tau} R(t)L(\mathbf{p}, t)dt, \tag{2}$$

where τ is the total integration time.

Light transport equation for C-ToF imaging: Let $L_\theta(\mathbf{p}, t)$ be the radiance incident at pixel $\mathbf{p}$ due to light emitted from the source along direction θ. $L_\theta(\mathbf{p}, t)$ is given as:

$$L_\theta(\mathbf{p}, t) = \beta(\mathbf{p}, \theta) I\left(\theta, t - \frac{\Gamma(\mathbf{p}, \theta)}{c}\right), \tag{3}$$

where $\Gamma(\mathbf{p}, \theta)$ is the length of the path taken (through the scene) by the ray emitted in direction θ and arriving at $\mathbf{p}$. The constant c is the speed of light. $\beta(\mathbf{p}, \theta)$ is the light transport coefficient between direction θ and pixel $\mathbf{p}$; it is defined as the fraction of emitted intensity that reaches the sensor.

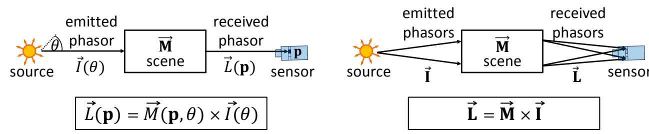
The total received radiance $L(\mathbf{p}, t)$ is the integral of contributions from the set of all outgoing directions Ω:

$$L(\mathbf{p}, t) = \int_\Omega L_\theta(\mathbf{p}, t) d\theta = \int_\Omega \beta(\mathbf{p}, \theta) I\left(\theta, t - \frac{\Gamma(\mathbf{p}, \theta)}{c}\right) d\theta\,. \tag{4}$$

This is the *light transport equation for C-ToF imaging*. It expresses the temporal radiance profiles received at a pixel in terms of the emitted radiance $I(t)$ and the scene properties (light transport coefficients and path lengths). Since it is scene dependent, in general, $L(\mathbf{p}, t)$ does not have a compact analytic form (as a function of t). $L(\mathbf{p}, t)$ is a combination of light coming along multiple paths, which cannot be easily separated and analyzed for recovering scene properties. Also, capturing the entire time profile requires long acquisition times.

A compact representation of radiance: $L(\mathbf{p}, t)$ is the impulse response of the scene, i.e., the temporal radiance received at the sensor if the scene is illuminated with an temporal impulse illumination. If, however, the scene is illuminated with sinusoidally varying illumination at a fixed frequency, the radiance at every point and every direction in space (including at the sensor) will also vary sinusoidally with the same frequency. This is because $L(\mathbf{p}, t)$ is an integral of shifted and scaled emitted radiance functions $I(t)$ (Eq. 4), and sinusoids are closed under scaling, shifting and integration. Since all the sinusoids are of the same frequency, we can factor the frequency out, and represent the radiance at any point $\mathbf{x}$ in space (including the sensor) along any direction θ by a single complex number, or phasor $\vec{L}(\mathbf{x}, \theta) = L(\mathbf{x}, \theta)e^{j\phi(\mathbf{x}, \theta)}$, where L is the amplitude and ϕ is the phase of the sinusoid [1]. $j = \sqrt{-1}$ is the complex square-root of unity. We call $\vec{L}$ the *phasor radiance*, short for phasor representation of radiance.

[1]Since light is non-negative, the sinusoidal modulation functions have a non-zero offset L_{DC}. The corresponding phasor representation is a 2-tuple: $[L_{DC}, \vec{L}_\omega]$, where L_{DC} is the DC and $\vec{L}_\omega$ is the oscillating component. For the intensity to be non-negative, $L_{DC} \geq |\vec{L}_\omega|$, where $|.|$ is the modulus operator that returns the magnitude of the complex number.



(a) Single emitted-received ray pair (b) All emitted-received ray pairs

Fig. 2. **Signal processing view of phasor light transport.** (a) Rays emitted from the source and received at the sensor are represented by single phasors. The scene transforms every emitted phasor into a received phasor. The transformation is linear (multiplication by the light transport coefficient for the emitted-received ray pair). (b) Light transport between all the emitted and received rays can be compactly represented as a matrix multiplication.

Phasor light transport: The scene can be considered as a system that transforms the phasor radiance emitted by the source (by modulating its phase and magnitude) into phasor radiance received by a sensor pixel. The transformation can be expressed as a multiplication of the emitted phasor by another phasor - the light transport coefficient between the emitted-received ray pair. This is illustrated in Figure 2 (a). As has been shown recently [O'Toole et al. 2014], the light transport between all the emitted and received rays can be represented as a matrix multiplication:

$$\vec{\mathbf{L}} = \vec{\mathbf{M}}\vec{\mathbf{I}}, \tag{5}$$

where $\vec{\mathbf{L}}$ is the array of phasor radiances received at sensor pixels and $\vec{\mathbf{I}}$ is the array of phasor radiances emitted by the source along different directions (we use bold upper-case letters to denote arrays and matrices). This is shown in Figure 2 (b). We call Eq. 5 the phasor light transport equation and $\vec{\mathbf{M}}$ the *phasor light transport matrix* of the scene [2]. Phasor and conventional light transport matrices are related as:

$$\vec{\mathbf{M}}(\mathbf{p}, \theta) = \mathbf{M}(\mathbf{p}, \theta)e^{-j\omega\frac{\Gamma(\mathbf{p}, \theta)}{c}}, \tag{6}$$

where $\mathbf{M}$ is the light transport matrix for conventional imaging and ω is the modulation frequency. Note that the phasor light transport matrix is a function of the modulation frequency ω. For DC component ($\omega = 0$), phasor transport matrix is the same as the conventional light transport matrix.

The phasor light transport equation expresses light transport in C-ToF imaging (Eq. 4) for a given modulation frequency as a linear, matrix multiplication. This simplifies light transport analysis in C-ToF imaging, especially the study of how light transport depends on the modulation frequencies. From a practical standpoint also, the phasor representation naturally lends itself to C-ToF imaging. This is because the phasor radiance received at every pixel has only two unknowns (phase and magnitude), which can be captured directly by C-ToF sensors operating at a single frequency with only two measurements. This forms the basis of the techniques presented in the paper, which require taking as few as four and three measurements for transport-robust shape recovery and direct-global separation, respectively.

4. PHASOR REPRESENTATION OF LIGHT TRANSPORT EVENTS

Light transport events can be categorized into three basic groups based on the phasor transformations that they induce, as illustrated

[2]The matrix representation assumes that the space of light rays has been discretized along spatial and angular dimensions.

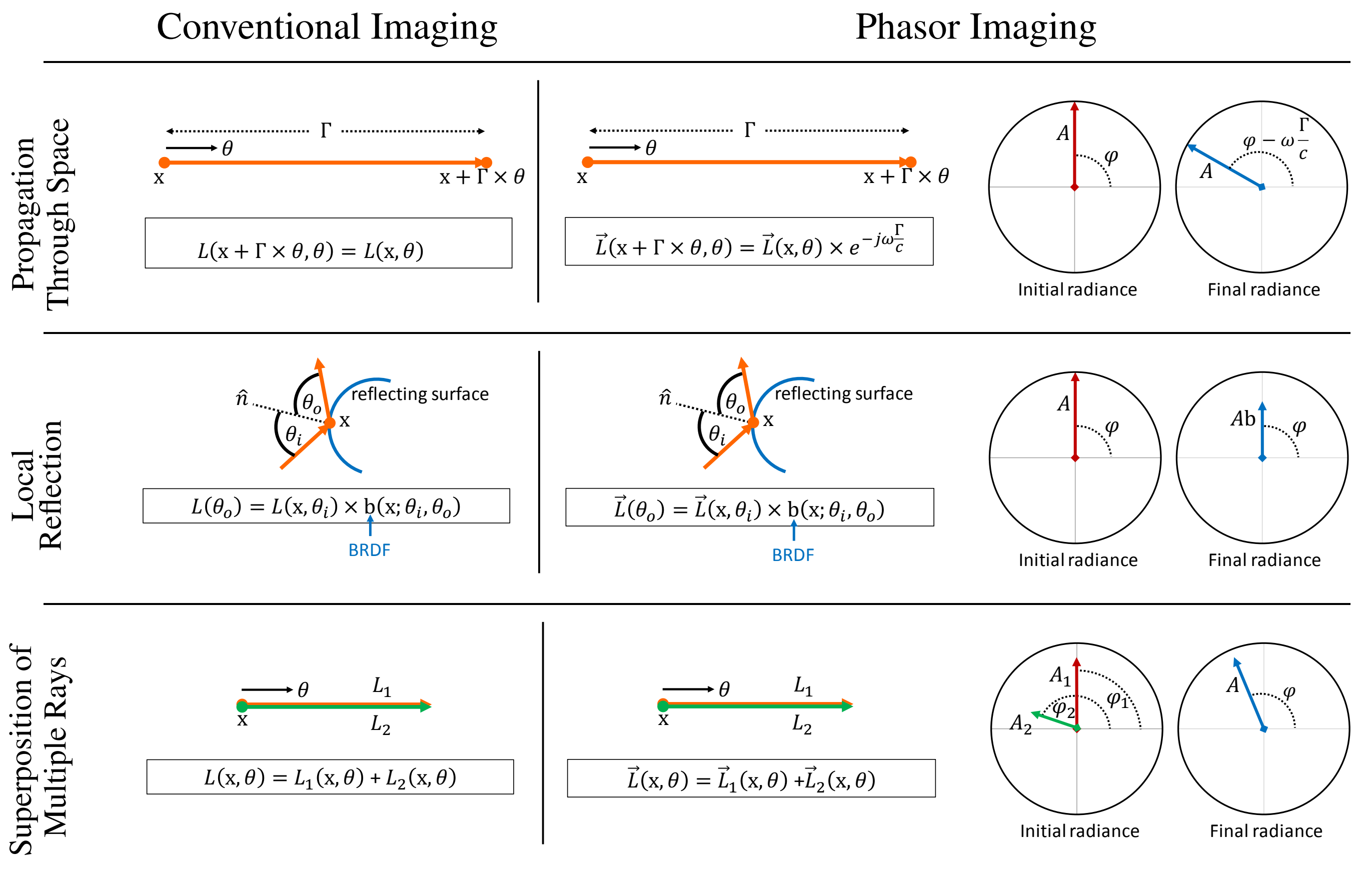


Fig. 3. **Phasor representation of light transport events.** Using phasors, all the light transport events can be represented by linear operations on complex numbers. (Top row) Propagation of a light ray through space changes the phase of the phasor radiance. The amount of change is proportional to both the distance traveled and the modulation frequency. This is represented by multiplication of the initial radiance with a phasor of unit amplitude. (Middle row) Local reflection and scattering events change only the amplitude of the radiance. These events are represented by multiplication with phasors having zero phase. This is similar to conventional imaging. (Bottom row) Multiple rays at the same point in space traveling in the same direction can be represented by a single ray whose radiance is the complex sum of the radiance of individual rays.

in Figure 3. First, events that change the phase of the radiance (propagation through space). Second, events that change only the magnitude of the radiance (local reflection and scattering). Third, the superposition event where multiple phasors are added to give a resultant phasor. In the following, we consider these individually.

Propagation Through Space: Propagation through free space changes the phase of the radiance, while the magnitude is conserved. Let $\vec{L}(\mathbf{x}, \theta)$ be the phasor radiance at a point $\mathbf{x}$ in space along the direction θ. Then, the radiance after propagating through a distance Γ is given as:

$$\vec{L}(\mathbf{x} + \Gamma\theta, \theta) = \vec{L}(\mathbf{x}, \theta) \times e^{-j\omega\frac{\Gamma}{c}}, \tag{7}$$

where ω is the modulation frequency. Propagation through participating media changes both the magnitude and the phase:

$$\vec{L}(\mathbf{x} + \Gamma\theta, \theta) = \vec{L}(\mathbf{x}, \theta) \times e^{-\left(\sigma\Gamma + j\omega\frac{\Gamma}{c}\right)}, \tag{8}$$

where σ is the medium's extinction coefficient. Note that *the amount of phase change* $\Delta\phi = \omega\frac{\Gamma}{c}$ *is proportional to both the modulation frequency* ω *and the travel distance* Γ.

Local Reflection and Scattering: Local reflection at a surface point changes only the magnitude of the radiance:

$$\vec{L}(\mathbf{x}, \theta_\mathbf{o}) = \vec{L}(\mathbf{x}, \theta_\mathbf{i}) \times \mathbf{b}(\mathbf{x}; \theta_\mathbf{i}, \theta_\mathbf{o}), \tag{9}$$

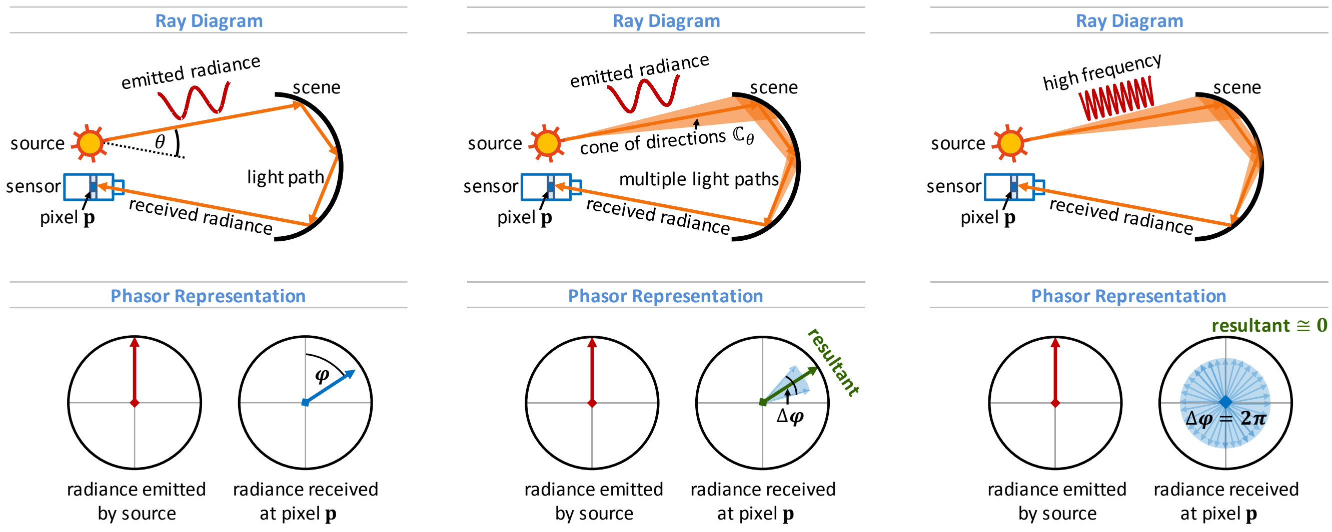
where $\mathbf{b}(\mathbf{x}; \theta_\mathbf{i}, \theta_\mathbf{o})$ is the BRDF term [3] at point $\mathbf{x}$ for incoming light direction $\theta_\mathbf{i}$ and outgoing light direction $\theta_\mathbf{o}$. Local scattering has the same effect as reflection, with the scattering term (product of scattering albedo and the scattering phase function) replacing the BRDF term.

Superposition of Multiple Rays: Multiple light rays traveling in the same direction through the same point can be represented as a single ray whose radiance is the phasor sum of individual radiances:

$$\vec{L}(\mathbf{x}, \theta) = \sum_i \vec{L}_i(\mathbf{x}, \theta), \tag{10}$$

where $\vec{L}_i(\mathbf{x}, \theta)$ are the individual radiances, and $\vec{L}(\mathbf{x}, \theta)$ is the total radiance. Due to phasor summation, the magnitude of the total radiance may be lesser than the sum of the individual magnitudes, i.e. $|\vec{L}(\mathbf{x}, \theta)| \leq \sum_i |\vec{L}_i(\mathbf{x}, \theta)|$. The resultant magnitude can be zero as well, even if all the initial radiances have non-zero magnitudes. This is different from conventional imaging where sum of non-zero radiances is strictly positive.

[3] The foreshortening effect is subsumed within the BRDF term.



(a) Global light transport for a single ray
(b) Global light transport for a local ray neighborhood
(c) Global light transport with high frequency illumination

Fig. 4. **Vanishing global light transport for high modulation frequency.** (a) A single indirect light path $\mathbb{P}(\mathbf{p},\theta)$ between outgoing direction θ and pixel $\mathbf{p}$. The phasor radiance received at the sensor along $\mathbb{P}(\mathbf{p},\theta)$ is given by rotating and attenuating the emitted phasor radiance. The angle of rotation ϕ is proportional to the length of $\mathbb{P}(\mathbf{p},\theta)$. (b) A set of indirect light paths $\mathbb{P}(\mathbf{p},\mathbb{C}_\theta)$ in a small neighborhood of $\mathbb{P}(\mathbf{p},\theta)$. All the paths end at $\mathbf{p}$. The phasor radiances along paths $\mathbb{P}(\mathbf{p},\mathbb{C}_\theta)$ can be assumed to have constant amplitudes and linearly varying phases, and thus form a circular sector in the phasor diagram. The angle of the sector $\Delta\phi$ is proportional to the modulation frequency ω. The total global radiance is the resultant of the individual phasors. (c) The light source emits high frequency illumination. The individual phasor radiances span the complete circle, and the resultant (total global radiance) is zero.

5. FREQUENCY DEPENDENCE OF PHASOR LIGHT TRANSPORT

Consider the phasor light transport equation (Eq. 5). We can decompose the incident sensor radiance as the sum of the direct component $\vec{\mathbf{L}}^d_\omega$ and the global component $\vec{\mathbf{L}}^g_\omega$, where the direct component is the light reaching the sensor after single reflection and the global component is the light reaching the sensor after multiple reflections (or scattering) events:

$$\vec{\mathbf{L}}_\omega = \vec{\mathbf{L}}^d_\omega + \vec{\mathbf{L}}^g_\omega = \vec{\mathbf{M}}^d_\omega \vec{\mathbf{I}}_\omega + \vec{\mathbf{M}}^g_\omega \vec{\mathbf{I}}_\omega \,. \tag{11}$$

$\vec{\mathbf{M}}^d_\omega$ and $\vec{\mathbf{M}}^g_\omega$ are the direct and global components of the light transport matrix $\vec{\mathbf{M}}_\omega$, respectively, for modulation frequency ω.

PROPOSITION 1. **Vanishing high-frequency global light transport:** *For a broad range of scenes, if the frequency ω is higher than a threshold ω_{thresh}, the global component vanishes:*

$$\vec{\mathbf{L}}^g_\omega = \vec{\mathbf{M}}^g_\omega \vec{\mathbf{I}}_\omega = 0 \;\; for \;\; \omega \geq \omega_{thresh} \,. \tag{12}$$

This is the key observation underlying our work. It is a consequence of the fact that typically, global radiance is temporally smooth, and can be assumed to be bandlimited. In the following, we provide an intuition behind the above observation by using phasor representations of light transport events. A frequency-domain proof is given in Section 5.1.

Intuition: Consider a light path $\mathbb{P}(\mathbf{p},\theta)$ involving multiple inter-reflections, starting at the light source in direction θ, and ending at a sensor pixel $\mathbf{p}$. An example light path is shown in Figure 4 (a). The radiance $\vec{L}_\theta(\mathbf{p})$ received at $\mathbf{p}$ along $\mathbb{P}(\mathbf{p},\theta)$ is given by:

$$\vec{L}_\theta(\mathbf{p}) = \vec{\mathbf{M}}(\mathbf{p},\theta)\vec{I}, \tag{13}$$

where $\vec{I}$ is the emitted radiance [4] and $\vec{\mathbf{M}}(\mathbf{p},\theta)$ is the light transport coefficient for the path $\mathbb{P}(\mathbf{p},\theta)$. Since $\mathbb{P}(\mathbf{p},\theta)$ involves propagation and reflection, $\vec{L}_\theta(\mathbf{p})$ is given by rotating and attenuating the emitted phasor radiance (Figure 4 (a)), as described in Section 4.

Next, consider the set of light paths $\mathbb{P}(\mathbf{p},\mathbb{C}_\theta)$ in a local neighborhood of $\mathbb{P}(\mathbf{p},\theta)$ that start in a cone of directions $\mathbb{C}_\theta$ around θ, and end at $\mathbf{p}$. This is illustrated in Figure 4 (b). The magnitudes of the light transport coefficients $|\vec{\mathbf{M}}(\mathbf{p},\theta)| = \mathbf{M}(\mathbf{p},\theta)$ can be assumed to be approximately constant in a small light path neighborhood. This assumption forms the basis of methods that use high spatial frequency illumination for separating light transport components [Nayar et al. 2006] and performing transport-robust shape recovery [Gu et al. 2011; Chen et al. 2008; Gupta and Nayar 2012; Couture et al. 2014]. The phases $\phi(\mathbf{p},\theta) = arg\left(\vec{\mathbf{M}}(\mathbf{p},\theta)\right)$ can be assumed to be linearly varying as a function of θ. This can be shown by considering the first order Taylor's expansion of the phases $\phi(\mathbf{p},\theta)$. See the supplementary technical report for a proof.

Thus, the individual received radiances $\vec{L}^g_\theta(\mathbf{p}) = \vec{\mathbf{M}}(\mathbf{p},\theta)\vec{I}$ have constant amplitudes and linearly varying phases, and sweep out a circle sector. From Eq. 7, the angle $\Delta\phi$ of the sector is:

$$\Delta\phi = \omega\frac{\Delta\Gamma(\mathbf{p},\theta)}{c}, \tag{14}$$

where $\Delta\Gamma(\mathbf{p},\theta)$ is the range of the lengths of paths $\mathbb{P}(\mathbf{p},\mathbb{C}_\theta)$. The total global radiance $\vec{L}^g_{\mathbb{C}_\theta}(\mathbf{p})$ is the resultant phasor of all the individual phasors (Eq. 10). Its magnitude is given by:

$$|\vec{L}^g_{\mathbb{C}_\theta}(\mathbf{p})| = 2Q\frac{\sin\left(\frac{\Delta\phi}{2}\right)}{\Delta\phi}, \tag{15}$$

[4] For ease of exposition, we assume an isotropic source, i.e., $\vec{I}(\theta) = \vec{I}$.

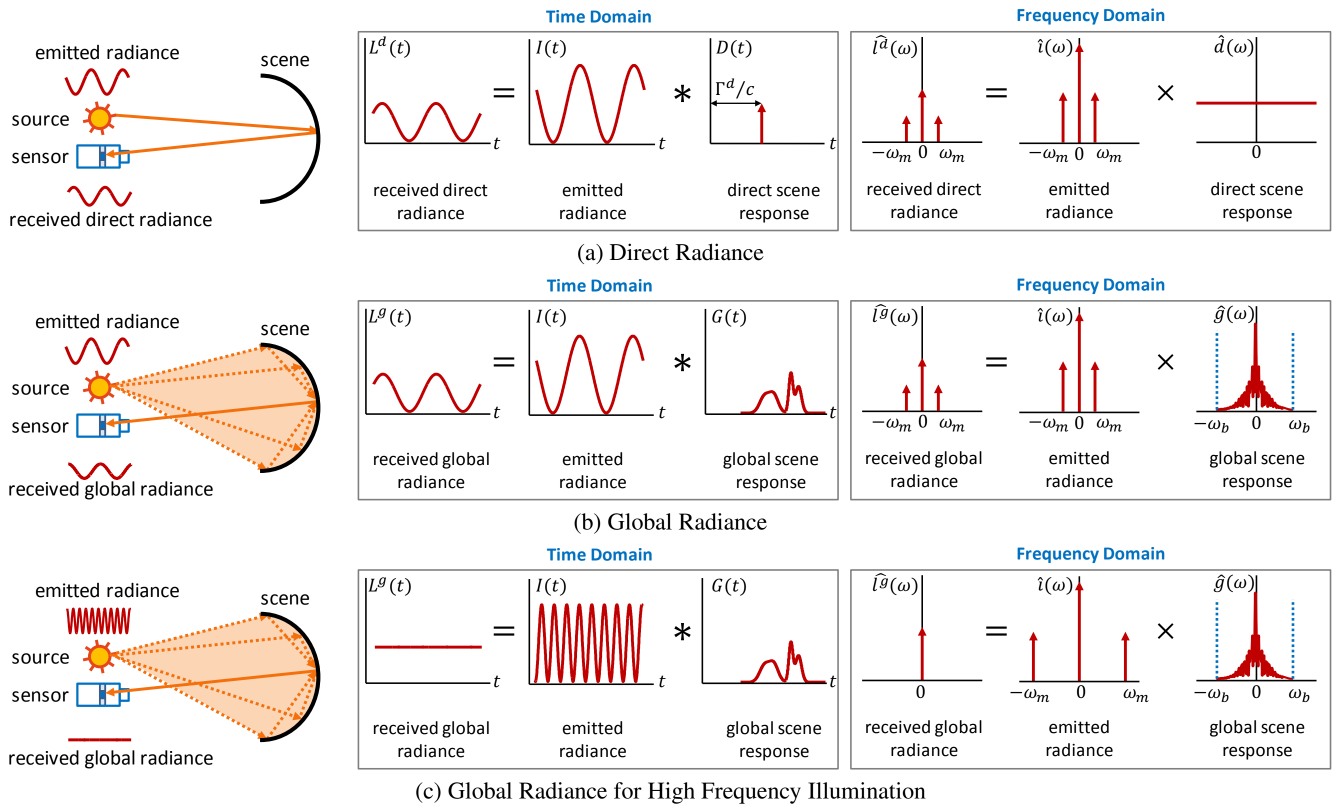


Fig. 5. **Frequency domain analysis of C-ToF light transport.** (a) In time-domain, the direct radiance received at the sensor is given by convolving the emitted signal and the direct scene response, which is a dirac delta function. The Fourier transform of a delta function is constant. Thus the direct radiance has all the frequencies that are present in the emitted radiance. (b) The global radiance is the convolution of the emitted radiance and the global scene response. For most real-world scenes, the global scene response is temporally smooth, and thus, bandlimited. If the bandlimit of the global scene response is ω_b, the global radiance is also bandlimited by ω_b. (c) If the emitted radiance is a sinusoid with frequency $\omega > \omega_b$, the global radiance contains only the DC component.

where Q is the sum of magnitudes of the individual phasors. The derivation is given in the supplementary technical report. $|\vec{L}_{\mathbb{C}_\theta}(\mathbf{p})|$ is a monotonically decreasing function of the sector angle $\Delta\phi$ for $0 \leq \Delta\phi \leq 2\pi$. Since $\Delta\phi$ is proportional to the modulation frequency ω (Eq. 14), as ω increases, $\Delta\phi$ increases, and the resultant magnitude decreases. If $\omega = \frac{2\pi c}{\Delta\Gamma(\mathbf{p},\theta)}$, $\Delta\phi = 2\pi$, and the magnitude of the global radiance $|\vec{L}^g_{\mathbb{C}_\theta}(\mathbf{p})| = 0$ [5].

5.1 Frequency Domain Proof Of Vanishing High-Frequency Global Transport

Let the temporally varying light intensity emitted from the source be given by $I(t)$. The direct radiance received at a pixel $\mathbf{p}$ is given by $L^d(t) = \alpha I(t - \phi)$, where α encapsulates the scene albedo and intensity fall-off. $\phi = \frac{\Gamma^d}{c}$ is the temporal shift due to travel of light and Γ^d is the length of the direct light path for pixel $\mathbf{p}$. We can write $L^d(t)$ as a convolution:

$$L^d(t) = I(t) * \alpha\, \delta\left(t - \frac{\Gamma}{c}\right), \tag{16}$$

where $\delta()$ is the dirac delta function. This is illustrated in Figure 5 (a). We define $D(t) = \alpha\delta(t - \frac{\Gamma}{c})$ as the direct scene response. $D(t)$ is the direct radiance received if the scene is illuminated with a temporal impulse function $\delta(t)$. Eq. 16 can be expressed in the frequency domain as:

$$\hat{l^d}(\omega) = \hat{i}(\omega) \times \hat{d}(\omega)\,, \tag{17}$$

where $\hat{l^d}(\omega)$, $\hat{i}(\omega)$ and $\hat{d}(\omega)$ are Fourier transforms of $L^d(t)$, $I(t)$ and $D(t)$, respectively. Since $D(t)$ is a dirac delta function, magnitude of $\hat{d}(\omega)$ is constant. This is illustrated in Figure 5 (a).

The global scene response $G(t)$ is defined as the global radiance received if the scene is illuminated with a temporal impulse $\delta(t)$. Similar to the direct component, the global component $L^g(t)$ is:

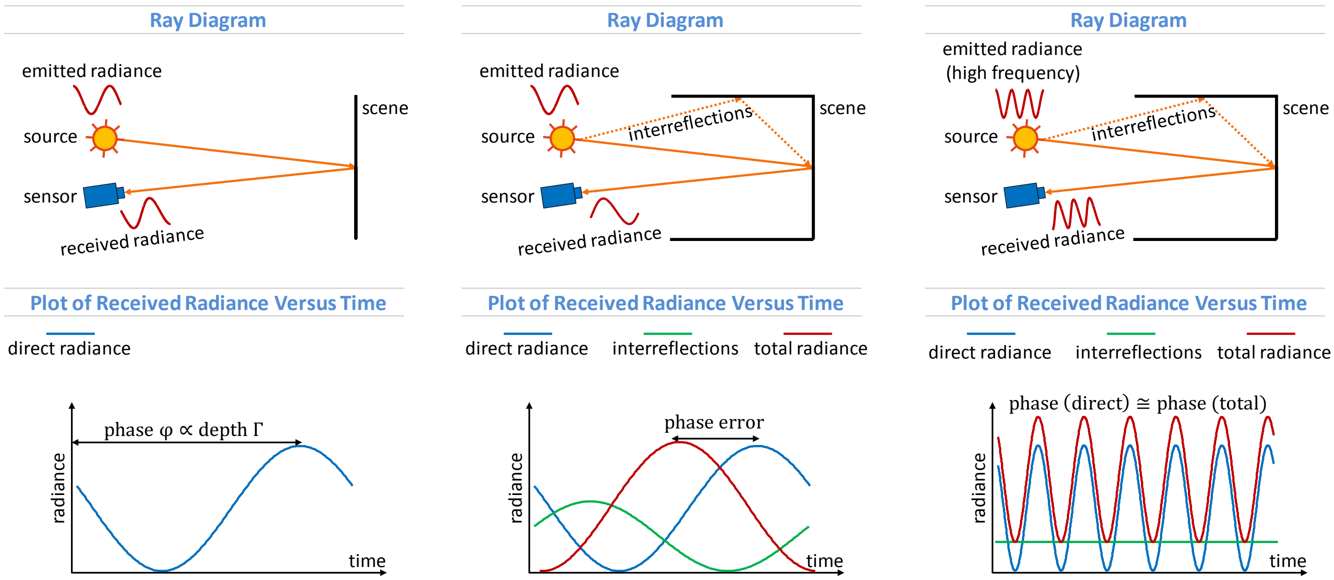
$$L^g(t) = I(t) * G(t). \tag{18}$$

In frequency domain, the above equation is expressed as:

$$\hat{l^g}(\omega) = \hat{i}(\omega) \times \hat{g}(\omega)\,, \tag{19}$$

where $\hat{l^g}(\omega)$ and $\hat{g}(\omega)$ are Fourier transforms of $L^g(t)$ and $G(t)$, respectively. This is illustrated in Figure 5 (b).

[5]Strictly speaking, $|\vec{L}^g_{\mathbb{C}_\theta}(\mathbf{p})| \approx 0$. This is because the assumptions (local constancy of light transport magnitudes and local linearity of light transport phases) hold approximately. As ω increases beyond $\frac{2\pi c}{\Delta\Gamma(\mathbf{p},\theta)}$, we can apply the above analysis in smaller light path neighborhoods (narrower cone $\mathbb{C}_\theta$), which improves the approximation.



(a) Direct radiance only (b) Direct and global radiance (c) Using high frequency illumination

Fig. 6. **Effect of global illumination on shape recovery using C-ToF imaging.** (a) Scene depths are computed using C-ToF imaging by measuring the phase of the received radiance relative to the emitted radiance. (b) Due to global radiance, a sensor pixel receives light along multiple light paths. These light paths have different lengths, and hence different phases as compared to the direct reflection path. The phase of the total radiance is different from the correct phase (direct radiance phase), resulting in incorrect depth. (c) For high frequency illumination, the phase of the total radiance is the same as that of the direct radiance, and can be used for accurate depth recovery.

For most real-world scenes, $G(t)$ is temporally smooth, and hence can be assumed to be bandlimited, i.e., there exists a frequency ω_b, called the *global transport bandlimit*, such that $\hat{g}(\omega) = 0 \; \forall \; \omega > \omega_b$. Thus, if the emitted radiance $I(t)$ is a sinusoid with modulation frequency ω larger than the global transport bandlimit ω_b, the oscillating component of the global radiance, $\vec{\mathbf{L}}^g_\omega$, is zero. Then, the global radiance has only a constant DC term (due to the DC term of the emitted radiance). This is illustrated in Figure 5 (c).

Other global illumination effects: While we have used interreflections for the analysis so far, the results and the proposed techniques are applicable to scenes with a broad range of global illumination effects such as subsurface scattering, volumetric scattering and diffusion. For each of them, the global scene response is typically smooth, and thus, bandlimited. By choosing a modulation frequency higher than the bandlimit, the global radiance can be made temporally constant.

5.2 How High Is The Frequency Bandlimit?

The global transport bandlimit ω_b depends on the scene geometry and material properties, as well as the global illumination effect. For volumetric scattering and diffuse interreflections, ω_b is typically low. On the other hand, for specular interreflections, ω_b is relatively high. Scene size is also a factor in determining ω_b. For large scenes, the indirect light paths have a large range of path lengths, resulting in a low bandlimit ω_b. For smaller scenes, indirect light paths have a smaller range of path lengths, and thus $G(t)$ has a higher bandlimit.

As a rule of thumb, ω_b is $1 - 10$ times $\frac{c}{\xi}$, where ξ is the geometric scale of the scene (in meters). For instance, for a room scale scene of size 3.0 meters, ω_b is approximately 100 MHz. - 1.0 GHz., depending on the material properties (higher bandlimit for more specular scenes). For table-top scenes of size approximately 50 centimeters, ω_b is approximately 600 MHz. to 6 GHz.

Arbitrary modulation functions: So far, we have considered sinusoidal modulation functions. In general, any modulation function can be decomposed into its Fourier components, and the presented analysis applied to each component separately. If the lowest frequency component of the function (except the DC) is higher than the global transport bandlimit, the global radiance at all the non-DC frequencies is zero. Thus, we get the following result:

RESULT 1. *If the lowest frequency component (except DC) of the modulation function is higher than the global transport bandlimit ω_b, the global radiance received at the sensor is temporally constant.*

For instance, if the emitted radiance $I(t)$ is a square wave with a period more than $\frac{2\pi}{\omega_b}$, all the frequency components of $I(t)$ are higher than ω_b. In this case, the global radiance received at the sensor contains only the DC component, and is temporally constant.

6. TRANSPORT ROBUST DEPTH RECOVERY

Effect of global light transport on depth recovery: C-ToF systems recover scene depths by measuring the phase $\phi = \omega \frac{2\Gamma}{c}$ of the radiance received at the sensor, where Γ is the scene depth. This is illustrated in Figure 6 (a). Depth is computed from the recovered phase as $\Gamma = \frac{c\phi}{2\omega}$. Due to global illumination effects such as interreflections (multi-path interference) and scattering, a sensor pixel may receive light along multiple light paths. These paths have different lengths, and hence light received along these paths have different phases as compared to the direct reflection path. Consequently, the phase of the total radiance (sum of direct and global components) is different from the correct phase, as shown in Figure 6 (b). The resulting depth errors are systematic, and can be orders of magnitude larger than the random errors due to noise.

6.1 Micro ToF Imaging

We now present our technique for mitigating depth errors due to global illumination. The basic idea is simple, and relies on the observation that global transport vanishes at high frequencies (Proposition 1). Let the scene be illuminated by a light source with intensity varying sinusoidally at frequency ω. In phasor notation, the intensity is given by the 2-tuple $[I_{DC}, \vec{I_\omega}]$, where I_{DC} is the DC component and $\vec{I_\omega}$ is the oscillating component. The direct radiance $L^d(\mathbf{p})$ received at a pixel $\mathbf{p}$ is given by the 2-tuple:

$$L^d(\mathbf{p}) = [D(\mathbf{p})\, I_{DC}, \vec{D}_\omega(\mathbf{p})\, \vec{I_\omega}]\,, \tag{20}$$

where $D(\mathbf{p})$ and $\vec{D}_\omega(\mathbf{p})$ are the DC and oscillating terms of the direct radiance for a light source with unit intensity. Similarly, the global radiance $L^g(\mathbf{p})$ is:

$$L^g(\mathbf{p}) \;=\; [G(\mathbf{p})\, I_{DC}, \vec{G}_\omega(\mathbf{p})\, \vec{I_\omega}]\,. \tag{21}$$

As shown in the previous section, if $\omega > \omega_b$, the oscillating term of the global radiance $\vec{G_\omega}(\mathbf{p}) = 0$. Then, the total radiance is:

$$L(\mathbf{p}) \left[(D(\mathbf{p}) + G(\mathbf{p}))\; I_{DC}, \vec{D}_\omega(\mathbf{p})\, \vec{I_\omega} \right]\,. \tag{22}$$

Since the global component of the radiance manifests only as a constant offset, it does not influence the phase. Thus, for high frequency illumination, *the phase of the total radiance is the same as that of the direct radiance*, and can be used for accurate depth recovery. This is shown in Figure 6 (c).

Wrapped phase problem and unambiguous depth range: The phase $\phi = \omega \frac{2\Gamma}{c}$ is computed by using inverse trigonometric functions (e.g., arctan, arcsine, arccosine) [Payne et al. 2010], which have a range of 2π. Consequently, the set of scene depths $\Gamma + n\frac{\pi c}{\omega}$ for any integer n will all have the same recovered phase, leading to depth ambiguities. This is called the wrapped phase problem. It limits the maximum depth range R_{max} in which scene depths can be measured unambiguously. R_{max} is inversely proportional to modulation frequency ω, and is given by $R_{max} = \frac{\pi c}{\omega}$ [Lange 2000]. For example, for $\omega = 2\pi \times 1500$ MHz. [6], R_{max} is only 10 centimeters. While it is possible to unwrap high-frequency phases using a low-frequency phase [Jongenelen et al. 2010], if there is global illumination, the low frequency phase is inaccurate. This causes unwrapping errors, resulting in erroneous shape.

This presents a tradeoff between achieving a large depth range, and robustness to global illumination. On one hand, higher modulation frequencies are robust to global illumination effects. On the other hand, using high frequencies result in depth ambiguities. How can we measure accurate scene depths in a large range when only high temporal frequencies are used?

Fortunately, it is possible to estimate a low frequency phase from multiple high frequency phases. This is a standard problem in interferometry [Gushov and Solodkin 1991; Takeda et al. 1997], structured light based triangulation [Gupta and Nayar 2012] and time-of-flight imaging [Jongenelen et al. 2010; Jongenelen et al. 2011].

[6] ω is the angular modulation frequency, which is 2π times the modulation frequency.

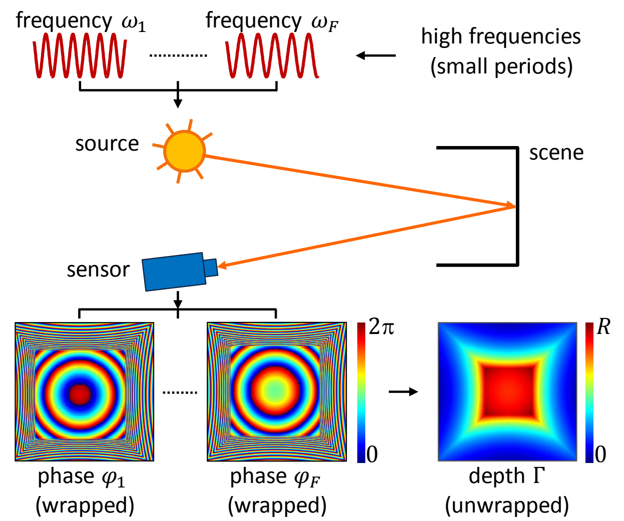


Fig. 7. **Micro ToF imaging.** The proposed Micro ToF imaging technique consists of illuminating the scene sequentially with multiple high frequency sinusoids, and computing phases corresponding to each of them. Theoretically, two high frequencies are sufficient. If all the frequencies are sufficiently high, global illumination does not introduce errors in the phases. The individual phases have depth ambiguities. Unambiguous depth is recovered by unwrapping the phases, which can be done either analytically or by building a look-up table.

There are both numerical and analytical solutions available which can be implemented efficiently.

Algorithm: The proposed technique involves illuminating the scene sequentially with multiple high frequency sinusoids, and computing phases corresponding to each of them. Since all the emitted signals have *micro* (small) periods, the technique is called *Micro ToF Imaging*. This is illustrated in Figure 7. Let the set of frequencies used be $\Omega = [\omega_1, \ldots, \omega_F]$. For each frequency ω_f, $1 \leq f \leq F$, the sensor measures the correlation of the received radiance with the sensor exposure function $R_f(t)$, which is also a sinusoid of the same frequency as the emitted light. The phasor representation of $R_f(t)$ is $\vec{R_f} = R_f e^{-j\psi_f}$, where R_f and ψ_f are the magnitude and phase of the exposure function. The measured brightness $B_f(\psi)$ is a sinusoid as a function of ψ_f (see Appendix A for derivation):

$$B_f(\psi_f) = O_f + A_f \cos(\phi_f - \psi_f)\,. \tag{23}$$

$B_f(\psi_f)$ is a function of three unknowns, ϕ_f, O_f and A_f. Phases ϕ_f, $1 \leq f \leq F$ encode the scene depth, and can be recovered by taking three measurements for each of the F frequencies, while varying the exposure function phase, $\psi_f = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$. The unambiguous depth Γ can then be computed from the wrapped phases ϕ_f and the frequencies ω_f either analytically (using the Gushov-Solodkin (G-S) algorithm [Gushov and Solodkin 1991; Jongenelen et al. 2010]) or by building a look-up table:

$$\Gamma = \mathcal{T}\left(\phi_1, \omega_1, \ldots, \phi_F, \omega_F\right). \tag{24}$$

6.2 Number of Measurements

Since there are three unknowns (offset, amplitude and phase) for every frequency (Eq. 23), in general, there are $3F$ unknowns if F frequencies are used. Thus, Micro ToF imaging with F frequencies requires taking $3F$ measurements.

However, if the set of frequencies lie in a narrow band, the offsets and the amplitudes can be assumed to be approximately the same for all the frequencies. In this case, the number of unknowns is $F+2$; F phases, one offset and one amplitude. These unknowns can be estimated by taking $F+2$ measurements where three measurements are taken for the first frequency $\{B_1(0), B_1(\frac{2\pi}{3}), B_1(\frac{4\pi}{3})\}$, and one measurement is captured for every subsequent frequency $\{B_2(0), \ldots, B_F(0)\}$. The offset, amplitude and the first phase are computed from the first three measurements. Using the computed offset and amplitude, remaining phases are computed from the remaining measurements.

How many frequencies are needed? Theoretically, phases computed for $F=2$ appropriately chosen high frequencies are sufficient for estimating the scene depths unambiguously in any desired depth range. Thus, since the number of measurements is $F+2$, we get the following result:

RESULT 2. *Four images are theoretically sufficient for transport-robust and unambiguous depth recovery using Micro ToF imaging.*

In practice, more frequencies and measurements per frequency may be needed due to limited dynamic range of the sensor, limited frequency resolution of the light source and the sensor and noise. Depending on the scene brightness, light source strength and sensor noise levels, in our simulations and experiments, we use $2-4$ frequencies, and $3-4$ images per frequency.

6.3 Frequency Selection for Micro ToF Imaging

What frequencies should be used for Micro ToF imaging? For any choice of F frequencies, a given scene depth Γ_i is encoded with a vector V_i of measured brightness values. The number of elements in V_i is equal to the total number of measurements. Ideally, scene depths and intensity vectors should have a one-to-one mapping and depths can be recovered without error from the measured intensity vectors. However, due to various source of noise in the measurements, depth estimations can be erroneous. We define the error function $E_{ij} = e^{-||V_i - V_j||_2}$ between vectors V_i and V_j. E_{ij} is proportional to the probability of vector V_j being incorrectly decoded as Γ_i, and vice versa. For a given frequency set Ω, the mean weighted error function $E(\Omega)$ is defined as:

$$E(\Omega) = \sum_{\Gamma_i, \Gamma_j \in [0, \ldots, R_{max}]} |\Gamma_i - \Gamma_j| E_{ij} \,. \tag{25}$$

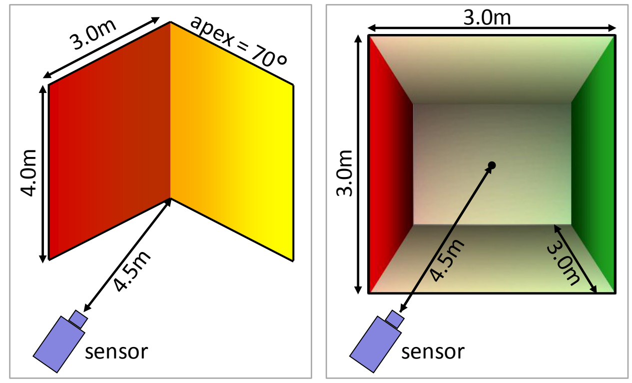
$E(\Omega)$ is the average expected depth error if the frequency set Ω is used. In order to minimize the depth error, the optimal set of frequencies should minimize the error function:

$$\Omega* = \arg\min_{\Omega} E(\Omega), \;\; \omega_f \in [\omega_{min}, \omega_{max}] \text{ for } 1 \leq f \leq F \,, \tag{26}$$

where $[\omega_{min}, \omega_{max}]$ is the range of values from which the frequencies are chosen. This is a constrained F dimensional optimization problem. We used the simplex search method implemented in MATLAB optimization toolbox for solving this. Note that since we use a search based procedure, the computed frequencies may not be theoretically optimal. In practice, we have found that the frequency set computed by running the method for $> 100,000$ iterations achieves stable depth results.

6.4 Simulations

In the following, we show depth recovery results using simulations for two different scene geometries.



(a) V-groove (b) Cornell Box

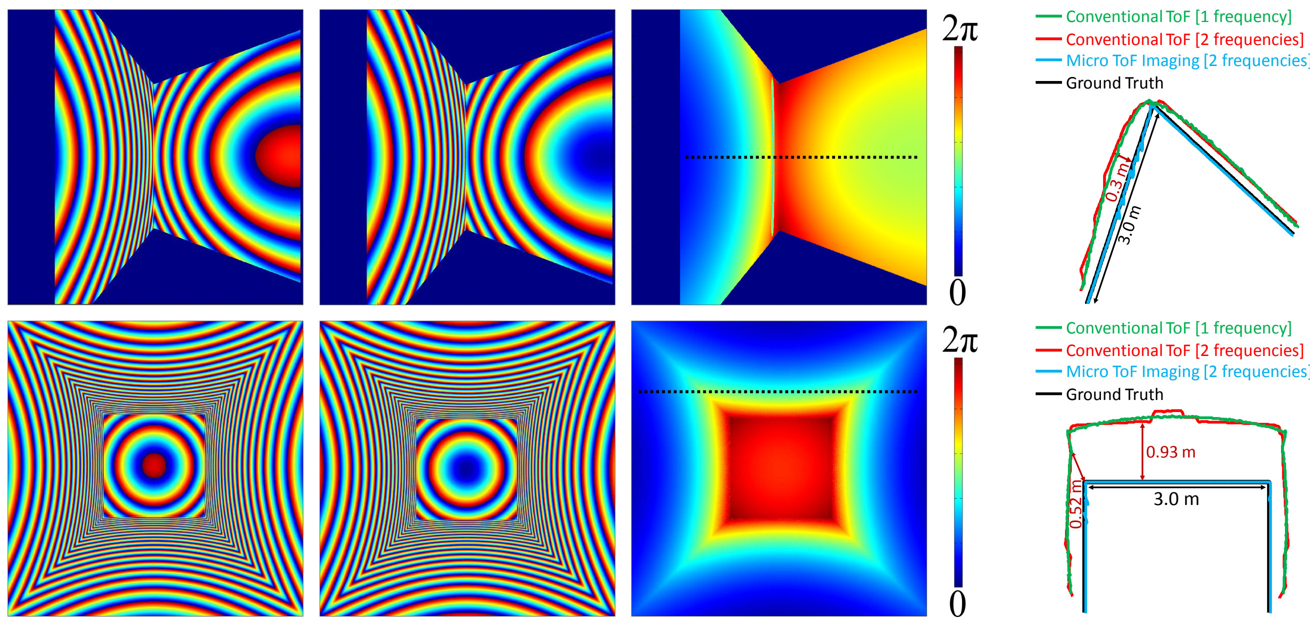
Fig. 8. **Simulation settings.** (a) A v-groove with two planes of size $3m \times 4m$ each. The angle between the planes is $70°$. (b) Cornell box with faces of size $3m \times 3m$. The sensor is at a distance of 4.5 m from the scenes.

Simulation setup: The setups are illustrated in Figure 8. The first scene is a *v-groove*, with an apex angle of $70\,\text{deg}$. Both faces are rectangles of dimension 4×3 meters. The second scene is the *Cornell box*. Each face is a square of side 3 meters. The sensor and source (co-located) are 4.5 meters from the scenes.

Simulation of input images: The input images were generated by discretizing the scene into small patches and simulating forward phasor light transport. We assumed the scenes to be Lambertian because for Lambertian scenes, the radiances can be computed using an analytic closed-form expression. This is similar to conventional imaging where the radiosity equation can be solved in a closed form manner for Lambertian scenes [O'Toole et al. 2014]. In all our simulations, the affine noise model [Hasinoff et al. 2010] was used for the sensor, with both scene independent read noise and scene dependent shot noise added to the captured images.

Simulation parameters and comparisons: For both scenes, we performed Micro ToF imaging using two frequencies $1063, 1034$ MHz., which were computed using the frequency selection procedure (Section 6.3). We compare Micro ToF with two different conventional ToF techniques. The single frequency conventional ToF technique uses a frequency of 10 MHz., so that the unambiguous depth range is more than the scene depths, and no phase unwrapping is required. Since the depth resolution of ToF techniques is directly proportional to the modulation frequency [Lange 2000], this technique achieves low depth resolution. The dual-frequency conventional ToF technique uses one high (1063 MHz.) and one low frequency (10 MHz.). The high frequency provides high resolution, and the low frequency is used for unwrapping. For each frequency, four measurements were captured, corresponding to the exposure function phases $\psi = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$. In order to compensate for the low SNR achieved by the single frequency ToF technique, we applied temporal averaging to its input images so that the random depth errors due to noise are approximately the same for all three techniques. The difference in the results is due to the structured errors caused by interreflections.

Results: Figures 9 (a-b) show phase maps for the two frequencies used for Micro ToF. Both the phase maps have ambiguities. Figures 9 (c) show the unwrapped phase map, which is used to compute unambiguous depths. Figures 9 (d) show the comparison of depths along horizontal scan-lines (shown in (c)). In both the conventional techniques, the low frequency phase is inaccurate, resulting in large depth errors (mean errors of 204 and 207 millimeters



(a-b) Recovered phase maps for Micro ToF (c) Unwrapped phase (d) Recovered Shape

Fig. 9. **Simulation results for shape recovery using Micro ToF.** (a-b) Phase maps for two high frequencies used for Micro ToF. Both the phase maps have ambiguities. (c) The unwrapped phase map, which is used to compute unambiguous depths. (d) Comparison of depths along horizontal scan-lines with single frequency and dual (one high and one low) frequency conventional ToF techniques. In both the conventional techniques, the low frequency phases are inaccurate, resulting in large depth errors (mean errors of >200 and >500 millimeters for the v-groove and the Cornell box, respectively). Micro ToF imaging achieves accurate shape, with mean errors of 6.6 and 3.2 millimeters for the v-groove and the Cornell box, a two orders of magnitude improvement.

for the v-groove and 534 and 538 millimeters for the Cornell box). Micro ToF imaging achieves accurate shape, with mean errors of 6.6 and 3.2 millimeters for the v-groove and the Cornell box, respectively, a two orders of magnitude improvement over conventional techniques.

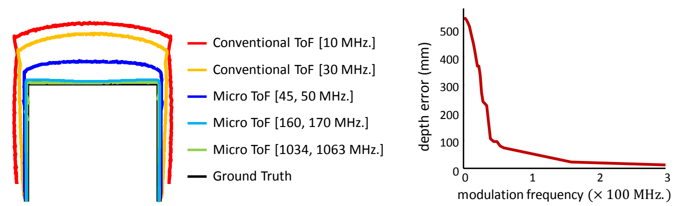
6.5 Error Analysis for Depth Computation

If ω is less than the global transport bandlimit ω_b, the oscillating term of the global radiance $\vec{G_\omega}(\mathbf{p})$ may not be zero. This will result in errors in phase recovery (and hence, depth estimation). As derived in Appendix B, the phase error ϵ_ϕ is given by (for brevity, we have dropped the argument $\mathbf{p}$):

$$\epsilon_\phi = \phi - acos\left(\frac{D\cos\phi + G^r_\omega \cos\phi_G}{\sqrt{D^2 + {G^r_\omega}^2 + 2DG^r_\omega\cos(\phi - \phi_G)}}\right), \quad (27)$$

where $\phi_G = arg(\vec{G}_\omega(\mathbf{p}))$ is the phase and $G^r_\omega = |\vec{G}_\omega(\mathbf{p})|$ is the magnitude of the global radiance at frequency ω. As ω increases, the magnitude of the global radiance $G^r_\omega \to 0$, and thus, $\epsilon_\phi \to 0$.

Figure 10 (a) shows the comparison of shapes recovered for the Cornell-box using different frequencies. Figure 10 (b) shows the mean depth error vs. frequency. Single frequency conventional ToF technique was used for frequencies less than 30 MHz. Micro ToF technique was used for frequencies more than 30 MHz. For each depth computation using Micro ToF, two high frequencies were used. Depth errors are plotted for the mean of the two frequencies. As the frequencies are increased, depth error approaches zero and the reconstructed shape approaches the ground-truth.



(a) Reconstructions for different frequencies (b) Mean error vs. frequency

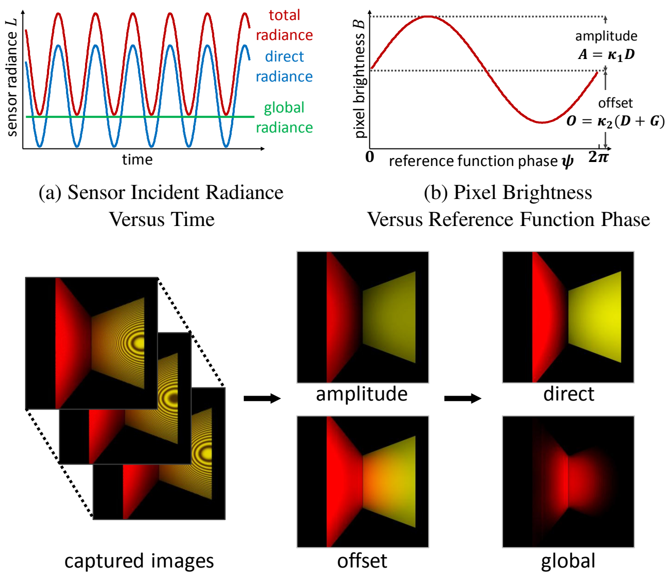
Fig. 10. **Effect of modulation frequency on shape recovery.** (a) Shapes recovered using different frequencies, for single frequency conventional and Micro ToF imaging (using two frequencies). As the frequencies are increased, reconstructed shape approaches the ground-truth. (b) Mean depth errors vs. frequency. Single frequency conventional ToF technique was used for frequencies less than 30 MHz. Micro ToF technique was used for frequencies more than 30 MHz. For Micro ToF, two frequencies were used. Depth errors are plotted for the mean of the two frequencies.

7. FAST SEPARATION OF DIRECT AND GLOBAL IMAGE COMPONENTS

In this section, we present a technique for separating the direct and global light transport components. The technique requires capturing as few as three measurements at a single high frequency. The measurements needed for the separation algorithm are a subset of the measurements taken for depth recovery (Section 6). Hence, separation can be achieved as a by-product of depth estimation.

7.1 Separation Algorithm

If the scene is illuminated with sinusoidally varying intensity, the direct and global radiance, as well as the total radiance are all sinu-



(a) Sensor Incident Radiance Versus Time

(b) Pixel Brightness Versus Reference Function Phase

(c) Steps of Direct-Global Separation Algorithm

Fig. 11. **Direct-global separation algorithm using high frequency illumination.** (a) If the scene is illuminated with high-frequency illumination, the global radiance is temporally constant. (b) The brightness B measured by the sensor is a sinusoid as a function of the phase ψ of the sensor exposure function. Since the global radiance is temporally constant, the amplitude of the sinusoid depends only on the direct radiance. (c) The direct and global components can be computed by measuring the offset and the amplitude of the sinusoid $B(\psi)$ at every pixel.

soids of the same frequency. In general, it is difficult to separate the direct and global radiance components from the total radiance due to the inherent ambiguity; a given sinusoid can be expressed as sum of two sinusoids of the same frequency in infinitely many ways.

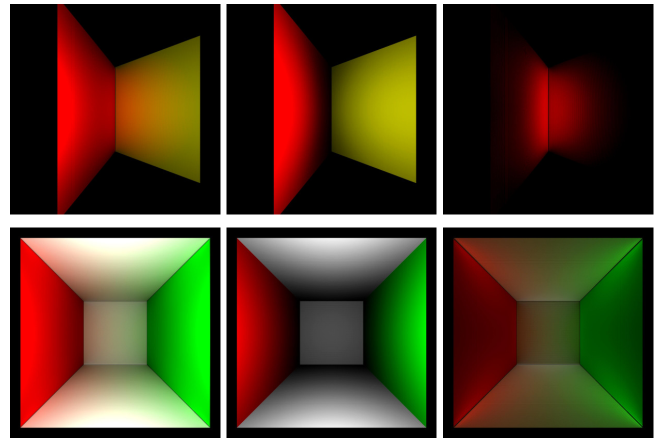
However, recall from Eq. 22 that for high frequency illumination, the global radiance manifests only as a DC offset. This forms the basis of our direct-global separation approach, and is illustrated in Figure 11 (a). Our goal is to separately recover $D(\mathbf{p})$ and $G(\mathbf{p})$, which are the direct and global components resulting from a light source with temporally constant, unit intensity.

Let the sensor exposure function be a sinusoid that is represented by the 2-tuple $R_{DC}, \vec{R}_\omega$. Let $\psi = arg(\vec{R}_\omega)$ be the phase of the exposure function. The correlation measurement $B(\psi)$ recorded at the sensor (see Appendix A for derivation) is given by:

$$B(\psi) = \underbrace{\tau(D+G)R_{DC}I_{DC}}_{\text{offset } O} + \underbrace{\tau\frac{DR_\omega I_\omega}{2}}_{\text{amplitude } A}\cos(\phi - \psi)\,, \quad (28)$$

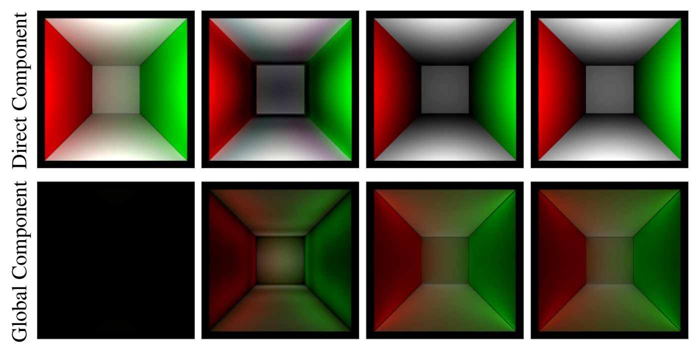
where $\phi = arg(\vec{D}_\omega(\mathbf{p}))$ is the phase of the direct radiance, and τ is the sensor integration time. $B(\psi)$ is a sinusoid with three parameters, the offset $O = \tau(D + G)R_{DC}I_{DC}$, the amplitude $A = \tau\frac{DR_\omega I_\omega}{2}$ and the phase ϕ, as shown in Figure 11 (b). The three parameters can be recovered by taking three correlation measurements. Since the constants $\tau, R_\omega, I_\omega, R_{DC}, I_{DC}$ are known, the direct and global components are recovered from the estimated offset and the amplitude, as shown in Figure 11 (c):

$$D = \frac{2A}{\tau R_\omega I_\omega}\,, \quad G = \frac{O}{\tau R_{DC}I_{DC}} - D. \quad (29)$$



(a) Scene image (b) Direct component (c) Global component

Fig. 12. **Simulation results for direct-global separation.** (a) Scene image. (b) Direct and (c) global components for the v-groove and the Cornell-box scenes, computed using the algorithm in Section 7. Notice the color-bleeding between different planes in the global component due to interreflections, and the direct component decreasing with increasing depth due to intensity fall-off.



(a) $\omega = 10$ MHz. (b) $\omega = 125$ MHz. (c) $\omega = 5.3$ GHz. (d) Ground Truth

Fig. 13. **Effect of modulation frequency on direct-global separation.** (a) For low frequencies (e.g., 10 MHz.), the direct component is overestimated and the global component is underestimated. (b) For 125 MHz., the separation is qualitatively correct, but ringing artifacts are noticeable, specially around the edges of the cube. (c) If the frequency is higher than the bandwidth of the global scene response, accurate separation is achieved. (d) Ground-truth separation.

RESULT 3. *Three images captured at a single high frequency are theoretically sufficient for separating the direct and global components of light transport.*

In practice, more measurements may need to be captured if the measurement noise is high. In our simulations and experiments, we use $3 - 4$ measurements.

7.2 Simulations

Figure 12 shows the direct and global components estimated for the two simulated scenes used in the previous section. Four measurements were taken for each of the examples. For the v-groove, notice that the global component is high near the corner, and decreases away from the corner. In the global component of the Cornell box, notice the color bleeding around the edges due to interreflections.

7.3 Error Analysis for Direct-Global Separation

Similar to the error analysis for depth estimation (Section 6.5), ω being less than the global transport bandlimit ω_b may result in erroneous direct-global separation. As derived in Appendix B, the estimation errors ϵ_D and ϵ_G for direct and global components, respectively, are given by:

$$\epsilon_D = D - \sqrt{D^2 + {G^r_\omega}^2 + 2DG^r_\omega \cos(\phi - \phi_G)}\,, \qquad (30)$$

$$\epsilon_G = \sqrt{D^2 + {G^r_\omega}^2 + 2DG^r_\omega \cos(\phi - \phi_G)} - D\,, \qquad (31)$$

where $\phi_G = arg(\vec{G}_\omega(\mathbf{p}))$ is the phase and $G^r_\omega = |\vec{G}_\omega(\mathbf{p})|$ is the magnitude of the global radiance at frequency ω. As ω increases, $G^r_\omega \rightarrow 0$, and thus, $\epsilon_D, \epsilon_G \rightarrow 0$.

Figure 13 shows the effect of modulation frequency on the results of direct-global separation. If a low modulation frequency (10 MHz.) is used, the direct component is over-estimated and the global component is under-estimated. As frequency increases, the global radiance decreases, and the separation accuracy increases. For 125 MHz., the estimation errors are lower, but the resulting images have ringing artifacts. At 5300 MHz., the result is close to the ground truth.

8. HARDWARE PROTOTYPE AND RESULTS

Our hardware prototype is based on the PMDTechnologies CamBoard Nano, a low-cost commercially available C-ToF imaging system. It is shown in Figure 14. In order to operate the system at various modulation frequencies, we used an external signal generator to provide the modulation signal instead of the on-board signal generator. Our light source is an array of 650 nm laser diodes [7], driven using an iC-Haus constant current driver. With this setup, we can achieve a maximum modulation frequency of 125 MHz. The modulation signals achieved by our setup are nearly sinusoids. As discussed in Section 5 though, this is not a strict requirement.

Results of depth recovery using Micro ToF imaging: The scene consists of a fixed wall and a movable wall arranged so that they form a concave v-groove, as illustrated in Figure 15 (a). The concave shape produces interreflections between the two walls. The amount of interreflections depends on the apex angle Υ, which can be changed by moving the right wall. Both walls are made of white, nearly diffuse material. The size of the walls is approximately $2\,m \times 2\,m$ each, and the sensor is placed at a distance of $5\,m$ from the corner of the groove. For Micro ToF imaging, we use two high frequencies of 125 and 108 MHz, computed using the frequency selection procedure [8].

We compare the results of Micro ToF with the single frequency conventional ToF technique. The frequency is chosen so that the unambiguous depth range is larger than the scene depths. In order to compensate for the low SNR achieved by the conventional ToF technique, we applied averaging to its input images so that the random perturbations due to noise are similar for both techniques (conventional and Micro ToF). The difference in the results is due to the structured errors caused by interreflections. Depth computed using the conventional ToF technique has mean errors of 87, 70 and 57 millimeters, for $\Upsilon = 45^\circ$, $\Upsilon = 60^\circ$ and $\Upsilon = 90^\circ$, respectively. Micro ToF achieves reconstructions that have $1 - 2$ orders of magnitude lower errors (mean errors of 2.8, 6.7 and 6.2 millimeters).

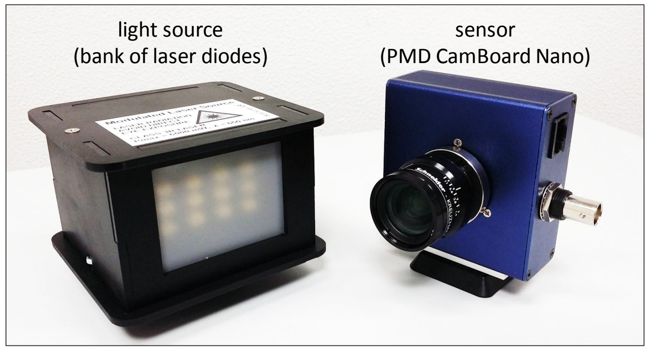


Fig. 14. **Image acquisition setup.** Our hardware prototype is based on the PMDTechnologies CamBoard Nano, a C-ToF sensor. An array of 650 nm laser diodes acts as the light source.

Figure 16 shows the performance of conventional and Micro ToF techniques as a function of the modulation frequency. For conventional ToF, we performed reconstructions using a single frequency in the range of $1 - 25$ MHz. For Micro ToF, we performed reconstructions using two frequencies in the range $[\omega - 2, \omega + 2]$ MHz. for $25 < \omega < 120$ MHz. The ground truth was achieved by measuring the scene distances using a measuring tape.

We compute the reconstructed apex angle by fitting two planes to the reconstructed shape, and computing the angle between them. As the frequency increases, the reconstruction error decreases and the apex angle approaches the ground truth. For frequencies higher than 100 MHz., the reconstruction error is less than 5 millimeters and the apex angle is within 2 degrees of the ground truth.

Results of direct-global separation: Figure 17 shows the direct-global separation results for the v-grooves of different apex angles. The modulation frequency used was 124 MHz. Three images were used for the separation in each case. As the apex angle increases (from top to bottom), the amount of interreflections reduce, and the global component decreases.

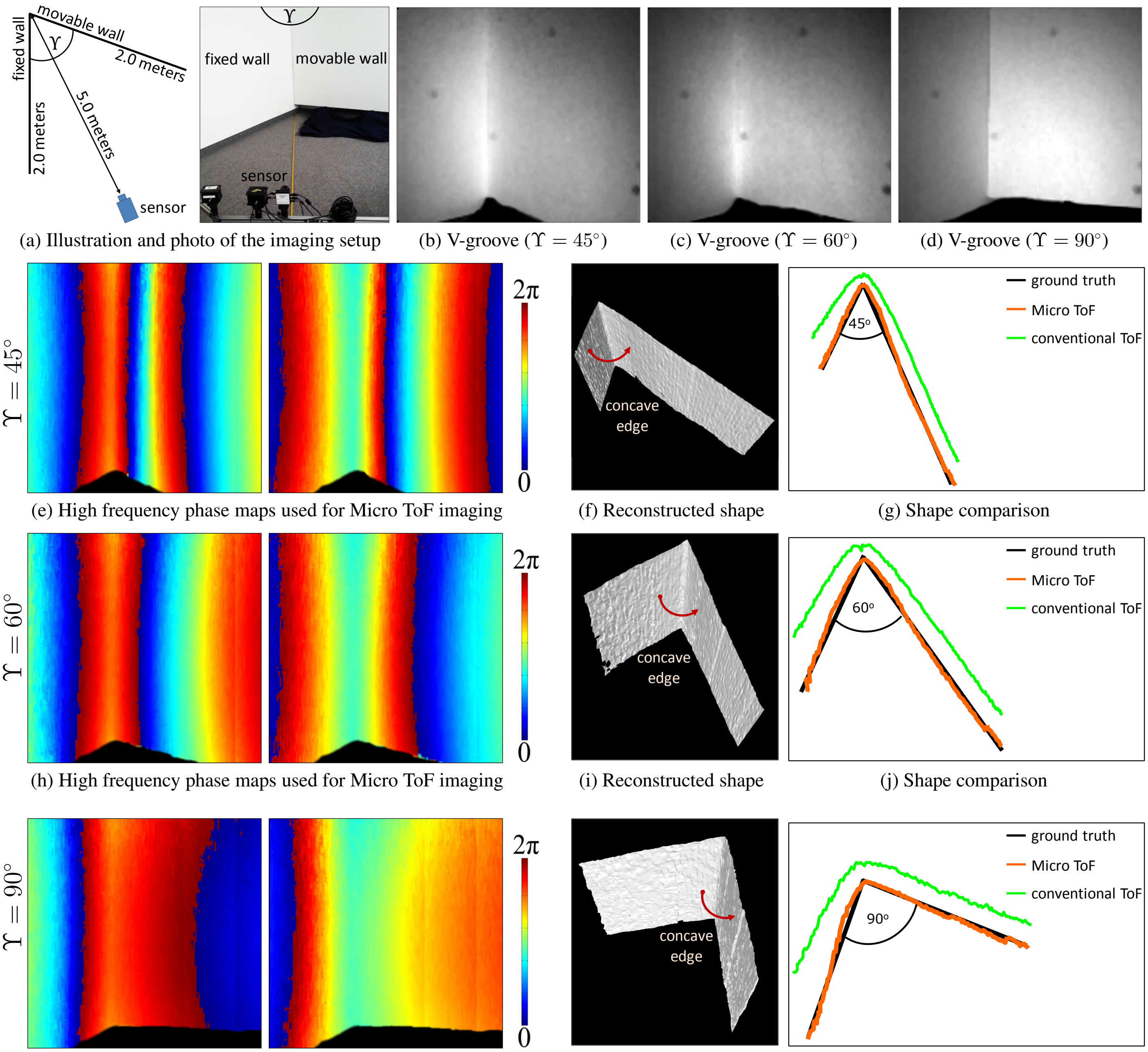
9. DISCUSSION AND LIMITATIONS

In this paper, we proposed phasor imaging, a tool for light transport analysis in C-ToF imaging, which can inspire novel imaging techniques in the future. Using this framework, we studied the (temporal) frequency dependence of light transport, and showed that global transport vanishes at high modulation frequencies. Based on this observation, we present techniques for transport-robust shape recovery and for separation of direct and global components. Since the presented techniques require few images and have low computational cost, we believe they can be incorporated into future ToF imaging systems. In the following, we discuss the limitations of our techniques.

Scope and limitations: Our techniques assume that the global light transport is temporally smooth. While this assumption holds for a broad range of scenes, for scenes with high-frequency light transport such as mirror interreflections, the presented techniques are prone to errors. For such scenes, shape recovery techniques that

[7]The size of the array is significantly smaller than the modulation wavelength. Hence, the array of diodes is assumed to be a single light source.

[8]For frequency selection, we set $\omega_{max} = 125$ MHz., the maximum frequency achievable by the imaging hardware, and ω_{min} to be $0.8\omega_{max}$. Although setting ω_{min} to an even higher value may achieve more robustness to global illumination, in practice, the difference $\omega_{max} - \omega_{min}$ needs to be above a threshold due to the limited intensity resolution and dynamic range of the sensor and the light source.



(a) Illustration and photo of the imaging setup (b) V-groove ($\Upsilon = 45°$) (c) V-groove ($\Upsilon = 60°$) (d) V-groove ($\Upsilon = 90°$)

(e) High frequency phase maps used for Micro ToF imaging (f) Reconstructed shape (g) Shape comparison

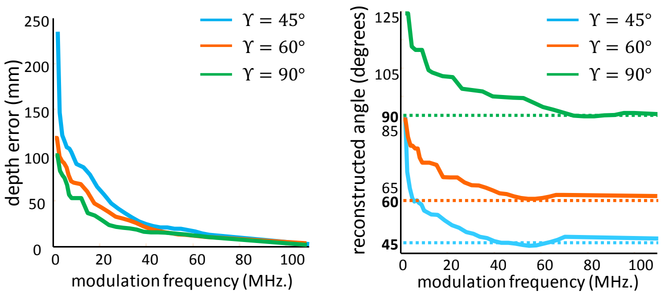
(h) High frequency phase maps used for Micro ToF imaging (i) Reconstructed shape (j) Shape comparison

(k) High frequency phase maps used for Micro ToF imaging (l) Reconstructed shape (m) Shape comparison

Fig. 15. **Experimental results of depth estimation on v-groove scenes.** (a) The scene consists of a fixed wall and a movable wall arranged at an angle so that they form a v-groove. The apex angle Υ can be changed by moving the right wall. The size of the walls is approximately $2\,m \times 2\,m$ each, and the sensor is placed at a distance of $5\,m$ from the corner. (b-d) Images captured by the PMD sensor, for the v-groove in three different configurations, $\Upsilon = 45°$, $\Upsilon = 60°$ and $\Upsilon = 90°$. For Micro ToF imaging, we use two high frequencies of 125 and 108 MHz. (e, h, k) Recovered phase maps for the two frequencies used in Micro ToF. All the phase maps have ambiguities. (f, i, l) The unambiguous shapes reconstructed by unwrapping the two high frequency phases. (g, j, m) Comparison of shapes reconstructed using Micro ToF and conventional ToF techniques. Shape computed using the conventional ToF technique has mean errors of 87, 70 and 57 millimeters, for $\Upsilon = 45°$, $\Upsilon = 60°$ and $\Upsilon = 90°$, respectively. Micro ToF achieves reconstructions that have $1 - 2$ orders of magnitude lower errors (mean errors of 2.8, 6.7 and 6.2 millimeters).

assume temporally sparse light transport are better suited [Godbaz et al. 2008; Dorrington et al. 2011; Kirmani et al. 2013].

Our direct-global separation technique can separate the direct radiance from relatively low-frequency global radiance, by capturing only three images. Techniques which can separate both low and



(a) Depth error (b) Reconstructed apex angle

Fig. 16. **Accuracy of shape recovery vs. modulation frequency for the v-groove scenes.** (a) Mean reconstruction error vs. the modulation frequency. (b) Reconstructed apex angle vs. the modulation frequency. The apex angle was computed by fitting two planes to the reconstructed shape, and computing the angle between them. For frequencies higher than 100 MHz., the reconstruction error is less than 5 millimeters and the apex angle is within 2 degrees of the ground truth.

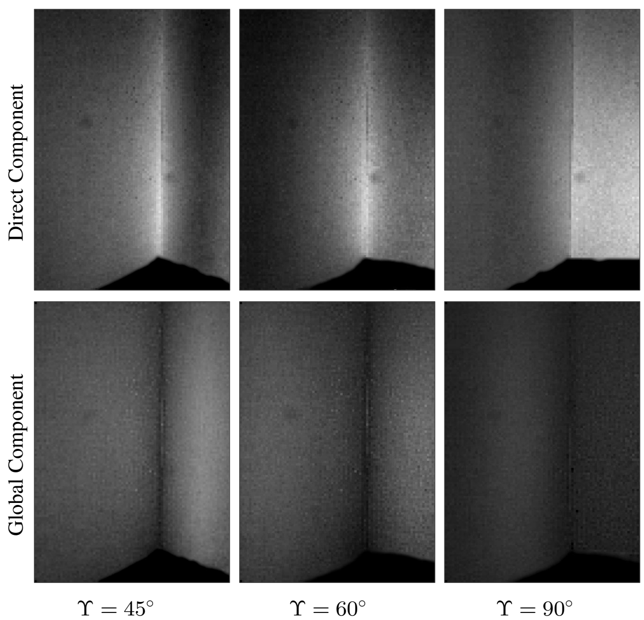


Fig. 17. **Results of direct-global separation for the v-groove scenes.** The direct and global components computed for the v-grooves. The modulation frequency used was 124 MHz. As the apex angle increases (from left to right), the global component decreases.

high frequency global transport [O'Toole et al. 2014], albeit by capturing more images, can be used for scenes with caustics and specular interreflections. It may be possible to develop hybrid scene-dependent algorithms (for both shape recovery and light transport analysis) where scene characteristics (low or high frequency light transport) determine the reconstruction technique to be used. This forms a promising direction of future work.

From a practical standpoint, for the proposed techniques to achieve accurate results, the modulation frequencies achieved by the system should be higher than the global transport bandlimit of the scene. The higher the modulation frequency, the larger the range of scenes (in terms of geometric scale and material properties) on which the proposed techniques are applicable. Although there are sensors and lasers that can achieve GHz. frequencies, they are expensive and require large acquisition times [Kirmani et al. 2009; Velten et al. 2013]. LEDs and PMDs are low cost sources and sensors that can achieve high SNR in real time, but due to various practical considerations such as power requirement, current devices are limited to approximately 150 MHz. With these devices, our techniques are restricted to large scale (room-size) scenes with relatively smooth reflectance.

Future outlook on hardware devices: Fortunately, high frequency LEDs have been actively researched with the goal of achieving high bandwidth optical communication networks. Recently, several research groups have demonstrated LEDs that can achieve modulation frequencies of multiple GHz. [Akbulut et al. 2001; Chen et al. 1999; Walter et al. 2009; Heinen et al. 1976; Wu et al. 2010], with a low power requirement. On the other hand, a new kind of PMD sensor based on MSM-technology (metal-semiconductor-metal) has recently been proposed that can potentially achieve > 10 GHz. modulation frequencies [Buxbaum et al. 2002; Schwarte 2004]. With these advances, it would be possible to apply the proposed techniques to a much larger class of scenes - scenes at centimeter/millimeter scale and comprising a broad range of reflectance properties. An additional motivation for achieving higher frequencies is that the depth resolution achieved by ToF sensors is proportional to the frequency used [Lange 2000]. Thus, higher frequencies can increase both the depth resolution and accuracy of ToF based depth sensing systems.

Generalization of light sources: Although the analysis and derivations were performed for a single point light source, the imaging framework proposed in the paper is applicable to extended light sources, arrays of light sources with each source potentially having a different phase and amplitude, or spatially modulated sources for performing spatio-temporal analysis of light transport [O'Toole et al. 2014].

APPENDIX

A. SENSOR CORRELATION MEASUREMENT

Let the sensor's exposure function be $R\cos(\omega t - \psi)$ (phasor representation: $\vec{R} = Re^{-j\psi}$). Let the incident radiance be $L\cos(\omega t - \phi)$ (phasor representation: $\vec{L} = Le^{-j\phi}$). The measured brightness B_ω is a function of ψ, and is given by the correlation between the exposure function and incident radiance:

$$B_\omega(\psi) = \int_0^\tau \left(R\cos(\omega t - \psi)\right)\left(L\cos(\omega t - \phi)\right) dt = \tau \frac{LR}{2}\cos(\phi - \psi), \quad (32)$$

where τ is the sensor integration time, and is assumed to be a integral multiple of the modulation period $\frac{2\pi}{\omega}$.

DC component: If both the radiance and the exposure function have a DC component as well, the measured brightness also has a DC component. Let the DC components of the exposure function and the radiance be R_{DC} and L_{DC}, respectively. The DC component of the measured brightness is given as:

$$B_{DC} = \int_0^\tau R_{DC} L_{DC} dt = \tau L_{DC} R_{DC} \,. \quad (33)$$

The total brightness is the sum of B_ω and B_{DC}:

$$B(\psi) = \underbrace{\tau L_{DC} R_{DC}}_{offset} + \underbrace{\tau \frac{LR}{2}}_{amplitude} \cos(\underbrace{\phi}_{phase} - \psi). \quad (34)$$

Thus, the measured brightness $B(\psi)$ is a sinusoid as a function of ψ, with three parameters - offset, amplitude and phase.

B. ERROR ANALYSIS

If ω is less than the global transport bandlimit ω_b, the oscillating term of the global radiance $\vec{G_\omega}(\mathbf{p})$ may not be zero. This will result in errors in phase recovery (for depth estimation) and direct-global separation. In the following, we derive the expression for the errors.

If $\vec{G_\omega}(\mathbf{p}) \neq 0$, the total radiance is given as:

$$L(\mathbf{p}) = \left[(D(\mathbf{p}) + G(\mathbf{p}))\ I_{DC}, \left(\vec{D}_\omega(\mathbf{p}) + \vec{G}_\omega(\mathbf{p}) \right)\ \vec{I_\omega} \right] . \quad (35)$$

From Appendix A, the correlation measurement taken by the sensor is given as:

$$\begin{aligned} B(\psi) &= \tau (D + G) R_{DC} I_{DC} + \\ & \tau \frac{R_\omega I_\omega}{2} \left(D \cos(\phi - \psi) + G^r_\omega \cos(\phi_G - \psi) \right), \quad (36) \end{aligned}$$

where $\phi_G = arg(\vec{G}_\omega(\mathbf{p}))$ is the phase and $G^r_\omega = |\vec{G}_\omega(\mathbf{p})|$ is the magnitude of the global radiance at frequency ω. Note that $0 \leq G^r_\omega \leq G$. The above equation can be rewritten as:

$$\begin{aligned} B(\psi) &= \underbrace{\tau (D + G) R_{DC} I_{DC}}_{\text{offset } O} + \\ & \underbrace{\tau \frac{R_\omega I_\omega}{2} \sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)}}_{\text{amplitude } A} \cos(\phi_T - \psi), \end{aligned}$$

where $\phi_T = acos \left(\frac{D \cos\phi + G^r_\omega \cos\phi_G}{\sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)}} \right)$ is the phase of the total radiance. Note that the measurement B is still a sinusoid, with offset $\hat{O} = \tau (D + G) R_{DC} I_{DC}$, amplitude $\hat{A} = \tau \frac{R_\omega I_\omega}{2} \sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)}$, and phase ϕ_T.

Phase error: The estimated phase (phase of the total radiance) is ϕ_T. The phase error ϵ_ϕ is given by the different between ϕ_T and ϕ, the true phase (phase of the direct radiance):

$$\epsilon_\phi = \phi - acos \left(\frac{D \cos\phi + G^r_\omega \cos\phi_G}{\sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)}} \right) \quad (37)$$

Direct-global errors: By using Eqs. 29, the direct and global components are estimated from $\hat{O}$ and $\hat{A}$:

$$\hat{D} = \sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)}\ , \quad (38)$$

$$\hat{G} = D + G - \sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)}\ . \quad (39)$$

The estimation errors ϵ_D and ϵ_G for direct and global components, respectively, are given by:

$$\epsilon_D = D - \sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)}\ , \quad (40)$$

$$\epsilon_G = \sqrt{D^2 + {G^r_\omega}^2 + 2 D G^r_\omega \cos(\phi - \phi_G)} - D\ . \quad (41)$$

As ω increases, the magnitude of the global radiance $G^r_\omega \rightarrow 0$, and thus, $\epsilon_\phi, \epsilon_D, \epsilon_G \rightarrow 0$.

REFERENCES

AI, X., NOCK, R., RARITY, J. G., AND DAHNOUN, N. 2011. High-resolution random-modulation CW lidar. *Applied Optics 50,* 22.

AKBULUT, M., CHEN, C. H., HARGIS, M. C., WEINER, A. M., MELLOCH, M. R., AND WOODALL, J. M. 2001. Digital communications above 1 Gb/s using 890-nm surface-emitting light-emitting diodes. *IEEE Photonics Technology Letters 13,* 1.

BUSCK, J. AND HEISELBERG, H. 2004. High accuracy 3D laser radar. *Proc. SPIE 5412*.

BUTTGEN, B., EL MECHAT, M.-A., LUSTENBERGER, F., AND SEITZ, P. 2007. Pseudonoise optical modulation for real-time 3-d imaging with minimum interference. *IEEE Transactions on Circuits and Systems I 54,* 10.

BUXBAUM, B., SCHWARTE, R., RINGBECK, T., GROTHOF, M., AND LUAN, X. 2002. Msm-pmd as correlation receiver in a new 3d-ranging system. In *Proc. SPIE Vol. 4546*.

CARNEGIE, D. A., MCCLYMONT, J. R. K., JONGENELEN, A. P. P., B. DRAYTO AND, A. A. D., AND PAYNE, A. D. 2011. Design and construction of a configurable full-field range imaging system for mobile robotic applications. *Lecture Notes in Electrical Engineering 83*.

CHEN, C. H., HARGIS, M., WOODALL, J. M., MELLOCH, M. R., REYNOLDS, J. S., YABLONOVITCH, E., AND WANG, W. 1999. Ghz bandwidth gaas light-emitting diodes. *Applied Physics Letters 74,* 21.

CHEN, T., SEIDEL, H. P., AND LENSCH, H. 2008. Modulated phase-shifting for 3D scanning. In *Proc. IEEE CVPR*.

COUTURE, V., MARTIN, N., AND ROY, S. 2014. Unstructured light scanning robust to indirect illumination and depth discontinuities. *International Journal of Computer Vision*.

DORRINGTON, A. A., GODBAZ, J. P., CREE, M. J., PAYNE, A. D., AND STREETER, L. V. 2011. Separating true range measurements from multi-path and scattering interference in commercial range cameras. In *Proc. SPIE 7864*.

FERRIERE, R., CUSSEY, J., AND DUDLEY, J. 2008. Time-of-flight range detection using lowfrequency intensity modulation of a cw laser diode: application to fiber length measurement. *Optical Engineering 47,* 9.

FREEDMAN, D., KRUPKA, E., SMOLIN, Y., LEICHTER, I., AND SCHMIDT, M. 2014. SRA: Fast removal of general multipath for tof sensors. In *Proc. ECCV (to appear)*.

FUCHS, S. 2010. Multipath interference compensation in time-of-flight camera images. In *IEEE ICPR*.

FUCHS, S., SUPPA, M., AND HELLWICH, O. 2013. Compensation for multipath in tof camera measurements supported by photometric calibration and environment integration. In *Computer Vision Systems*. Vol. 7963. Springer Berlin Heidelberg.

GODBAZ, J., CREE, M., AND DORRINGTON, A. 2008. Mixed pixel return separation for a full-field ranger. In *International Conference Image and Vision Computing New Zealand*.

GODBAZ, J. P., CREE, M. J., AND DORRINGTON, A. A. 2009. Multiple return separation for a full-field ranger via continuous waveform modelling. In *Proc. SPIE 7251*.

GODBAZ, J. P., CREE, M. J., AND DORRINGTON, A. A. 2012. Closed-form inverses for the mixed pixel/multipath interference problem in amcw lidar. In *Proc. SPIE 8296*.

GU, J., KOBAYASHI, T., GUPTA, M., AND NAYAR, S. 2011. Multiplexed illumination for scene recovery in the presence of global illumination. In *Proc. IEEE ICCV*.

GUPTA, M., AGRAWAL, A., VEERARAGHAVAN, A., AND NARASIMHAN, S. G. 2013. A practical approach to 3d scanning in the presence of interreflections, subsurface scattering and defocus. *International Journal of Computer Vision*.

GUPTA, M. AND NAYAR, S. K. 2012. Micro phase shifting. In *Proc. IEEE CVPR*.

GUPTA, M., TIAN, Y., NARASIMHAN, S. G., AND ZHANG, L. 2009. (De) Focusing on global light transport for active scene recovery. In *Proc. IEEE CVPR*.

GUSHOV, V. I. AND SOLODKIN, Y. N. 1991. Automatic processing of fringe patterns in integer interferometers. *Optics and Lasers in Engineering 14,* 4-5.

HASINOFF, S. W., DURAND, F., AND FREEMAN, W. T. 2010. Noise-optimal capture for high dynamic range photography. In *Proc. IEEE CVPR*.

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

HEINEN, J., HUBER, W., AND HARTH, W. 1976. Light-emitting diodes with a modulation bandwidth of more than 1 ghz. *Electronics Letters 12,* 21.

JIMENEZ, D., PIZARRO, D., MAZO, M., AND PALAZUELOS, S. 2012. Modelling and correction of multipath interference in time of flight cameras. In *IEEE CVPR*.

JONGENELEN, A. P. P., BAILEY, D. G., PAYNE, A. D., DORRINGTON, A. A., AND CARNEGIE, D. A. 2011. Analysis of errors in tof range imaging with dual-frequency modulation. *IEEE Transactions on Instrumentation and Measurement 60,* 5.

JONGENELEN, A. P. P., CARNEGIE, D., PAYNE, A., AND DORRINGTON, A. 2010. Maximizing precision over extended unambiguous range for tof range imaging systems. In *IEEE Instrumentation and Measurement Technology Conference (I2MTC)*.

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. *SIGGRAPH Asia*.

KIRMANI, A., BENEDETTI, A., AND CHOU, P. 2013. SPUMIC: Simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods. In *IEEE International Conference on Multimedia and Expo (ICME)*.

KIRMANI, A., HUTCHISON, T., DAVIS, J., AND RASKAR, R. 2009. Looking around the corner using transient imaging. In *IEEE ICCV*.

KOECHNER, W. 1968. Optical ranging system employing a high power injection laser diode. *IEEE Trans. aerospace and electronic systems 4,* 1.

LANGE, R. 2000. 3d time-of-flight distance measurement with custom solid-state image sensors in cmos-ccd-technology. *PhD Thesis*.

LANGE, R. AND SEITZ, P. 2001. Solid state time-of-flight range camera. *IEEE J. Quantum Electronics 37,* 3.

NAIK, N., ZHAO, S., VELTEN, A., RASKAR, R., AND BALA, K. 2011. Single view reflectance capture using multiplexed scattering and time-of-flight imaging. *SIGGRAPH Asia*.

NAYAR, S. K., KRISHNAN, G., GROSSBERG, M. D., AND RASKAR, R. 2006. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph. 25,* 3.

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 SIGGRAPH (to appear)*.

OTOOLE, M., RASKAR, R., AND KUTULAKOS, K. N. 2012. Primal-dual coding to probe light transport. *ACM SIGGRAPH*.

PANDHARKAR, R., VELTEN, A., BARDAGJY, A., LAWSON, E., BAWENDI, M., AND RASKAR, R. 2011. Estimating motion and size of moving non-line-of-sight objects in cluttered environments. In *IEEE CVPR*.

PAYNE, A. D., DORRINGTON, A. A., AND CREE, M. J. 2010. Illumination waveform optimization for time-of-flight range imaging cameras. In *Proc. SPIE 8085*.

REDDY, D., RAMAMOORTHI, R., AND CURLESS, B. 2012. Frequency-space decomposition and acquisition of light transport under spatially varying illumination. In *Proc. ECCV*.

SCHWARTE, R. 2004. Breakthrough in multi-channel laser-radar technology providing thousands of high-sensitive lidar receivers on a chip. In *Proc. SPIE Vol. 5575*.

SCHWARTE, R., XU, Z., HEINOL, H., OLK, J., KLEIN, R., BUXBAUM, B., FISCHER, H., AND SCHULTE, J. 1997. New electro-optical mixing and correlating sensor: Facilities and applications of the photonic mixer device. In *Proc. SPIE (3100)*.

STEVEN M. SEITZ, YASUYUKI MATSUSHITA, K. N. K. 2005. A theory of inverse light transport. In *Proc. IEEE ICCV*.

TAKEDA, M., GU, Q., KINOSHITA, M., TAKAI, H., AND TAKAHASHI, Y. 1997. Frequency-multiplex fourier-transform profilometry: A single-shot three-dimensional shape measurement of objects with large height discontinuities and/or surface isolations. *Applied Optics 36,* 22.

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 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. *SIGGRAPH*.

WALTER, G., WU, C., THEN, H., FENG, M., AND HOLONYAK, N. 2009. Tilted-charge high speed (7 ghz) light emitting diode. *Applied Physics Letters 94,* 23.

WU, C. H., WALTER, G., THEN, H. W., AND FENG, M. 2010. Design and layout of multi ghz operation of light emitting diodes. In *GaAs MANTECH Conference*.

WU, D., O'TOOLE, M., VELTEN, A., AGRAWAL, A., AND RASKAR, R. 2012. Decomposing global light transport using time of flight imaging. In *IEEE CVPR*.

WU, D., WETZSTEIN, G., BARSI, C., WILLWACHER, T., O'TOOLE, M., NAIK, N., DAI, Q., KUTULAKOS, K., AND RASKAR, R. 2012. Frequency analysis of transient light transport with applications in bare sensor imaging. In *ECCV*.

Supplementary Technical Report for Phasor Imaging: A Generalization of Correlation-Based Time-of-Flight Imaging.

In this technical report, we provide derivations supporting the content in the paper submission titled "Phasor Imaging: A Generalization of Correlation-Based Time-of-Flight Imaging".

1 Local Linearity of Light Transport Phases

Consider a light path $\mathbb{P}(\mathbf{p}, \theta)$ taken by a light ray emitted in direction θ from the light source, through the scene, and reaching the sensor pixel p. We assume the path to be piecewise linear. The last linear segment of the path is between a scene point, say X, and pixel p.

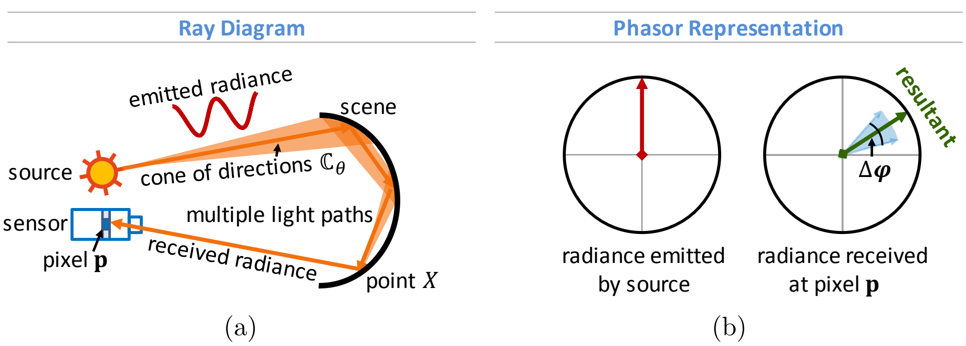


Figure 1: **Light transport in a local path neighborhood.** (a) A set of indirect light paths $\mathbb{P}(\mathbf{p}, \mathbb{C}_\theta)$ in a local path neighborhood. All the paths end at sensor pixel $\mathbf{p}$. (b) For a sufficiently small neighborhood, the phasor radiances along paths $\mathbb{P}(\mathbf{p}, \mathbb{C}_\theta)$ can be assumed to have constant amplitudes and linearly varying phases. Thus, the phasors form a sector. The angle of the sector $\Delta\phi$ is proportional to the modulation frequency ω. The total indirect radiance is given by the resultant sum of the individual phasors.

Consider the light paths in a local neighborhood of $\mathbb{P}(\mathbf{p}, \theta)$ that start in a narrow cone of directions $\mathbb{C}_\theta$ around θ and end at p. This is illustrated in Figure 1. All the light paths share the last segment, between X and p. Since all the paths end at the same pixel, for brevity, we drop the argument $\mathbf{p}$ in the rest of the section, i.e., $\mathbb{P}(\theta) = \mathbb{P}(\mathbf{p}, \theta)$. Let $\phi(\mu)$ be the phase of the light incident on the sensor along the path $\mathbb{P}(\mu)$, for $\mu \in \mathbb{C}_\theta$. The Taylor series expansion of the function $\phi(\mu)$ around θ is given as:

$$\phi(\mu) = \phi(\theta) + \frac{\phi^{[1]}(\theta)}{1!}(\mu - \theta) + \frac{\phi^{[2]}(\theta)}{2!}(\mu - \theta)^2 + \cdots , \quad (1)$$

where $n!$ denotes the factorial of n and $\phi^{[n]}(\theta)$ denotes the n^{th} derivative of function ϕ evaluated at θ.

Let the total length of the path $\mathbb{P}(\mu)$ be $\Gamma(\mu)$. Assuming the phase of all the emitted rays to be the same, $\phi(\mu) = \frac{2\pi\omega}{c}\Gamma(\mu)$, where ω is the modulation frequency. Taking the n^{th} derivative of both sides, we get:

$$\phi^{[n]}(\mu) = \frac{2\pi\omega}{c}\Gamma^{[n]}(\mu) . \quad (2)$$

Substituting the above in right hand side of Eq. 3, we get:

$$\phi(\mu) = \phi(\theta) + \frac{2\pi\omega}{c}\frac{\Gamma^{[1]}(\theta)}{1!}(\mu - \theta) + \frac{2\pi\omega}{c}\frac{\Gamma^{[2]}(\theta)}{2!}(\mu - \theta)^2 + \cdots . \quad (3)$$

If the cone $\mathbb{C}_\theta$ is small and modulation frequency ω is high, we can consider the first order approximation of the above [1]:

$$\phi(\mu) \approx \phi(\theta) + \frac{2\pi\omega}{c}\frac{\Gamma^{[1]}(\theta)}{1!}(\mu - \theta) . \quad (4)$$

After reorganizing, the above can be written as:

$$\phi(\mu) \approx c_1 + c_2\mu , \quad (5)$$

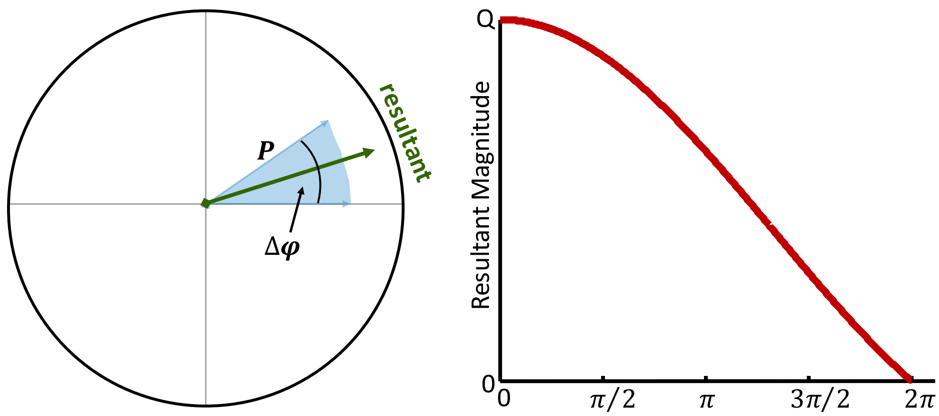
where c_1 and c_2 are constants. Thus, in a local light path neighborhood, the phases $\phi(\mu)$ of the light paths vary approximately linearly.

2 Resultant Phasor Magnitude

In the following, we derive the magnitude of the resultant of a set of phasors of equal magnitude, so that their phases are uniformly (and continuously) distributed within a range $[0, \Delta\phi]$. Let the phasors be $\vec{P}(\phi) = Pe^{j\phi}$, for $\phi \in [0, \Delta\phi]$. Let Q be the total magnitude, i.e., the integral of the magnitudes of the phasors:

$$Q = \int_0^{\Delta\phi} |\vec{P}(\phi)| d\phi = P\Delta\phi. \quad (6)$$

[1] We assume that path-lengths are not constant in a neighborhood, and $\Gamma^{[1]}(\theta) \neq 0$. An exception is if (a) all the scene points (except X) form an ellipsoid, and (b) point X and the source lie on the foci of the ellipsoid, and (c) the inside of the ellipsoid is a mirror so that only single bounce paths reach X from the source. In this special case, all the light paths have the same lengths, and $d^{[1]}(\theta) = 0$.



(a) (b)

Figure 2: **Resultant phasor magnitude versus the sector angle.** (a) A set of phasors of equal magnitude P so that their phases are uniformly and continuously distributed with a range $[0, \Delta\phi]$. The integral of magnitude of the individual phasors is Q. (b) The magnitude of the resultant decreases monotonically as a function of the sector angle $\Delta\phi$.

This is illustrated in Figure 2 (a). It follows that the magnitude of each individual phasor is given by $P = \frac{Q}{\Delta\phi}$. The resultant phasor $\vec{R}$ is given by integrating the individual phasors:

$$\begin{aligned} \vec{R} &= \frac{Q}{\Delta\phi}\left(\int_0^{\Delta\phi}\cos(\phi)d\phi + j\int_0^{\Delta\phi}\sin(\phi)d\phi\right) \\ &= \frac{Q}{\Delta\phi}\left(\sin(\Delta\phi) + j\left(1-\cos(\Delta\phi)\right)\right). \end{aligned} \quad (7)$$

The magnitude $|\vec{R}|$ is given as:

$$|\vec{R}| = 2Q\frac{\sin\left(\frac{\Delta\phi}{2}\right)}{\Delta\phi}. \quad (8)$$

$|\vec{R}|$ is a monotonically decreasing function of the sector angle $\Delta\phi$. This is illustrated in Figure 2 (b).