

MIT Open Access Articles

A light transport model for mitigating multipath interference in Time-of-flight sensors

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Naik, Nikhil et al. "A Light Transport Model for Mitigating Multipath Interference in Time-of-Flight Sensors." 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7-12 June, 2015, Boston, Massachusetts, USA, IEEE, 2015.

As Published: <http://dx.doi.org/10.1109/CVPR.2015.7298602>

Publisher: Institute of Electrical and Electronics Engineers (IEEE)

Persistent URL: <http://hdl.handle.net/1721.1/110641>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



A Light Transport Model for Mitigating Multipath Interference in Time-of-flight Sensors

Nikhil Naik^{1,2} Achuta Kadambi¹ Christoph Rhemann² Shahram Izadi²
Ramesh Raskar¹ Sing Bing Kang²

¹MIT Media Lab ²Microsoft Research

{naik, kadambi, raskar}@mit.edu {chrheman, shahrami, sbkang}@microsoft.com

Abstract

Continuous-wave Time-of-flight (TOF) range imaging has become a commercially viable technology with many applications in computer vision and graphics. However, the depth images obtained from TOF cameras contain scene dependent errors due to multipath interference (MPI). Specifically, MPI occurs when multiple optical reflections return to a single spatial location on the imaging sensor. Many prior approaches to rectifying MPI rely on sparsity in optical reflections, which is an extreme simplification. In this paper, we correct MPI by combining the standard measurements from a TOF camera with information from direct and global light transport. We report results on both simulated experiments and physical experiments (using the Kinect sensor). Our results, evaluated against ground truth, demonstrate a quantitative improvement in depth accuracy.

1. Introduction

Time-of-flight (TOF) cameras represent an increasingly popular method to obtain depth maps at full framerates of 30 FPS. The recent release of the TOF-based Kinect sensor [47] will result in the wide-spread adaption of TOF cameras for various traditional and non-traditional applications [35, 23, 31, 34, 24, 2, 44] in computer vision, graphics and physics. In this paper we restrict our discussion to continuous-wave TOF cameras, which calculate depth by measuring the phase difference between an emitted and received optical signal. In Section 3 we provide an overview of the technology and a more extensive overview can be found in [7, 21].

Even though TOF cameras represent an attractive technology for generating depth images in real-time, they suffer from many sources of error [9]. We focus on perhaps the main source of error: the problem of multi-path interference (MPI). The MPI problem occurs when multiple reflections of light in a scene return to a single pixel on the cam-

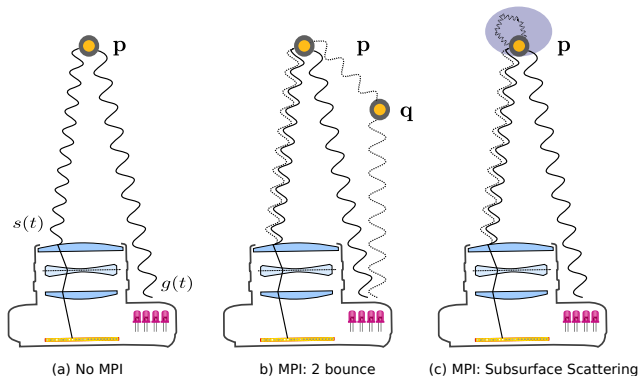


Figure 1. Multipath interference in time of flight cameras. In (a) the typical operation of TOF cameras is shown. A single optical path returns to the image sensor. In (b) an indirect bounce from point q to p causes two optical paths to return to the image sensor. In (c), subsurface scattering causes many optical paths to return to the sensor (one possible path is shown).

era sensor. This problem is particularly significant for TOF sensors, where depth calculations assume a single optical reflection.

Figure 1 illustrates three different measurement scenarios for a TOF sensor. In Figure 1a, the standard operation is shown, where a single reflection from point p strikes the image sensor. Figure 1b and Figure 1c illustrate cases of MPI where multiple optical reflections strike the imaging sensor. The former considers a simple two-bounce reflection and the latter subsurface scattering, which cannot be described by a discrete model.

Most prior MPI correction algorithms assume sparsity in optical reflections. Since sparsity is a weak approximation for the complexity of scattering, the space of sparse solutions has, to date, been restricted to toy scenes [3, 10, 8]. Other techniques rely on optical computing, which involves significant optical and electronic modifications to the camera [40, 32].

In this paper we cast the problem of MPI into the realm

Paper	Multipath Type	Solution Type	Quantitative Analysis	Hardware Modifications
Fuchs [11]	Continuous	Iterative	Limited	None
Dorrington et al. [8]	2-sparse	Iterative	Limited	Frequency sweep
Godbaz et al. [12]	2-sparse	Closed-form	None	Frequency sweep
Kadambi et al. [32]	K -sparse	Iterative	None	Custom code
Kirmani et al. [33]	K -sparse	Iterative	Simulations only	Frequency sweep
Freedman et al. [10]	K -sparse	Iterative	Simulations only	None
Jiminez et al. [29]	K -sparse	Iterative	Limited	None
O’Toole et al. [40]	Continuous	None	None	Extensive
Gupta et al. [19]	Continuous	Closed-form	Yes	Extensive
This Paper	Continuous	Closed-form	Yes	External projector

Table 1. This paper proposes a closed form approach to correcting multipath, by relying on light transport information.

of direct-global illumination. Recall that direct illumination represents optical paths from source, to scene point to camera (Figure 1a). In contrast, global illumination represents optical paths that reach the scene point through indirect means (Figure 1b and Figure 1c). The crux of this paper is as follows: we develop a closed-form solution that combines the measurements from standard TOF sensors with direct and global radiance maps (cf. Nayar et al. [38]).

1.1. Contributions

Our key contribution is a novel image formation model for TOF sensors that incorporates direct-global separation into phase and amplitude calculations. We use this model to significantly reduce errors caused by multipath interference. Another important contribution is our quantitative evaluation of the benefit of our technique using ground truth data.

Our approach uses both the direct and global components of radiance to modify the traditional depth calculation in time of flight sensing. Direct and global components can be acquired with the use of a projector [38]. For comparison, O’Toole et al. [40] use an involved, optical setup to filter out the scattered radiance. In addition, where most previous methods provide limited or no quantitative analysis of their results [11, 12, 32, 40], we report errors for all scenes in the paper by evaluating our technique against ground truth results (e.g., from an accurate laser scanner).

2. Related Work

2.1. Multipath Interference Correction

We summarize the important related publications on MPI correction in table 1. Fuchs [11] and Jiminez et al [29] solve for the MPI-free phase using an optimization problem and a forward model for TOF imaging. Dorrington et al. [8], Godbaz et al. [12] and Kirmani et al. [33] model MPI using two bounce approximation and propose iterative or closed form solutions using multiple frequency measurements (or ‘frequency sweep’). Bhandari et al. [4, 3] and

Freedman et al. [10] propose iterative solutions for generalized MPI using multiple frequency measurements. These solutions based on K -sparse models (table 1) are tied to the sparsity assumption, which leads to errors when scene returns are not sparse (e.g., subsurface scattering). Kadambi et al. [32] introduce a ‘coded’ illumination technique for TOF imaging and demonstrate results for MPI correction on scenes containing translucent objects. This technique, while promising, requires modifications in hardware, and precludes real-time performance.

We now turn to an emerging research trend, which incorporates light transport information to correct for MPI. Our work is of this flavor and is closely related to two recent papers by O’Toole et al. [40] and Gupta et al. [19]. Building on [41], O’Toole et al. [40] propose a hardware solution to block the global component of light transport during the capture process, thus eliminating MPI. O’Toole’s optical computing approach is accurate, but requires extensive optical modifications. The work by Gupta et al. [19] centers on a theoretical justification of temporal light transport. Specifically, Gupta shows that the global light transport vanishes at high frequencies. In related work, Gupta et al. [18, 15] have also tackled the problem of structured light scanning in presence of global illumination using coded illumination patterns. For a more complete overview, we encourage the reader to review the work by O’Toole et al. and Gupta et al.

2.2. Light Transport Acquisition

Since our work ties light transport acquisition to MPI correction, we provide a brief overview of this work. Seitz et al. [46] introduced the problem of ‘inverse light transport’, where the goal is to decompose an image into a sum of n -bounce images. Nayar et al. [38] propose a fast method for separating the direct component of light transport from the global component using high frequency illumination patterns. Ng et al. [39] introduce a stratified light transport matrix which allows computation of the inverse light transport using matrix multiplications. Chandraker

et al. [6] treat inverse light transport acquisition as a dual of forward rendering and propose an algorithm that needs only matrix-vector multiplication. Reddy et al. [45] introduce a method to decompose the light transport in a scene into direct (zero-th bounce), near-range (subsurface scattering and local inter-reflections) and far-range (diffuse inter-reflections) components. An exciting approach by O’Toole et al. [42] exploits coaxial optical setups to separate distinct bounces of optical paths.

Light transport has been successfully applied to a variety of practical problems, including rendering of complex phenomena [16, 49, 50], imaging in poor visibility environments [17], and recovery of surface reflectance [26, 35, 28, 27]. In addition, light transport techniques have been very successful for the challenging problem of 3-D scanning of translucent objects [25, 20] and high quality structured light scanning [36, 13, 37, 14].

While there are several light transport techniques to consider, we incorporate the method by Nayar et al. [38]. We believe it remains the simplest method to acquire and separate the direct and global components of light transport; it requires few images (two in theory) and, importantly, no calibration.

3. Continuous Wave TOF Measurements

A continuous-wave TOF camera (hereafter, TOF camera) is able to measure the phase delay of optical paths and obtain depth through the following relation:

$$z = \frac{c\varphi}{4\pi f_M}, \quad (1)$$

where f_M is the modulation frequency of the camera and c is the speed of light. To estimate φ with high precision, a TOF camera contains an active illumination source that is strobed according to a periodic illumination signal. In this paper we consider standard implementations (e.g. MS Kinect) where the emitted signal takes the form of a sinusoid

$$g(t) = \cos(f_M t). \quad (2)$$

Suppose the light is reflected from a scene point \mathbf{p} . Then at the sensor plane, the received optical signal can be written as

$$s(t; \mathbf{p}) = \alpha^{\mathbf{P}} \cos(f_M t + \varphi^{\mathbf{P}}) + \beta^{\mathbf{P}}, \quad (3)$$

where $\alpha^{\mathbf{P}}$ is the attenuation in the projected radiance and $\beta^{\mathbf{P}}$ is the intensity of ambient light. Note the time-invariance of parameters $\alpha^{\mathbf{P}}$, $\varphi^{\mathbf{P}}$, and $\beta^{\mathbf{P}}$ —it is assumed that they are constant within the exposure time. A TOF camera computes the cross-correlation of the emitted and received signals, which evaluates to

$$c(\tau; \mathbf{p}) = s(t; \mathbf{p}) \otimes g(t) = \frac{\alpha^{\mathbf{P}}}{2} \cos(f_M \tau + \varphi^{\mathbf{P}}) + \beta^{\mathbf{P}}. \quad (4)$$

Note that the same information from Equation 3—amplitude, phase, and ambient light—is present in the cross-correlation function. However, the cross-correlation equation is in the domain of τ , which is easier to sample (in fact, the TOF sensor directly reads $c(\tau; \mathbf{p})$ from the sensor). As our technique does not incorporate spatial parametrization, hereafter we consider $c(\tau)$ as the received signal.

In order to recover the phase and amplitude from the received signal, the “four bucket principle” is typically used (see [8, 3]). In this method, the TOF camera evenly samples $c(\tau)$ four times over the length of a period, for instance, $\tau = [0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}]^T$. Then the calculated phase can be written as

$$\tilde{\varphi} = \arctan\left(\frac{c(\tau_4) - c(\tau_2)}{c(\tau_1) - c(\tau_3)}\right), \quad (5)$$

and calculated amplitude as

$$\tilde{\alpha} = \frac{1}{2} \sqrt{(c(\tau_4) - c(\tau_2))^2 - (c(\tau_1) - c(\tau_3))^2}. \quad (6)$$

This method is a closed-form technique—in the absence of factors such as multipath interference and sensor noise, the calculated variables $\tilde{\varphi}$ and $\tilde{\alpha}$ are equal to ground truth values φ and α .

4. Multipath Interference

To calculate accurate depth it is necessary to obtain the time of flight of the direct light path. However, the light received from a TOF sensor is a combination of direct and global illumination [38]. Multipath interference refers to the undesirable mixing of global illumination with direct illumination, which, in a TOF camera, causes a deviation in the measured phase wrt. the direct phase only. We now provide the mathematical intuition behind MPI, starting with the most general case.

4.1. Generalized MPI

In the presence of global illumination, the received signal includes contributions from illumination reflected from multiple scene points. Specifically, we denote illumination that has reflected from K locations in the scene as a K -bounce sinusoid with phase φ_K and amplitude α_K . In this context, the received signal is a composite of multi-bounce sinusoids and can be written as

$$c(\tau) = \underbrace{\frac{\alpha_0}{2} \cos(f_M \tau + \varphi_0)}_{\text{Direct Contribution}} + \underbrace{\sum_{i=1}^K \frac{\alpha_i}{2} \cos(f_M \tau + \varphi_i)}_{\text{Global Contribution}} + \beta, \quad (7)$$

where the problem is not constrained to finite-dimensions, i.e., $K = \infty$ is possible. Substituting Equation 7 into Equations 5 and 6 results in the following expressions for the

measured phase and amplitude

$$\tilde{\varphi} = \arctan \left(\frac{\alpha_0 \sin \varphi_0 + \sum_{i=1}^K \alpha_i \sin \varphi_i}{\alpha_0 \cos \varphi_0 + \sum_{i=1}^K \alpha_i \cos \varphi_i} \right), \quad (8)$$

$$\tilde{\alpha}^2 = \alpha_0^2 + \sum_{i=1}^K \alpha_i^2 + 2 \underbrace{\sum_{i=0}^K \sum_{j=0}^K \alpha_i \alpha_j \cos(\varphi_i - \varphi_j)}_{i \neq j}. \quad (9)$$

Not surprisingly, the measured phase $\tilde{\varphi}$ is not equal to the phase contribution from the direct bounce φ_0 . As a result, the measured depth will be distorted.

4.2. Approximating Global Illumination

In this paper, we approximate global illumination using only the lowest order, indirect bounce (i.e., $K = 1$). This simplification has been exploited in previous techniques for multipath interference correction (e.g., [11, 8, 12, 33]). Under this assumption, the received signal can be modeled as

$$c(\tau) = \underbrace{\frac{\alpha_D}{2} \cos(f_M \tau + \varphi_D)}_{\text{Direct}} + \underbrace{\frac{\alpha_G}{2} \cos(f_M \tau + \varphi_G)}_{\text{Approximate Global}} + \beta, \quad (10)$$

where subscripts D and G explicitly denote parameters of direct and approximate global illumination. The total projected radiance α_T is equal to $\alpha_D + \alpha_G$, where $0 \leq \alpha_D, \alpha_G, \leq 1$ and the total radiance is equal to unity.

After substituting the measurement from Equation 10 into Equations 5 and 6, we obtain

$$\tilde{\varphi} = \arctan \left(\frac{\alpha_D \sin \varphi_D + \alpha_G \sin \varphi_G}{\alpha_D \cos \varphi_D + \alpha_G \cos \varphi_G} \right), \quad (11)$$

$$\tilde{\alpha}^2 = \alpha_D^2 + \alpha_G^2 + 2\alpha_D \alpha_G \cos(\varphi_D - \varphi_G). \quad (12)$$

4.3. Direct-Global Separation for MPI Correction

Accurately estimating φ_D would make it possible to ameliorate MPI and obtain a robust measurement of depth. Recall that the fast method from Nayar et al. [38] can be used to obtain the radiance of direct and global illumination. For instance, in the case of generalized MPI (Section 4.1), direct-global separation allows us to obtain radiance values α_D and $\sum_{i=1}^K \alpha_i$. However, these two measurements are not helpful in the context of general MPI, where Equations 8 and 9 consist of $2N+2$ variables, of which only α_0 is known.

We now turn to the case of approximate global illumination, where we use the lowest-order indirect bounce to approximate contributions from global illumination. Under this approximation, direct-global separation returns values for α_D and α_G . Notice that between Equations 11 and 12 there are four variables, two of which are provided by the

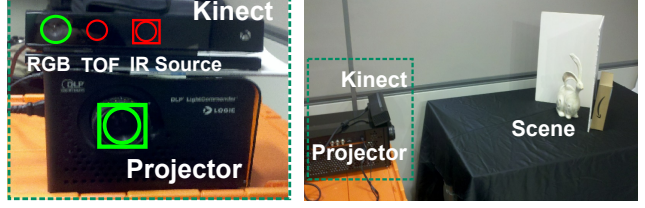


Figure 2. We demonstrate our method with no hardware modifications to the TOF hardware. Our capture setup (left) consists of a Kinect One sensor and a DLP Lightcommander Infrared projector. The capture setup is placed ~ 1 meter away from the scene (right).

Nayar method. Given that there are two equations and only two unknown variables, the direct phase can be computed in closed-form as:

$$\hat{\varphi}_D = \arctan \left(\frac{\alpha_D \gamma + \alpha_G (\sin(\varphi_D - \varphi_G) + \gamma \cos(\varphi_D - \varphi_G))}{\alpha_D + \alpha_G (\cos(\varphi_D - \varphi_G) - \gamma \sin(\varphi_D - \varphi_G))} \right), \quad (13)$$

where $\gamma = \tan(\tilde{\varphi})$. In summary, relaxing the problem to the second bounce approximation provides a closed-form solution for ameliorating multipath interference. In the remainder of the paper, we evaluate this closed-form correction through simulations and captured data.

5. Implementation

We first describe our simulation setup, followed by details on the physical setup.

5.1. Simulations

To perform simulations, we use the time of flight simulator described by Gupta et al. [19]. Given certain scene parameters, for instance, camera and scene specifications, the simulator outputs the raw measurements from a TOF camera in the presence of scattering.¹ Following from Section 3, we obtain four phase samples of the cross-correlation (cf. Equation 4), which are then used to obtain amplitude and phase images, derived using Equations 5 and 6.

Recall that our proposed algorithm requires both the direct and global components of light transport. To obtain such values in simulation we place constraints on the number of bounces. Specifically, we follow the convention from [19], where direct and global simulations constrain the maximum number of bounces to zero and four, respectively. The ground truth depth map is obtained by applying, in order, Equation 5 and then Equation 1, to the measurements from direct illumination only.

To establish the simulated experiments as a reasonable proxy for Kinect measurements, we incorporate the noise model described by Hasinoff et al [22]. In particular, we

¹Specifically, we use the following camera parameters: a spatial resolution of 512×512 pixels, a focal length of 368 pixels and an illumination frequency of 120 MHz.

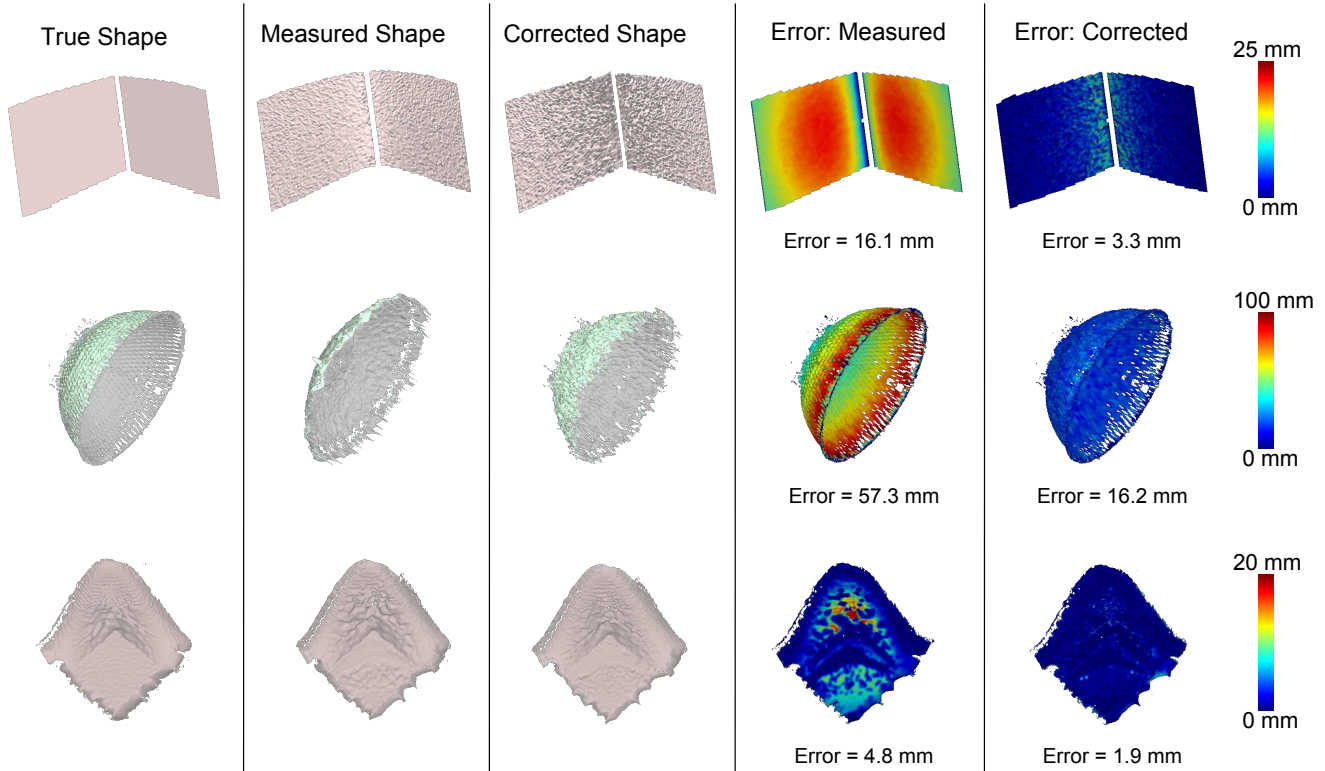


Figure 3. Our MPI correction algorithm reduces the root mean squared error on a variety of objects (simulated experiments). The first column shows the ground truth point cloud for a given object. The second column depicts the measured point cloud, simulated using rendering software. The output of our algorithm is shown in the third column. Finally, error comparisons with ground truth are shown in the fourth and fifth columns.

consider two types of additive noise: (i) scene-dependent shot noise, and (ii) scene-independent read-out noise.

5.2. Physical Implementation

All physical experiments are performed using the Microsoft Kinect One sensor, which provides amplitude and phase measurements (we use a modulation frequency of 120 MHz). To separate the direct and global components of light transport we follow the Nayar method [38], which requires high-frequency coded illumination. Instead of modifying the Kinect One to project such patterns, we use an external projector. Since the Kinect sensor has an IR-pass, VIS-block filter, we use the DLP Lightcommander projector, which can project patterns at near infrared wavelengths (850 nm).² Figure 2 summarizes the experimental setup.

We must also remark that the internal light source of the Kinect One sensor is much brighter (~ 8 times) than the external projector. Therefore, it is necessary to perform radiometric calibration. To encourage reproducibility, further details on the capture process (including radiometric calibration and details on projected patterns) can be found in

supplemental material.

Ground Truth Datasets: A key contribution of this paper lies in the evaluation of our technique in the context of ground truth datasets. To capture ground truth for the corner scene in Figure 4, we use the Kinect sensor to capture each planar component of the corner separately. Then, we fit a plane through each point cloud using least-squares, which are then fused together to obtain the ground truth 3D model. For complex scenes, as in Figure 6, we obtain ground truth by capturing a scan of the object using a multi-stripe, triangulation, laser scanner.³

6. Experimental Results

We divide our results into simulations and physical experiments. For all scenes, we evaluate our corrected depth maps using ground truth data. Following standard protocol (e.g., [48]), we report errors using the root-mean-squared error (RMSE) metric. This metric provides less weight to outliers.

²Projector website: <http://www.ti.com/tool/dlplightcommander>

³Laser Scanner Website: <http://www.nextengine.com/assets/pdf/scanner-techspecs.pdf>

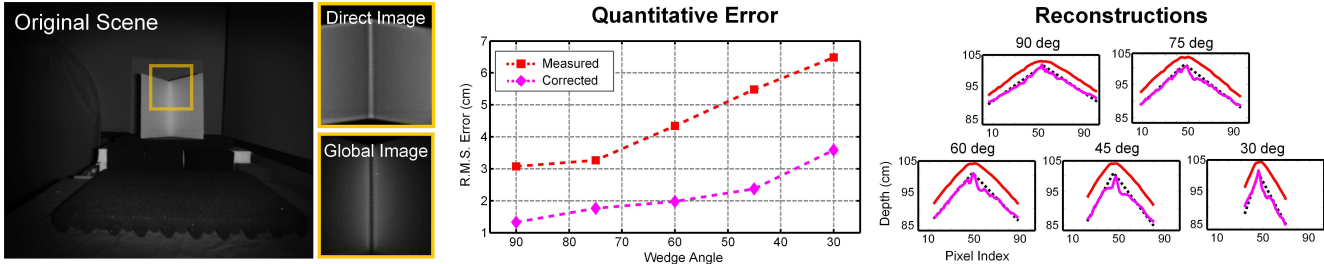


Figure 4. Correcting MPI in scenes with corners (physical experiments). Given an input scene containing a corner, we use the Nayar method to separate out direct and global components (zoom crop shown). Our correction algorithm shows a reduction in root mean squared error for corners of varying angles. For each angle, we plot a cross-section of the reconstruction. The dashed line represents ground truth, while the red and magenta lines represent measured and corrected cross-sections.

6.1. Simulations

To validate the algorithm in a controlled setting we perform simulated experiments. Ground-truth 3-D models—obtained from an online database—are illustrated in the first column of Figure 3. In particular we select three Lambertian objects: a 90 degree corner, a concave bowl, and an eigenfunction of the wave equation (i.e. the Matlab logo).

To obtain the measurements, we use a raytracing approach, as detailed in Section 5.1. The measurements are shown in the second column of Figure 3, and the error caused by MPI can be observed. In particular, MPI is most significant for the corner and bowl scenes. Both scenes encourage multipath interference in the form of interreflections—between the planes (for the corner) and inside the concave region (for the bowl). To summarize the 3-D measurements: the corner is now rounded, the bowl has been drastically flattened, and the features of the Matlab logo are slightly less distinctive.

The third column of Figure 3 shows the point cloud after applying the proposed MPI-correction algorithm. Notably, the corrected corner regains part of its original sharpness, and the bowl returns to its original geometry. Quantitative results are shown in the final two columns of Figure 3. In all cases the corrected RMSE is lower than the measured RMSE.

The simulated results also reveal a drawback of our proposed technique: the MPI-corrected results are noisy, jagged, and overall lack the smoothness of the measurements (compare columns two and three of Figure 3). However, we believe the MPI-corrected results are preferable. First, the overall error is reduced after MPI correction. Moreover, errors from additive noise can be corrected by using smoothing filters, already a common element of the depth sensing pipeline (see supplement for a smoothed result). For comparison, MPI errors cannot be solved by simple filtering (e.g., the bowl in Figure 3).

In summary, the simulated results suggest that a proper physical implementation of our technique should reduce the

RMSE on real scenes.

6.2. Experiments with Real Scenes

The results from physical experiments closely mirror the simulations, in the sense that the MPI-corrected results reduce the error in depth maps for each scene we tested. We now describe results for three separate scenes that exhibit different forms of MPI. Specifically, this includes MPI from interreflections, subsurface scattering, and geometric surface variations.

Corner Scene: Figure 4 illustrates a simple corner scene, where interreflections contribute to MPI. Following from our method, we first separate the original scene radiance into direct and global components. Note that, as expected, the direct component is brightest at the center of the corner and at the outer edges, which is where the global component is at a minimum. To make the problem a bit more interesting, we apply our correction algorithm on corners with different angles. As illustrated in Figure 4, at all angles, the measured depth is overestimated and the measured profile lacks sharpness. After applying our correction algorithm, at all angles, we report reduced error in the depth map (see Figure 4). For all angles, the corrected shape is most accurate at the center and the outer edges of the corner, since the global component is much weaker as compared to the direct component.

Our proposed technique works best on wide corners (with angles greater than 60 degrees). Note that the corner problem becomes harder as the angle becomes increasingly acute due to increasing amounts of global illumination.

Subsurface Scattering Scene: As illustrated in Figure 5, this test scene consists of a block of wax placed in front of a cardboard background. The wax material exhibits strong subsurface scattering, and therefore the expected MPI is of the form described earlier in Figure 1c. Note that subsurface multipath is a continuous phenomena—it is not possible to

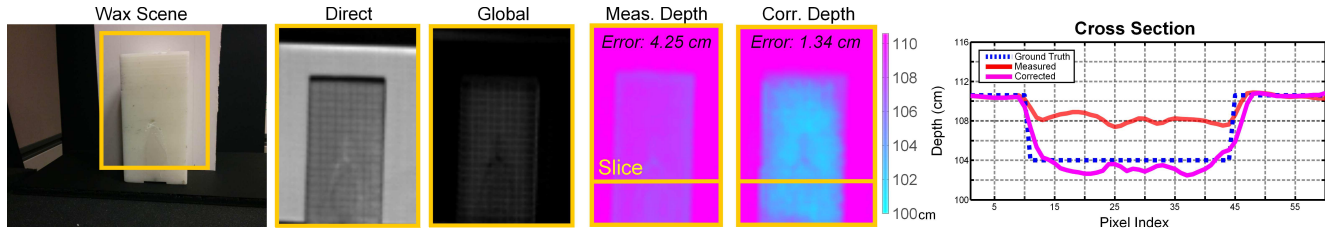


Figure 5. MPI-correction reduces error on scenes with subsurface scattering (physical experiment). The original scene (left) consists of a wax candle. The direct and global components are obtained following the Nayar method. The measured depth of the wax candle is overestimated due to the subsurface scattering. After applying our correction algorithm, we observe a 70 percent reduction in RMSE error.

provide a finite K for the number of optical reflections [43]. After separation with the Nayar method, we note that (as expected), the wax is significantly brighter than the cardboard in the global image.

After applying our correction technique to the data, we observe that the measured depth has an RMSE of about 4.25 cm, while the corrected depth has an RMSE of about 1.34 cm. Note that the error that remains in the corrected depth map is close to the depth precision of the TOF sensor (this value is about 0.5 cm [30]).

Geometric Scene: We evaluate our method on a plastic facial mannequin, which exhibits relatively complex geometry. The ground truth point cloud—obtained with a laser scanner—is illustrated in Figure 6. Since the face is made of plastic it lacks significant subsurface scattering. Therefore, we expect that the MPI should arise from concave sections in the surface geometry.

The measured shape in Figure 6 exhibits speckle like artifacts, which at first glance look like random noise. However, these artifacts arise because the plastic face has fine pits in the surface, which causes similar MPI artifacts to those observed in the large bowl from Figure 3. After applying our correction technique, the recovered shape is more representative of the surface. Quantitatively, the correction reduces the RMSE error by 39 percent. After correction, the distribution of residual error is uniform over the surface geometry, and at 5.35 mm, the error is very nearly at the depth precision of the TOF sensor [30].

7. Discussion

In summary, we present a light transport technique to mitigate MPI. Instead of relying on sparsity—an unrealistic assumption—we use the data from direct-global separation to recover depth. Our system architecture employs off-the-shelf hardware components, specifically a Kinect and a projector. For all tested scenes, the corrected point clouds demonstrate a reduction in RMSE error.

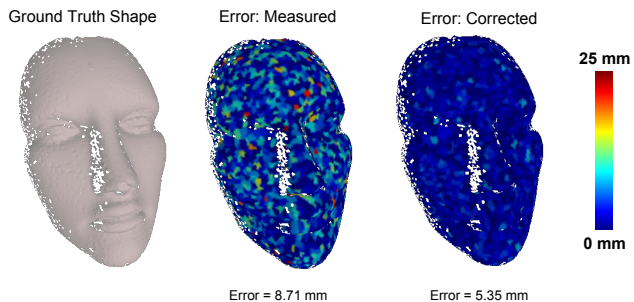


Figure 6. MPI-correction reduces error on scenes with complex geometry (physical experiment). The original scene consists of a plastic facial mannequin. At (left) is a laser scan of the object, used as ground truth. The measured shape overestimates depth in several areas. We observe a 39 percent reduction in RMSE error after applying MPI-correction.

Approximating Global Transport: The proposed technique relies on approximating global light transport with the first indirect bounce. This is a common simplification that has been exploited in prior work in light transport (see Bimber et al. [5]). Intuitively, the contribution of higher order bounces rapidly diminishes due to a combination of surface absorbance, the inverse square law, and Lambert’s law (see supplementary material for details).

Numerical Stability: Our closed form solution for MPI correction (Eq. 13) depends on the phase difference between the direct and approximate global bounce i.e. ($\varphi_D - \varphi_G$). The scene characteristics and modulation frequency determine the amplitude of this term. For low modulation frequencies, this term can get buried in the noise floor in case of small scenes, thus introducing errors in our correction. In experiments, we observe that our method is robust to noise at the modulation frequency of 120 MHz. Please see the supplement for the perturbation analysis of this problem.

Real-time Performance: Correcting MPI at full framerates (i.e. 30 Hz) is not a contribution of this paper. The key

challenge lies in the time required to separate the direct-global components of scene radiance. Specifically, while the Nayar method requires, in theory, only two photographs, we must take 25 photographs to compensate for projector bleed and defocus. In a commercial setting, a more robust implementation could be achieved by using improved projection systems or incorporating real-time direct-global separation as described in [1].

7.1. Comparisons

Our prototype combines the advantages of a light transport based approach with an implementation realized in commodity hardware. For comparison, the novel solution by O’Toole et al. [40] is tied to a particular hardware architecture, which is a combination of transient imaging hardware, beamsplitters and modified projective optics. The measurement model for transient imaging [32, 23] precludes real-time performance. In contrast, our technique relies on the Nayar model of direct-global separation [38], which has been shown—in separate work—to work in real-time [1].

We believe the recent work by Gupta et al. [19] represents a more fair comparison. In their paper they use a slightly modified time of flight sensor to generate dual-frequency measurements. To resolve MPI, they contribute novel theory, which shows that global light transport is negligible at high modulation frequencies. However, for tabletop scenes, this vanishing property requires GHz modulation frequencies. As a result, their physical experiments are constrained to large, meter-size scenes. In comparison, our method is not constrained to large scenes or high modulation frequencies. We provide results on a variety of objects, including specimens with subsurface scattering or complex geometry.

7.2. Limitations

Our method reduces error on every tested scene, but cannot be implemented in the Kinect pipeline without introducing additional limitations. In particular, to acquire the global and direct components, we use a separate projector, requiring the use of radiometric calibration. Due to the quality of the projector, the we oversample beyond Nayar’s theoretical rule (25 photographs instead of 2). In addition, the FOV of the projector is narrow, at approximately 43 degrees. Taken together, these three limitations on the projector restrict the current setup to small, static scenes; however we are confident that real-time implementations are possible with improved projective capture setups now available for direct-global separation (cf. [1]).

8. Concluding Remarks

We propose a new computational photography technique to generate higher quality 3-D scans than the stan-

dard Kinect (in the context of RMSE error). By coupling lightweight optical complexity with a closed-form, mathematical solution, the proposed technique takes a step toward scalable MPI correction.

Acknowledgments. We would like to thank Philip Chou, Mike Sinclair, Hrvoje Benko and the Microsoft Kinect product team for their helpful comments on theory and hardware implementation. We also acknowledge the help of all the members of the Interactive Visual Media group at Microsoft Research.

References

- [1] S. Achar, S. T. Nuske, and S. G. Narasimhan. Compensating for motion during direct-global separation. 2013. 8
- [2] A. Bhandari, C. Barsi, R. Whyte, A. Kadambi, A. J. Das, A. Dorrington, and R. Raskar. Coded time-of-flight imaging for calibration free fluorescence lifetime estimation. pages 2–5, 2014. 1
- [3] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. *Optics Letters*, 39(6):1705–1708, 2014. 1, 2, 3
- [4] A. Bhandari, A. Kadambi, R. Whyte, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Multifrequency time of flight in the context of transient renderings. In *ACM SIGGRAPH 2013 Posters*, page 46. ACM, 2013. 2
- [5] O. Bimber, D. Iwai, G. Wetzstein, and A. Grundhöfer. The visual computing of projector-camera systems. In *Computer Graphics Forum*, volume 27, pages 2219–2245, 2008. 7
- [6] M. Chandraker, J. Bai, T.-T. Ng, and R. Ramamoorthi. On the duality of forward and inverse light transport. *IEEE TPAMI*, 33(10):2122–2128, 2011. 3
- [7] C. Dal Mutto, P. Zanuttigh, and G. M. Cortelazzo. *Time-of-Flight Cameras and Microsoft Kinect*. Springer, 2012. 1
- [8] A. A. Dorrington, J. P. Godbaz, M. J. Cree, A. D. Payne, and L. V. Streeter. Separating true range measurements from multi-path and scattering interference in commercial range cameras. In *IS&T/SPIE Electronic Imaging*, pages 786404–786404, 2011. 1, 2, 3, 4
- [9] S. Foix, G. Alenya, and C. Torras. Lock-in time-of-flight (tof) cameras: a survey. *IEEE Sensors Journal*, 2011. 1
- [10] D. Freedman, Y. Smolin, E. Krupka, I. Leichter, and M. Schmidt. SRA: Fast removal of general multipath for ToF sensors. In *ECCV*, pages 234–249, 2014. 1, 2
- [11] S. Fuchs. Multipath interference compensation in time-of-flight camera images. In *ICPR*, 2010. 2, 4
- [12] J. P. Godbaz, M. J. Cree, and A. A. Dorrington. Closed-form inverses for the mixed pixel/multipath interference problem in amcw lidar. In *IS&T/SPIE Electronic Imaging*, 2012. 2, 4
- [13] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar. Multiplexed illumination for scene recovery in the presence of global illumination. In *ICCV*, pages 691–698, 2011. 3
- [14] M. Gupta. Shape from scatter. *Computer Vision: A Reference Guide*, pages 721–724, 2014. 3

- [15] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan. A practical approach to 3d scanning in the presence of interreflections, subsurface scattering and defocus. *IJCV*, 102(1-3):33–55, 2013. [2](#)
- [16] M. Gupta and S. G. Narasimhan. Legendre fluids: a unified framework for analytic reduced space modeling and rendering of participating media. In *ACM SIGGRAPH/Eurographics SCA*, 2007. [3](#)
- [17] M. Gupta, S. G. Narasimhan, and Y. Y. Schechner. On controlling light transport in poor visibility environments. In *CVPR*, pages 1–8. IEEE, 2008. [3](#)
- [18] M. Gupta and S. Nayar. Micro phase shifting. In *CVPR*, pages 813–820, June 2012. [2](#)
- [19] M. Gupta, S. K. Nayar, M. Hullin, and J. Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. Technical report, Jun 2014. [2](#), [4](#), [8](#)
- [20] M. Gupta, Q. Yin, and S. K. Nayar. Structured light in sunlight. In *ICCV*, pages 545–552, 2013. [3](#)
- [21] M. Hansard, O. Choi, S. Lee, and R. Horaud. *Time-of-Flight Cameras*. Springer, 2013. [1](#)
- [22] S. W. Hasinoff, F. Durand, and W. T. Freeman. Noise-optimal capture for high dynamic range photography. In *CVPR*, pages 553–560, 2010. [4](#)
- [23] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM TOG*, 32(4):45, 2013. [1](#), [8](#)
- [24] F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *CVPR*, 2014. [1](#)
- [25] M. B. Hullin, M. Fuchs, I. Ihrke, H.-P. Seidel, and H. P. Lensch. Fluorescent immersion range scanning. *ACM TOG*, 27(3):87–87, 2008. [3](#)
- [26] M. B. Hullin, J. Hanika, B. Ajdin, H.-P. Seidel, J. Kautz, and H. Lensch. Acquisition and analysis of bispectral bidirectional reflectance and reradiation distribution functions. *ACM TOG*, 29(4):97, 2010. [3](#)
- [27] M. B. Hullin, I. Ihrke, W. Heidrich, T. Weyrich, G. Damberg, and M. Fuchs. Computational fabrication and display of material appearance. In *Eurographics State of the Art Reports*, pages 137–153, 2012. [3](#)
- [28] M. B. Hullin, H. Lensch, R. Raskar, H.-P. Seidel, and I. Ihrke. Dynamic display of brdfs. In *Computer Graphics Forum*, pages 475–483, 2011. [3](#)
- [29] D. Jiménez, D. Pizarro, M. Mazo, and S. Palazuelos. Modeling and correction of multipath interference in time of flight cameras. *Image and Vision Computing*, 32(1):1–13, 2014. [2](#)
- [30] A. Kadambi, A. Bhandari, and R. Raskar. 3d depth cameras in vision: Benefits and limitations of the hardware. In *Computer Vision and Machine Learning with RGB-D Sensors*, pages 3–26. Springer, 2014. [7](#)
- [31] A. Kadambi, A. Bhandari, R. Whyte, A. Dorrington, and R. Raskar. Demultiplexing illumination via low cost sensing and nanosecond coding. In *ICCP*, 2014. [1](#)
- [32] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM TOG*, 32(6):167, 2013. [1](#), [2](#), [8](#)
- [33] A. Kirmani, A. Benedetti, and P. A. Chou. Spumic: Simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods. In *IEEE International Conference on Multimedia and Expo*, pages 1–6, 2013. [2](#), [4](#)
- [34] N. Naik, C. Barsi, A. Velten, and R. Raskar. Estimating wide-angle, spatially varying reflectance using time-resolved inversion of backscattered light. *JOSA A*, 31(5):957–963, 2014. [1](#)
- [35] N. Naik, S. Zhao, A. Velten, R. Raskar, and K. Bala. Single view reflectance capture using multiplexed scattering and time-of-flight imaging. In *ACM TOG*, volume 30, page 171, 2011. [1](#), [3](#)
- [36] S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal. Structured light in scattering media. In *ICCV*, pages 420–427, 2005. [3](#)
- [37] S. K. Nayar and M. Gupta. Diffuse structured light. In *ICCP*, pages 1–11. IEEE, 2012. [3](#)
- [38] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast separation of direct and global components of a scene using high frequency illumination. In *ACM TOG*, volume 25, pages 935–944. ACM, 2006. [2](#), [3](#), [4](#), [5](#), [8](#)
- [39] T.-T. Ng, R. S. Pahwa, J. Bai, T. Q. Quek, and K.-H. Tan. Radiometric compensation using stratified inverses. In *ICCV*, 2009. [2](#)
- [40] M. O’Toole, F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos. Temporal frequency probing for 5D transient analysis of global light transport. *ACM TOG*, 33(4), 2014. [1](#), [2](#), [8](#)
- [41] M. O’Toole, J. Mather, and K. N. Kutulakos. 3d shape and indirect appearance by structured light transport. In *CVPR*, 2014. [2](#)
- [42] M. O’Toole, R. Raskar, and K. N. Kutulakos. Primal-dual coding to probe light transport. *ACM Trans. Graph.*, 31(4):39, 2012. [3](#)
- [43] R. Raskar and J. Davis. 5D time-light transport matrix: What can we reason about scene properties. *Int. Memo*, 2008. [7](#)
- [44] D. Raviv, C. Barsi, N. Naik, M. Feigin, and R. Raskar. Pose estimation using time-resolved inversion of diffuse light. *Optics express*, 22(17):20164–20176, 2014. [1](#)
- [45] D. Reddy, R. Ramamoorthi, and B. Curless. Frequency-space decomposition and acquisition of light transport under spatially varying illumination. In *ECCV*, 2012. [3](#)
- [46] S. M. Seitz, Y. Matsushita, and K. N. Kutulakos. A theory of inverse light transport. In *ICCV*, 2005. [2](#)
- [47] J. Sell and P. O’Connor. The xbox one system on a chip and kinect sensor. *IEEE Micro*, 34(2):44–53, Mar 2014. [1](#)
- [48] T. Weise, H. Li, L. Van Gool, and M. Pauly. Face/off: Live facial puppetry. In *ACM SIGGRAPH/Eurographics SCA*, pages 7–16. ACM, 2009. [5](#)
- [49] D. Wu, G. Wetzstein, C. Barsi, T. Willwacher, M. O’Toole, N. Naik, Q. Dai, K. Kutulakos, and R. Raskar. Frequency analysis of transient light transport with applications in bare sensor imaging. In *ECCV*, pages 542–555. 2012. [3](#)
- [50] L.-Q. Yan, M. Hašan, W. Jakob, J. Lawrence, S. Marschner, and R. Ramamoorthi. Rendering glints on high-resolution normal-mapped specular surfaces. *ACM TOG*, 33(4):116, 2014. [3](#)