



Optics Letters

April 2014, Volume 39, Issue 06.

<http://dx.doi.org/10.1364/OL.39.001705>

Resolving Multi-path Interference in Time-of-Flight Imaging via Modulation Frequency Diversity and Sparse Regularization

Ayush Bhandari

Massachusetts Institute of Technology

Achuta Kadambi

Massachusetts Institute of Technology

Refael Whyte

University of Waikato and

Massachusetts Institute of Technology

Christopher Barsi

Massachusetts Institute of Technology

Micha Feigin

Massachusetts Institute of Technology

Adrian Dorrrington

University of Waikato

Ramesh Raskar

Massachusetts Institute of Technology

Abstract

Time-of-flight (ToF) cameras calculate depth maps by reconstructing phase shifts of amplitude-modulated signals. For broad illumination or transparent objects, reflections from multiple scene points can illuminate a given pixel, giving rise to an erroneous depth map. We report here a sparsity regularized solution that separates K interfering components using multiple modulation frequency measurements. The method maps ToF imaging to the general framework of spectral estimation theory and has applications in improving depth profiles and exploiting multiple scattering.

Keywords: Depth imaging, multi-path interference, sparse regularization, time-of-flight imaging.

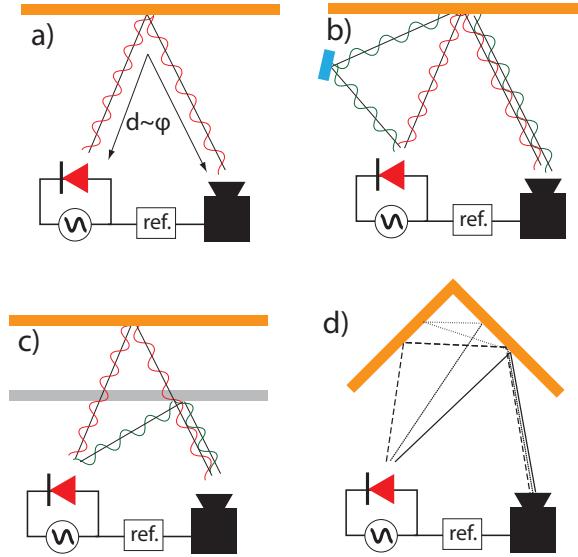


Figure 1: (a) ToF principle: the phase delay of an emitted AMCW wave proportionally encodes the distance of the reflecting object. (b) Mirror-like and (c) semi-transparent reflections produce MPI at a given camera pixel and yields an incorrect phase. (c) A complicated scene with severe MPI.

Optical ranging and surface profiling have widespread applications in image-guided surgery [5], gesture recognition [4], remote sensing [1], shape measurement [7], and novel phase imaging [17]. Generally, the characteristic wavelength of the probe determines the resolution of the image, making time-of-flight (ToF) methods suitable for macroscopic scenes [10, 18, 22]. Although ToF sensors can be implemented with impulsive sources, commercial ToF cameras rely on the continuous wave approach: the source intensity is modulated at radio frequencies (\sim 10s of MHz), and the sensor reconstructs the phase shift between the reflected and emitted signals. Distance is calculated by scaling the phase by the modulation frequency (Fig. 1 (a)). This method, amplitude modulated continuous wave (AMCW) ToF, offers high SNR in real time.

However, AMCW ToF suffers from multipath interference (MPI) [3, 8, 9, 12, 14–16, 20, 21]. Consider, for example, the scenes in Figs. 1 (b,c). Light rays from multiple reflectors scatter to the observation point. Each path acquires a different phase shift, and the measurement consists of the sum of these components. The recovered phase, therefore, will be incorrect. Such “mixed” pixels contain depth errors and arise whenever global lighting effects exist. In some cases (Fig. 1 (d)), the measurement comprises a continuum of scattering paths. This can be improved with structured light or mechanical scanning [6, 11], but these are limited by the source resolution. Computational optimization [13, 19] schemes rely on radiometric assumptions and have limited applicability.

Here, we resolve MPI via sparse regularization of multiple modulation frequency measurements. The formulation allows us to recast this problem into the general framework of spectral estimation theory [23]. This contribution generalizes the two-component, dual-frequency approach [8, 15, 16], beyond which the two-component optimization methods fail. Thus, our method here has two significant benefits. First, we separate MPI from direct illumination to produce improved depth maps. Second, we resolve MPI into its components, so that we can characterize and exploit multiple scattering phenomena. The procedure has two steps: (1) record a scene with multiple modulation frequencies and (2) reconstruct the MPI components using a sparsity constraint.

Consider first the single-component case. Mathematically, the camera emits the normalized time-modulated intensity $s(t)$ ¹ and detects a signal $r(t)$:

$$s(t) = 1 + s_0 \cos(\omega t), t \in \mathbb{R} \quad (1a)$$

$$r(t) = \Gamma(1 + s_0 \cos(\omega t - \phi)). \quad (1b)$$

Here, s_0 and $\Gamma \in [0, 1]$ are the signal modulation depth and the reflection amplitude, respectively, ω is the modulation frequency, and ϕ is the phase delay between the reference waveform $s(t)$ and the delayed version $r(t)$. For a co-located source and detector, the distance to the object from the camera is given by the relation $d = c\phi/2\omega$, where c is the speed of light.

Electronically, each pixel acts as a homodyne detector, measuring the cross-correlation between the reflected signal and the reference. Denoting the complex conjugate of $f \in \mathbb{C}$ by f^* , the cross-correlation of two functions f and g is

$$\mathcal{C}_{f,g}(\tau) \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} f^*(t + \tau) g(t) dt. \quad (2)$$

Note that infinite limits are approximately valid when the integration window $2T$ is such that $T \gg \omega^{-1}$. A shorter time window produces residual errors, but this is easily avoidable in practice. The pixel samples the cross-correlation at discrete times τ_q :

$$m_\omega[q] \stackrel{\text{def}}{=} \mathcal{C}_{s,r}(\tau_q) \stackrel{(2)}{=} \Gamma \left(1 + \frac{s_0^2}{2} \cos(\omega \tau_q + \phi) \right). \quad (3)$$

Using the “4 Bucket Sampling” technique [10], we calculate the estimated reflection amplitude and the

¹Here, we consider continuous wave imaging and hence the sinusoidal model, but the discussion is generally applicable to any periodic function.

phase, $\tilde{\Gamma}, \tilde{\phi}$, using four samples $\tau_q = \pi q / 2\omega$ with $q = 0, \dots, 3$:

$$\tilde{\Gamma} = \sqrt{(m_\omega[3] - m_\omega[1])^2 + (m_\omega[0] - m_\omega[2])^2} / s_0^2, \quad (4a)$$

$$\tan \tilde{\phi} = \left(\frac{m_\omega[3] - m_\omega[1]}{m_\omega[0] - m_\omega[2]} \right). \quad (4b)$$

Therefore, we associate a complex value, z_ω , with a pixel measurement:

$$z_\omega = \tilde{\Gamma} e^{j\tilde{\phi}(\omega)}. \quad (5)$$

Note that these results are formally equivalent to wavefront reconstruction via phase-shifting digital holography [25].

When multiple reflections contribute to a single measurement, the return signal comprises a sum. In phasor notation, for K components,

$$r(t) = C_0 + \sum_{k=0}^{K-1} \Gamma_k e^{j(\omega t - \phi_k(\omega))}, \quad (6)$$

where C_0 is a constant, $\phi_k(\omega) = 2d_k\omega/c$, and $\{d_k\}_{k=0}^{K-1}$ are K depths at which the corresponding reflection takes place. The reflection amplitude of the k^{th} surface is Γ_k . Each pixel records

$$m_\omega^K[q] = C_0 + \frac{s_0^2}{2} e^{j\omega\tau_q} \sum_{k=0}^{K-1} \Gamma_k e^{j\phi_k(\omega)}. \quad (7)$$

Importantly, for a given modulation frequency ω_0 (ignoring a constant DC term), $m_{\omega_0}^K[\tau_q] \propto \exp j\omega_0\tau_q$, i.e., there is no variation with respect to individual depth components $\{\Gamma_k(\omega), \phi_k\}_{k=0}^{K-1}$ [3], regardless of the sampling density. Equivalently, the camera measurement,

$$z_\omega^{(K)} = \tilde{\Gamma}(\omega) e^{j\tilde{\phi}(\omega)} = \sum_{k=0}^{K-1} \Gamma_k(\omega) e^{j\phi_k(\omega)} \quad (8)$$

is now a complex sum of K reflections, which cannot be separated without independent measurements. Thus, at a given frequency, the measured phase, and hence the depth, is a nonlinear mixture of all interfering components.

Our method separates these components by recording the scene with equi-spaced frequencies $\omega = n\omega_0$ ($n \in \mathbb{N}$) and acquiring a set of measurements \mathbf{z} :

$$\mathbf{z} = \left[z_{\omega_0}^{(K)}, z_{2\omega_0}^{(K)}, \dots, z_{N\omega_0}^{(K)} \right]^\top. \quad (9)$$

The forward model can be written compactly in vector-matrix form as $\mathbf{z} = \Phi \mathbf{g} + \boldsymbol{\sigma}$, where $\Phi \in \mathbb{C}^{N \times K}$ is identified as a Vandermonde matrix,

$$\Phi = \begin{pmatrix} e^{j\omega_0\phi_0} & e^{j\omega_0\phi_1} & \dots & e^{j\omega_0\phi_{K-1}} \\ e^{j2\omega_0\phi_0} & e^{j2\omega_0\phi_1} & \dots & e^{j2\omega_0\phi_{K-1}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{jN\omega_0\phi_0} & e^{jN\omega_0\phi_1} & \dots & e^{jN\omega_0\phi_{K-1}} \end{pmatrix}, \quad (10)$$

$\mathbf{g} = [\Gamma_0, \dots, \Gamma_{K-1}]^\top \in \mathbb{R}^{K \times 1}$, and $\boldsymbol{\sigma}$ represents zero-mean Gaussian i.i.d. noise, which controls the error ε_0 in our reconstruction algorithm. Our goal is to estimate the phases $\boldsymbol{\phi} = [\phi_0, \dots, \phi_{K-1}]^\top \in \mathbb{R}^{K \times 1}$ and the reflection amplitude vector \mathbf{g} .

To recover these quantities, first note the similarity between Φ and an oversampled $N \times L$ discrete Fourier transform (DFT) matrix Ψ , with elements $\Psi_{nl} = \exp(jnl/L)$. If $L \gg K$, the discretization of Ψ is small enough to assume that the columns of Φ are contained in Ψ . We can also define a vector $\mathbf{g}' \in \mathbb{R}^{L \times 1}$, whose elements are zero except for K reflection amplitudes $\{\Gamma_k\}_{k=0}^{K-1}$, such that $\mathbf{z} = \Psi \mathbf{g}'$. We use the (K)-sparsity of \mathbf{g}' to regularize the problem:

$$\underbrace{\|\mathbf{z} - \Psi \mathbf{g}'\|_{\ell_2}^2}_{\text{Data-Fidelity}} < \varepsilon_0 \quad \text{such that} \quad \underbrace{\|\mathbf{g}'\|_{\ell_0}}_{\text{Sparsity}} = K, \quad (11)$$

where the ℓ_p -norm is $\|\mathbf{x}\|_{\ell_p}^p \stackrel{\text{def}}{=} \sum_n |x_n|^p$. The case of $p \rightarrow 0$ is used to define $\|\mathbf{g}'\|_{\ell_0}$ as the number of nonzero elements of \mathbf{g}' . Eq. 11 demands a least-squares solution to the data-fidelity problem $\|\mathbf{z} - \Psi \mathbf{g}'\|_{\ell_2}^2$ up to some error tolerance ε_0 , with the constraint that we accommodate up to K nonzero values of \mathbf{g}' .

The sparsity of \mathbf{g}' arises from two underlying assumptions. First, we do not consider the case of volumetric scattering, which would preclude discrete reflections and require a different parametrization (e.g., through the diffusion coefficient). Second, we ignore the contributions of inter-reflections between scattering layers, as their amplitudes fall off quickly. They could be incorporated, into our formulation, with the result of changing the sparsity of \mathbf{g}' from K to K' , where $K' - K$ is the number of inter-reflections considered.

We solve Eq. 11 via orthogonal matching pursuit (OMP), which is an iterative algorithm that searches for the best-fit projections (in the least-squares sense) of the coefficients onto an over-complete dic-

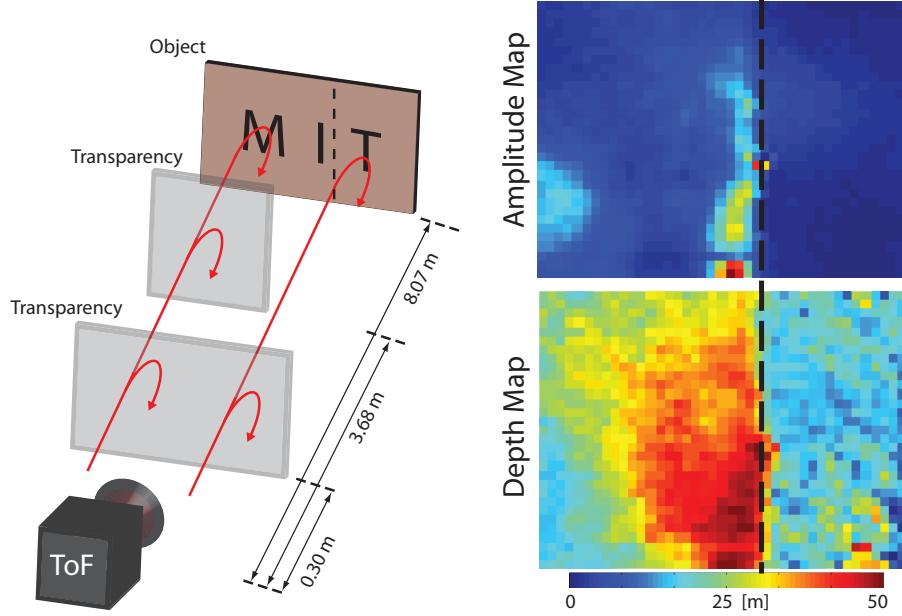


Figure 2: Left: experimental setup. Two transparencies block the left side of the camera (for a three-component measurement), and one transparency blocks the right (two-component measurement). Right: measured amplitude and depth at $\omega = 3\omega_0$. Dashed line indicates edge of second transparency.

tional. We input Ψ and measurements \mathbf{z} into the algorithm. The outputs are the set of reflection coefficients Γ_k and their positions in \mathbf{g}' . With the position of each Γ_k reconstructed, the corresponding phases ϕ_k are recovered through the elements of Ψ : $\phi_k = (jn)^{-1} \log(\Psi_{nl_k}) = l_k/L$, where l_k is the location of Γ_k in \mathbf{g}' .

We verify this theory with the experimental setup shown in Fig. 2. A PMD19k-2 160×120 sensor array is controlled by a Stratix III FPGA. Analog pixel values are converted to 16-bit unsigned values by an ADC during the pixel readout process. Eight 100 mW Sony SLD 1239JL-54 laser diodes illuminate the scene. The lasers are placed symmetrically around the detector for a coaxial configuration. The base frequency modulation is $f_0 = \omega_0/(2\pi) = 0.7937$ MHz, and the integration time is 47 ms. The scene consists of three layers. Farthest, at 8.1 m, is an opaque wall with gray-scale text (“MIT”) printed on it. Closest, at 0.3 m is a semi-transparent sheet. Between the two layers is another semi-transparent sheet that covers only the left half of the field of view. Therefore, the left-hand side records three bounces and the right only two. All three layers are within the depth of field of the camera to avoid mixed pixels from blurring.

Depth and amplitude maps acquired at a specific frequency are shown in Fig. 2. Due to MPI,

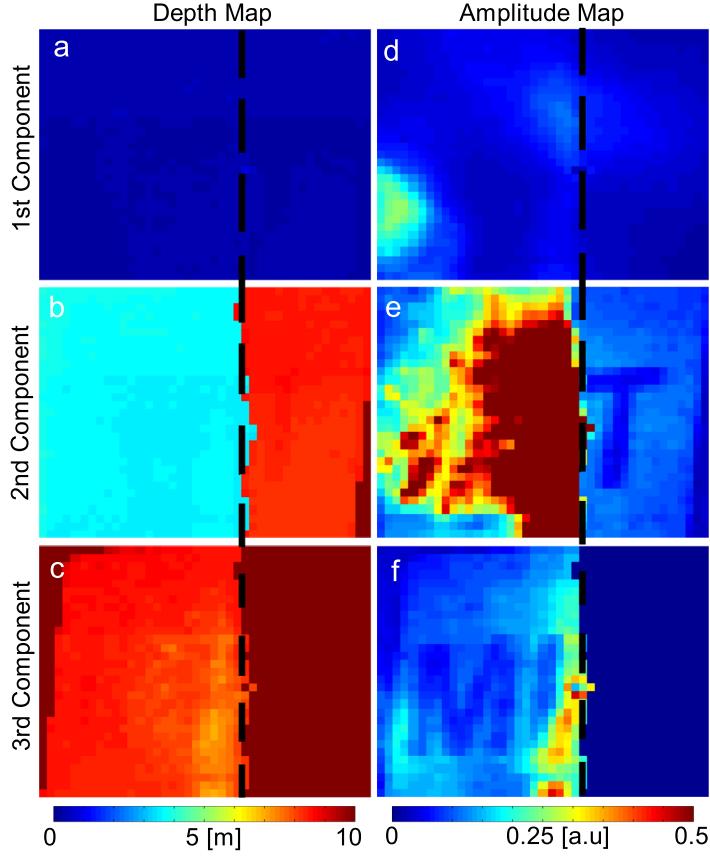


Figure 3: Reconstructed amplitudes and depths via sparse regularization. Dashed lines indicate edge of second transparency.

the measured depths do not correspond to any physical layer in the scene. All depth and amplitude information from the three scene layers is mixed nonlinearly into a set of composite measurements (pixels) and cannot be recovered.

We repeat the acquisition 77 times, with modulation frequencies spaced 0.7937 MHz apart and input these data into the OMP algorithm with $K = 3$. The reconstruction, shown in Fig. 3, shows each depth correctly recovered. The closest depth map (Fig. 3 (a), first transparency) is constant. The second map (Fig. 3 (b)) contains two depths: the second transparency on the LHS and the wall on the RHS. The third depth map contains the wall depth on the LHS (Fig. 3 (c)). The third-bounce amplitude (Fig. 3 (f)) is zero where there are only two layers (RHS). The depth here is therefore undefined, though we set the distance to be 10 m to avoid random fluctuations. Further, the text is recovered properly in the amplitude maps corresponding to the correct depths (Figs. 3 (e,f)). Note that accurate depths are recovered even in the presence of strong specularity (Fig. 3 (e)).

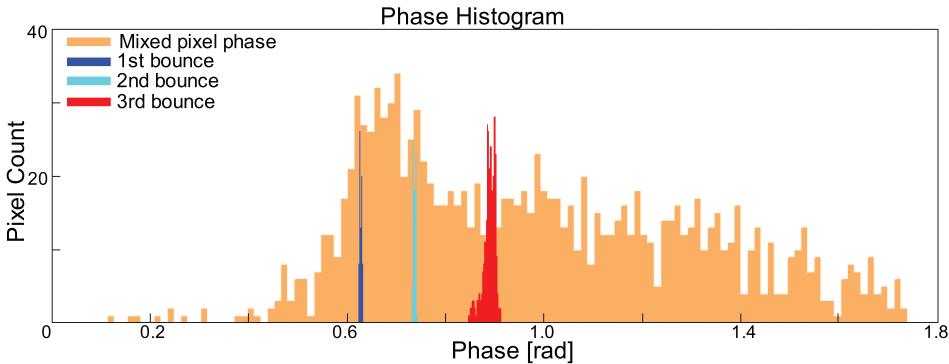


Figure 4: Phase histogram for reconstructed and measured depth maps. Reconstructed phases cluster around the correct depths, whereas the measured depth map has a wide variance across the entire range.

A phase histogram is shown in Fig. 4. The histogram from the single frequency measurement in Fig. 1 varies from 0.6 to 1.8 rad. Recovered phases are centered around the ground truth values. The third-phase variance is wider because OMP computes the first two components, leaving little residual energy, so that several columns in Ψ can minimize the least-squares error.

In principle, the technique can be extended to any number of bounces, provided enough modulation frequencies are used (though a first-principles derivation is beyond the scope of this contribution). In practice, however, the reflected amplitudes decrease with increasing component number, so that higher-order components diminish in importance. Furthermore, OMP need not assume a number of components that is the same as that of the physical implementation. If the assumed number is greater than the physical number, OMP will reconstruct all the physical components, with higher-order ones having an amplitude on order of the system noise. Conversely, if the assumed number is less than the physical number, OMP will recover the strongest reflections.

Therefore, the method is a generalization of global/direct illumination separation and can decompose different elements of global lighting. This is useful not only for improved depth accuracy, but also imaging in the presence of multiple scatterers such as diffuse layers, sediment, turbulence, and turbid media, as well as in places where third-component scattering must be extracted [24]. Furthermore, because it is based on phase measurements, this technique can be mapped to multiple scattering in holography [2] by substituting optical frequency for the modulation frequency.

In conclusion, we implemented a multi-frequency approach for decomposing multiple depths for a ToF camera. The result is general and holds for any number of bounces, and it can be extended to non-harmonic signals [21]. Future work includes calculating bounds on measurements and resolution. The method can be incorporated with structured illumination and pixel correlations and for edge detection,

and refocusing. The result holds promise for mitigating and exploiting multipath for a wide variety of scenes.

References

- [1] M. C. Amann, T. Boch, R. Myllyla, M. Rioux, and M. Lescure. Laser ranging: a critical review of usual techniques for distance measurement. *Opt. Eng.*, 40:10–19, 2001.
- [2] J. J. Barton. Removing multiple scattering and twin images from holographic images. *Phys. Rev. Lett.*, 67:3106–3109, 1991.
- [3] Ayush Bhandari, Achuta Kadambi, Refael Whyte, Lee Streeter, Christopher Barsi, Adrian Dorrrington, and Ramesh Raskar. Multifrequency time of flight in the context of transient renderings. In *ACM SIGGRAPH 2013 Posters*, number 46, 2013.
- [4] Pia Breuer, Christian Eckes, and Stefan Müller. Hand gesture recognition with a novel IR time-of-flight range camera—a pilot study. In *Computer Vision/Computer Graphics Collaboration Techniques*, pages 247–260. Springer, 2007.
- [5] D. M. Cash, T. K. Sinha, W. C. Chapman, H. Terawaki, B. M. Dawant, R. L. Galloway, and M. I. Miga. Incorporation of a laser range scanner into image-guided liver surgery: surface acquisition, reigstration, and tracking. *Med. Phys.*, 30:1671–1682, 2003.
- [6] S. Y. Chen, Y. F. Li, and J. W. Zhang. Vision processing for realtime 3d data acquisition based on coded structured light. *IEEE Trans. Image Proc.*, 17:167–176, 2008.
- [7] Y. Cui, S. Schoun, D. Chan, S. Thrun, and C. Theobalt. 3d shape scanning with a time-of-flight camera. In *Proc. Computer Vision and Pattern Recognition*, 2010.
- [8] A.A. Dorrrington, 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.
- [9] D. Droeschel, D. Holz, and S. Behnke. Multi-frequency phase unwrapping for time-of-flight cameras. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1463–1469, 2010.
- [10] Sergi Foix, Guillem Alenya, and Carme Torras. Lock-in time-of-flight (tof) cameras: a survey. *IEEE Sensors Journal*, 11(9):1917–1926, 2011.
- [11] Active Sensor Plnanning for Multiview Vision Tasks. *Chen, S. Y.* Springer, 2008.
- [12] Mario Frank, Matthias Plaue, Holger Rapp, Ullrich Köthe, Bernd Jähne, and Fred A Hamprecht. Theoretical and experimental error analysis of continuous-wave Time-of-Flight range cameras. *Proc. SPIE Conf. on Vis. Commun. and Image Proc.*, 48(1):013602—013602, 2009.
- [13] S. Fuchs. Multipath interference compensation in time-of-flight camera images. In *Proc. Computer Vision and Pattern Recognition*, 2010.
- [14] John P Godbaz, Michael J Cree, and Adrian A Dorrrington. Understanding and ameliorating non-linear phase and amplitude responses in amcw lidar. *Remote Sensing*, 4(1):21–42, 2011.
- [15] John P Godbaz, Adrian A Dorrrington, and Michael J Cree. Understanding and ameliorating mixed pixels and multipath interference in amcw lidar. In *TOF Range-Imaging Cameras*, pages 91–116. Springer, 2013.
- [16] J.P. Godbaz, M.J. Cree, and A.A. Dorrrington. Closed-form inverses for the mixed pixel/multipath interference problem in amcw lidar. In *IS&T/SPIE Electronic Imaging*, pages 829618–829618, 2012.

- [17] J. C. Halimeh and M. Wegener. Time-of-flight imaging of invisibility cloaks. *Opt. Express*, 20:63–74, 2012.
- [18] Miles Hansard, Seungkyu Lee, Ouk Choi, and Radu Horaud. *Time-of-flight cameras: principles, methods and applications*. Springer, 2013.
- [19] D. Jimenez, D. Pizarro, M. Mazo, and S. Palazuelos. Modelling and correction of multipath interference in time of flight cameras. In *Proc. Computer Vision and Pattern Recognition*, 2012.
- [20] A. P P Jongenelen, D. G. Bailey, A. D. Payne, A. A. Dorrrington, and D. A. Carnegie. Analysis of errors in tof range imaging with dual-frequency modulation. *IEEE Trans. on Instrumentation and Measurement*, 60(5):1861–1868, 2011.
- [21] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. A. Dorrrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Trans. Graph.*, to appear.
- [22] Andreas Kolb, Erhardt Barth, Reinhard Koch, and Rasmus Larsen. Time-of-Flight sensors in computer graphics. In *Proc. Eurographics*, pages 119–134, 2009.
- [23] P. Stoica and R. L. Moses. *Introduction to Spectral Analysis*. Prentice Hall, 1997.
- [24] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, M. G. Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nat. Commun.*, 3:745, 2012.
- [25] T. Yamaguchi, I. and Zhang. Phase shifting digital holography. *Opt. Lett.*, 22:1268, 1997.

Affiliation:

Ayush Bhandari
Media Laboratory, 75 Amherst St.
Massachusetts Institute of Technology
02139, Cambridge. USA
E-mail: Ayush.Bhandari@googlemail.com, ayush@MIT.edu
URL: <http://mit.edu/~ayush>

Achuta Kadambi
Media Laboratory, 75 Amherst St.
Massachusetts Institute of Technology
02139, Cambridge. USA
E-mail: achoo@MIT.edu

Refael Whyte
School of Engineering, University of Waikato
Private Bag 3105
Hamilton 3240. New Zealand
E-mail: rzw2@students.waikato.ac.nz

Christopher Barsi
Media Laboratory, 75 Amherst St.
Massachusetts Institute of Technology
02139, Cambridge. USA
E-mail: cbarsi@MIT.edu

Micha Feigin
Media Laboratory, 75 Amherst St.
Massachusetts Institute of Technology
02139, Cambridge. USA
E-mail: michaf@MIT.edu

Adrian Dorrington
School of Engineering, University of Waikato
Private Bag 3105
Hamilton 3240. New Zealand
E-mail: adrian@waikato.ac.nz

Ramesh Raskar
Media Laboratory, 75 Amherst St.
Massachusetts Institute of Technology
02139, Cambridge. USA
E-mail: raskar@MIT.edu

Author's Manuscript
Appears with changes in Optics Letters.
Volume 39, Issue 06
April 2014

www.mit.edu/~ayush
<http://dx.doi.org/10.1364/OL.39.001705>
Submitted: 20XX-XX-XX
Accepted: 20XX-XX-XX