### LINEAR SYSTEMS APPROACH TO IDENTIFYING PERFORMANCE BOUNDS IN INDIRECT IMAGING

Adithya Kumar Pediredla<sup>\*</sup>

Nathan Matsuda<sup>†</sup>

Oliver Cossairt<sup>†</sup>

Ashok Veeraraghavan\*

\* Rice University † Northwestern University

#### ABSTRACT

Light scattering on diffuse rough surfaces was long assumed to destroy geometry and photometry information about hidden (non line of sight) objects making 'looking around the corner' (LATC) and 'non line of sight' (NLOS) imaging impractical. Recent work pioneered by Kirmani et al. [1], Velten et al. [2] demonstrated that transient information (time of flight information) from these scattered third bounce photons can be exploited to solve LATC and NLOS imaging. In this paper, we quantify the geometric and photometric reconstruction limits of LATC and NLOS imaging for the first time using a classical linear systems approach. The relationship between the albedo of the voxels in a hidden volume to the third bounce measurements at the sensor is a linear system that is determined by the geometry and the illumination source. We study this linear system and employ empirical techniques to find the limits of the information contained in the third bounce photons as a function of various system parameters.

#### 1. INTRODUCTION

We can identify and localize sounds originating around a corner but cannot do the same with light. This is due to the fact that the wavelength of sound is large relative to the occluding geometry and thus reflections are effectively fully specular for typical hard surfaces such as walls[3, 4]. Light on the other hand, with its much smaller wavelengths, undergoes random phase modulation from macroscopic occluders. Nonetheless, multi-bounce photons do travel from a hidden object to an around-the-corner observer. It has recently be demonstrated that temporal information can be exploited to resolve spatial information [1, 5, 6, 7, 8] hidden from direct view, enabling us to build imaging systems capable of 'looking around the corner' (LATC) and 'non line of sight' (NLOS) imaging.

While there is a good understanding on how to computationally focus the rays around corners [1, 2, 7, 9], very little is known about the effect of various parameters on the reconstruction. In this paper, we show that linear system analysis can be performed to quantitatively answer these questions and establish bounds on the reconstruction performance of a given system (number of transient imagers, temporal resolution, Signal to noise ratio (SNR), number of measurements).



**Fig. 1. Problem Setup:** A hidden object resides in a known cubic volume with side length 500 mm. A thick (fully occluding) wall separates the unknown object from the source (S) and the detector (D), which are bounded within a similar cubic volume of 500 mm on a side. A scatterer (W) allows photons to pass between the source-detector setup and the hidden object. All virtual sources (S') and virtual detectors (D') must lie on the surface of this scattering wall. We will computationally find the performance bounds for this setup.

#### 1.1. Problem Setup

We consider an example scenario of interest (See Figure 1). We assume unknown object (O) with unknown texture, geometry and reflectance is bounded by a known volume. We assume that the known volume is a cube 500 mm on a side. It is further assumed that a thick wall occludes the hidden object from the source and detector. The source (S) and the detector (D) are restricted within another bounded volume, again of the same size: i.e., a cube of size 500 mm. A scattering object (W) is simultaneously visible to the hidden object, the source and detector. All virtual sources and virtual detectors are assumed to lie on the surface of this scattering object, which may not be planar. The entire geometry of the challenge problem is illustrated in the Figure 1. We are interested in identifying performance bounds for geometric and photometric recovery given this scene geometry.



**Fig. 2.** Mutual coherence and spectral analysis of transient radiance transport matrix: We show the mutual coherence (lower is better) and singular values (higher is better) of transport matrix A for various configurations of temporal resolution and detector count with 80k measurements. (a-c) Notice that the transport matrix is well conditioned as the temporal resolution or the number of virtual detectors is increased. Therefore, systems with improved transient resolution and number of virtual detectors capture more LATC information.

#### 1.2. Contributions

- Upper bound on geometric reconstruction: We show that a point scatterer can be localized up to 10 microns using streak camera (2ps ~ .6mm), 100-500 microns for Single Photon Avalanche Diode (SPAD) (50 ps ~ 15 mm) and Intensified Charged Coupled Device (ICCD) (200 ps ~ 60 mm) and around 1 mm using Photonic Mixer Device (PMD) (1000 ps ~ 300 mm) with 80k measurements. This demonstrates that when the hidden scene is made up of individual point scatterers that are far enough from each other one can improve localization accuracy far beyond the limit suggested by the transient temporal resolution.
- Lower bound on photometric reconstruction: We empirically compute the lower bound on the albedo error. Any prior-information will increase performance.

#### 2. PRIOR WORK

Transient imaging [10, 11], whereby picosecond resolution timing information about light incident on a photosensor is encoded spatially, has long been used to mitigate the effect of scattering in biological microscopy [12]. More recently, with the advent of computational imaging techniques, Kirmani, et al. [1] demonstrated the possibility of third bounce imaging using pulsed laser source and picosecond resolution photodetector. Subsequent work by Velten, et al. [2] achieved sub-millimeter localization precision within a 40 cm cubic volume. A variety of work has explored alternative hardware configurations targeting cost or simplicity. Heide, et al. [7] demonstrated that third bounce imaging is feasible with significantly cheaper time-of-flight sensors, albeit with lower precision on the order of centimeters. Further sensor alternatives were proposed by Katz, et al. [13], who utilized a spatial light modulator to invert the phase modulation induced by the scattering wall, allowing the hidden object to be imaged by a conventional CCD camera - an approach suggested prior to

the availability of appropriate phase modulation and reconstruction techniques [14]. Laurenzis, et al. [6], [15] evaluate single-photon avalanche diode devices for use in third bounce imaging setups, including the possibility of shortwave infrared illumination. More generally, recent work in other imaging modalities has considered similar problems, such as detecting and reconstructing human forms from scattered RF signals [16],[17], photoacoustic wavefront manipulation for imaging in scattering biological tissue [18],[19], or utilizing volumetric random scatterers for synthetic aperture imaging [20].

In terms of establishing bounds, Kadambi, et al. [8] seek to determine resolution bounds on third-bounce imaging with time-of-flight cameras relative to the BRDF the scattering wall. Their resolution bounds are based on full-width-halfmaxima (FWHM) of the blur size. As the authors note, these bounds can be surpassed for sparse objects to arbitrarily high resolution, limited only by signal-to-noise ratio (SNR), and the sparsity of the object. In fact, for Velten et al. [2] paper, the predicted FWHM is 5mm, but the sub-mm resolution was achieved by using sparsity constraints.

#### 3. LINEAR SYSTEM ANALYSIS OF TRANSIENT RADIANCE TRANSPORT

For geometric (point localization) problems, the transient transport matrix A should have low mutual coherence [8]. This will ensure that the point scatterer we are trying to localize has a good margin between the point scatterer we are trying to localize, and the neighboring points. However, mutual coherence of the linear system only explains the effect of changing the configuration, but does not quantify the geometric error. In Section 4, we derive the geometric error. For photometric (voxel albedo recovery) problem, the transient transport matrix should be invertible.

The analysis of photometric reconstruction performance should be based on the mutual information (MI) or the meansquared-error (MSE) criteria. The MI criterion is preferred



**Fig. 3. Localization Error vs SNR:** Localization error is defined as the root-mean-squared distance between the points in the volume-of-confusion to the true point scatterer. (a) Notice that improved transient time resolution results in far larger improvement in localization accuracy than increased SNR. (b) Increasing the number of detectors has diminishing returns compared to improving SNR. Hence, the temporal resolution has more impact on localization than SNR, which inturn has more impact than number of detectors.

when computing bounds for an ensemble of hidden scenes with known statistical prior. In contrast, the MSE criterion is preferred when analyzing a specific (deterministic) hidden scene. Furthermore, the MSE lends itself to easier interpretation when compared to MI. Despite the differences between the MI and MSE criteria, the two metrics are both dependent on the singular values of transient light transport matrix A. For Gaussian noise case, the MSE on reconstruction of albedos is given by  $\sigma^2 \operatorname{Trace}((A'A)^{-1})$  [21], and MI is  $\frac{1}{2}\log\frac{|\sigma_x^2 A'A + \sigma^2 I|}{|\sigma^2 I|}$  [22], where  $\sigma^2$  is the variance of Gaussian noise and  $\sigma_x^2$  is the variance of the Gaussian prior on x. Note that both these metrics, MSE and MI are intimately related and crucially depend upon the eigenvalues of the matrix A'A. In this report, we only use MSE to quantify as the results are on empirical scenes, but given the similarity between MSE and MI the results should not change significantly based on the choice of MSE vs MI.

#### 4. GEOMETRIC RECOVERY (LOCALIZATION)

Given that there is a single point scatterer in a volume, we want to examine how well we can localize its position in space. For a given transport mechanism (i.e. matrices representing radiance, transient radiance, phase, and transient phase), signal-to-noise ratio, we can localize the point source down to a certain volume-of-confusion. We want the volumeof-confusion to be as small as possible, a size dependent on parameters surch as imaging system type, number of detectors, and detector signal-to-noise ratio. We define the localization error as the root-mean-squared distance between the points in the volume-of-confusion to the true point scatterer. We will quantify the localization error in this section.

## 4.1. How do transients improve localization over radiance?

Transients provide additional information when compared to radiance information alone, thereby reducing localization error. For a given virtual source-detector pair, the temporal information results in an ellipsoidal locus of points with virtual source and virtual detector at the focal points of the ellipsoid [2]. If the albedo is known, the radiance value provides additional constraints in the form of a hyperboloid. Using multiple detectors will result in multiple ellipsoids and hyperboloids. The location of the point scatterer is the intersection of these ellipses and hyperboloids. Even if the albedo of the point source is not known a-priori, we can still localize using the temporal information and then compute the albedo from the radiance information.

4.2. Reconstruction Using Transient Radiance Transport Using a transient camera and incoherent pulsed illumination, the relationship between the spatially varying albedo (x) and time-resolved sensor measurements (y) can be modeled using a linear system: y = Ax. Here, each column of matrix A has mT measurements, where m represents the number of sensors and T represents the number of time samples measured at each sensor. Each column of A should ideally be msparse. However, due to the time-jitter of the imaging system, each column of A will be a sum of m-Gaussian functions, with shifts representing the time-of-travel of the photons and standard deviation equal to the time jitter of the imaging system. Note that different configuration of A will have different number of measurements. Hence, for a fair comparison across configurations, we repeated the measurements (physically equivalent to increasing exposure duration or light intensity) so that the number of rows of A are fixed (either 80k or 1M based on the experiment).

To compute the localization performance of a point scatter via simulations, we have computed the transient response of the point scatterer at a given location. (Note that, even if the object is at a different location, the resulting localization error will not change for a given SNR. However, the signal strength will decrease.) We also have computed the transient responses of various voxels around the point scatterer. Repeating this process at multiple noise levels, we recovered the location of the point scatterer using the matched filter algorithm. By measuring the radius of the resulting point cloud at



Fig. 4. Photometric error vs # detectors: The detectors are distributed randomly over a unit area  $(1m \times 1m)$  around the source and we computed the square reconstruction error for SPAD and ICCD cameras. For any given tolerance on the reconstruction error, we need significantly less number of SPAD detectors compared to ICCD pixels, even though the temporal resolution of SPAD is just 4 times better than the ICCD camera

each noise level, we identify the associated localization error.

Figure 3(a) shows the effect of temporal resolution and SNR on localization error for 16 virtual detectors placed randomly on a  $1m \times 1m$  diffuse wall 1.5 m from the known volume. The temporal resolution is varied to simulate various commercially available transient cameras (PMD, ICCD, SPAD, STREAK). Figure 3(b) shows the effect of number of detectors and SNR on resolution for SPAD detectors placed randomly on a  $1m \times 1m$  diffuse wall 1.5 m from the known volume. Note that increasing the number of detectors decreases the localization error logarithmically.

# 5. PHOTOMETRIC RECOVERY (VOLUME OF SCATTERING SURFACES)

In the most general setting, if the unknown hidden object is a volume of scattering surfaces, how well can we recover the per-voxel albedo of the volumetric scatterer? While this is impractical (most real-world scenes will likely be surfaces), this setting assumes the least prior knowledge about the unknown scene and therefore will allow us to derive lower bounds on albedo recovery. We have empirically computed the lower bounds of a system by computing the squared error loss for different number of detectors.

**5.1. Reconstruction Using Transient Radiance Transport** To reconstruct the albedo of each voxel, we first assemble the A matrix by voxelizing the known volume (in to  $12 \times 12 \times 12$ grid). As the albedo is bounded between 0 to 1, we generated a random albedo field  $x \sim \mathcal{U}^{12^3}[0, 1]$ . We have computed the measurements at the virtual detectors and added Gaussian noise based on the desired SNR level. To reconstruct, we have used pseudo-inverse algorithm (cannot use OMP as the signal is not sparse). We computed the mean squared error per voxel of the reconstruction algorithm (which is a function of the singular values of A matrix) and repeated the procedure for hundred realizations of x. The results are shown in Figure 4. This experiment gives the lower bound on the albedo error. Any prior-information will increase the performance.

#### 6. CONCLUSIONS

We have proposed a linear systems approach to identify the performance bounds in indirect imaging. We studied the transient imaging approach in detail and gave a best-case bound for resolution using localization accuracy. We also studied the hidden surface reconstruction problem as a function of various system parameters such as temporal resolution, SNR, number of detectors, placement of detectors etc.

The bounds show the limitations of a given transient imager and also help us in determining the system-to-be-built for a task of interest. For example, from Figure 4, if the requirements are to build a system with mean-squared-error less than 0.05, and we have SPAD detectors with SNR of 10, we know that having more than 500 virtual detectors produces negligible improvement in the reconstruction, for any scene. Real world problems may come from a class of systems that are sparse in both spatial locations and image gradients, a scenerio in which bounds may be tighter but more difficult to calculate due to non-linear system constraints.

#### 6.1. Key Observations

- 1. Increasing temporal resolution improves localization: The localization improves with improved time resolution and SNR. However, increasing the temporal resolution improves localization more significantly compared to SNR (see Figure 3(a) for details).
- 2. Increasing the number of virtual detectors improves localization: As the number of virtual detectors increase, the resolution improves for any given SNR. However, as the SNR increases, the localization accuracy increases significantly in low SNR regimes and saturates in high SNR regimes. (see Figure 3(b)).
- 3. **Time resolution vs Number of virtual detectors:** Increasing the number of virtual detectors and their associated temporal resolution, will both improve performance. However, we have noticed that having a small increase in time-resolution can lead to significant decrease in the number of the virtual detectors needed for any given tolerance in reconstruction error. (see Figure 4)

## Acknowledgements

This work was supported in part by NSF CAREER Grant IIS-1453192, DARPA REVEAL grant HR0011-16-C-0028, ONR grant N00014-15-1-2735, and the Big-Data Private-Cloud Research Cyberinfrastructure MRI-award funded by NSF under grant CNS-1338099 and by Rice University.

#### 7. REFERENCES

- Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar, "Looking around the corner using transient imaging," in 2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009, pp. 159–166.
- [2] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G Bawendi, and Ramesh Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nature communications*, vol. 3, pp. 745, 2012.
- [3] Paul M Chaikin and Tom C Lubensky, *Principles of condensed matter physics*, vol. 1, Cambridge Univ Press, 2000.
- [4] Matthew R Francis, "Why Can You Hear Around Corners But Not See?," http://www.doublexscience.org/why-can-you-heararound-corners-but-not-see, 2011.
- [5] Ahmed Kirmani, Haris Jeelani, Vahid Montazerhodjat, and Vivek K Goyal, "Diffuse imaging: Creating optical images with unfocused time-resolved illumination and sensing," *IEEE Signal Processing Letters*, vol. 19, no. 1, pp. 31–34, 2012.
- [6] Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten, "Non-line-of-sight imaging using a time-gated single photon avalanche diode," *Opt. Express*, vol. 23, no. 16, pp. 20997–21011, Aug 2015.
- [7] Felix Heide, Lei Xiao, Wolfgang Heidrich, and Matthias B Hullin, "Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3222–3229.
- [8] Achuta Kadambi, Hang Zhao, Boxin Shi, and Ramesh Raskar, "Occluded imaging with time-of-flight sensors," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 2, pp. 15, 2016.
- [9] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos, "Temporal frequency probing for 5d transient analysis of global light transport," ACM Transactions on Graphics (ToG), vol. 33, no. 4, pp. 87, 2014.
- [10] Nils Abramson, "Light-in-flight recording by holography," Opt. Lett., vol. 3, no. 4, pp. 121–123, Oct 1978.
- [11] Nils Abramson, "Light-in-flight recording: high-speed holographic motion pictures of ultrafast phenomena," *Applied optics*, vol. 22, no. 2, pp. 215–232, 1983.

- [12] James G Fujimoto, CA Puliafito, R Margolis, A Oseroff, S De Silvestri, and EP Ippen, "Femtosecond optical ranging in biological systems," *Optics letters*, vol. 11, no. 3, pp. 150–152, 1986.
- [13] Ori Katz, Eran Small, and Yaron Silberberg, "Looking around corners and through thin turbid layers in real time with scattered incoherent light," *Nature photonics*, vol. 6, no. 8, pp. 549–553, 2012.
- [14] Isaac Freund, "Looking through walls and around corners," *Physica A: Statistical Mechanics and its Applications*, vol. 168, no. 1, pp. 49–65, 1990.
- [15] Martin Laurenzis, Frank Christnacher, Jonathan Klein, Matthias B Hullin, and Andreas Velten, "Study of single photon counting for non-line-of-sight vision," in *SPIE Sensing Technology+ Applications*. International Society for Optics and Photonics, 2015, pp. 94920K– 94920K.
- [16] Stephen D Liang, "Sense-through-wall human detection based on uwb radar sensors," *Signal Processing*, vol. 126, pp. 117–124, 2016.
- [17] Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand, "Capturing the human figure through a wall," ACM Trans. Graph., vol. 34, no. 6, pp. 219:1– 219:13, Oct. 2015.
- [18] Donald B Conkey, Antonio M Caravaca-Aguirre, Jake D Dove, Hengyi Ju, Todd W Murray, and Rafael Piestun, "Super-resolution photoacoustic imaging through a scattering wall," *Nature Communications*, vol. 6, 2015.
- [19] Puxiang Lai, Lidai Wang, Jian Wei Tay, and Lihong V Wang, "Photoacoustically guided wavefront shaping for enhanced optical focusing in scattering media," *Nature photonics*, vol. 9, no. 2, pp. 126–132, 2015.
- [20] Scott A Basinger, David Palacios, Marco B Quadrelli, and Grover A Swartzlander, "Optics of a granular imaging system (ie orbiting rainbows)," in SPIE Optical Engineering+ Applications. International Society for Optics and Photonics, 2015, pp. 96020E–96020E.
- [21] Yoav Y Schechner, Shree K Nayar, and Peter N Belhumeur, "Multiplexing for optimal lighting," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 8, pp. 1339–1354, 2007.
- [22] Dongning Guo, Shlomo Shamai, and Sergio Verdú, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.