Trapping Light for Time of Flight

Ruilin Xu
Columbia University
rxu@cs.columbia.edu

Mohit Gupta
University of Wisconsin-Madison
mohitg@cs.wisc.edu

Shree K. Nayar
Columbia University
nayar@cs.columbia.edu

Abstract

We propose a novel imaging method for near-complete, surround, 3D reconstruction of geometrically complex objects, in a single scan. The key idea is to augment a time-of-flight (ToF) based 3D sensor with a multi-mirror system, called a light-trap. The shape of the trap is chosen so that light rays entering it bounce multiple times inside the trap, thereby visiting every position inside the trap multiple times from various directions. We show via simulations that this enables light rays to reach more than 99.9% *of the surface of objects placed inside the trap, even those with strong occlusions, for example, lattice-shaped objects. The ToF sensor provides the path length for each light ray, which, along with the known shape of the trap, is used to reconstruct the complete paths of all the rays. This enables performing dense, surround 3D reconstructions of objects with highly complex 3D shapes, in a single scan. We have developed a proof-of-concept hardware prototype consisting of a pulsed ToF sensor, and a light trap built with planar mirrors. We demonstrate the effectiveness of the light trap based 3D reconstruction method on a variety of objects with a broad range of geometry and reflectance properties.*

1. Introduction

The ability to measure complete surround 3D shape of objects has a wide range of applications, from digital modeling of rare specimen and artifacts in museums, to motion capture for digital entertainment and bio-mechanics research, and to medical applications such as 3D scanning of anatomical features for surgical planning, customized orthopedic casts, and customized prosthetics. Single viewpoint 3D imaging systems can image only a part of an object's surface at a time. Furthermore, single viewpoint systems have limited capability in dealing with occlusions inherent in geometrically complex objects. These limitations can be addressed by capturing images from multiple perspectives, either by multiple cameras placed at different locations [27], or by rotating the object mechanically [26, 1]. These systems, while capable of measuring surround 3D shapes, require synchronization of multiple cameras, sophisticated stitching algorithms, and/or a slow capture process which prohibits scanning of dynamic scenes.

We propose light trap 3D imaging, a novel imaging technique that enables *single-scan, surround and dense* 3D imaging of geometrically complex objects. The proposed technique uses a time-of-flight (ToF) 3D sensor, and a *light trap* built by planar mirrors. The object to be scanned is placed inside the trap. Light rays emitted by the ToF sensor are *trapped* within the mirror system, i.e., they bounce several times within the trap. This creates a diverse set of light paths that sweep through the entire inside space of the trap. Furthermore, light rays visit every location inside the trap from several different directions, thus providing a continuous set of virtual view-points. This ensures that for a trap with the appropriate shape, light rays can reach more than 99.9% of the surface of geometrically complex objects with strong occlusions (e.g., a closely spaced lattice of spheres) and cavities, even after a relatively small number of bounces.

The use of mirrors for 3D reconstruction has been explored in the past, in conjunction with multi-view stereo [25] and structured light [12, 20]. The main challenge with mirror based multi-view methods is the need to solve the correspondence problem. Reflection of a ray at every mirror surface creates a new virtual sensor view. Solving the correspondence problem requires assigning every pixel to a specific virtual view, i.e., determining the number of bounces that light entering the pixel has gone through. Unfortunately, this labeling problem is intractable for general shapes, especially those with self-occlusions, given the exponentially large number of virtual views. In contrast, our approach does not need to solve the labeling problem. This is because a ToF sensor is used, which can estimate the path length of each ray. Since the shape of the light trap is known a priori, the entire path of each ray can be computed, including the number of bounces. This enables measuring the 3D shape of objects with complex 3D geometry including strong self-occlusions, from a single scan of a ToF sensor.

Since rays bounce multiple times inside the trap, a ray emitted from the sensor may visit a scene point multiple

2575-7075/18 $31.00 © 2018 IEEE
DOI 10.1109/CVPR.2018.00651

IEEE computer society

times. This creates multiple light paths from the sensor to the same scene point, each with a different length. For a continuous wave ToF sensor [19, 24], the estimated path length is the linear combination of these lengths, which is different from the true length (length of the shortest path). This results in an incorrectly estimated 3D shape. This problem is similar to that of multi-path interference in ToF imaging [7, 9, 15, 13]. Fortunately, we can mitigate the effect of multi-path interference by using pulsed ToF sensors [8, 18, 16], which consider only the first return for every ray, corresponding to the first intersection of the ray with the scene (shortest path). We demonstrate the proposed method via a hardware prototype consisting of a pulsed ToF system, and a light trap made with planar mirrors. We developed a calibration procedure for measuring the shape of the trap in order to trace the path of every ray, and designed algorithms for recovering the 3D shape of objects inside the light trap. We show near-complete 3D reconstructions of several objects with complex geometries and reflectance.

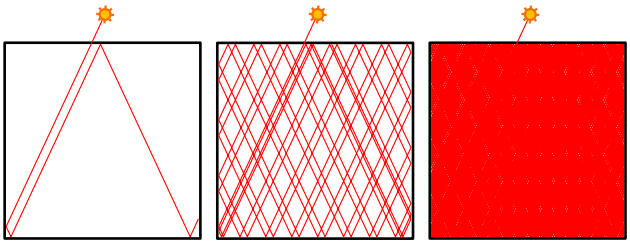
2. Related Work

Catadioptric imaging systems: Catadioptric imaging systems are used to augment cameras with additional mirrors in order to increase the number of effective view-points. Han and Perlin [10] proposed a kaleidoscopic system with multiple mirrors for reflectance (BRDF) measurements. The mirrors were used to generate a large number of views of the target object, covering the entire hemisphere of directions surrounding it. Gluckman and Nayar [5, 6] developed a catadioptric stereo system with multiple mirrors and a single camera. Mirrors have also been used to generate multiple view-points in the context of multi-view 3D reconstruction [2, 22, 23, 28, 4, 25, 11, 21]. As mentioned earlier, the main limitation of mirror-based multi-view 3D reconstruction approaches is the need to solve the correspondence and labeling problems, which is computationally expensive, and impossible for shapes with occlusions and cavities.

3D reconstruction with multiple or moving cameras: One popular method for surround 3D scene reconstruction, especially for large scenes, is to use multiple color or depth cameras [17, 14]. Another approach is to mount the object of interest on a rotating platform [26, 1], or move the camera around the object [3]. These systems require multi-camera synchronization, calibrating for radiometric responses of different cameras, or capturing several images while moving the object/sensor. In contrast, our goal is to perform surround 3D scanning of objects in a single scan.

3. Trapping Light via Mirrors

Consider the imaging scenario shown in Figure 1. For ease of exposition, the illustration is shown in 2D. A collection of planar mirrors is arranged as the sides of a polygon, with the mirror sides facing inwards. The polygon is closed, except for a small opening as an entrance for light rays. A small opening ensures that the probability of the light ray escaping the closed polygon is small. A collimated light source emits light into the opening, and is incident on one of the mirror facets. We assume that each mirror is perfect, without any absorption. In that case, light will continue to be reflected from the mirror facets. Figure 1 demonstrates the paths traced by the ray after 5 bounces, 50 bounces, and 500 bounces, respectively.



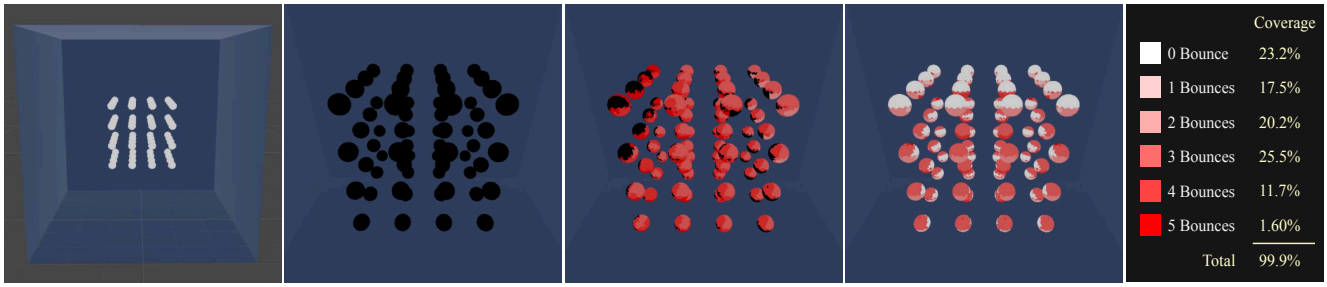
(a) 5 bounces (b) 50 bounces (c) 500 bounces

Figure 1: A collimated light source emits a light ray into a light trap made of mirrors. (a-c) Paths traced by the ray after 5 bounces, 50 bounces and 500 bounces, respectively. After a sufficient number of reflections, nearly all locations in the trap are visited multiple times from different directions.

What are good light trap designs? For light traps with appropriate shapes, a light ray entering the trap visits *every location* in the trap after a sufficient number of reflections[1], as shown in Figure 1c. However, in order to recover surround 3D shape of complex shapes, not only must light rays be able to reach all (or most) of the locations in the trap, they must be able to reach every location multiple times, from different directions. This ensures dense coverage within a trap, but also dense coverage of each point from multiple directions. The latter property is critical for reconstructing highly complex shapes with severe self-occlusions and strong cavities, for which certain scene points can be reached only along a narrow set of directions. For instance, consider an extreme case, where the scene is a lattice of small, closely spaced spheres. A natural question to ask is: *What is the shape of the light trap that will enable light rays to reach all (or as many as possible) scene points in such scenes with strong occlusions?*

While a detailed theoretical analysis of the shapes of mirror traps that lead to complete coverage for all possible shapes is beyond the scope of this paper, we perform an empirical analysis based on 3D simulations (implemented in Unity3D). Specifically, we consider several potential light

[1] An exception is a regular polygon-shaped light trap, with light incident at an angle of $\theta/2$ with respect to the normal of a facet, where θ is the angle between two faces of the polygon. For example, for a square (4-gon), if light is incident at an angle of $45°$, then it will continue to go around the light trap along the same path.



(a) (b) (c) (d) (e)

Figure 2: **Simulation for testing surface coverage of different trap shapes.** (a) Simulation setup. A $4 \times 4 \times 4$ lattice of spheres is placed in a cubic light trap. (b) Initial state of the simulation. (c) As the simulation runs, a color code is assigned to points on the surface of spheres, depending on the number of bounces the light ray undergoes before hitting the points. (d) End result of the simulation. (e) 99.9% of the lattice's surface area is visited after 5 bounces.

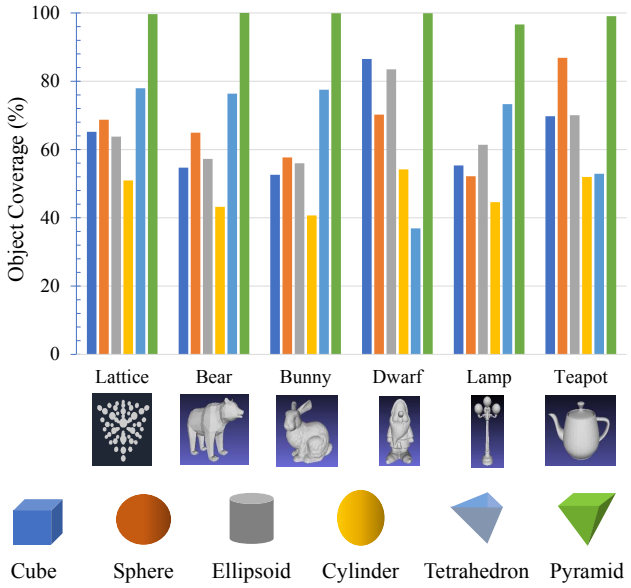


Figure 3: **Surface coverage comparison.** We compared the coverage of 6 different trap shapes (cube, sphere, ellipsoid, cylinder, tetrahedron, and pyramid), for 6 objects (lattice, bear, bunny, dwarf, lamp and teapot). The pyramid-shaped light trap performs the best on all the test objects, reaching $> 99\%$ coverage on all shapes, with 3 bounces per ray.

trap shapes; for each candidate trap shape, we compute the percentage of the reachable surface locations for a set of objects with different shapes of varying complexity. The simulation consists of a point light source, a light trap built from ideal planar mirrors with 100% reflectance, and the target object placed inside the trap. The light source emits 1500×1500 light rays in a field-of-view (FOV) chosen so that all light rays are able to enter the light trap. The configuration of the simulation setup is shown in Figure 2a. An object consisting of a $4 \times 4 \times 4$ lattice of spheres is placed in a cubic light trap. The spheres have a radius of $5mm$ each and are $10mm$ apart from each other. Figure 2b shows the initial state of the simulation – no sphere has been visited by any light rays, hence the black surface color. As the simulation runs, light visits the surface of the spheres. A color code is assigned depending on the number of bounces the light ray undergoes before hitting the sphere's surface. Figure 2c shows a snapshot during the duration of a simulation. The end of the simulation is shown in Figure 2d. As shown in Figure 2e, 99.9% of the lattice's surface area (combined surface area of 64 spheres) has been visited, after 5 bounces.

We compared six different light trap shapes, namely, cube, sphere, cylinder, ellipsoid, tetrahedron, and pyramid, for six different objects, namely, lattice, bear, bunny, dwarf, lamp, and teapot. Coverage results for each object and trap shape are shown in Figure 3. The pyramid-shaped light trap performs the best on all the test objects among different trap shapes. With a pyramid-shape trap, light rays emitted from a single point source can visit $> 99\%$ of the points on the surface of the tested objects, even those with complex shapes, with ≤ 3 bounces of each ray. Recovering shape of such complicated objects (e.g., lattice) may be challenging with conventional 3D shape reconstruction methods.

4. Recovering 3D Shape with a Light Trap

In this section, we discuss the proposed 3D reconstruction method using a light trap. As discussed earlier in Section 1, mirror-based multi-view 3D imaging systems need to determine the number of reflections that each light ray undergoes before reaching the object of interest. Effectively, these systems need to "unroll" every light path, from the sensor to the object. This is only possible if the shape of the object is known, but to estimate the shape, we need to first label every light path with its number of bounces. This is a chicken-and-egg problem. While the labeling can be performed for relatively simple, convex shapes, it is not always possible for objects with occlusions or cavities.

One way to solve the labeling problem is to measure the length of each light path (from the sensor to the scene

point). Fortunately, ToF sensors can estimate the path lengths by measuring the time taken for the emitted light to travel to the scene point and back to the sensor. The path length is computed by the product of the travel time, and the known speed of light.

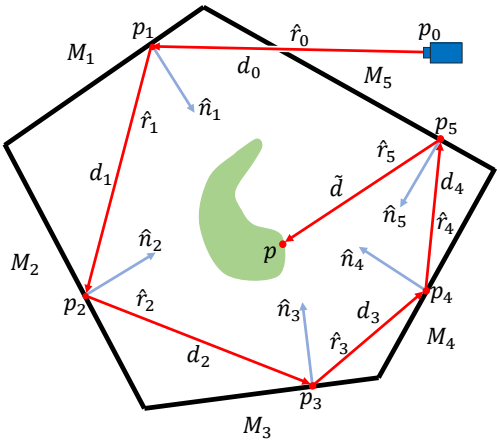


Figure 4: **3D imaging using light trap ToF sensor:** A diffuse object is placed inside a light trap. A light ray emitted by a ToF sensor at p_0 reaches scene point p after multiple bounces. The ToF sensor measures the total length of the light path from p_0 to p, which, along with known shape of the trap, can be used to determine the 3D coordinates of p.

4.1. Time of Flight Based Light Trap

Consider a light trap with perfect mirrors, as shown in Figure 4. Suppose a ToF sensor is placed at point p_0. The sensor emits a light ray L along direction $\hat{r}_0$, which hits mirror M_1 at point p_1. The light reflected from M_1 hits mirror M_2 at point p_2, and so on. Suppose light hits mirror M_K at point p_K, before finally reaching point p on the surface of a diffuse object placed inside the trap. Let d_k be the distance between points p_k and p_{k+1}. Then, d_{total}, the total length of the light path from the sensor location p_0 to p is given as:

$$d_{total} = \sum_{k=0}^{K-1} d_k + \tilde{d}\,, \tag{1}$$

where $\tilde{d}$ is the distance between p (point on the surface) and p_K (the last mirror point on the light path). If d_{total} is measured, the 3D location of point p can be estimated by tracing the light path starting from the sensor. The end point of the path, at length d_{total}, is the 3D location of p. However, the ToF sensor cannot directly calculate d_{total}. In fact, it cannot perform depth calculation unless the light ray reaches the scene point p and then trace back to point p_0. What path will the light ray follow, back from point p to the sensor location p_0?

Lemma: Suppose a light ray emitted from p_0 along direction $\hat{r}_0$ arrives at a diffuse object point p, either directly, or indirectly (after one or more specular reflections). Then, the shortest path a ray reflected at p can return to p_0 along direction $-\hat{r}_0$ is by retracing the incident path.

Proof: Suppose there were a shorter return path from point p to p_0 along the final direction $-\hat{r}_0$. In that case, light would have taken this path to reach point p because of the reciprocity of light propagation. This leads to a contradiction, thus proving the lemma.

4.2. Path Length and Depth Computation

The above lemma states that a light ray retraces its path from the sensor to a scene point. Let the total path length measured by the sensor be d_{meas}. Then the path length d_{total} from the sensor to the scene point (Eq. 1) is given as:

$$d_{total} = \frac{d_{meas}}{2} = \sum_{k=0}^{K-1} d_k + \tilde{d}\,. \tag{2}$$

The length of each ray segment d_k can be calculated as $d_k = p_{k+1} - p_k$, where the location p_{k+1} is given as $p_{k+1} = d_k \hat{r}_k + p_k$, and $\hat{r}_k$ is the unit vector along the direction of ray reflected from the k^{th} mirror at point p_k. $\hat{r}_k$ can be computed by using Snell's law of specular reflection:

$$\hat{r}_k = \hat{r}_{k-1} - 2\,(\hat{n}_k \cdot \hat{r}_{k-1})\,\hat{n}_k\,. \tag{3}$$

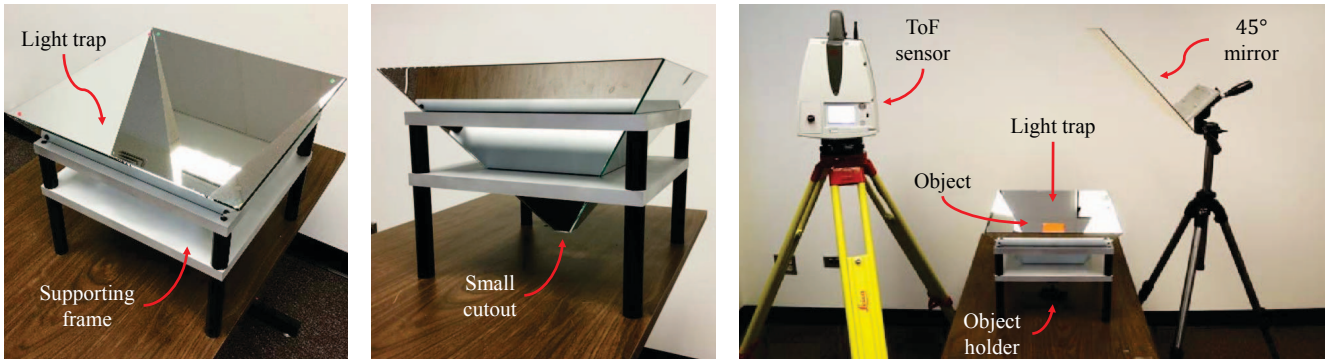
Thus, given the initial sensor location p_0 and initial ray direction $\hat{r}_0$, we can recursively compute the locations p_k and directions $\hat{r}_k$, for all $k \in [1 \ldots K]$, by using the above equations. We can use the computed p_k and $\hat{r}_k$ to trace the light path, starting from the sensor, up to length d_{total}, as computed in Eq. 2. This is shown in Figure 4. As discussed earlier, the end point of the light path, p_{end} is the location of the surface point p. See Alg. 1 for the detailed algorithm. Note that for the operation on line 8 of Alg. 1, we can find the intersections of the rays with the mirror surfaces of the light trap because we know (a) the shape of the trap, and (b) the relative pose of the trap with respect to the camera by calibration, as discussed in Section 5.3.

5. Hardware Prototype

We have developed a hardware prototype system for light trap based 3D imaging. In the following section, we discuss the hardware design choices, as well as details regarding the calibration procedure.

5.1. Mirror Light Trap

The simulation results shown in Section 3 suggested that a pyramid-shaped light trap can achieve maximal coverage for a wide range of shapes. Therefore, we developed our prototype using a pyramid-shaped light trap built by four planar mirrors. We leave the pyramid base open, which allows light to enter the trap with full intensity. The absence



(a) (b) (c)

Figure 5: **Hardware prototype of light trap based 3D imaging system.** (a-b) Design of the light trap and the supporting frame. (c) Complete setup with the ToF sensor (left), light trap and an object of interest placed on the object holder through the cutout of the trap (middle), and the additional 45° planar mirror (right).

Algorithm 1 Depth Computation

Require: p_0, initial sensor location
$\hat{r}_0$, initial ray direction

procedure COMPUTEDEPTH
  for each d_{meas} **do**
    $p_{end} \leftarrow TracePath(p_0, \hat{r}_0, \frac{d_{meas}}{2})$
  end for
end procedure

function TRACEPATH($p_k, \hat{r}_k, d_{remain}$)
  $ray \leftarrow (p_k, \hat{r}_k)$
  $p_{k+1} \leftarrow$ intersection of *ray* and *light trap*
  $d_k = p_{k+1} - p_k$
  if $d_{remain} > d_k$ **then**
    $d_{remain} \leftarrow d_{remain} - d_k$
    $\hat{r}_{k+1} = \hat{r}_k - 2\,(\hat{n}_{k+1} \cdot \hat{r}_k)\,\hat{n}_{k+1}$
    TRACEPATH($p_{k+1}, \hat{r}_{k+1}, d_{remain}$)
  end if
  $p_{end} \leftarrow d_{remain}\hat{r}_k + p_k$
  return p_{end}
end function

of a base limits the total number of reflections for most rays to 3. Fortunately, as the simulation results showed, only three reflections per light ray are sufficient to achieve $> 99\%$ coverage for a wide range of objects.

We used planar mirrors with protected aluminum metal coating with thickness of $3.0mm$ to construct the trap. The mirrors have a reflectance of $> 90\%$ for wavelengths from $400-2000nm$. Each mirror is cut to the shape of an equilateral triangle with a side length of $400mm$. A steel frame is also constructed to hold all four mirrors to form a four-sided pyramid with a height of $282.84mm$ (see Figure 5a and Figure 5b). The object to be scanned is placed inside the trap supported on a thin rod which is inserted via a small cutout hole at the tip of the pyramid (Figure 5b).

5.2. ToF Sensor

We used a ScanStation P40 3D ToF Laser Scanner from Leica. This is a pulsed ToF sensor, which uses infra-red light of wavelength $1550nm$. The 3D measurement accuracy is $3mm$ at a range of $50m$. In addition to capturing a high quality depth map, the scanner also captures corresponding color (RGB) data.

Since the light trap is constructed to face vertically upward, and the ToF sensor must ideally be mounted horizontally, with its optical axis parallel to the ground, an additional planar mirror (with identical characteristics as the mirrors used to make the trap) is placed at a 45° angle facing the ToF sensor above the light trap. See Figure 5c for the complete hardware setup.

5.3. Light Trap Calibration Procedure

The light trap must be calibrated to estimate the position and orientation (surface normal) of each of the four planar mirrors, as shown in Figure 6. To this end, we place small dot-shaped markers near every corner of each mirror, as shown in Figure 6a. Next, we obtain a 3D reconstruction (point cloud) of the empty trap with the markers. The point cloud is manually processed and each point cloud patch that corresponds to a marker is labeled, as shown in Figure 6b. Using the reconstructed locations of the markers, we perform an initial calculation of each mirror's position and surface normal (16 parameters in total) (see Figure 6c).

The initial estimate of the mirror shapes is refined by using a cube as a reference calibration object. Specifically, using the known shape of the cube, we perform a 16-parameter Levenberg-Marquardt (LM) optimization recursively, from which the mirror parameters as well as a calibrated cube reconstruction is obtained (Figure 6d).

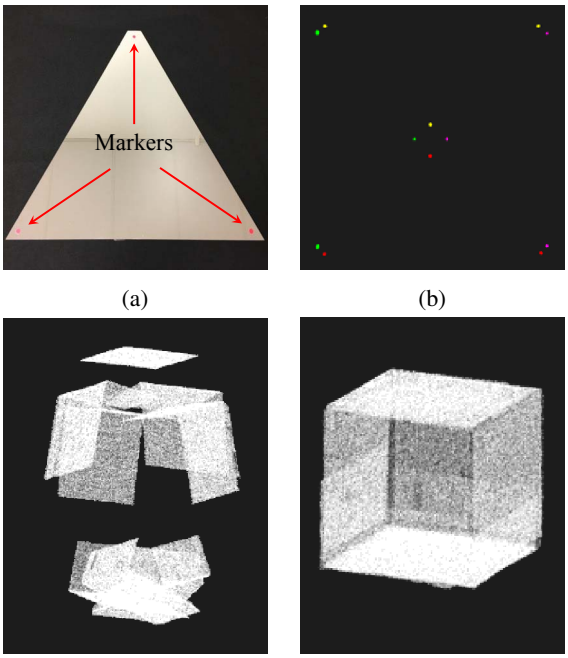


Figure 6: **Calibration of the light trap based imaging system.** (a) Small dot-shaped markers are placed near each corner of each mirror. (b) Point cloud patches that correspond to each of the markers on the mirrors. (c) A cube reconstruction with initial mirror parameters. (d) Cube reconstruction using LM optimization.

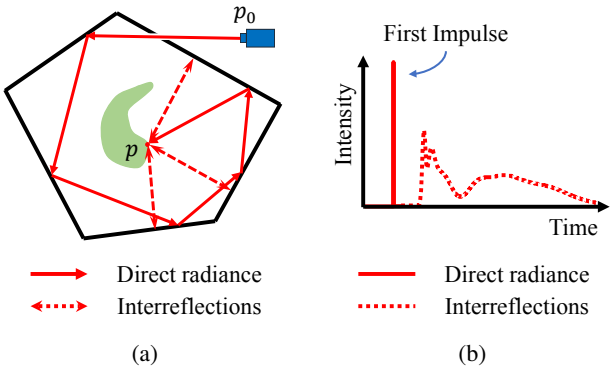


Figure 7: (a) Point p is illuminated by the direct ray from p_0 to p (denoted as solid red arrows). It is also illuminated by several indirect rays (denoted as dashed red double arrows) that travel from p to all mirrors from which p can be seen and back to p. These indirect rays result in longer path lengths back to the sensor. In case of continuous-wave ToF sensors, longer path lengths result in erroneous phase-shift and hence incorrect depth measured. (b) Pulsed ToF sensors consider only the first impulse (shortest path back to the sensor), thus mitigate the effect of interreflections.

5.4. Dealing with Multi-path Interference

Since rays bounce multiple times inside the trap, a scene point may be reached multiple times by a single emitted light ray from direct and indirect paths, as shown in Figure 7. This is similar to multi-path interference problem in ToF imaging, where an object is illuminated not only directly by the light source, but also indirectly by other scene points.

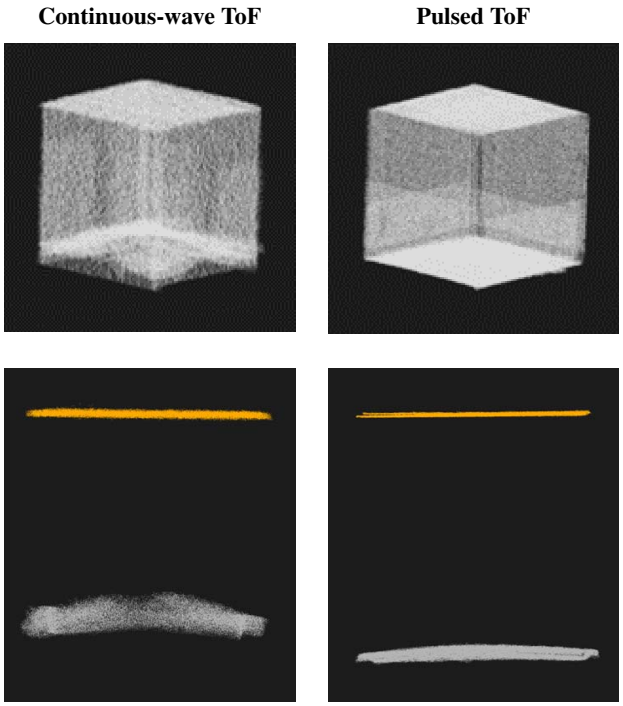


Figure 8: **Comparison between continuous-wave ToF and pulsed ToF for a cube**. (Top row) complete reconstructions of a cube. (Bottom row) The top and bottom surfaces of the cube. In the continuous-wave reconstruction result, the bottom surface of the cube is pushed in and curved upward due to multipath interference. Pulsed ToF sensor mitigates the errors due to multipath interference, and produces accurate reconstructions.

This results in multiple light paths from the sensor to the same scene point, each with a different length. For continuous-wave ToF sensors, the estimated path length (computed from the phase-shift between the received waveform and the emitted waveform) is the linear combination of all the path lengths, which is different from the true length. This results in erroneous phase-shift, and hence, incorrect depths [9]. Pulsed ToF sensors, on the other hand, mitigate this problem by considering only the first impulse of the returning signal. Since the direct light path has the shortest distance back to the sensor, it always corresponds to the first returning pulse; other paths are longer, and thus, correspond to larger travel times, as shown in Figure 7b. In order to demonstrate the robustness of a pulsed ToF sensor to multi-path interference, we compared its performance with a continuous-wave ToF scanner (Faro x330). Comparison results between the continuous-wave and pulsed ToF sensor are shown and discussed in Section 6.

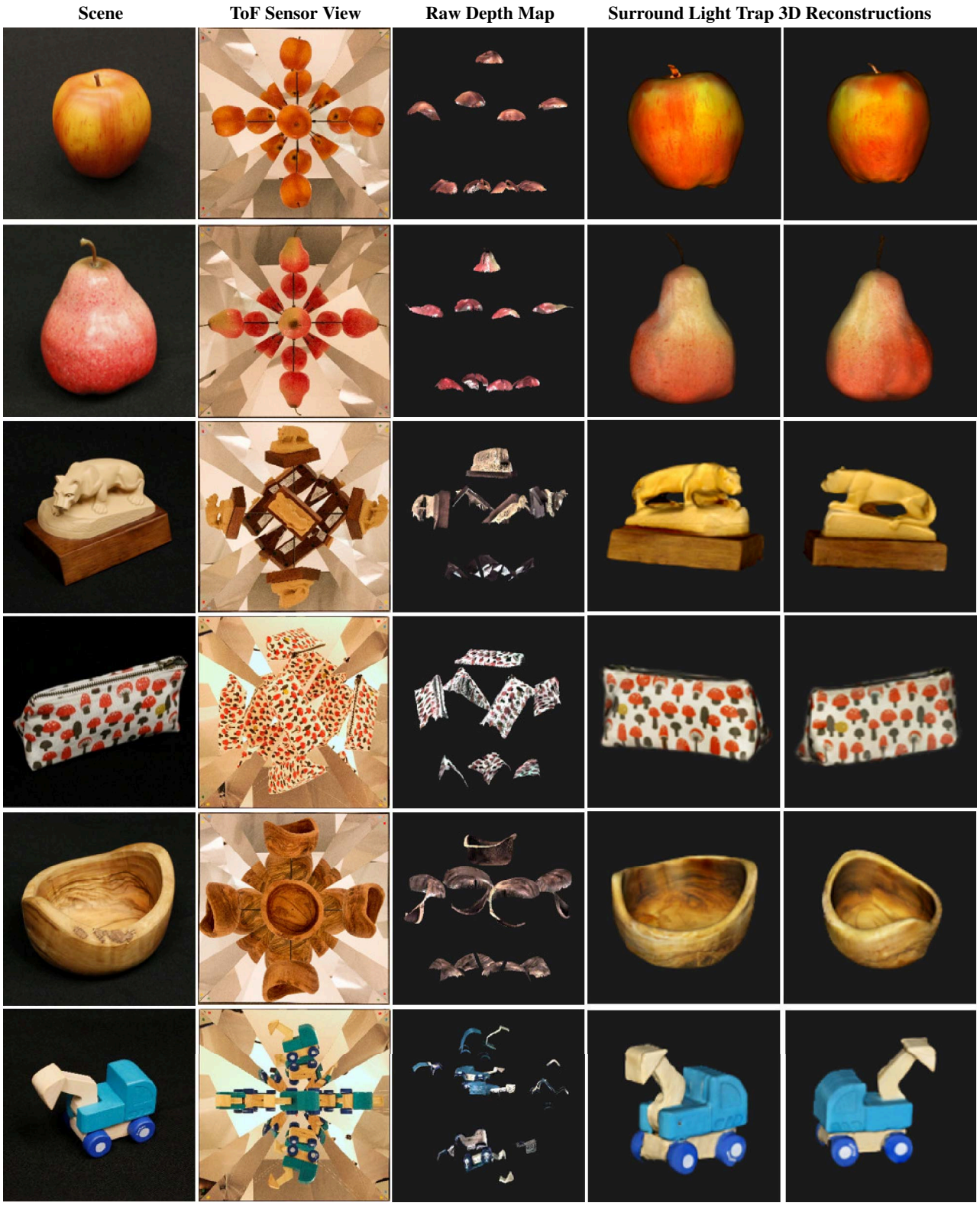


Figure 9: **Dense, surround 3D reconstructions using the light trap based imaging system.** In each row, we show, from the left to right, (1) a photo of the scanned object, (2) ToF camera view, (3) "raw" depth map recovered by the ToF camera, (4-5) two different perspective views of the reconstructed object.

	Top Thickness (mm)		*Bottom Thickness (mm)*	
Continuous -wave ToF (FARO x330)	Max.	3.01775	Max.	8.26323
	Avg.	0.421061	Avg.	1.89455
	Sigma	0.397341	Sigma	1.47339
Pulsed ToF (Leica P40)	*Top Thickness (mm)*		*Bottom Thickness (mm)*	
	Max.	1.33144	Max.	3.80526
	Avg.	0.225517	Avg.	0.861381
	Sigma	0.316642	Sigma	0.669976

Table 1: **Comparison between continuous-wave ToF and pulsed ToF sensors**.

6. Experimental Results

Continuous-wave vs Pulsed ToF: In Section 5.4, we discussed how the multi-path problem influences the continuous-wave ToF and pulsed ToF differently. To investigate this, we performed a comparison for a cube, using our pulsed ToF scanner, and a continuous-wave ToF scanner. The comparison result is shown in Figure 8.

In the continuous-wave reconstruction result, the bottom surface of the cube is pushed in and curved upward, as compared with the result from the pulsed ToF. For quantitative comparisons, we extracted the top and bottom surfaces from the reconstructed results for both scanners and measured, as shown in Table 1. The quantitative results confirm the conclusion discussed in Section 5.4 that if a scene point can be seen by multiple reflections of itself, its depth is estimated incorrectly with a continuous-wave scanner. On the other hand, pulsed ToF scanner significantly mitigates the multi-path problem, and recovers accurate reconstructions.

3D Reconstruction of geometrically complex objects: Figure 9 shows examples of several objects scanned using our hardware prototype. Consider the lion and wooden truck results in Figure 9. Both are highly complex objects with severe self occlusions (i.e., the joint between the truck arm and truck body) and strong cavities (i.e., the abdomen region of the lion), which are difficult to reconstruct via conventional methods. However, since the proposed imaging system results in dense coverage of each scene point from multiple directions, the quality of 3D shape recovery of highly complex objects is ensured.

7. Trade-offs and Limitations

Scan time: One practical limitation of the proposed approach is the scan time. While it is possible to build a large light trap to accommodate large-scale objects such as human bodies, the minute-long scan time makes it difficult to capture dynamic human body motions. This limitation may be resolved by using a full-frame flash ToF camera.

Achievable Spatial Resolution: Theoretically, the resolution (total number of 3D points) achieved by the light-trap is equal to the number of rays emitted by the 3D sensor. However, in practice, only a fraction of the rays reach the surface of the scanned object and return to the sensor, thus lowering the effective resolution. The fraction depends on several factors, including the shape of the object and the trap, as well as their relative sizes. For example, only 16.6935% of the rays return back to the sensor in the $4 \times 4 \times 4$ lattice of spheres (Figure 2), and 19.8339% for the bunny (Figure 3). Therefore, in contrast to the conventional methods in which 100% of the emitted rays return to the camera, the resolution of the proposed approach is object- and trap-dependent.

One of the advantages of the proposed approach is that multiple light rays can reach the same position from different directions, thus increasing the probability of reaching difficult-to-reach areas, such as cavities. However, for simple objects, this feature can lower the overall resolution since multiple rays may reach the same scene point, resulting in repeated and redundant measurements.

Effect of BRDF on Reconstruction: The proposed approach requires a relatively strong retro-reflective component in the object BRDF. This requirement is similar as other active 3D imaging methods (e.g., structured light, ToF), where the sensor and source are located in close proximity. However, due to the extra interreflections of the mirror trap, this problem may be amplified in the proposed approach because of the possibility of 'indirect' returns that are stronger than the first (direct) return, especially for specular objects.

8. Conclusion and Future Works

Scanning vs. full-frame flash sensors: Our current setup uses a scanning based pulsed ToF sensor which emits one light ray at a time. A next step is to apply the proposed concept to full-frame flash ToF cameras that can capture 3D depth maps of scenes at high speeds, without requiring mechanical scanning. With full-frame flash ToF sensors, multiple light rays can simultaneously enter the light trap, potentially leading to depth ambiguities due to multi-path interference. Resolving these ambiguities will aid surround 3D reconstructions of dynamic objects, and forms a promising avenue of future work.

Optimality of the trap shape: Although we demonstrated via simulations that the pyramidal trap achieves near 100% coverage for several objects, it is not provably optimal in terms of coverage. Developing a better theoretical understanding and deriving formal proofs towards the optimality of a particular trap shape, and designing traps which yield better coverage, is a promising future research direction.

Acknowledgement: This work was supported in parts by ONR awards N00014-14-1-0741; N00014-16-1-2152; N00014-15-1-2902; and DARPA award HR011-16-C-0025.

References

[1] NextEngine. http://www.nextengine.com. [Online; accessed 11-November-2017]. 1, 2

[2] S. Bangay and J. D. Radloff. Kaleidoscope configurations for reflectance measurement. In *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, pages 161–170, 2004. 2

[3] Y. Cui, S. Schuon, D. Chan, S. Thrun, and C. Theobalt. 3d shape scanning with a time-of-flight camera. In *Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1173–1180, 2010. 2

[4] K. Forbes, F. Nicolls, G. D. Jager, and A. Voigt. Shape-from-silhouette with two mirrors and an uncalibrated camera. *Computer Vision ECCV 2006 Lecture Notes in Computer Science*, pages 165–178, 2006. 2

[5] J. Gluckman and S. Nayar. Planar catadioptric stereo: Geometry and calibration. In *Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 22–28, 1999. 2

[6] J. Gluckman and S. Nayar. Rectified catadioptric stereo sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):224–236, 2002. 2

[7] J. P. Godbaz, A. A. Dorrington, and M. J. Cree. Understanding and ameliorating mixed pixels and multipath interference in amcw lidar. In *TOF Range-Imaging Cameras*, pages 91–116, 07 2013. 2

[8] B. Goldstein and G. Dalrymple. Gallium arsenide injection laser radar. In *Proceedings of the IEEE*, volume 55 of *2*, pages 181–188, 02 1967. 2

[9] M. Gupta, S. Nayar, M. Hullin, and J. Martin. Phasor imaging: A generalization of correlation based time-of-flight imaging. *ACM Transactions on Graphics*, 34(5), 2015. 2, 6

[10] J. Y. Han and K. Perlin. Measuring bidirectional texture reflectance with a kaleidoscope. In *Proceedings of ACM SIGGRAPH 2003 Papers*, pages 741–748, 2003. 2

[11] B. Hu, C. M. Brown, and R. C. Nelson. Multiple-view 3-d reconstruction using a mirror. Technical report, University of Rochester, Department of Computer Science, May 2005. 2

[12] I. Ihrke, I. Reshetouski, A. Manakov, A. Tevs, M. Wand, and H.-P. Seidel. A kaleidoscopic approach to surround geometry and reflectance acquisition. In *Proc. Computational Cameras and Displays Workshop*, 2012. 1

[13] D. Jimenez, D. Pizarro, M. Mazo, and S. Palazuelos. Modelling and correction of multipath interference in time of flight cameras. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '12, pages 893–900, Washington, DC, USA, 2012. IEEE Computer Society. 2

[14] H. Joo, H. Liu, L. Tan, L. Gui, B. Nabbe, I. Matthews, T. Kanade, S. Nobuhara, and Y. Sheikh. Panoptic studio: A massively multiview system for social motion capture. In *Proceedings of 2015 IEEE International Conference on Computer Vision*, pages 161–170, 2015. 2

[15] 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. In *ACM Transactions on Graphics*, volume 32, pages 167:1–167:10, 11 2013. 2

[16] A. Kilpela, R. Pennala, and J. Kostamovaara. Precise pulsed time-of-flight laser range finder for industrial distance measurements. *Review of Scientific Instruments*, 72:2197–2202, 04 2001. 2

[17] Y. M. Kim, C. Theobalt, J. Diebel, J. Kosecka, B. Miscusik, and S. Thrun. Multi-view image and tof sensor fusion for dense 3d reconstruction. In *Proceedings of 2009 IEEE 12th International Conference on Computer Vision Workshops*, pages 1542–1549, 2009. 2

[18] W. Koechner. Optical ranging system employing a high power injection laser diode. In *IEEE Transactions on Aerospace and Electronic Systems*, volume AES-4 of *1*, pages 81–91, 01 1968. 2

[19] R. Lange. *3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in. CMOS/CCD-Technolog*. PhD thesis, Department of Electrical Engineering and Computer Science at University of Siegen, 2000. 2

[20] D. Lanman, D. Crispell, and G. Taubin. Surround structured lighting for full object scanning. pages 107–116, Aug 2007. 1

[21] A. Manakov, J. F. Restrepo, O. Klehm, R. Hegedüs, E. Eisemann, H.-P. Seidel, and I. Ihrke. A reconfigurable camera add-on for high dynamic range, multi-spectral, polarization, and light-field imaging. *ACM Transactions on Graphics*, 32(4):47:1–47:14, July 2013. 2

[22] H. Mitsumoto, S. Tamura, K. Okazaki, N. Kajimi, and Y. Fukui. 3-d reconstruction using mirror images based on a plane symmetry recovering method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(9):941–946, 1992. 2

[23] D. W. Murray and P. A. Beardsley. Range recovery using virtual multi-camera stereo. *Computer Vision and Image Understanding*, 61(2):285–291, 1995. 2

[24] J. M. Payne. An optical distance measuring instrument. *Review of Scientific Instruments*, 44:304–306, 03 1973. 2

[25] I. Reshetouski, A. Manakov, H.-P. Seidel, and I. Ihrke. Three-dimensional kaleidoscopic imaging. In *Proceedings of 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 353–360, 2011. 1, 2

[26] D. Shark and T. Townsend. Matter and Form. https://matterandform.net/scanner. [Online; accessed 11-November-2017]. 1, 2

[27] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics*, 24(3):765–776, 2005. 1

[28] X. Ying, K. Peng, R. Ren, and H. Zha. Geometric properties of multiple reflections in catadioptric camera with two planar mirrors. In *Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2010. 2