Agile Depth Sensing Using Triangulation Light Curtains

Joseph R. Bartels[1], Jian Wang[1,2], William "Red" Whittaker[1], and Srinivasa G. Narasimhan[1]
[1]The Robotics Institute, Carnegie Mellon University
[2]Snap Research
{josephba, red, srinivas}@cs.cmu.edu jwang4@snap.com

Abstract

Depth sensors like LIDARs and Kinect use a fixed depth acquisition strategy that is independent of the scene of interest. Due to the low spatial and temporal resolution of these sensors, this strategy can undersample parts of the scene that are important (small or fast moving objects), or oversample areas that are not informative for the task at hand (a fixed planar wall). In this paper, we present an approach and system to dynamically and adaptively sample the depths of a scene using the principle of triangulation light curtains. The approach directly detects the presence or absence of objects at specified 3D lines. These 3D lines can be sampled sparsely, non-uniformly, or densely only at specified regions. The depth sampling can be varied in real-time, enabling quick object discovery or detailed exploration of areas of interest. These results are achieved using a novel prototype light curtain system that is based on a 2D rolling shutter camera with higher light efficiency, working range, and faster adaptation than previous work, making it useful broadly for autonomous navigation and exploration.

1. Introduction

A light curtain is a safety device that detects nearby obstacles (e.g., a human) to stop the operation of a machine. Light curtains are ubiquitous. They are used in garage doors and elevators to stop the doors from closing when an object blocks them. They are used on factory floors around dangerous machinery. Light curtains operate on a simple principle that an object is detected when it obstructs the line of sight between a source and a sensor. Light curtains are highly reliable, yet are very simple devices. However, light curtain systems are specifically customized for each machine and task, hindering their broad use for vision and robotics.

Recently, Wang et al. [20] extended this principle to generate light curtains along any ruled surface using a line sensor and a line source. Here, an obstacle is detected when it intersects both the plane of illumination generated from the line source and the imaging plane captured by the line sensor. The illumination and imaging planes are rotated using steerable mirrors at different velocities to sweep out an arbitrary ruled curtain surface in 3D. Such triangulating light curtains are highly flexible and can be useful for obstacle detection and avoidance for autonomous navigation.

In this paper, we build upon the idea of triangulation light curtains [20] and develop a general framework for agile depth sensing for vision and robotics. Note that the triangulation light curtain intersects the scene along 3D lines. Thus, instead of capturing 3D data of the scene in the entire volume of interest (using say, a LIDAR [16] or Kinect [22]), this framework allows us to flexibly sample the depths along 3D lines in a scene over time. The depth sampling detects the presence or absence of objects at these locations in real-time without any additional computation. The depth sampling could be sparse, non-uniform (including random), or dense only at specified 3D surfaces. The sparse sampling can be used to adaptively increase the spatial density of the depths only in the regions of interest as specified by an application. Alternatively, objects in the scene can be discovered quickly by initial random sampling followed by adaptive sampling of depths. The depth sampling can be rapidly varied over time depending on the task at hand.

Our agile depth sensing framework has several advantages over traditional depth sensing that uses a fixed acquisition strategy independent of the scene. First, we show that it is possible to capture small, thin, and fast moving objects that are difficult for low frequency uniform angular resolution LIDAR, or low spatial resolution Kinect-like sensors. Example objects include thin wires or meshes, or balls thrown at high speed. We also demonstrate fast and dense 3D capture of objects far away from the sensor when they are barely visible in a LIDAR point cloud, allowing for better detection, recognition and tracking of such objects. Our framework also allows a robot to explore a region of interest based on initial sparse depth estimates. By continuously sampling the scene as the robot moves, it is possible to simultaneously detect obstacles and map the scene.

For videos and supplemental material please see our website: http://www.cs.cmu.edu/~ILIM/agile_depth_sensing

To achieve these results, we present a novel design for triangulation light curtains that uses a 2D camera and a mirror-steered laser line. The rolling shutter of the 2D camera and the rotating laser line triangulate at a set of 3D lines in the scene forming a light curtain. By controlling the pixel clock and steering the laser mirror, it is possible to generate arbitrary ruled surfaces as in Wang et al. [20]. However, because we use a rapid 2D rolling shutter sensor, we achieve significant advantages over the line sensor design: (a) our light curtains have a refresh rate of 60 fps (more than 10x speedup) allowing us to change the curtains rapidly and adaptively for the first time, (b) our system is more light efficient and achieves similar range with less collected light because the optics in front of a 2D sensor are not limited by the size of a steering mirror, and (c) the system has fewer moving parts and is more reliable. Our system of triangulation light curtains and depth sampling works outdoors at ranges up to $20-30\mathrm{m}$ and indoors at up to $50\mathrm{m}$ (something that Kinect [22] cannot do). For this, we propose a new method to suppress ambient illumination.

In summary, our work enables real-time adaptation for agile depth sensing tasks ranging from human robot interaction [11], to robot manipulation [10], path planning [6] and navigation [4]. Our system is easily reconfigurable and can greatly impact robotics and manufacturing applications.

2. Related Work

Safety light curtains are used for ensuring safe operation of automated machinery near humans but since they typically only protect a planar region, 2D scanning LIDAR units are used for more complex settings. Safety laser scanners detect objects radially out to a maximum range from the device and can be configured to trigger a safety event if something enters a programmed 2D safety zone [17].

The ability of triangulation light curtains to image light from a single depth in a scene is a geometric form of depth gating. Temporal depth gating [2, 7] also images light from a specific depth but does so by imaging a pulsed laser with a synchronized high-speed gated shutter camera. By emitting a short laser pulse and then briefly opening the shutter once the time has elapsed for the light to travel to the depth of interest and back (usually pico- to nano-seconds), the camera will only receive light from this depth of interest. Another depth selective imaging method uses time-of-flight cameras and on-chip modulation to reject light outside a specified depth range [18]. Push-broom stereo selectively images depths by only processing pixel disparity pairs that have matching features [3], and enables high-speed obstacle detection and avoidance on UAVs.

Existing work in robust depth scanning has shown that imaging points [12] and lines of light [8, 13] perform much better in the presence of global and ambient light than full frame methods. Synchronized imaging of epipolar-aligned

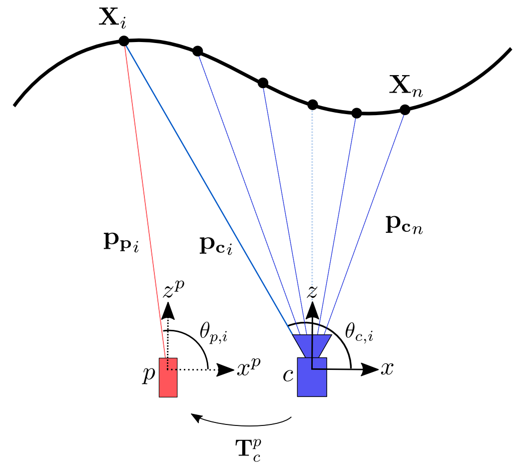


Figure 1: Top-view of the geometry of a triangulation light curtain using a 2D camera. Given that a 3D point, $\mathbf{X}_i$, is on the intersection of the camera plane, $\mathbf{p}_{\mathbf{c}_i}$, and the desired curtain surface, the required projected light plane, $\mathbf{p}_{\mathbf{p}_i}$, and angle, $\theta_{p,i}$, to image that point can be found geometrically. As the camera plane changes (indicated by the faint lines), the required light planes are calculated for each new point.

planes of light with planes of imaging enable video frame rate depth imaging of scenes with high levels of robustness to ambient light and scattering [14, 1].

To reduce the amount of data processed from 3D LIDAR scanners and depth cameras, advanced pointcloud filtering methods have been developed that intelligently and adaptively filter entire pointclouds and process only virtual regions of interest around objects [5]. Recent advances in beam steering technology and MEMS mirrors have enabled the adaptive sampling of scenes with LIDAR point scanners by "zooming in" on certain regions of interest while sampling the rest of the scene at a lower resolution [19, 9]. This type of adaptive sampling can reduce pointcloud processing and enable more intelligent sampling of the scene so that fewer points are sensed and immediately filtered out.

3. Light Curtains with 2D Cameras

Wang et al. showed that light curtains could be imaged by steering a line imager with a galvomirror, but the nature of that design limited the frame rate to 5.6fps and used a small lens to fit the line sensor's field of view onto the rotating galvomirror, which reduced its light efficiency and range [20]. Both of these issues can be improved by using a 2D camera and larger aperture lens.

Instead of imaging along a plane like a mirror-steered line imager, a 2D camera images the scene with a 2D set of discrete rays that are defined by the optics and pixel array. Light curtains can fundamentally be generated by imaging any given set of light rays with any set of camera rays that intersect at a curve. With an aligned plane of imaging and plane of light from a light sheet projector, as used in [20], the intersection curve is a line, but when a light plane is imaged with a set of rays from a 2D camera, the intersection is not restricted to a line and can now be a 3D curve. The position and orientation of the light plane to the camera also affects the orientation and shape of the intersection. This means that any lens and light-plane orientation can be used to image light curtains but that the resulting shape of the curtain can be a complex 3D surface. For lenses with no distortion and parallel imaging and illumination rotation axes as in [20], a ruled surface is generated, but for large-distortion lenses or highly-skewed rotation axes the surface is highly-curved. Regardless of the configuration, calibration of the light sheet projector and camera can guarantee the 3D location of every point in the generated curtains.

In this work, we choose to use low-distortion optics and parallel rotation axes to capture vertical ruled surface light curtains, which enables simple and intuitive curtain design in just two-dimensions. This vertical extrusion of a 2D profile to form a 3D light curtain is very useful, as it matches how most objects in the world protrude from the ground (e.g. pedestrians, buildings, cars, etc) and enables the capture of the front of these objects with a single curtain.

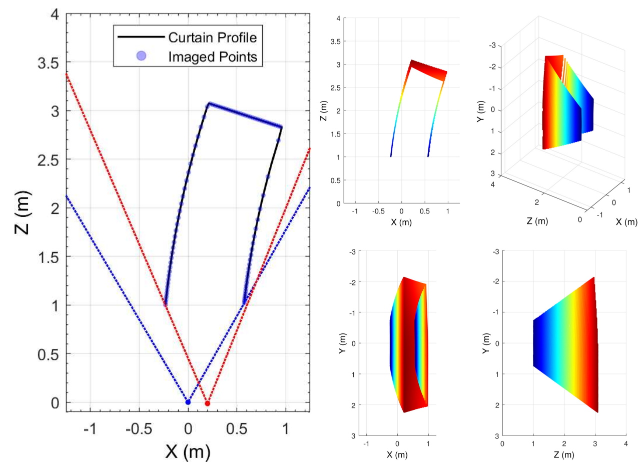
3.1. Light Curtain Design

To design a light curtain, the intersection of the camera pixel rays and the desired light curtain surface must be found. For vertical ruled surfaces, we assume that the pixel rays of a given camera column are coplanar and that the light sheet projector emits a true plane of light and that its rotation axis is parallel to columns of the camera. These assumptions can generally be enforced by using a low distortion lens, careful design of the light sheet projector optics, and precise alignment of the projector axis to the camera columns. With these assumptions, light curtains can be designed in two-dimensions by looking at the rays on the xz-plane, as shown in Figure 1. Given that a 3D point in the camera frame, $\mathbf{X}_i$, lies on the light curtain surface, the camera plane, $\mathbf{p}_{\mathbf{c}i}$, going through this point is found by creating a plane from $\mathbf{X}_i$ and two points that lie on the camera plane's rotation axis. The required laser plane is then found in a similar manner by first projecting $\mathbf{X}_i$ into the frame of the projector to get

$$\mathbf{X}_i^p = \mathbf{T}_c^p \mathbf{X}_i, \tag{1}$$

where $\mathbf{T}_c^p$ is the transformation matrix that converts points in the camera frame to the projector frame found through calibration. The point $\mathbf{X}_i^p$ is then used with two points on the projector rotation axis to find the projected light plane, $\mathbf{p}_{\mathbf{p}_i}^p$, which is in the frame of the projector. This plane in the camera frame is then found by

$$\mathbf{p}_{\mathbf{p}i} = {\mathbf{T}_c^p}^\top \mathbf{p}_{\mathbf{p}_i^p}. \tag{2}$$



(a) 2D Profile **(b)** 3D Profile

Figure 2: Light curtain design with 2D cameras. (a) Light curtains are designed with a 2D camera by finding the points at the intersection of the camera rays with the desired curtain profile. This produces a non-uniformly spaced set of points on the curtain surface. Field-of-view of the camera and light sheet projector are shown by dotted blue and red lines, respectively. (b) Due to small optical distortion and alignment error, the actual 3D profile of the curtain is not a perfect extrusion of the designed curtain into 3D, as shown by the difference in curtain shape along each axis. For example, note the varying depth of the front of the curtain in the z-axis. The blue to red coloring indicates depth along the z-axis from $0-3$m.

To image a curtain profile, the desired profile is first discretized into a series of m points uniformly distributed along the curtain profile, where m is approximately the number of columns of the camera. Line segments are then formed between adjacent points. The ray representing each column is then checked for intersection with the line segments following the approach in [15] to produce a series of points, $\mathbf{X}_i \ .. \ \mathbf{X}_n$, on the column rays that lie on the light curtain profile, as shown in Figure 2.

It is possible that some of the camera rays will not intersect the desired curtain profile or that the design points will be outside the field of view of the light sheet projector or camera. In these cases the points are marked invalid and not used. If the design point is valid, it is transformed into the frame of the light sheet projector using (1) and the galvomirror angle needed to create a light sheet that will travel through the design point is calculated using

$$\theta_{p,i} = \operatorname{atan2}\left(x_{z,i}^p, x_{x,i}^p\right). \tag{3}$$

For an ideal system with true planes of illumination and imaging, the 3D points of the surface are defined by the intersection line of the two planes. However, for a real system, the lens has a small amount of distortion and there may be some small alignment error between the camera and light sheet projector, so the true points of intersection with the light sheet plane for pixels in a given column will not be coplanar and can be found by calculating the intersection of each pixel ray with the light sheet plane, using known calculation methods for ray-plane intersection. This means that the actual 3D surface of the light curtain can vary from the designed profile throughout the camera field of view as shown in Figure 2. However, depending on the severity of the distortion and misalignment, the curtain profile is fairly consistent towards the middle of the field-of-view and only changes significantly towards the edges.

4. Fast Curtains with Rolling Shutter Cameras

Light curtain imaging requires the imaging plane to change and intersect the light curtain surface where desired. For a 2D camera, this can be done by only imaging a select region-of-interest on the imager, but this method is not fast enough to enable agile light curtain imaging.

A much quicker method of imaging the curtain is to use the rolling shutter of a 2D CMOS imager to move the imaging plane. The rolling shutter changes the imaging plane rapidly at a uniform speed specified by the pixel clock of the imager. This characteristic enables light curtain imaging at the full frame rate of the rolling shutter camera. Imaging a curtain is as simple as commanding the light sheet projector to the angle necessary to project the light sheet at the point defined by the intersection of the active row of the 2D camera and the light curtain profile. By synchronizing the motion of the rolling shutter with the motion of the light sheet, light curtains forming any ruled surface can be imaged.

For each frame of the rolling shutter camera an image of the captured curtain is produced. While imaging light curtains, the camera captures both laser light and ambient light. If the ambient light is low enough, (i.e. indoor imaging), the image from the camera can be directly thresholded to produce a mask indicating the detected points. However, in many circumstances the captured ambient light is much greater than the captured laser light and the curtain cannot be detected (i.e. outdoors in sunlight). A narrow band-pass filter significantly reduces the captured ambient light, but the remaining light limits detection sensitivity.

4.1. Ambient Subtraction

To increase the performance in the presence of ambient light, we follow the example of [20], where both an ambient image and laser+ambient image were captured at each light curtain position and then subtracted to produce an image with just the laser light. This enabled great ambient performance but required that the same camera plane be imaged twice. This can be done with a 2D camera and selectable ROI, but not with a single rolling shutter frame capture. To solve this, we developed a method using adjacent columns of a captured image to perform ambient subtraction.

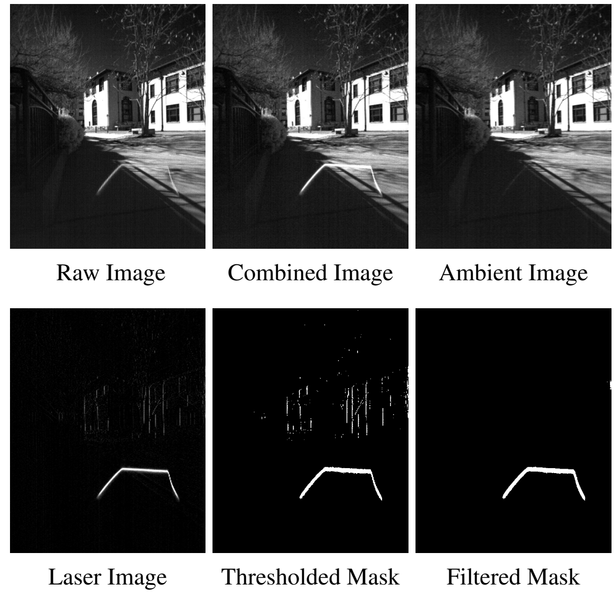


Figure 3: Our ambient light suppression method captures an image where the even columns of the image are with the laser on and the odd columns are with the laser off. Separating these columns and upsampling them forms a full resolution ambient-only image and a combined image. Subtracting the ambient image from the combined image produces a laser-only image, which is then thresholded to find the light curtain detections. Filtering this mask provides robustness to large intensity gradients and upsampling artifacts.

Our method sacrifices curtain resolution and captures a raw image where even columns are captured with the laser on and odd columns with the laser off to get an ambient only image and a combined image, as shown in Figure 3. We then interpolate the two images to form full resolution images and subtract the ambient image from the combined image to get an image with just the laser light, which we call the curtain image. This image is then thresholded to provide a mask indicating the imaged detection points. In areas of high ambient light, this technique may still produce errors at locations of high intensity gradient as shown by the faint edges in the laser image and vertical thin lines in the thresholded mask image in Figure 3. We filter out these artifacts with a thin erosion/dilation filter.

Depending on the synchronization precision of the rolling shutter with the motion of the light sheet projector, there may be slight bleed through of the laser light onto neighboring columns which shows up in the ambient image. This bleed-through will reduce the performance of the device since it is subtracted from the combined image. With a precise synchronization, this light can be limited to a few bits, so that it does not greatly affect performance.

4.2. Limitations

By using a rolling shutter camera, we surrender the uniform sampling of the light curtain profile that dual galvomirrors provides in [20], due to the discrete nature of the camera pixel array. This leads to a situation where there may be a gap in the curtain profile that the light curtain device cannot image and occurs when the rays of the camera are similar in direction to the curtain surface. This effect is shown in the top-right portion of the curtain profile in Figure 2. One other disadvantage of using a rolling shutter camera is that each plane of the camera can only be imaged once in a given frame. If a camera plane intersects multiple curtain segments (e.g., a zig-zag), one of the segments must be chosen for the imaged curtain and sequential curtains must be imaged to capture all the points along a given ray. The galvomirror also limits the maximum frequency and amplitude of the curtains that can be imaged. If it is commanded to move too quickly (greater than a 100Hz step function) then it will lag behind the commanded position and will produce a different curtain than desired. In practice this requires that some curtains be smoothed before imaging.

5. Hardware Prototype

The hardware prototype is comprised of a light sheet projector and a rolling shutter camera. The light sheet projector contains a custom-designed line laser module using a 1W 830nm laser diode (Thorlabs LD830-MA1W) that is then collimated and shaped into a line with a 45° Powell Line Lens (Thorlabs PL0145). This laser line is then projected onto and steered with a galvomirror (Thorlabs GVS001). The line laser module is aligned and mounted to the galvomirror in an aluminum mount that enables the precise collinear alignment of the laser line with the galvomirror's axis of rotation. The mirror has dimensions of 14.5mm $\times$ 8.5mm and has a 50° optical scan angle.

The rolling shutter camera (IDS UI-3240CP-NIR-GL) was fitted with a low distortion C-mount lens (Kowa LMVZ4411) with 70°(h) $\times$ 60°(v) field of view. We operate the camera in 2x binned mode for a resolution of 640 $\times$ 512 to increase the signal of the received light and reduce the noise. We placed a 12nm bandpass filter centered at 830nm between the lens and the image sensor to reduce the amount of collected ambient light. The camera is aligned to the galvomirror such that the rows of the camera are parallel with the galvomirror's axis of rotation. The rotated camera was then placed at a fixed baseline of 200mm from the galvomirror's rotation axis. A micro-controller (Teensy 3.2) is used to synchronize the camera, the laser, and the galvomirror. A color 2D helper camera is used for visualizing the light curtains and detected results in the scene by projecting the light curtain to its view.

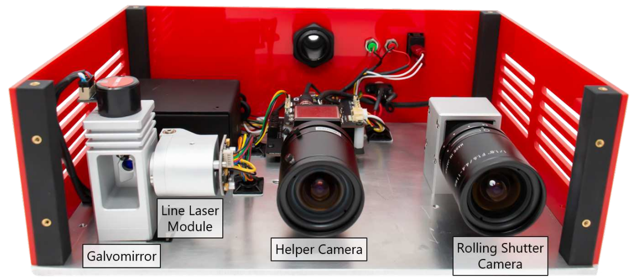


Figure 4: The hardware prototype is comprised of a light sheet projector, rolling shutter camera, and helper camera. The light sheet projector uses a custom line laser module to project a laser line onto a galvomirror which then steers the light sheet onto the light curtain surface. The rolling shutter camera is aligned and rotated 90° so that the rolling shutter and rows of the camera are parallel with the galvomirror axis, with a baseline of 200mm. The helper camera is used for visualization of the light curtains only.

5.1. Calibration

The performance of light curtains depends on precise calibration of the camera and light sheet projector. First, the camera intrinsics were determined with [21] and then the extrinsic calibration of the light sheet projector and camera was found by imaging a set of light planes projected by the light sheet projector onto a planar wall. A checkerboard target of known dimensions was then attached to the wall and imaged with the calibrated camera to get the known 3D coordinates of points on each imaged laser line. This was repeated with the same set of planes at several depths to fully define each plane of light. Best fit plane equations were then found for each set of points using weighted least squares where the weights were the normalized intensity values. Then, given the equations for the planes, the location of the galvomirror axis with respect to the camera was found by a least squares fit of a line to the intersection of all the planes. We then fit a function to the relationship of the plane angle (with respect to the galvomirror axis) and the light sheet projector angle, which is then used to determine the galvomirror position needed for a given design point.

5.2. Capture Process

In [20], the light sheet and camera plane angles were commanded by the positions of galvomirrors, but with a 2D camera, the camera plane angle is defined by the optics and the active line of pixels on the imager. For a rolling shutter camera, the active line of pixels is defined by the speed of the rolling shutter and the time since the start of the frame. The speed of the rolling shutter is determined by the readout rate of the pixels, known as the pixel clock. The maximum time that a given line is active is found by dividing the number of pixels on the line by the pixel clock. At the maximum 60 frames per second capture rate of the camera, the maximum active exposure time of a line is ≈ 15µs.

To image a light curtain, a host computer transmits the required galvomirror positions to the micro-controller which then triggers the camera to start the frame capture. The micro-controller then sequentially commands the galvomirror positions and laser power in lock step with the timed progression of the rolling shutter. This process repeats for each successive light curtain and a different light curtain can be imaged every frame at the full frame rate of the camera.

5.3. Performance

Working Range: At maximum frame rate, our device has a maximum working range of 20m while imaging a whiteboard in approximately 50klx of ambient light and a working range of 50m indoors. This is similar to the device in [20], but with only 15% of the 100µs exposure time used in that device. When our device is configured for an exposure of 100µs the device can image over 40 meters outdoors in similar conditions, but at this range the detection ability is actually more limited by the resolution of the camera.

Speed & Resolution: Our device is capable of imaging 60 different light curtains per second. This speed and flexibility enables agile and dynamic light curtains that can be used to intelligently sample the scene. Figure 5 shows the results of imaging various types of light curtains both indoors and outdoors and are just a sample of the different types of curtains that can be imaged with our device. Images show the light curtain surface and detections projected into the helper camera's view for visualization on the scene. The rapid capture rate and resolution of our device enables the imaging of small and fast objects as they pass through light curtains, as shown in Figure 6a-6b. The resolution of light curtains provide increased detail over scanning LIDAR devices and can enable enhanced object recognition and critical detection of small objects (e.g. wires, branches, etc). This is especially noticeable when imaging thin structures or objects at a distance as shown in Figure 6c-6f. Please see the supplementary material for videos of the device's capabilities.

6. Depth Sensing Strategies

Light curtains do not provide depth of the entire volume at once like traditional depth sensors, but rather enable selective depth sampling of the volume along a given curtain. Uniform sampling by sweeping a plane or other constant shape through a scene provides results similar to traditional sensors, but is inefficient since most scenes are not uniformly structured. A possibly more efficient but nondeterministic method is random sampling of the volume by fitting a curtain to randomly generated points within the volume. By changing the random curtains, different parts of the scene are imaged and an estimate of the full volume

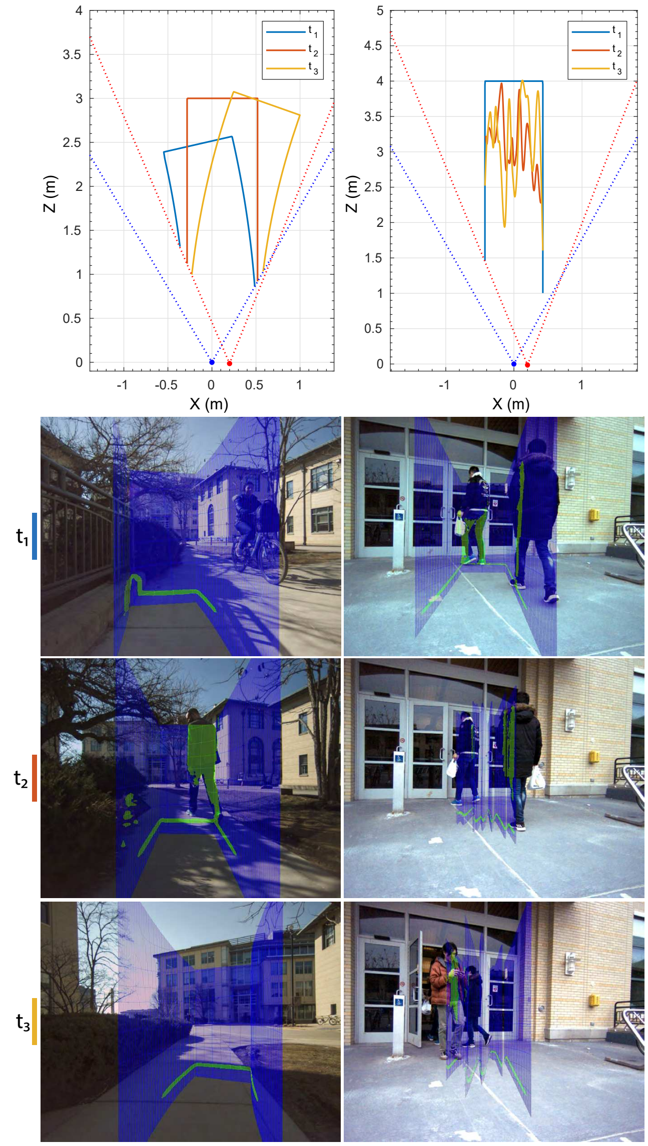


Figure 5: Light curtains captured with our prototype can be imaged and changed up to 60 times per second to quickly scan different curtains through a volume. The images shown here are from the 2D helper camera's view with the light curtain surface rendered in blue and detections rendered in green. The agile nature of these curtains enables many applications including the rapid checking of planned paths (left), and detecting obstacles entering and within a safety zone by rapidly alternating between checking the border of the safety-zone and the area within it (right).

is produced quicker than uniform scanning. The sampling process could be even more efficient if the curtains dynamically changed based on the current estimate of the scene.

Adaptive Depth Imaging: The ability to specify depths of interest at high-rates enables intelligent depth imaging of an

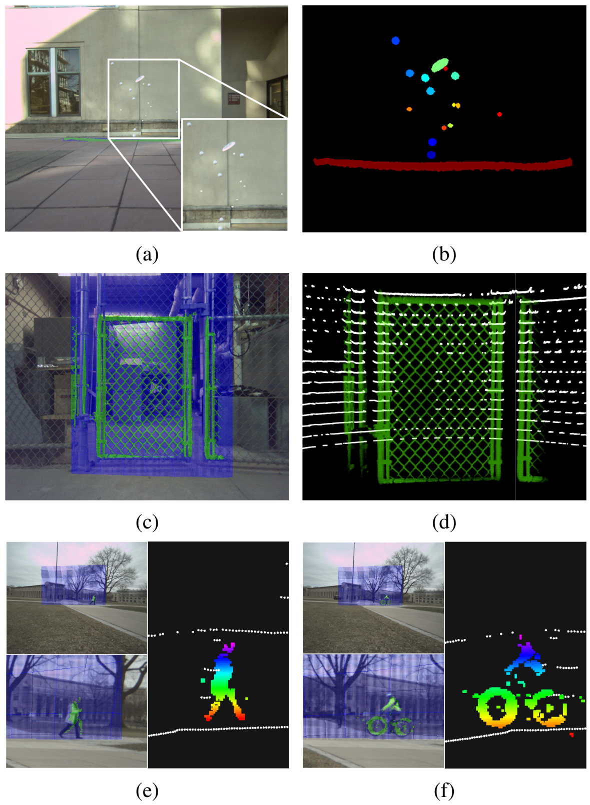


Figure 6: Our device captures small, fast objects and provides more detail than scanning LIDAR devices. (a) Several composited images show the view and detected location (b) of small 30-70mm diameter balls and a 265mm diameter Frisbee that were thrown through a planar curtain 5m away. (c) A planar curtain images a thin wire fence and reconstructs the fence mesh (d) at much higher resolution (green points) than a Velodyne VLP-16 scanning LIDAR (white points) at a distance of 1.5m away. The robust ambient performance and resolution of the light curtain enable it to create high resolution height maps of objects at ranges of 15m outdoors (e-f). At this range, a static VLP-16 only senses the objects as a few points, shown in white. The colors indicate the height of the object $0 - 1.75\text{m}$.

environment based on the current knowledge of the scene. For example, when a device first enters into the scene, it has no knowledge of it's environment, but by quickly scanning a volume of interest with light curtains it can generate a coarse estimate of the locations of objects in the scene, and a new curtain can be designed to image around these points of interest. This process is then rapidly repeated to produce an accurate map of the objects in the scene. Figure 7 shows an example where random curtains were used to initialize the curtain by randomly sampling a volume of interest within 3m of the light curtain. These curtains detected several objects, which the device then used as design points to fit a new curtain to the front surfaces of the objects.

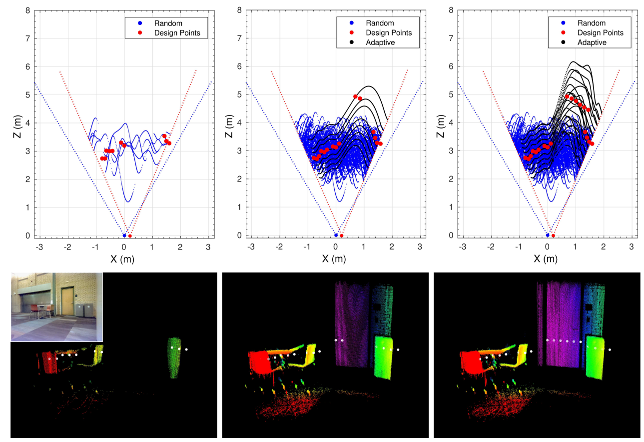


Figure 7: Adaptive depth imaging of a table and chairs scene is performed by first randomly sampling a volume of interest to find objects. Then, the detected points at the front of these objects (indicated by red dots in plots and white dots in bottom row) are used to design a set of light curtains which is then imaged to refine and discover more of the scene. As time continues, the curtains eventually discover and image the back wall of the scene.

Within a few milliseconds, a set of 10 additional curtains were designed that were scaled versions of this curtain that would cover the area directly in front of and behind the detected scene points to refine and discover more of the scene. By interleaving a few random curtains with the adaptive curtains, the device can continue checking the scene for any changes and sample the rest of the scene at low resolution. The design process for the experiment in Figure 7 and all other adaptive sensing experiments projected the detected points to the xz-plane and used the closest point within a $\approx 2.5°$ angular region as the design point for that region. By splitting the entire plane into these uniformly spaced angular regions, the set of design points were determined.

Depth Imaging Comparison: A comparison of depth imaging methods using different light curtain types was performed by capturing the same scene with each curtain type and comparing the coverage of the scene at specified times. The methods included plane sweeping, random sampling, and adaptive depth imaging. For plane sweeping, the planes were designed to fully image the 4.25m scene in 1.0s with a depth resolution of 0.2m. The adaptive curtains were tuned to sense the front surfaces of detected objects at a high resolution and the rest of the scene at a low resolution. Figure 8, shows the results of this comparison. Once initialized with 0.25s of random detections, the discovery nature of the adaptive curtain enabled it to sense near the detected objects and not waste time sensing empty areas at the front of the scene. This approach quickly covered the interesting parts of the scene in less than 0.25s. Given the time, plane sweep

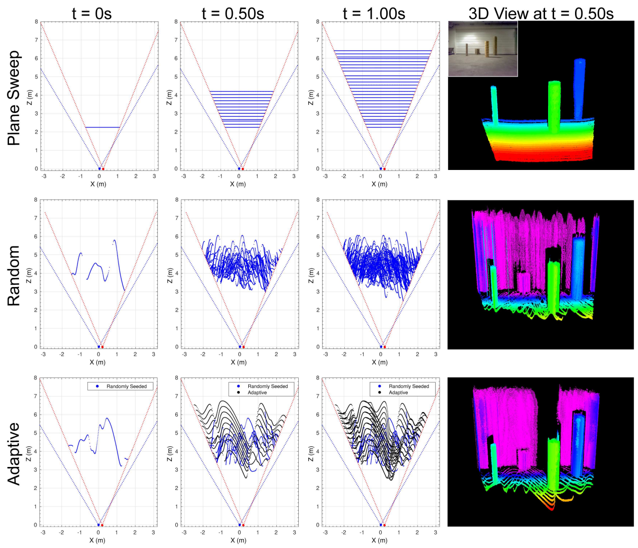


Figure 8: Depth imaging of a scene using planar curtains, random curtains, and adaptive curtains, shows that adaptive curtains intelligently cover more of the scene in less time than sweeping a plane or random sampling of the scene. After initialization with random sampling, the adaptive curtains discovered the interesting parts of the scene in 0.25s and then continued to refine the map.

curtains can provide high resolution and complete coverage, but with limited time the random and adaptive curtains can image more of the scene in less time. For example, the plane sweep curtains could have been configured to image the entire scene in 0.5s but only at 0.4m depth resolution, which is much less than the other methods.

Discovery and Mapping using Adaptive Light Curtains: Adaptive light curtains can be used to discover and map a scene from a moving platform. We used a small robot with on-board localization to move the light curtain device through a cluttered highbay scene. As the robot progressed through the scene, adaptive light curtains discovered the scene structure and continuously adapted to image the newly detected objects from each frame. Rather than using a fixed depth sensing strategy, the light curtains intelligently sampled regions around the detected objects at high-resolution and sampled the rest of the scene with random curtains at a lower resolution. For our mapping experiments, a set of 5 uniquely random curtains were interleaved with every 10 adaptive curtains. Figure 9 shows a few instances of the curtains fitting to objects in the scene as well as a generated map of the full route. Please refer to the supplementary material to see a video of these results.

7. Discussion

We believe that our work opens a new direction of research in adaptive 3D sensing. Over the lifetime of computer vision research, we have been mainly restricted to 3D sensors that capture entire volumes independent of the scene. Never before have we had the rapid ability to selectively and adaptively sense only the parts of an unknown moving scene that are necessary for object discovery, tracking, navigation, and mapping. This work is the first to introduce and demonstrate that light curtains can be used for much more than detection of obstacles at static surface locations. Our 2D camera light curtain implementation can complement the vast majority of vision and robotics research that already include one or more 2D cameras. Future research could include real-time 3D tracking of moving objects as well as active sensing for localization and mapping.

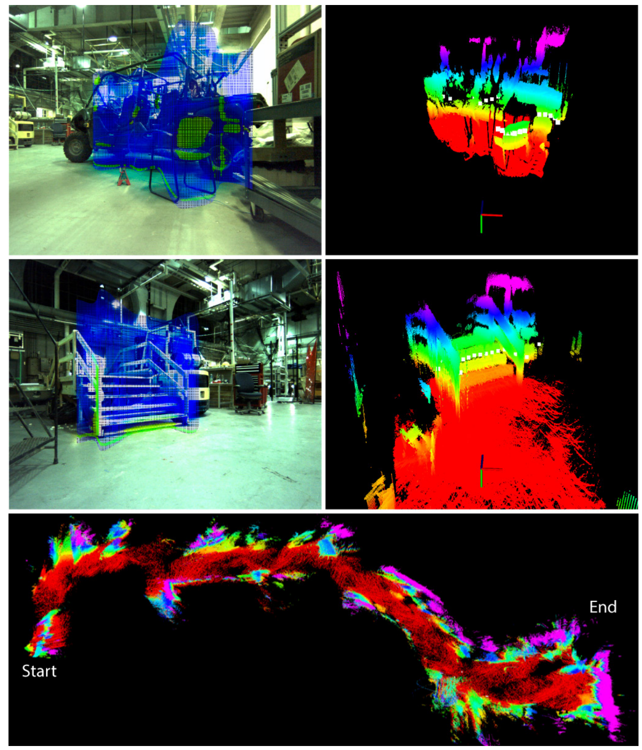


Figure 9: As the light curtain device was moved through a highbay scene, adaptive light curtains discovered the scene and adapted to image the surfaces in the scene. Notice that the curtains (shown in blue) fit tightly around the objects in the scene. When moved throughout the environment, the detected points on the curtains (shown in green) also mapped the scene (right column). The white points on the 3D maps show the design points for the set of adaptive curtains. The bottom row shows a top-down view of the map generated from the entire route.

Acknowledgements. This work was sponsored by the Defense Advanced Research Projects Agency Reveal Grant HR00111620021 and the National Science Foundation Grant CNS-1446601 and 1900821.

References

[1] Supreeth Achar, Joseph R. Bartels, William L. 'Red' Whittaker, Kiriakos N. Kutulakos, and Srinivasa G. Narasimhan. Epipolar time-of-flight imaging. *ACM Trans. Graph.*, 36(4):37:1–37:8, July 2017. 2

[2] Ian M Baker, Stuart S Duncan, and Jeremy W Copley. A low-noise laser-gated imaging system for long-range target identification. In *Infrared Technology and Applications XXX*, volume 5406, pages 133–145. International Society for Optics and Photonics, 2004. 2

[3] Andrew J Barry and Russ Tedrake. Pushbroom stereo for high-speed navigation in cluttered environments. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3046–3052. IEEE, 2015. 2

[4] Joydeep Biswas and Manuela Veloso. Depth camera based indoor mobile robot localization and navigation. In *2012 IEEE International Conference on Robotics and Automation*, pages 1697–1702, May 2012. 2

[5] Jaebum Choi, Simon Ulbrich, Bernd Lichte, and Markus Maurer. Multi-target tracking using a 3d-lidar sensor for autonomous vehicles. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 881–886, Oct 2013. 2

[6] Dmitri Dolgov, Sebastian Thrun, Michael Montemerlo, and James Diebel. Path planning for autonomous vehicles in unknown semi-structured environments. *The International Journal of Robotics Research*, 29(5):485–501, 2010. 2

[7] Yoav Grauer and Ezri Sonn. Active gated imaging for automotive safety applications. In *Video Surveillance and Transportation Imaging Applications 2015*, volume 9407, page 94070F. International Society for Optics and Photonics, 2015. 2

[8] Mohit Gupta, Qi Yin, and Shree K Nayar. Structured light in sunlight. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 545–552, 2013. 2

[9] Abhishek Kasturi, Veljko Milanovic, Bryan H Atwood, and James Yang. Uav-borne lidar with mems mirror-based scanning capability. In *Laser Radar Technology and Applications XXI*, volume 9832, page 98320M. International Society for Optics and Photonics, 2016. 2

[10] Michael Krainin, Brian Curless, and Dieter Fox. Autonomous generation of complete 3d object models using next best view manipulation planning. In *2011 IEEE International Conference on Robotics and Automation*, pages 5031–5037, May 2011. 2

[11] Alessandro De Luca and Fabrizio Flacco. Integrated control for phri: Collision avoidance, detection, reaction and collaboration. In *2012 4th IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 288–295, June 2012. 2

[12] Minsky Marvin. Microscopy apparatus, Dec. 19 1961. US Patent 3,013,467. 2

[13] Srinivasa G. Narasimhan, Shree K. Nayar, Bo Sun, and Sanjeev J. Koppal. Structured light in scattering media. In *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, ICCV '05, pages 420–427, Washington, DC, USA, 2005. IEEE Computer Society. 2

[14] Matthew O'Toole, Supreeth Achar, Srinivasa G. Narasimhan, and Kiriakos N. Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM Trans. Graph.*, 34(4):35:1–35:13, July 2015. 2

[15] William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical recipes in C++*, volume 3. Cambridge University Press, 2007. 3

[16] Brent Schwarz. Lidar: Mapping the world in 3d. *Nature Photonics*, 4(7):429, 2010. 1

[17] SICK. *microScan3 - The new generation of safety laser scanners*, 2 2018. 2

[18] Ryuichi Tadano, Adithya Kumar Pediredla, and Ashok Veeraraghavan. Depth selective camera: A direct, on-chip, programmable technique for depth selectivity in photography. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 3595–3603, Dec 2015. 2

[19] Zaid Tasneem, Dingkang Wang, Huikai Xie, and Koppal Sanjeev. Directionally controlled time-of-flight ranging for mobile sensing platforms. In *Proceedings of Robotics: Science and Systems*, Pittsburgh, Pennsylvania, June 2018. 2

[20] Jian Wang, Joseph Bartels, William Whittaker, Aswin C Sankaranarayanan, and Srinivasa G Narasimhan. Programmable triangulation light curtains. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 19–34, 2018. 1, 2, 3, 4, 5, 6

[21] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22, 2000. 5

[22] Zhengyou Zhang. Microsoft kinect sensor and its effect. *IEEE MultiMedia*, 19(2):4–10, Feb 2012. 1, 2