Dual Structured Light 3D using a 1D Sensor

Jian Wang[†], Aswin C. Sankaranarayanan[†],
Mohit Gupta[‡], and Srinivasa G. Narasimhan[†]

[†]Carnegie Mellon University, [‡]University of Wisconsin-Madison
{jianwan2,saswin}@andrew.cmu.edu,
{mohitg}@cs.wisc.edu, {srinivas}@cs.cmu.edu

Abstract. Structured light-based 3D reconstruction methods often illuminate a scene using patterns with 1D translational symmetry such as stripes, Gray codes or sinusoidal phase shifting patterns. These patterns are decoded using images captured by a traditional 2D sensor. In this work, we present a novel structured light approach that uses a 1D sensor with simple optics and no moving parts to reconstruct scenes with the same acquisition speed as a traditional 2D sensor. While traditional methods compute correspondences between columns of the projector and 2D camera pixels, our 'dual' approach computes correspondences between columns of the 1D camera and 2D projector pixels. The use of a 1D sensor provides significant advantages in many applications that operate in short-wave infrared range (0.9 - 2.5 microns) or require dynamic vision sensors (DVS), where a 2D sensor is prohibitively expensive and difficult to manufacture. We analyze the proposed design, explore hardware alternatives and discuss the performance in the presence of ambient light and global illumination.

Keywords: Structured light, Dual photography

1   Introduction

Structured light (SL) [9] is one of the most popular techniques for 3D shape acquisition. An SL system uses active illumination, typically via a projector, to obtain robust correspondences between pixels on the projector and a camera, and subsequently, recovers the scene depth via triangulation. In contrast to passive techniques like stereo, the use of active illumination enables SL systems to acquire depth even for textureless scenes at a low computational cost.

The simplest SL method is point scanning [7], where the light source illuminates a single scene point at a time, and the camera captures an image. Correspondence between camera and projector pixels is determined by associating the brightest pixel in each acquired image to the pixel illuminated by the projector. However, this approach requires a large number (N^2) of images to obtain a depth map with $N \times N$ pixels. In order to reduce the acquisition time, stripe scanning technique was proposed where the light source emits a planar sheet of light [25, 1, 5]. Consider a scene point that lies on the emitted light plane. Its depth can be estimated by finding the intersection between the light plane,
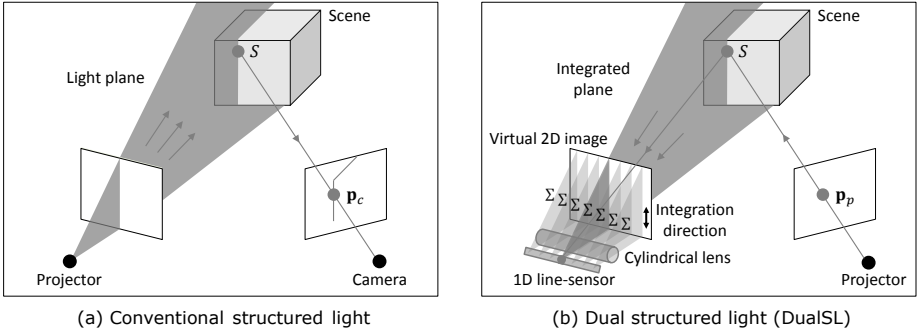
(a) Conventional structured light   (b) Dual structured light (DualSL)

Fig. 1. DualSL compared with traditional SL. Depth from SL can be obtained by performing ray-plane triangulation. For traditional SL, the ray is from camera pixel, and plane is formed by center of projection and a column of the projector. In DualSL, ray is from projector pixel, and plane is formed by a line sensor pixel with cylindrical optics.

and the ray joining the camera center and the camera pixel. This is illustrated in Figure 1(a). We can further reduce the acquisition time by using more sophisticated temporal coding techniques; for example, binary codes [19], Gray codes [22, 12] and sinusoidal phase shifting [26].

Underlying all these methods is the idea that, for a calibrated camera-projector pair, we only need to measure disparity, i.e., a 1D displacement map. Thus, we need to perform coding along only one dimension of the projector image plane, thereby achieving significant speed-ups over point-scanning systems. For example, several structured light patterns have a 1D translational symmetry, i.e., in the projected patterns, all the pixels within a column (or a row) have the same intensities.[1] This is illustrated in Figure 1(a). *For such patterns with 1D translational symmetry, conventional structured light systems can be thought of as using a 1D projector, and a 2D sensor.*

In this paper, we present a novel SL design called *DualSL* (Dual Structured Light) that uses *a 2D projector and a 1D sensor, or a line-sensor*. DualSL comprises of a novel optical setup where pixels on the line-sensor integrates light along columns of the image focused by the objective lens, as shown in Figure 1(b). As a consequence, the DualSL design can be interpreted as the optical dual [23] of a traditional SL system, i.e., we *find correspondences between columns of the camera and pixels on the projector*. In contrast, in conventional SL, we find correspondences between pixels of the camera and columns of the projector.

---

[1] An exception is 'single-shot' structured light techniques that use patterns with 2D intensity variations, for example, sparse 2D grid of lines [20], 2D color encoded grids [21], 2D pseudo-random binary code [27], and 2D random dots (used in the first generation Microsoft Kinect depth sensing cameras [16]).

Why use a 1D sensor for structured light? The use of a line-sensor, instead of a 2D sensor, can provide significant advantages in many applications where a 2D sensor is either expensive or difficult to obtain. For example, the typical costs for sensors in shortwave infrared (SWIR; 900nm-2.5μm) is \$0.10 per pixel [8]; hence, a high-resolution 2D sensor can be prohibitively expensive. In this context, a system built using a 1D line-sensor, with just a few thousand pixels, can have a significantly lower cost. A second application of DualSL is in the context of dynamic vision sensors (DVS) [14], where each pixel has the capability of detecting temporal intensity changes in an asynchronous manner. It has been shown that the use of a DVS with asynchronous pixels can reduce the acquisition time of line striping based structured light by up to an order of magnitude [4, 15]. However, the additional circuit at each pixel for detecting temporal intensity changes and enabling asynchronous readout leads to sensors that are inherently complex and have a poor fill-factor (around 8.1% for commercially available units [6, 11]), and low resolution (e.g., 128×128). In contrast, a 1D DVS sensor [18] can have a larger fill-factor (80% for the design in [18]), and thus, a significantly higher 1D resolution (e.g., 2048 pixels), by moving the per-pixel processing circuit to the additional space available both above and below the 1D sensor array.

Our contributions are as follows:

– **SL using a line-sensor.** We propose a novel SL design that utilizes a line-sensor and simple optics with no moving parts to obtain the depth map of the scene. This can have significant benefits for sensing in wavelength regimes where sensors are expensive as well as sensing modalities where 2D sensors have low fill-factor, and thus poor resolution (e.g., dynamic vision sensors).
– **Analysis.** We analyze the performance of DualSL and show that its performance in terms of temporal resolution is the same as a traditional SL system.
– **Validation via hardware prototyping.** We realize a proof-of-concept hardware prototype for visible light to showcase DualSL, propose a procedure to calibrate the device, and characterize its performance.

2   DualSL

In this section, we describe the principle underlying DualSL, and analyze its performance in terms of the temporal resolution of obtaining depth maps.

2.1   Design of the sensing architecture

The optical design of sensing architecture, adapted from [28], is shown in Figure 2. The setup consists of an objective lens, a cylindrical lens, and a line-sensor. The line-sensor is placed on the image plane of the objective lens, so that the scene is perfectly in focus along the axis of the line-sensor. A cylindrical lens is placed in between the objective lens and the sensor such that its axis is aligned
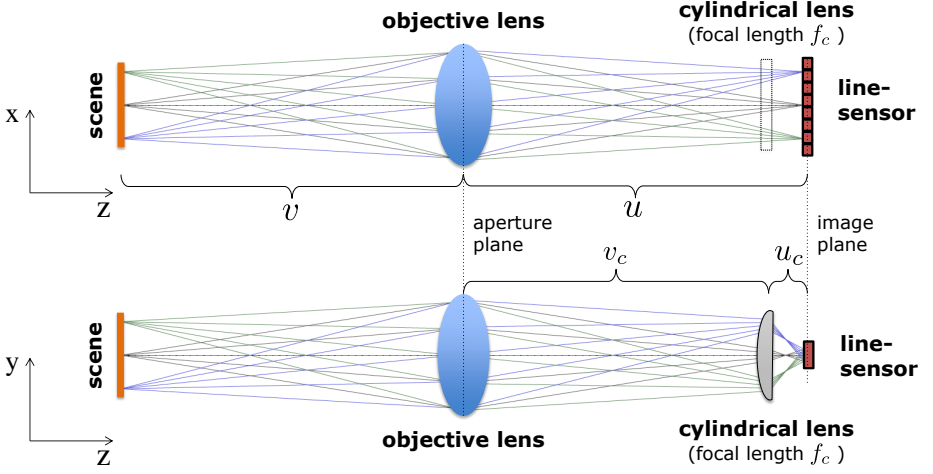
Fig. 2. Design of the sensing architecture visualized as ray diagrams along two orthogonal axes. The line-sensor is placed at the image plane of the objective lens. A cylindrical lens is placed in front of the line-sensor such that its axis is aligned with that of the line-sensor. The cylindrical lens does not perturb light rays along the x-axis (top-row); this results in the scene being in focus along the x-axis. Along the y-axis (bottom-row), the cylindrical lens brings the aperture plane into focus at the image plane. Hence, the scene is completely defocused along the y-axis, i.e., each line-sensor pixel integrates light along the y-axis.

with that of the line-sensor. The cylindrical lens does not perturb light rays along the x-axis (axis parallel to its length). This results in the scene being in focus along the x-axis. Along the y-axis (perpendicular to the length of the cylindrical lens), the position and focal length of the cylindrical lens are chosen to ensure that its aperture plane is focused at the image plane. Hence, the scene is completely defocused along the y-axis, i.e., each line-sensor pixel integrates light along the y-axis. This is illustrated in Figure 2(bottom-row). Further, for maximum efficiency in gathering light, it is desirable that the aperture of the objective lens is magnified/shrunk to the height of the line-sensor.

Determining the parameters of the cylindrical lens. The focal length, f_c, of the cylindrical lens and its distance to the line-sensor, u_c, can be derived from the desiderata listed above. Given the aperture diameter of the objective lens D, the sensor-lens distance u, the height H of the line-sensor pixels, and the length of the line-sensor L, we require the following constraints to be satisfied:

$$\frac{1}{u_c} + \frac{1}{u - u_c} = \frac{1}{f_c} \qquad \text{(focusing aperture plane onto image plane)}$$

$$\frac{D}{H} = \frac{u - u_c}{u_c} \qquad \text{(magnification constraints)}$$

Putting them together, we can obtain the following expressions for u_c and f_c:

$$u_c = \frac{H}{D+H} u, \quad f_c = \frac{HD}{(D+H)^2} u. \tag{1}$$

The last parameter to determine is the height of the cylindrical lens which determines the field-of-view along the axis perpendicular to the line-sensor. For a symmetric field-of-view, we would require the height of the cylindrical lens to be greater than $\frac{D}{D+H}(L+H)$.

Remark. It is worth noting here that line-sensors are often available in form-factors where the height of the pixels H is significantly greater than the pixel pitch. For example, for the prototype used in this paper, the pixel height is 1mm while the pixel pitch is 14μm. This highly-skewed pixel aspect ratio allows us to collect large amounts of light at each pixel with little or no loss of resolution along the x-axis. Further, such tall pixels are critical to ensure that the parameters' values defined in (1) are meaningful. For example, if $D \approx H$, then $u_c = u/2$ and $f_c = u/4$. Noting that typical values of flange distances are 17.5mm for C-mount lenses and 47mm for Nikkor lenses, it is easily seen that the resulting values for the position and the focal length are reasonable.

Scene to sensor mapping. The sensor architecture achieves the following scene-to-sensor mapping. First, the objective lens forms a virtual 2D image of the scene, $I(m,n)$.[2] Second, the effect of the cylindrical lens is to completely defocus the virtual image along one direction. Hence, the measurements on the line-sensor are obtained by projecting the virtual image along a direction perpendicular to the axis of the line-sensor. Specifically, each pixel of the line-sensor integrates the virtual image along a line, i.e., the measurement made by a pixel x on the line sensor is the integration of intensities observed along a line with a slope b/a:

$$i(x) = \int_\alpha I(x + a\alpha, b\alpha) d\alpha.$$

The slope b/a is controlled by the axis of the line-sensor/cylindrical lens. An important feature of this design is that the *pre-image* of a line-sensor pixel is a plane. Here pre-image of a pixel is defined as the set of 3D world points that are imaged at that pixel; for example, in a conventional perspective camera, the pre-image of a pixel is a ray. As we will demonstrate shortly, this property will be used by the DualSL system for acquiring 3D scans of a scene.

2.2   3D scanning using the DualSL setup

We obtain 3D scans of the scene by obtaining correspondences between projector and line-sensor pixels. Suppose that pixel (m, n) of the projector corresponds to

---

[2] For simplicity, the image formation is described while ignoring the effects of pixelation and quantization.

pixel x on the line-sensor. We can obtain the 3D scene point underlying this correspondence by intersecting the pre-image of the projector pixel, which is a line in 3D, with the pre-image of the line-sensor pixel — a plane in 3D. As long as the line and the plane are not parallel, we are bound to get a valid intersection, which is the 3D location of the scene point. For simplicity, we assume that the projector and line-sensor are placed in a rectified left-right configuration; hence, we can choose $a = 0$ and $b = 1$ to integrate vertically (along "columns") on the virtual 2D image.

Obtaining projector-camera correspondences. The simplest approach for obtaining correspondences is to illuminate each pixel on the projector sequentially, and capturing an image with the 1D sensor for each projector pixel location. For each projector pixel, we can determine its corresponding line-sensor pixel by finding the pixel with the largest intensity. Assuming a projector with a resolution of $N \times N$ pixels, we would need to capture N^2 images from the line-sensor. Assuming a line-sensor with N pixels, this approach requires N^3 pixels to be read out at the sensor.

We can reduce the acquisition time significantly by using temporal coding techniques, similar to the use of binary/Gray codes in traditional SL (see Figure 3). In the DualSL setup, this is achieved as follows. We project each row of every SL pattern sequentially.[3] Given that the row has N pixels, we can find projector-camera correspondences using a binary/Gray code with $\log_2(N)$ projector patterns. For this scanning approach, we would require to read out $\log_2(N)$ frames for each row. Given that the projector has N rows, and the sensor has N pixels, a total of $N^2 \log_2(N)$ pixels will need to be read out at the sensor.

2.3   Analysis of temporal resolution

We now show that the temporal resolution of DualSL is the same as that of a traditional SL setup for synchronous-readout sensor, i.e. a conventional sensor. For this analysis, we assume that the goal is to obtain an $N \times N$-pixels depth map. The temporal resolution is defined in terms of the time required to acquire a single depth map. We assume that in a sensor, all pixels share one analog-to-digital converter (ADC). We also assume that the bottleneck in all cases is the ADC rate of the camera. This assumption is justified due to the operating speed of projectors (especially laser projectors) being many orders of magnitude greater than cameras. Hence, the number of pixels to be read out divided by ADC rate is temporal resolution of the system.

Line striping. The simplest instance of a traditional SL setup is to illuminate projector columns, one at a time. For each projector column, we read out an

---

[3] It would be desirable to scan along epipolar lines of the projector-camera setup, since this avoids illuminating multiple scene points that lie on the pre-image of the same sensor pixel. However, this would be hard in non-rectified setups and hence, row-scanning is a simpler and effective alternative.
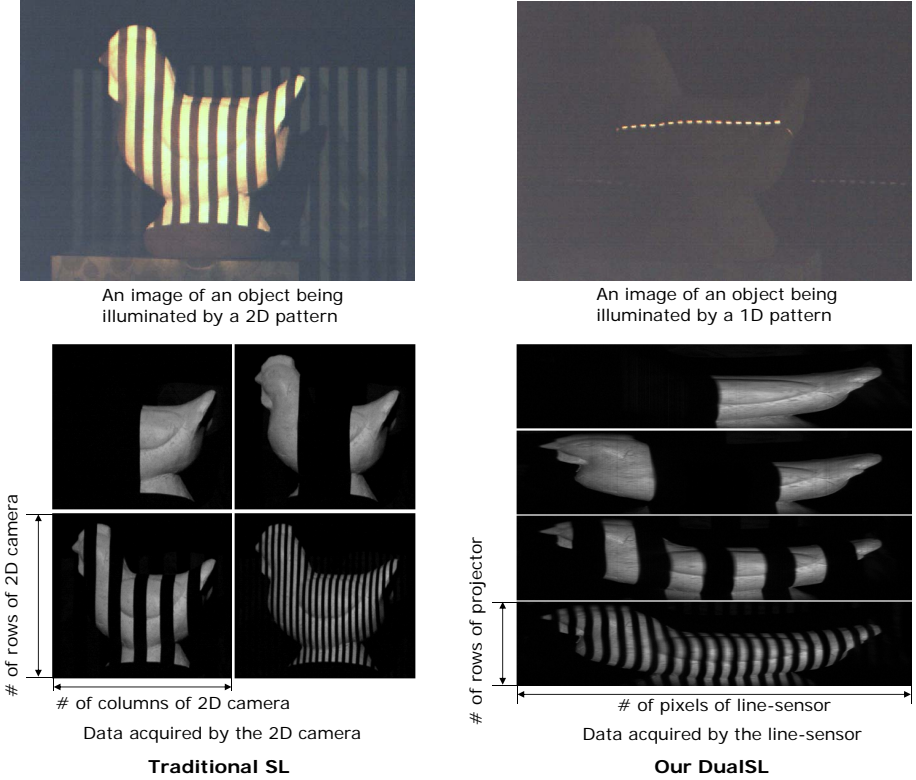
An image of an object being
illuminated by a 2D pattern

An image of an object being
illuminated by a 1D pattern

# of rows of 2D camera

# of columns of 2D camera

Data acquired by the 2D camera

Traditional SL

# of rows of projector

# of pixels of line-sensor

Data acquired by the line-sensor

Our DualSL

Fig. 3. 3D scanning using (left) traditional SL and (right) DualSL. For each setup, we show (top-row) the scene being illuminated by the projector, observed using an auxiliary camera, as well as (bottom-row) measurements made by the cameras. Here the auxiliary camera is not required for triangulation; it is used only for visualization purposes. For the traditional SL setup, these are simply images acquired by the 2D sensor. For DualSL, we stack together 1D measurements made for the same Gray code into individual image.

$N \times N$-pixels image at the camera. Hence, for a total of N projector columns, we read out N^3 pixels at the ADC. DualSL has an identical acquisition time, equivalent to the readout of N^3 pixels per depth map, when we sequentially scan each projector pixel.

Binary/Gray codes. As mentioned earlier, by scanning one projector-row at a time and using binary/Gray temporal codes, we can reduce the acquisition time to the readout of $N^2 \log_2(N)$ pixels per depth map. This readout time is identical to the amount required for a traditional SL system when using binary/Gray coding of the projector columns, where $\log_2(N)$ images, each with $N \times N = N^2$ pixels, are captured.

In essence, with appropriate choice of temporal coding at the projector, the acquisition time and hence, the temporal resolution, of the DualSL is identical to that of a traditional SL system.

For asynchronous readout using a DVS sensor, the temporal resolution is determined by the minimum time between two readouts. Current 1D DVS sensors typically support approximately one million readouts per second [2]. This would be the achievable limit with a DualSL system using DVS.

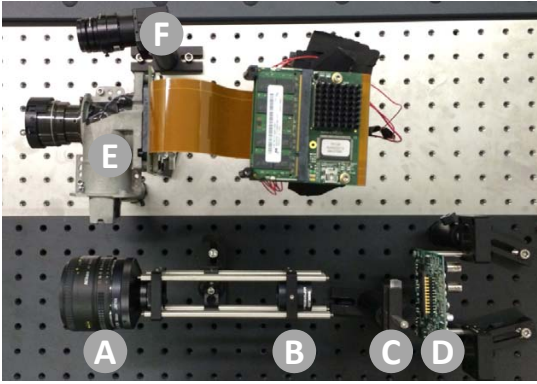### 2.4   DualSL as the optical dual of traditional SL

Consider a traditional SL system involving a 1D projector and a 2D image sensor. Recall that all pixels in a single projector columns are illuminated simultaneously. Let us consider the light transport matrix $\mathcal{L}$ associated with the columns of the projector and pixels on the 2D sensor. Next, let us consider the optical setup whose light transport is $\mathcal{L}^\top$. Under principles of Helmholtz reciprocity, this corresponds to the dual projector-camera system where the dual-projector has the same optical properties as the camera in the traditional SL system, and the dual-camera has the same optical properties as the projector. Hence, the dual-camera integrates light along the planes originally illuminated by the 1D projector in the traditional setup. This dual architecture is the same as that of the DualSL setup and enables estimation of the depth map (and the intensity image) as seen by a camera with the same specifications as the projector.

3   Hardware prototype

In this section, we present the specifications of our DualSL hardware prototype, shown in Figure 4(a). The hardware prototype consists of a 50mm F/1.8 objective lens, a 15mm cylindrical lens, a Hamamatsu S11156-2048-01 line-sensor, and a DMD-based projector built using DLP7000 development kit and ViALUX STAR CORE optics. The line-sensor has 2048 pixels, each of size $14\mu m \times 1$mm. The projector resolution is 768×1024 pixels.
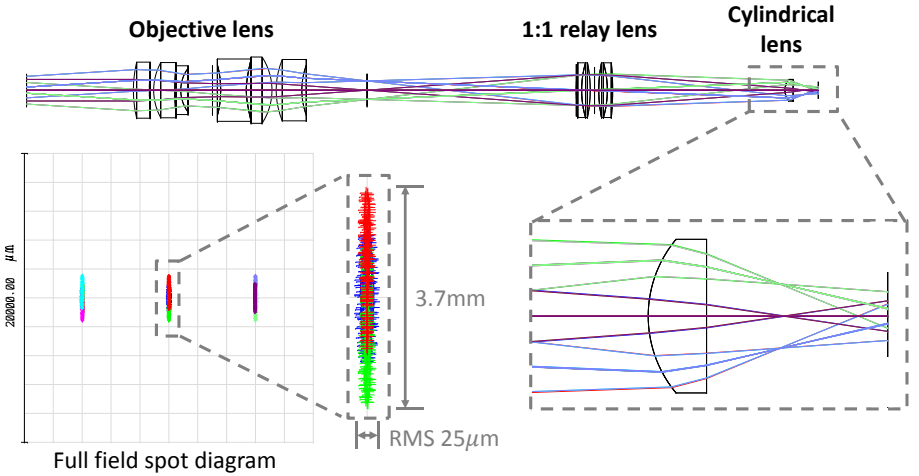
We implemented a slightly different optical design from the ray diagrams shown in Figure 2. To accommodate the cylindrical lens in the tight spacing, we used a 1:1 relay lens to optically mirror the image plane of the objective lens. This provided sufficient spacing to introduce the cylindrical lens and translational mounts to place it precisely. The resulting schematic is shown in Figure 4(b). Zemax analysis of this design shows that the spot-size has a RMS-width of $25\mu m$ along the line-sensor and 3.7mm perpendicular to the line-detector. Given that the height of the line-detector pixel is 1mm, our prototype loses 73% of the light. This light loss is mainly due to sub-optimal choice of optical components.

Calibration. The calibration procedure is a multi-step process with the eventual goal of characterizing the light ray associated with each projector pixel, and the plane associated with each line-sensor pixel.

	Objective lens
A	Nikkor 50mm F/1.8
B	**Relay lens**
	Thorlabs MAP10100100-A
C	**Cylindrical lens**
	Thorlabs LJ1636L2-A f15mm
D	**Line-sensor**
	Hamamatsu S11156-2048-01
E	**Projector**
	DLP7000 DMD and
	ViALUX STAR CORE optics
F	**2D camera helper**
	Point Grey FL3-U3-13E4C-C (not
	required for triangulation in DualSL)

(a) Hardware prototype with list of components

(b) Analysis of spot size

Fig. 4. Our hardware prototype. We used a 1:1 relay lens to mirror the image plane of the objective lens to provide more space to position the cylindrical lens. The spot-size of the resulting setup has a RMS-width of 25μm along the axis of the line-sensor and 3.7mm across.

- We introduce a helper 2D camera whose intrinsic parameters are obtained using the MATLAB Camera Calibration Toolbox [3].
- The projector is calibrated using a traditional projector-camera calibration method [13]. In particular, we estimate the intrinsic parameters of the projector as well as the extrinsic parameters (rotation and translation) with respect to the helper-camera's coordinate system.
- To estimate the plane corresponding to each line-sensor pixel, we introduce a white planar board in the scene with fiducial markers at corners of a rectangle

of known dimensions. The helper-camera provides the depth map of the planar
board by observing the fiducial markers.
- The projector illuminates a pixel to the board which is observed at the line-
  sensor, thereby providing one 3D-1D correspondence, where the 3D location
  is computed by intersection of projector's ray and the board.
- This process is repeated multiple times by placing the board in different poses
  and depths to obtain more 3D-1D correspondences. Once we obtain sufficiently
  many correspondences, we fit a plane to the 3D points associated with each
  pixel to estimate its pre-image.

As a by-product of the calibration procedure, we also measure deviations of
the computed depth from the ground truth obtained using the helper-camera.
The root-mean-square error (RMSE) over 2.5 million points with depth ranges
from 950mm to 1300mm (target is out-of-focus beyond this range) was 2.27mm.

4   Experiments

We showcase the performance of DualSL using different scenes. The scenes were
chosen to encompass inter-reflections due to non-convex shapes, materials that
produce diffuse and specular reflectances as well as subsurface scattering. Figure
5 shows 3D scans obtained using traditional SL as well as our DualSL prototype.
The traditional SL was formed using the helper-camera used for calibration. We
used Gray codes for both systems. To facilitate a comparison of the 3D scans
obtained from the two SL setups, we represented the depth map as seen in the
projector's view, since both systems shared the same projector. Depth maps
from both traditional SL and DualSL were smoothened using a 3×3 median
filter. We computed the RMSE between the two depth maps for a quantitative
characterization of the difference of the two depth maps. Note that, due to
differences in view-points, each depth map might have missing depth values at
different locations. For a robust comparison, we compute the RMSE only over
points where the depth values in both maps were between 500mm and 1500mm.

For the chicken and ball scenes (rows 1 and 2 in Figure 5), both systems
get good results and the average difference is smaller than 2mm. For the box
scene (row 3), the average difference is only slightly larger in spite of complex
geometry of the scene. We fit planes to four different planar surfaces, and the
mean deviation from the fitted planes was 0.45mm with the average distance
to camera being 1050mm. The porcelain bowl scene (row 4), which has strong
inter-reflections, and wax scene (row 5), which exhibits subsurface scattering,
have strong global components. The depth maps generated by DualSL in both
cases are significantly better than that of the traditional SL. This is because
traditional SL illuminates the entire scene, and in contrast, DualSL illuminates a
line at a time, thereby reducing the amount of global light. The exact strength of
global illumination for a general scene, however, depends on the light transport.
For instance, it may be possible to construct scenes where the amount of global
light is smaller for conventional SL over DualSL, and vice-versa. However, a

formal analysis is difficult because global illumination is scene dependent. Here, the depth map recovered by traditional SL has many "holes" because of missing projector-camera correspondences and the removal of depth values beyond the range of (500, 1500)mm.

Figure 6 shows the 3D scans of the five scenes in Figure 5 visualized using MeshLab. We observe that DualSL can capture fine details on the shape of the object. We can thus conclude that DualSL has a similar performance as traditional SL for a wide range of scenes, which is immensely satisfying given the use of a 1D sensor.

5 Discussion

DualSL is a novel SL system that uses a 1D line-sensor to obtain 3D scans with simple optics and no moving components, while delivering a temporal resolution identical to traditional setups. The benefits of DualSL are most compelling in scenarios where sensors are inherently costly. To this end, we briefly discuss the performance of DualSL under ambient and global illumination as well as discuss potential applications of DualSL.

Performance of DualSL under ambient illumination. Performance of SL systems often suffers in the presence of ambient illumination, in part due to potentially strong photon noise. We can measure the effect of ambient illumination using signal-to-noise ratio (SNR), the ratio of the intensity observed at a camera pixel when a scene point is directly illuminated by the projector to the photon noise caused by ambient illumination. The larger the value of SNR, the less is the effect of ambient illumination on the performance since we can more reliably provide thresholds to identify the presence/absence of the direct component. We ignore the presence of global components for this analysis.

The hardware prototype used in this paper uses a DMD-based projector which attenuates a light source, spatially, using a spatial light modulation to create a binary projected pattern. In a traditional SL system, since we read out each camera pixel in isolation, SNR can be approximated as $\frac{P}{\sqrt{A}}$ where P and A are the brightness of the scene point due to the projector and ambient illumination, respectively [10]. Unfortunately, due to integration of light at each line-sensor pixel, SNR of DualSL drops to $\frac{P}{\sqrt{NA}} = \frac{1}{\sqrt{N}} \frac{P}{\sqrt{A}}$ where N is the number of pixels that we sum over. This implies that DualSL is significantly more susceptible to the presence of ambient illumination when we use attenuation-type projectors. This is a significant limitation of our prototype. An approach to address this limitation is to use a scanning laser projector which concentrates all of its light onto a single row of the projector. As a consequence, SNR becomes $\frac{NP}{\sqrt{NA}} = \sqrt{N} \frac{P}{\sqrt{A}}$. In contrast, traditional SL has no gain from using scanning laser projector because it needs projector to illuminate the entire scene.

A more powerful approach is to avoid integrating light and instead use a mirror to scan through scene points in synchrony with a scanning laser projector. Here, we are optically aligning the line-sensor and the illuminated projector pixels to be on epipolar line pairs and is similar to the primal-dual coding system
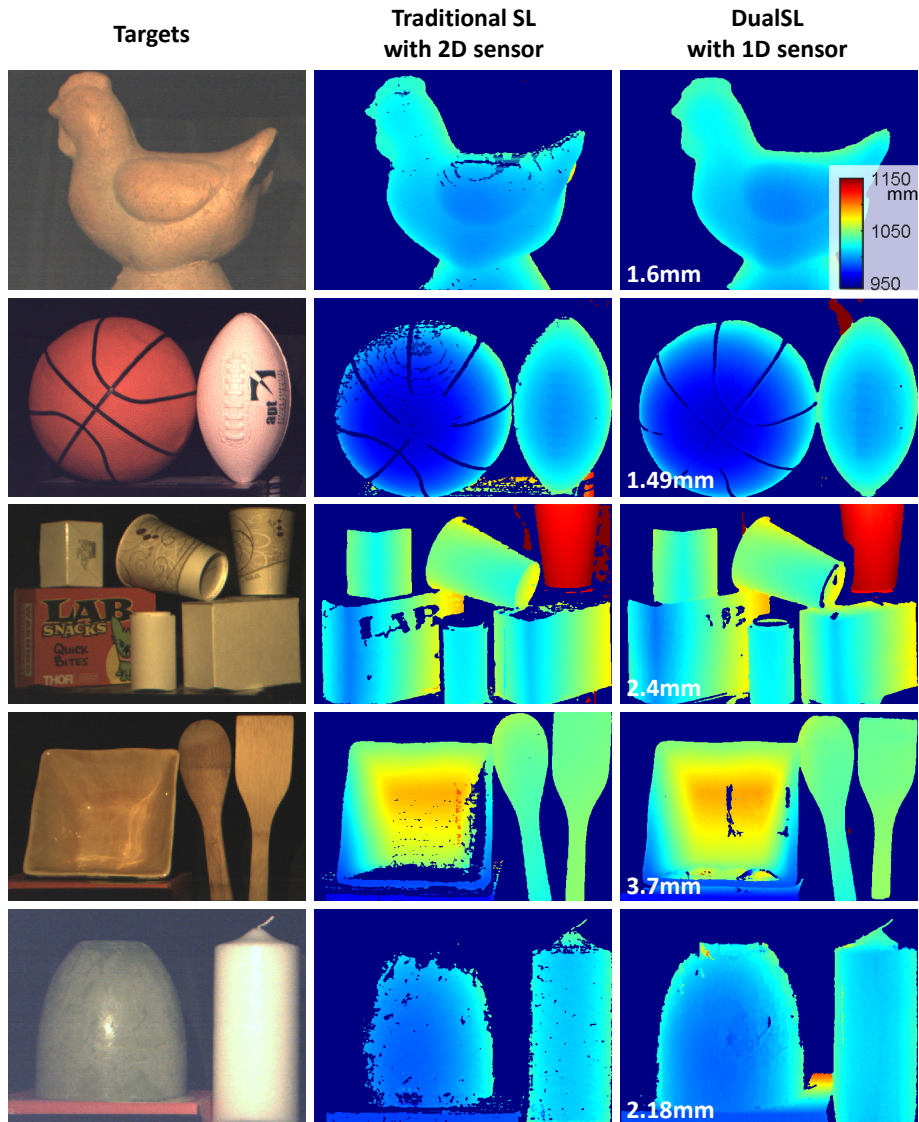
Fig. 5. Depth maps of five scenes obtained using traditional SL and DualSL. Objects in scenes range from simple convex diffuse materials to shiny and translucent materials exhibiting global illumination. Traditional SL is realized by projector and the 2D camera helper here. The left column is a photograph of the target acquired using the 2D camera helper (used only for visualization). The middle and right columns show depth maps obtained by the traditional SL system and DualSL system, respectively. Both depth maps are shown in projector's viewpoint. The number overlaid on DualSL's depth map indicates the average difference between the two depth maps.
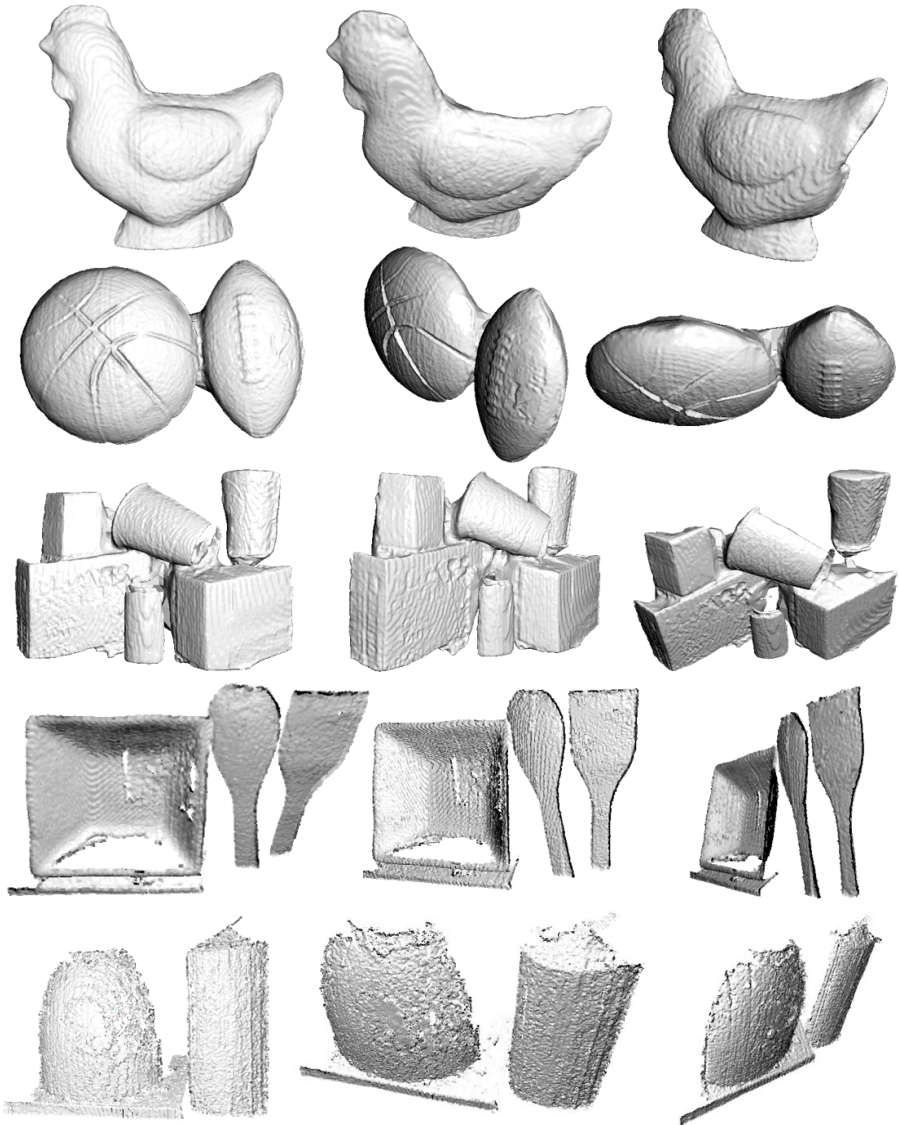
Fig. 6. 3D reconstructions of scenes scanned by DualSL.

of [17]. This enables acquisition of 3D scans that are highly robust to global and ambient illumination.

Performance of DualSL under global illumination. Global illumination is often a problem when dealing with scenes that have inter-reflections, subsurface- and volumetric-scattering. Similar to ambient illumination, global illumination also leads to loss in performance in decoding at the camera. In a traditional SL

system, it is typical that half-the-scene points are illuminated and hence, at any camera pixel, we can expect to receive contributions to the global component from half the scene element. In DualSL, even though we only illuminate one projector-row at a time (and hence, fewer illuminated scene points), each camera pixel integrates light along a scene plane which can significantly increase the amount of global light observed at the pixel. While the results of bowl scene and wax scene in Figure 5 are promising, a formal analysis of the influence of global illumination on the performance of DualSL is beyond the scope of this paper.

Applications of SWIR DualSL. Imaging through volumetric scattering media often benefits via the use of longer wavelengths. SWIR cameras, which operate in the range of 900nm to 2.5μm, are often used in such scenarios (see [24]). The DualSL system design for SWIR can provide an inexpensive alternative to otherwise-costly high-resolution 2D sensors. This would be invaluable for applications such as autonomous driving and fire-fighting operations, where depth sensitivity can be enhanced in spite of fog, smog, or smoke.

High-speed depth imaging using DVS line-sensors. The asynchronous readout underlying DVSs allows us to circumvent the limitations imposed by the readout speed of traditional sensors. Further, the change detection circuitry in DVSs provides a large dynamic range ($\sim$120 dB) that is capable of detecting very small changes in intensity even for scenes under direct sunlight. This is especially effective for SL systems where the goal is simply detecting changes in intensity at the sensor. The MC3D system [15] exploits this property to enable high-speed depth recovery even under high ambient illumination; in particular, the system demonstrated in [15] produces depth maps at a resolution of 128 $\times$ 128 pixels, due to lack of commercial availability of higher resolution sensors. In contrast, a DualSL system using the line-sensor in [2] would produce depth maps of 1024 $\times$ 1024 pixels at real-time video rates. Further, the DualSL system would also benefit from higher fill-factor at the sensor pixels (80% versus 8%).

Active stereo using DualSL. Another interesting modification to the DualSL setup is to enable active stereo-based 3D reconstruction. The envisioned system would have two line-sensors, with its associated optics, and a 1D projector that illuminates one scene plane at a time. By establishing correspondences across the line-sensors, we can enable 3D reconstructions by intersecting the two pre-images of sensor-pixels, which are both planes, with the plane illuminated by the projector. Such a device would provide very high-resolution depth maps (limited by the resolution of the sensors and not the projector) and would be an effective solution for highly-textured scenes.

Acknowledgments

References

1. Agin, G.J., Binford, T.O.: Computer description of curved objects. IEEE Transactions on Computers 100(4), 439–449 (1976)
2. Belbachir, A.N., Schraml, S., Mayerhofer, M., Hofstatter, M.: A novel hdr depth camera for real-time 3d 360-degree panoramic vision. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 425–432 (2014)
3. Bouguet, J.: Camera calibration toolbox for matlab. `http://www.vision.caltech.edu/bouguetj/calib_doc/` (2015)
4. Brandli, C., Mantel, T.A., Hutter, M., Höpflinger, M.A., Berner, R., Siegwart, R., Delbruck, T.: Adaptive pulsed laser line extraction for terrain reconstruction using a dynamic vision sensor. Frontiers in neuroscience 7(EPFL-ARTICLE-200448) (2014)
5. Curless, B., Levoy, M.: Better optical triangulation through spacetime analysis. In: Proc. of International Conference on Computer Vision (ICCV). pp. 987–994 (1995)
6. Delbruck, T.: Frame-free dynamic digital vision. In: International Symposium on Secure-Life Electronics, Advanced Electronics for Quality Life and Society. pp. 21–26 (2008)
7. Forsen, G.E.: Processing visual data with an automaton eye. In: Pictoral Pattern Recognition (1968)
8. Gehm, M.E., Brady, D.J.: Compressive sensing in the EO/IR. Applied optics 54(8), C14–C22 (2015)
9. Geng, J.: Structured-light 3d surface imaging: a tutorial. Advances in Optics and Photonics 3(2), 128–160 (2011)
10. Gupta, M., Yin, Q., Nayar, S.K.: Structured light in sunlight. In: Proc. of International Conference on Computer Vision (ICCV). pp. 545–552 (2013)
11. iniLabs: Dvs128 specifications. `http://inilabs.com/products/dynamic-and-active-pixel-vision-sensor/davis-specifications/` (2015)
12. Inokuchi, S., Sato, K., Matsuda, F.: Range imaging system for 3-d object recognition. In: Proc. of International Conference on Pattern Recognition (ICPR). vol. 48, pp. 806–808 (1984)
13. Lanman, D., Taubin, G.: Build your own 3D scanner: 3D photograhy for beginners. In: ACM SIGGRAPH 2009 courses. pp. 30–34 (2009)
14. Lichtsteiner, P., Posch, C., Delbruck, T.: A 128×128 120 db 15 μs latency asynchronous temporal contrast vision sensor. IEEE Journal of Solid-State Circuits 43(2), 566–576 (2008)
15. Matsuda, N., Cossairt, O., Gupta, M.: MC3D: Motion contrast 3d scanning. In: IEEE International Conference on Computational Photography (ICCP) (2015)
16. Microsoft: Kinect for xbox 360. `https://en.wikipedia.org/wiki/Kinect#Kinect_for_Xbox_360` (2010)
17. O'Toole, M., Achar, S., Narasimhan, S.G., Kutulakos, K.N.: Homogeneous codes for energy-efficient illumination and imaging. ACM Trans. on Graph. 34(4), 35 (2015)
18. Posch, C., Hofstätter, M., Matolin, D., Vanstraelen, G., Schön, P., Donath, N., Litzenberger, M.: A dual-line optical transient sensor with on-chip precision timestamp generation. In: International Solid-State Circuits Conference. pp. 500–618 (2007)

19. Posdamer, J., Altschuler, M.: Surface measurement by space-encoded projected beam systems. Computer graphics and image processing 18(1), 1–17 (1982)
20. Proesmans, M., Van Gool, L.J., Oosterlinck, A.J.: One-shot active 3d shape acquisition. In: Proc. of International Conference on Pattern Recognition (ICPR). pp. 336–340 (1996)
21. Sagawa, R., Ota, Y., Yagi, Y., Furukawa, R., Asada, N., Kawasaki, H.: Dense 3d reconstruction method using a single pattern for fast moving object. In: Proc. of International Conference on Computer Vision (ICCV). pp. 1779–1786 (2009)
22. Sato, K., Inokuchi, S.: Three-dimensional surface measurement by space encoding range imaging. Journal of Robotic Systems 2, 27–39 (1985)
23. Sen, P., Chen, B., Garg, G., Marschner, S.R., Horowitz, M., Levoy, M., Lensch, H.: Dual photography. ACM Trans. on Graph. 24(3), 745–755 (2005)
24. Sensors Unlimited Inc: Swir image gallery. `http://www.sensorsinc.com/gallery/images` (2016)
25. Shirai, Y., Suwa, M.: Recognition of polyhedrons with a range finder. In: Proc. of International Joint Conference on Artificial Intelligence. pp. 80–87 (1971)
26. Srinivasan, V., Liu, H.C., Halioua, M.: Automated phase-measuring profilometry: A phase mapping approach. Applied Optics 24(2), 185–188 (1985)
27. Vuylsteke, P., Oosterlinck, A.: Range image acquisition with a single binary-encoded light pattern. IEEE Trans. Pattern Anal. Mach. Intell. 12(2), 148–164 (1990)
28. Wang, J., Gupta, M., Sankaranarayanan, A.C.: LiSens — A scalable architecture for video compressive sensing. In: IEEE International Conference on Computational Photography (ICCP) (2015)