Structured light with a million light planes per second

Dhawal Sirikonda[†], Praneeth Chakravarthula[‡], Ioannis Gkioulekas[§], Adithya K Pediredla[†]
[†]Dartmouth College, [‡]UNC Chapel Hill, [§]Carnegie Mellon University

Abstract—We introduce a structured light system that enables full-frame 3D scanning at speeds of $1000\,\mathrm{fps}$, four times faster than the previous fastest systems. Our key innovation is the use of a custom acousto-optic light scanning device capable of projecting two million light planes per second. Coupling this device with an event camera allows our system to overcome the key bottleneck preventing previous structured light systems based on event cameras from achieving higher scanning speeds—the limited rate of illumination steering. Unlike these previous systems, ours uses the event camera's full-frame bandwidth, shifting the speed bottleneck from the illumination side to the imaging side. To mitigate this new bottleneck and further increase scanning speed, we introduce adaptive scanning strategies that leverage the event camera's asynchronous operation by selectively illuminating regions of interest, thereby achieving effective scanning speeds an order of magnitude beyond the camera's theoretical limit.

Index Terms—3D scanning, structured light, acousto-optics, event cameras

◆

1 INTRODUCTION

3D scanning systems acquire the detailed geometry of real-world objects, using methods such as time of flight, structured light, or photogrammetry. Though such systems are already commonplace for static scenes, nowadays there is an increasing demand for 3D scanning systems that can operate on dynamic scenes and provide real-time 3D information for accurate analysis of moving objects. For instance, such systems can enhance navigation and safety in autonomous driving and robotics, enable interactive experiences in virtual and augmented reality, or allow for inspection of fast-moving objects in industrial manufacturing. These applications require *fast* 3D scanning that minimizes motion blur and other motion-induced artifacts.

Fast 3D scanning necessitates hardware for both high-speed illumination (e.g., using Galvo mirror systems or MEMS mirror projectors), and high-speed imaging (e.g., using single-photon avalanche diodes or event sensors). It also requires efficient algorithms to process and filter data rapidly, handling tasks such as noise reduction, data fusion, and reconstruction. In scanning systems that utilize multiple sensors, synchronization also plays a crucial role in ensuring coherent data capture and integration from different sources. Unfortunately, fast 3D scanning methods are limited to rates below $1000\,\mathrm{fps}$, primarily due to constraints imposed by the scanning speeds of existing light scanning systems: Though fast imaging sensors such as event cameras exist, light scanning systems are limited to scan rates of a few kHz, falling well short of fully exploiting the fast sensor readout.

In this work, we push the envelope of achievable 3D scanning speeds, by designing an ultra-fast structured light (SL) system that combines an acousto-optic (AO) light scanning device and an event camera. The AO device comprises a pulsed laser and an ultrasonic transducer. The transducer generates an ultrasonic wave that sculpts traveling virtual cylindrical lenses in a water medium. These lenses focus the laser light onto a line. By steering these cylindrical lenses at MHz rates, we can scan the line over a 3D scene imaged by the event camera, enabling structured light scanning. The AO device uses a design that results in an order-of-magnitude cost reduction relative to prior such devices [1].

Our AO device achieves light scanning rates that are more than three orders of magnitude higher than the frame rate of the event camera. Thus, when used for structured light, our device shifts the bottleneck for achieving higher frame rates from the speed of light scanning system to the readout bandwidth of imaging sensors. To further improve speed beyond this bottleneck, we demonstrate an adaptive scanning method inspired by Muglikar et al. [2] that illuminates only regions of interest. Doing so allows capture $10\times$ faster than the theoretical full-frame limit of the event sensor. We demonstrate these capabilities experimentally by building a benchtop prototype, and using it to scan a variety of static and dynamic scenes.[1]

2 PRIOR WORK

Light scanning. Light scanning is a core component of active imaging technologies, including lidar [3], structured light [4–7], light-transport probing [8–10], motion contrast 3D [11], light curtains [12, 13], slope-disparity gating [14–16], and non-line-of-sight imaging [17–21]. We can broadly classify the light scanning methods in these technologies into mechanical and non-mechanical. Mechanical methods require moving parts such as rotating prisms [22, 23] or mirrors [24, 25], and micro-electromechanical systems (MEMS) [26, 27]. These methods are slow due to mechanical inertia. Non-mechanical methods require no moving parts, and include acousto-optic (AO) [28, 29] and electro-optic (EO) [30] devices, liquid crystal devices (LCDs) [31, 32], and optical phased arrays (OPAs) [33, 34]. LCDs are the slowest of these methods due to long settling times. OPAs are the fastest and can scan light at even GHz rates, but have low angular resolutions, and

[1] Project Page: https://dartmouth-risc-lab.github.io/aosl/

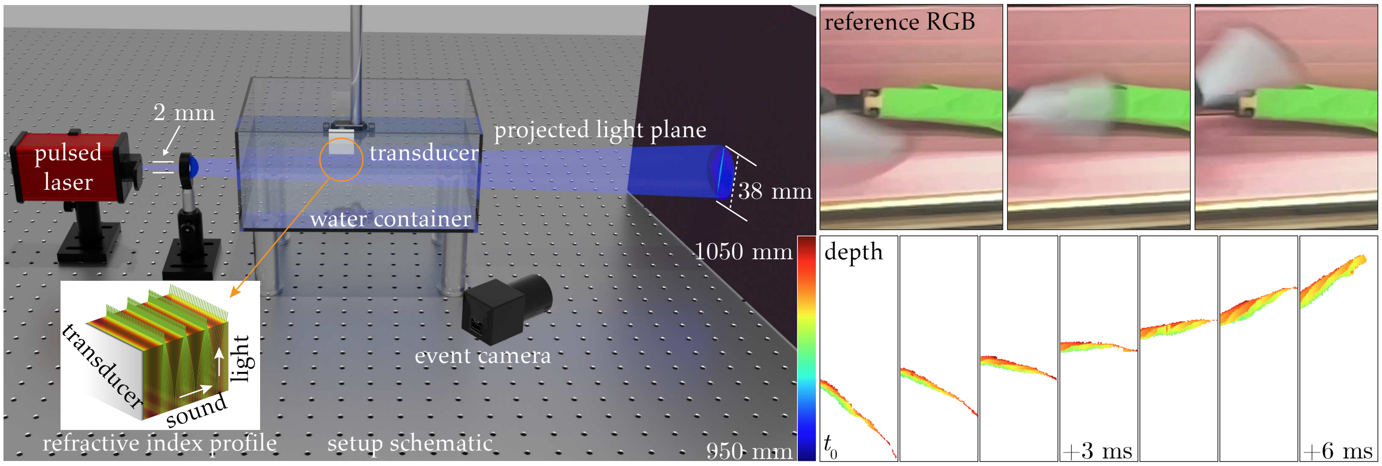


Fig. 1. We present a structured light technology that combines acousto-optic light steering with an event camera for high-speed full-frame scanning. Left: Schematic of our setup. We use an ultrasonic transducer to sculpt virtual gradient-index (GRIN) cylindrical lenses in water (the inset shows the refractive index profile focusing light). Coupling this setup with pulsed illumination, we can scan the imaged scene with a light plane at speeds three orders of magnitude faster than those of previous light scanning methods, allowing structured light operation at the camera's full-frame bandwidth. Top-right: RGB images, captured at $240\,\text{fps}$, of a scene comprising a fan that rotates at $1800\,\text{rpm}$. The images correspond to the approximate positions of the depth scans below. Bottom-right: Reconstructed depth frames of the rotating fan, scanned at $1000\,\text{fps}$.

TABLE 1. Scan speed comparison against recent works.

Work	Static	Dynamic
Muglikar et al. [35]	60 fps	60 fps
Matsuda et al. [11]	60 fps	60 fps
Dashpute et al. [36]	250 fps	30 fps
Ours	1 kfps	1 kfps

thus are unsuitable for structured light. AO and EO devices provide a good middle ground, as we discuss next.

Light scanning with AO and EO devices. Commercially available AO and EO devices include tunable filters [37], modulators [38], frequency shifters [39], and deflectors [40]. AO deflectors can serve as light scanning devices, but are limited to kHz scanning rates due to their reliance on Bragg's diffraction [41]. EO deflectors [42, 43] operate on a similar principle, and thus are similarly slow. Alternative AO light scanning devices are tunable acoustic gradient-index (TAG) lenses [44, 45] and ultrasonically-sculpted virtual optical waveguides [1, 46, 47]—our technology is an instance of the latter. TAG lenses can change the focus depth of an incident beam at kHz rates, but cannot scan it transversely.

Most closely related to our work, Pediredla et al. [1] demonstrated transverse scanning at MHz rates using ultrasonically-sculpted virtual optical waveguides, motivating our work. Compared to their scanning system, which phase-modulates the ultrasound for scanning, our system modulates the laser pulse. This adjustment allows us to use a *narrowband* amplifier, which is widely available and cost-effective (USD 250 instead of USD 23 000 for the broadband amplifier in Pediredla et al. [1]), reducing the system cost.

Lastly, Hullin et al. [48] used perturbation of surface waves to dynamically display BRDFs. However, their technique is limited to kHz range as high frequency waves experience significant damping.

Event cameras and 3D sensing. Event cameras perform sparse readouts of relative intensity changes exceeding a preset threshold, and thus can achieve much larger frame rates than conventional cameras that always perform full-frame readouts. This high speed has made event cameras increasingly popular in computer vision [49–52]. Though extensive research exists for 2D imaging applications of event cameras, such as full-frame reconstruction [53–58] and deblurring [59–64], 3D imaging applications are still nascent. Brandli et al. [65] combined an event camera with a laser line scanner for structured-light 3D scanning, leveraging the fact that swept-plane structured light requires very sparse sensor readouts—only a column per projected light plane. More recent works perform structured light by combining event cameras with laser point projectors [2, 11, 35, 66] or digital light processing projectors [67, 68]. In all these works, light scanning is slower than the frame rate of the event camera, and thus the bottleneck preventing faster 3D scanning. We remove this bottleneck by using a laser-line scanning system that combines acousto-optic lens sculpting with a pulsed laser, to achieve megahertz-rate light scanning—three orders of magnitude faster than the theoretical frame-rate limit of an event camera, thus turning the camera into the bottleneck towards faster structured light.

3 METHOD

Structured light (SL) methods triangulate depth by establishing stereo correspondences between an active light projector and a camera. In its simplest form, SL uses the projector to raster scan individual points [11, 35]. Using the epipolar geometry allows accelerating structured light while maintaining robust correspondences, by projecting a line (light plane in 3D) orthogonal to the epipolar lines. Methods using optical codes such as binary [69], Gray [70], sinusoidal [71], or XOR [4] reduce the number of projected light patterns, but are detrimental for the event camera case—approximately half the sensor pixels will be lit for each

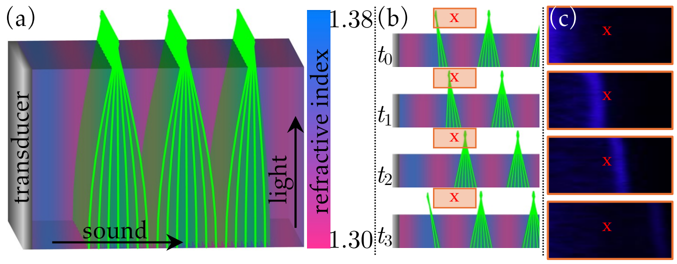


Fig. 2. (a) We show the refractive index profile created by the ultrasonic transducer and how light traveling from bottom to top focuses onto lines. (b) As the refractive index profile moves horizontally over time, the focused lines also move horizontally. (c) Captured images demonstrate the temporal movement of the focused line

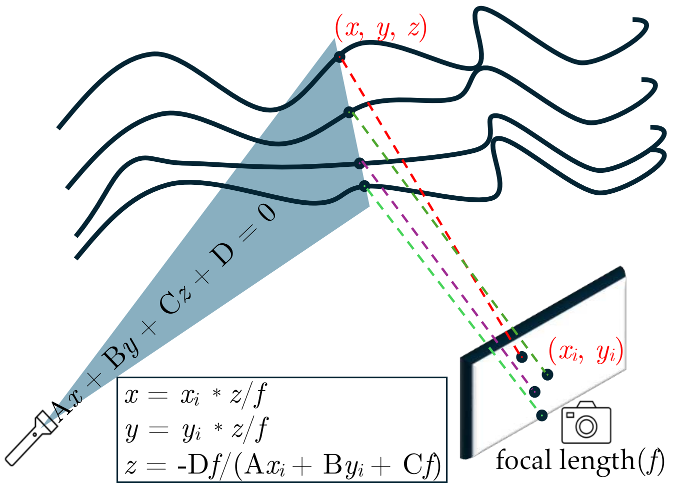


Fig. 3. Structured light scanning with swept light planes. The light scanning device sweeps a light plane across the imaged scene. The reflected light triggers events at pixels along a vertical curve on the event camera. For each such pixel, depth is triangulated by intersecting a backprojected ray with the corresponding light plane.

projected pattern, increasing the number of generated events and thus slowing the scanning rate.

Motivated by these considerations, we opt for a system that uses line projection and an event camera for structured light. We build an acousto-optic line scanning prototype that can project two million lines per second (lps), exceeding the event camera theoretical detection rate ($\approx 10^6$ lps), and being orders of magnitude faster than alternative projection systems used for structured light, including Galvo mirrors [11, 36], MEMS mirrors [67, 68], and laser projectors [66]. Below we describe the physical principles governing the operation of the line scanning device, and the algorithmic principles for programmable light scanning.

Ultrasonic sculpting of steerable lenses. The refractive index of an optical medium is a function of its density. By controlling a medium's spatial density, we can control its spatially varying refractive index and convert the medium into a gradient-index (GRIN) lens. To this end, we use a planar acoustic transducer to create a pressure wave that sculpts inside a transparent medium (water) a cylindrical GRIN lens traveling at the speed of sound in that medium. Using this virtual cylindrical lens, we can focus a collimated beam of light onto a line that also moves at the speed of sound. We can then use this light beam for structured light.

If we apply sinusoidal voltage $V(t) = V_{\text{us}} \cos\left(2\pi f_{\text{us}} t\right)$ to the planar transducer at $x = 0$, the transducer will create ultrasound inducing a traveling pressure wave:

$$P(x,y,z,t) = P_0 + P_{\text{us}} \cos\left({}^{2\pi}\!/\!_{\lambda_{\text{us}}} x - 2\pi f_{\text{us}} t\right), \tag{1}$$

where x, y, z are Cartesian coordinates, t is time, P_0 is the pressure in the medium without ultrasound, P_{us} is proportional to the input voltage amplitude V_{us}, and λ_{us} and f_{us} are the wavelength and frequency (resp.) of the ultrasound output by the transducer. The pressure wave creates a proportional change in the refractive index of the medium, resulting in a time-varying refractive index profile:

$$n(x,y,z,t) = n_0 + n_{\text{us}} \cos\left({}^{2\pi}\!/\!_{\lambda_{\text{us}}} x - 2\pi f_{\text{us}} t\right), \tag{2}$$

where n_0 is the refractive index without ultrasound, and n_{us} is proportional to the pressure amplitude P_{us}. We drop y, z as the refractive index does not change along those axes. As Pediredla et al. [1] explain, each of the the convex lobes of $n(x,t)$—corresponding to $x \in (k\lambda_{\text{us}} - {}^{\lambda}\!/\!_{2}, k\lambda_{\text{us}} + {}^{\lambda}\!/\!_{2})$, $k \in \mathbb{Z}$—acts as a cylindrical GRIN lens that travels along the x axis at the speed of ultrasound, $c_{\text{us}} = f_{\text{us}}\lambda_{\text{us}}$.

A light beam passing through the medium will focus to a line—or series of lines if the beam width is larger than the ultrasound wavelength—traveling at speed c_{us}. Assuming for simplicity that the cylindrical lenses are aberration-free, the resulting intensity at the lens focal plane is

$$I(x,t) = \sum_{k} \delta(x + k\lambda_{\text{us}} - c_{\text{us}} t), \tag{3}$$

up to a scale factor proportional to the illumination. Fig. 2 visualizes the cylindrical lenses, focusing behavior of light rays, and their temporal dynamics.

Programmable control of light planes. The traveling cylindrical lens enables scanning a light plane, but does not provide a mechanism for controlling scanning speed. Consequently, the scanning rate of the light plane is faster than the capture rate of the event camera. We overcome this problem by using a laser that we pulse at a controllable frequency. If this frequency is the same as that of the ultrasound transducer, and assuming laser pulsation starts at $t = 0$, the light intensity at the focal plane will become:

$$I(x,t) = \sum_{k,l} \delta(x + k\lambda_{\text{us}} - l\,\underbrace{c_{\text{us}} T_{\text{us}}}_{\lambda_{\text{us}}})\,\delta(t \bmod T_{\text{us}}), \tag{4}$$

where l indexes laser pulses, and the inter-pulse time $T_{\text{us}} = f_{\text{us}}^{-1}$ equals the ultrasound period. Eq. (4) describes a pulse train that does not translate in space but flickers in time. Every l^{th} laser pulse illuminating the l^{th} period of the ultrasonic wave will generate a line at the origin ($x = 0$).

To move the lines spatially, we instead pulse the laser at a frequency slightly offset from the ultrasonic one. If α is the ratio of ultrasonic to laser frequency, Eq. (4) becomes:

$$I(x,t) = \sum_{k,l} \delta(x + k\lambda_{\text{us}} - l\alpha\lambda_{\text{us}})\,\delta(t \bmod \alpha T_{\text{us}}). \tag{5}$$

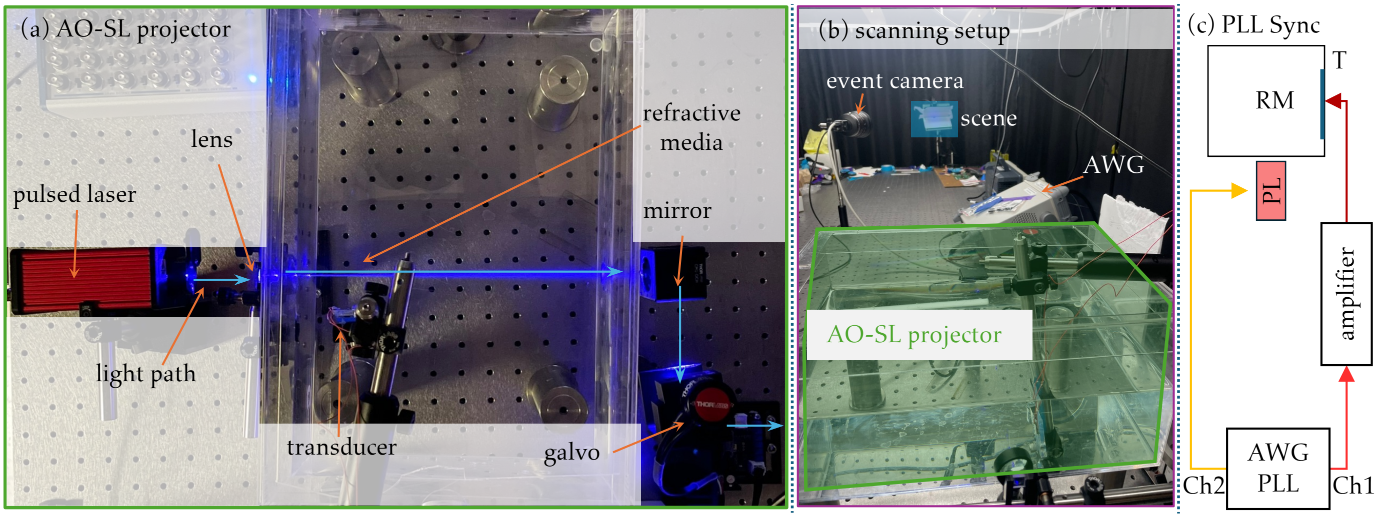


Fig. 4. Hardware implementation of the setup in Fig. 1. (a) AO light scanning device. A pulsed laser (PL) emits a beam that is expanded by a convex lens and directed into a refractive medium (RM). An ultrasonic transducer (T), perpendicular to the beam, generates acoustic waves that sculpt cylindrical lenses in the medium. These lenses focus the light into sweeping light planes, which are then redirected by a fixed mirror onto a Galvo mirror. During acousto-optic scanning, the Galvo mirror remains stationary. Conversely, when using the Galvo mirror for scanning, the transducer and pulse frequency are constant. (b) Scanning setup. The AO device from (a) is shown in green. The scene (in cyan) contains orthogonal planes for calibration. The laser and ultrasound are driven by an arbitrary waveform generator (AWG). (c) Phase-lock-loop (PLL) synchronization of the transducer with the pulsed laser, driven via Channels 1 and 2 (Ch1, Ch2) of the AWG. Channel 1's output is amplified using an amplifier (Amp). This PLL synchronization ensures accurate pulse placement with respect to the ultrasound.

Therefore, when the l^{th} laser pulse illuminates the l^{th} period of the ultrasonic wave, the focused line will appear at $x = (1-\alpha)l\lambda_{\text{us}}$ making the line move spatially from pulse to pulse. The line frequency is equal to the beat frequency $(1-\alpha)f_{\text{us}}$ between ultrasound and laser.

Besides serial line scanning, our laser pulsing design allows us to place lines at a non-serial set of locations x_l; $l \in \{1, 2, \ldots, L\}$, by emitting laser pulses at times

$$T_l = \frac{x_l + l\lambda_{\text{us}}}{lc_{\text{us}}}. \qquad (6)$$

This capability facilitates adaptive scanning of regions of interest, and in turn yet faster structured light rates (Sec. 5).

The above analysis assumed a collimated beam. If we use a diverging beam, the line scanning behavior remains the same, but the distance between the lines will increase. In our prototype, we use a diverging beam to increase the field of view of the structured light system (Fig. 1).

Comparison to Pediredla et al. [1]. Our acousto-optic laser line scanning system is inspired by that introduced by Pediredla et al., with two important differences: 1) Pediredla et al. focus on point scanning for lidar applications, whereas we focus on line scanning for structured light. Therefore, their system uses two linear transducers, whereas ours using just one, effectively replacing spherical GRIN waveguides with cylindrical ones. 2) Pediredla et al. control the position of the focused spot by modulating the phase of the ultrasound, whereas we control the position of the focused line by pulsing the laser. The approach of Pediredla et al. creates two challenges: First, it requires using an expensive broadband RF amplifier (ENI-300L, USD 23 000), as the ultrasonic wave is no longer a monotonic sinusoid. Second, it requires that the transducers operate at non-resonant frequencies, where they consume more power. By contrast, our approach allows using a much less expensive narrowband amplifier (USD 250) and operating at lower power.

Structured light. For structured light, we use our AO light scanning device together with an event camera to form a stereo pair. We orient the planar transducer of our device such that the projected light planes are as orthogonal to the corresponding epipolar lines on the event camera as possible. By synchronizing the AO device and event camera, we can perform structured light scanning using the classical swept-plane procedure [72, 73], which we visualize in Fig. 3: As the AO device scans a light plane, light reflected off the scanned object forms a vertical curve on the image plane that sweeps horizontally across the field of view, triggering sparse events (ideally one event per sensor row for each scan position, assuming perfect optics). From each such event, we can reconstruct a depth value for the corresponding pixel via triangulation, by backprojecting a ray and intersecting it with the light plane that triggered the event.

4 HARDWARE PROTOTYPE

In this section, we detail hardware components and the system implementation. Fig. 1 shows a schematic of the optical setup, and Fig. 4 shows our hardware prototype.

Light scanning device. To implement the acousto-optic light scanner, we use an ultrasonic transducer of size $12.5\,\text{mm} \times 12.5\,\text{mm}$ tuned to operate at a $2\,\text{MHz}$ frequency. We drive the transducer using a Siglent SDG 1032X arbitrary-waveform generator (AWG). We amplify the AWG output using a narrowband RF amplifier (rated at $50\,\text{W}$ and operating frequency $1\,\text{MHz}$ to $3\,\text{MHz}$) before applying it to the transducer. For illumination, we use a Thorlabs NPL45B

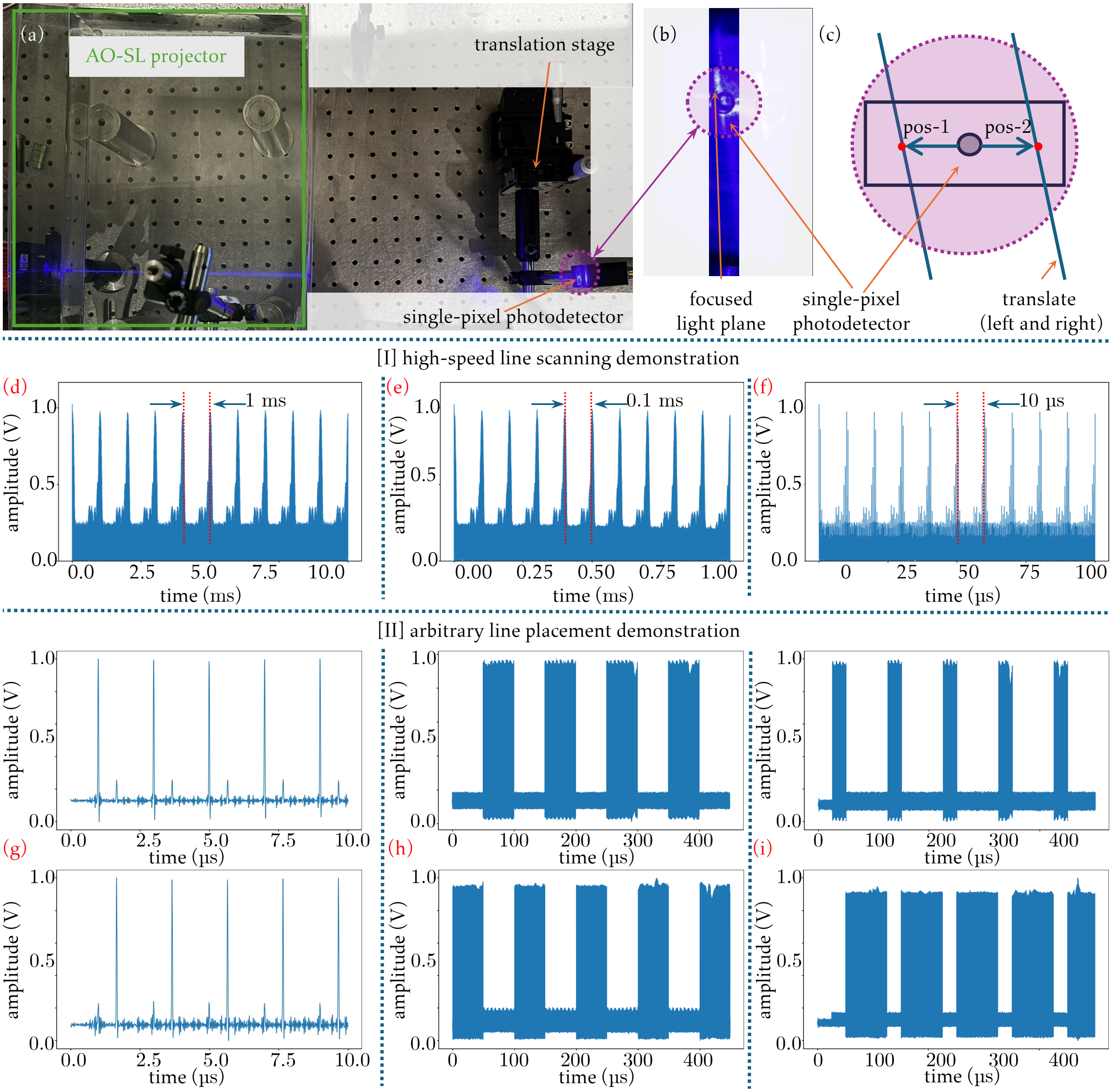


Fig. 5. Experimental validation of [I] high-speed line scanning and [II] arbitrary line placement. Our AO device is capable of translating the projected line at high speeds and positioning the lines at arbitrary spatial locations. (a) To validate these capabilities, we place a fast single-pixel photodetector (PD) mounted on a translation stage in the scene and monitor its output using an oscilloscope. (b) We show the line focused on the single-pixel photodetector surface. (c) We place the PD at two arbitrary positions (pos-1 and pos-2) where we intend to place arbitrary lines. This is equivalent to placing two PDs and monitoring their outputs. **[I]. High-speed scanning:** To demonstrate temporal scanning capability, the photodetector is fixed at a static location while the AO device scans the line periodically across it. (d)—(f) show oscilloscope outputs at scanning speeds of $1\,\mathrm{kfps}$, $10\,\mathrm{kfps}$, and $100\,\mathrm{kfps}$. The periodic peaks and their consistent temporal spacing confirm successful scanning at the target frequencies. **[II]. Arbitrary line placement:** To demonstrate arbitrary line placement, we place the PD at pos-1 or pos-2 and vary the line placement strategy. In (g)—(i), the top and bottom row show the oscilloscope outputs for pos-1 and pos-2, for various line placements: In (g), alternate lines are directed to pos-1 and pos-2, resulting in alternating signal peaks. In (h), 100 lines are directed to pos-1 followed by 100 lines to pos-2 producing a square waveform at $10\,\mathrm{kHz}$. In (i), a sequence of 50 lines at pos-1 and 150 lines at pos-2, results in a waveform with a duty cycle of $25\,\%$ at pos-1 and $75\,\%$ at pos-2.

pulsed laser that we synchronize and drive using the AWG to achieve line scanning. In the supplementary document, we provide detailed descriptions of the circuitry and the associated waveforms used for static and controlled steering of light planes, along with visualizations of their effects on illumination within the scanning system. We use a Thorlabs LA1274 $40\,\mathrm{mm}$ convex lens to diverge the beam enough to cover the scanned scene. For comparison with previous steering techniques, we include in our system a Thorlabs GVS211 single-axis Galvo mirror, which we keep stationary when we use our acousto-optic light scanning.

Event camera. We use the Prophesee EVK-4 event camera, equipped with Sony's IMX636 sensor. This sensor can process up to $1\,\mathrm{GEvents/s}$, with events recorded at a temporal resolution of $1\,\mu\mathrm{s}$ and spatial resolution of 1280×720 pixels. As the horizontal axis has the fastest readout and provides timestamping, we orient this axis parallel to the light plane. This alignment allows us to utilize the event camera's full bandwidth [35], corresponding to a maximum achievable frame rate of approximately $^{10^6}/_{720} = 1388\,\mathrm{Hz} \approx 1\,\mathrm{kHz}$. We replace the default lens of the EVK-4 camera with a Rokinon DS50M-C full-frame lens of focal length $50\,\mathrm{mm}$ and speed $f/1.5$ to improve light efficiency during scanning.

We observed empirically that, with default settings, the event camera fails to capture events when sweeping light planes at frequencies higher than $100\,\mathrm{Hz}$. To mitigate this problem, we follow Gallego et al. [74] and set the camera's parameters `bias-hpf` and `bias-refractory` to their maximum values and remove the `bias-off` parameter, allowing a single polarity to utilize the complete bandwidth of the camera.

Synchronization. To synchronize events with the light planes that trigger them, one option is to use the event camera's hardware trigger mechanism and synchronize with the planes' timestamps. However, at high-speed scanning when the event camera generates several events, we observed that this mechanism fails. We thus develop a software-based synchronization system: We project light planes continuously and record them with the event camera. We then use the event camera's timestamp for the left-most line and the corresponding AO light plane's timestamp for synchronization.

Geometric calibration. As our light scanning device is highly repeatable, we precalibrate scanned light planes through their linear intersections with two orthogonal reference planes, which we create using two LCD screens: We display blinking checkered patterns on the screens to calibrate the reference planes with respect to the event camera, then keep them turned off during structured light. We visualize this arrangement in the supplement. The procedure uses standard geometric calibration algorithms [72, 73].

Validation of high-speed line scanning. To validate that our AO system can achieve line scanning speeds exceeding $1\,\mathrm{kHz}$, we use a fast single-pixel photodetector (Thorlabs DET25A, bandwidth $2\,\mathrm{GHz}$), as in Fig. 5: We position the photodetector directly in front of the AO system, so that the detector will produce a sharp temporal response whenever the line sweeps the detector's active area. We use a high-speed oscilloscope to read out the detector response. Each instance of the line intersecting the detector results in an intensity peak in the detector response, and the temporal separation between these peaks equals the line scanning period. We show oscilloscope traces for three target scanning speeds, $1\,\mathrm{kfps}$, $10\,\mathrm{kfps}$, and $100\,\mathrm{kfps}$; the traces confirm that our AO system achieves these speeds(Fig. 5(d—f)).

TABLE 2. Impact of scan rate. (CD: chamfer distance in mm; F1: F1-score; Pr: precision; Re: recall.)

	fps	CD (↓)	Pr (↑)	Re (↑)	F1 (↑)
cat	10	1.48	0.762	0.910	0.830
	100	1.25	0.812	0.953	0.877
	200	1.67	0.738	0.956	0.833
	1000	4.56	0.146	0.981	0.253
	1000 (acc.)	4.30	0.524	0.827	0.641
dolphin	10	0.750	0.958	0.939	0.949
	100	1.273	0.961	0.817	0.883
	200	1.583	0.915	0.803	0.855
	1000	2.342	0.844	0.687	0.758
	1000 (acc.)	1.338	0.986	0.926	0.955

Validation of arbitrary line scanning. To validate that our AO system can place all lines at target locations, we placed the fast single-pixel photodetector on a translation stage, as in Fig. 5. We illuminated two lines in three different sequences: 1) alternately illuminate each line; 2) illuminate each line 100 times before alternating; 3) illuminate first line 50 and second line 150 times. The photodetector output confirms that our system achieves the target scanning sequences((Fig. 5(g—i)).

5 EXPERIMENTS AND ANALYSIS

In this section, we systematically evaluate the performance of our proposed structured light system on static and dynamic scenes. We show that our system can reconstruct 3D scenes at the maximum frame rate of the event camera at megapixel resolution. We also show an adaptive 3D scanning method that allows to selectively scan parts of the scene, thereby achieving a $10\times$ higher frame rate. We empirically show that structured light systems using Galvo mirrors cannot effectively perform adaptive 3D scanning.

Supplement. The supplement shows additional experimental and other results, including: 1) A video explaining our method. 2) Videos showing temporal sequences of depth scans at various scan rates. 3) A PDF document and short videos detailing setup implementation and calibration.

5.1 Static scene results

We first scan static scenes, to quantitatively and qualitatively evaluate the effect of scan rate on reconstruction performance. In Fig. 6, we use two figurines as test objects for scanning. We first used Galvo-based SL at a slow speed ($10\,\mathrm{fps}$) to compute ground truth for comparison. We then used the AO device to scan the object at $10\,\mathrm{fps}$, $100\,\mathrm{fps}$, $200\,\mathrm{fps}$, and $1\,\mathrm{kfps}$ scan rates. The figure shows the captured depth maps, and Tab. 2 provides quantitative metrics for reconstruction quality. In the supplement, we provide 3D visualizations of the reconstructed point clouds.

From Fig. 6, we observe that our system produces depth results nearly identical to those from the Galvo-based system at low scanning rates. As scanning rate increases, reconstruction quality progressively deteriorates due to two reasons: 1) Higher scan rates result in decreased exposure

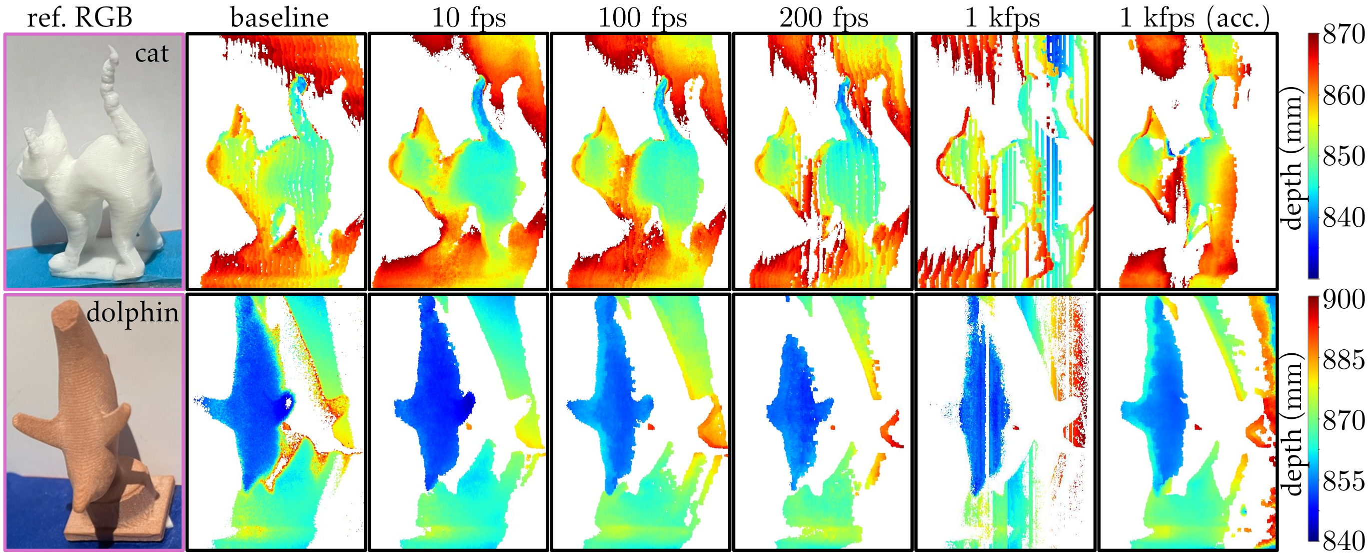


Fig. 6. Depth scanning of static scenes at different scan rates. We scan a cat (top) and a dolphin (bottom) figurine, using the proposed AO SL system, at scan rates of $10\,\mathrm{fps}$, $100\,\mathrm{fps}$, $200\,\mathrm{fps}$, and $1\,\mathrm{kfps}$. We use Galvo-based SL at $10\,\mathrm{fps}$ as the baseline for comparison. In Tab. 2, we quantify reconstruction quality. Our proposed AO SL system reconstructs depth with high fidelity, but quality worsens as frame rate increases. At high frame rates, the event camera drops events randomly as the number of events exceeds its readout bandwidth. Future event cameras with higher bandwidth can mitigate this problem. We emulate them by scanning the static scene several times and accumulating the events to account for dropped events.

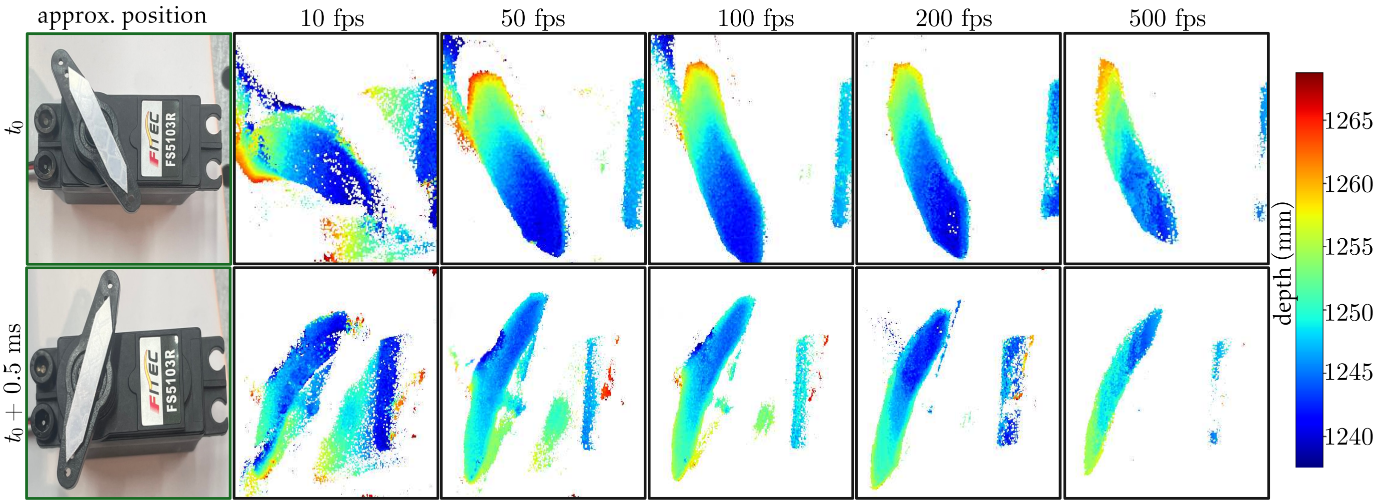


Fig. 7. Depth scanning of a servo motor rotating at $70\,\mathrm{rpm}$. Each row displays two frames (corresponding to times t_0 and $t_0 + 0.5\,\mathrm{s}$) from scans obtained at scan rates of $10\,\mathrm{fps}$, $50\,\mathrm{fps}$, $100\,\mathrm{fps}$, $200\,\mathrm{fps}$, and $500\,\mathrm{fps}$. At $10\,\mathrm{fps}$, the scan does not produce a correct depth map, as the rapid motion in the scene exceeds the scan rate. At $50\,\mathrm{fps}$, the scan provides a depth map that is more accurate but still exhibits motion blur (visible more clearly in the supplementary video). As scan rates increase to $500\,\mathrm{fps}$, the depth maps become progressively more accurate.

time and signal-to-noise ratio, in turn causing a lot more noise events to be triggered. 2) The linewidth from the AO device is around 10 pixels, thus each line triggers events at multiple sensor columns. At high scan rates, this increased number of events saturates the camera, making it randomly lose some events (columns) and causing the missing stripes. At $1\,\mathrm{kfps}$, the number of missing columns due to the second issue becomes significant. To mitigate this issue, we experimented with accumulating multiple frames ("acc." in Fig. 6 and Tab. 2), effectively simulating a higher-bandwidth camera. We observe that doing so produces complete depth scans. This effect is scene-dependent, and sparse scenes as in Fig. 1 do not suffer from reduced reconstruction quality due to limited readout bandwidth.

In Tab. 2, we quantify the 3D reconstruction fidelity using chamfer distance, precision, recall, and F1 score. As the frame rate increases, all metrics deteriorate as expected. With accumulation, the metrics improve significantly.

5.2 Dynamic scene results

To showcase the speed and characterize the performance of our structured light system, we use two types of dynamic

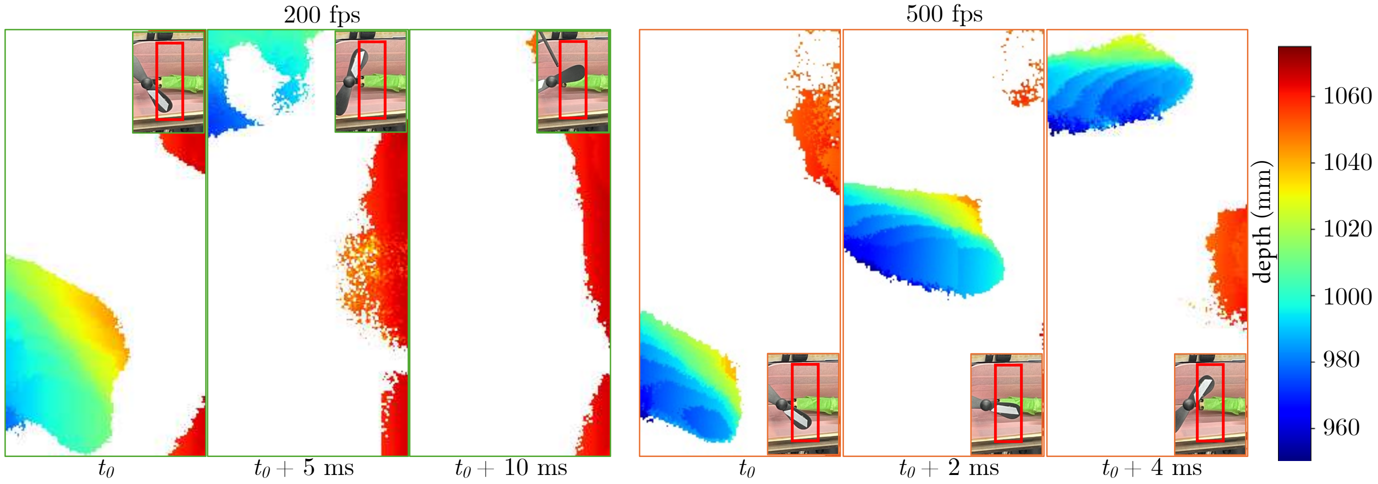


Fig. 8. Depth scanning of a fan rotating at 1800 rpm. The left column displays three depth frames captured at a scan rate of 200 fps. The right column shows the same for a higher scan rate of 500 fps. The depth maps show significantly more motion blur on the blade area at 200 fps than at 500 fps. The red regions are from the background calibration plane. The insets show the approximate orientation of the fan blade, and the area inside the red box is the scanned region.

scenes: 1) Periodic dynamic scenes, including a servo motor rotating at 70 rpm (Fig. 7) and a high-speed fan rotating at 1800 rpm (Fig. 8). 2) Non-periodic dynamic scenes, including zig-zag motion and left-right oscillation (Fig. 9).

Periodic dynamic scenes. We use these scenes to perform experiments with controllable motion and speeds, to assess motion blur and determine optimal scan speeds.

For the 70 rpm servo (blade diameter 35 mm), Fig. 7 shows results for five scan rates: 10 fps, 50 fps, 100 fps, 200 fps, and 500 fps. The servo is rotating clockwise, and we show depth scans of two frames separated by 0.5 s. As the fan blades are dark and we are scanning at very low exposures, we used retroreflective tape to increase reflectivity.

The RGB images show the orientation of the blade and the region where the reflectors are located. At 10 fps, the depth maps suffer from significant blur. At 50 fps, the depth maps improve significantly but still have motion blur (seen as increased blade thickness). The results become increasingly sharp as the scan rate increases. The video results in the supplement show continuous depth scans of the dynamic scene at all scan rates.

For the more challenging fan rotating at 1800 rpm (blade diameter 72 mm), Fig. 8 shows results at scan rates of 200 fps and 500 fps. The fan blade has considerably more motion blur at 200 fps than at 500 fps. The former is representative of the expected performance of the prior state of the art (Dashpute et al. [36], who reported scan rates of 250 fps).

Non-periodic dynamic scenes. In Fig. 9, we show results from experiments using hand-held objects undergoing various types of motion. Our system successfully acquires full-frame depth despite the fast motion. The supplement shows the full depth videos we captured with our system.

5.3 Adaptive depth scanning

As we mentioned in Sec. 4, the event sensor can capture only up to 1 k frames per second at megapixel resolution. In a dynamic scene, typically depth changes only at few pixels corresponding to the moving objects. Therefore, one way to increase scan rate beyond the sensor's theoretical full-frame rate is to perform adaptive scanning, where we only scan the regions where the depth has changed. Muglikar et al. [2] have previously demonstrated that adaptive scanning of only regions of interest improves spatial and temporal resolution. As in their work, we can illuminate only the important regions of the scene to reduce the event camera bandwidth and increase overall scan rate.

Our AO light scanning device can illuminate any arbitrary line at the same high rate (ultrasound frequency) as serial scanning, making it well-suited for adaptive scanning. Importantly, using Eq. (6), our AO device can perfectly distribute the full power of the light source at *only* the lines of interest (i.e., it is a *redistributive* line projector [9]). By contrast, using a Galvo mirror to perform adaptive scanning is slow, as the mirror is no longer operating at resonant mode. Additionally, due to the need for continuous mechanical rotation from one target line to another, a Galvo mirror ends up additionally illuminating the region between target lines (i.e., it is *not* a redistributive projector). This deficiency not only wastes light power, but also generates extraneous events, making adaptive scanning overall less effective.

To demonstrate adaptive scanning, we show in Fig. 10 a scene comprising two orange knobs at two different depths. We scanned the positions of only these knobs with the Galvo mirror device and the AO device at various scan rates. From the plot in Fig. 10, we observe that the Galvo mirror device produces accurate depth at low frame rates. However, as the frame rate increases, the depth error increases significantly. By contrast, the accuracy of the AO device remains fairly constant, even at a very high scan rate.

To better understand this performance difference, we show captured event data in Fig. 10(b), with color representing the time stamp of the events. In the supplement, we provide the event stream videos to show the behavior more clearly. We also detail the light pulse timing control used to achieve controlled oscillation between the desired light planes at specified fps in the supplementary material. From

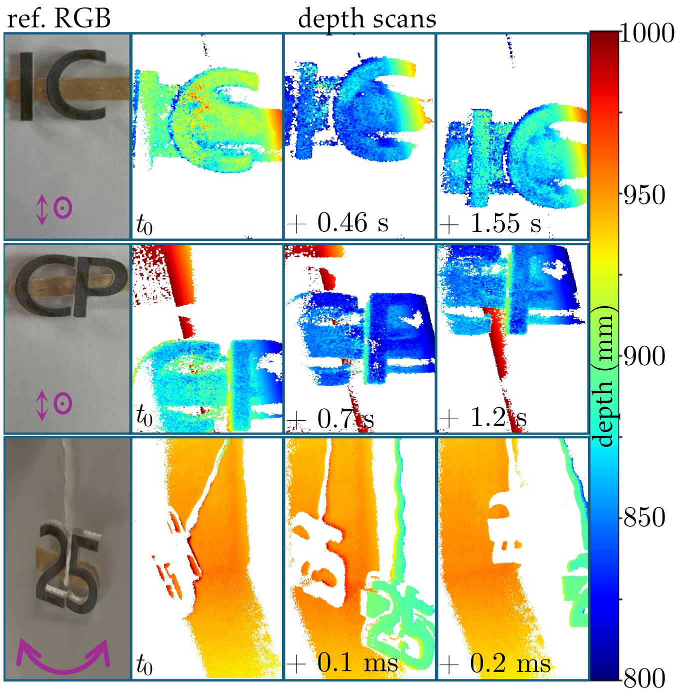


Fig. 9. Non-periodic dynamic scenes: Depth scans of characters "IC", "CP", and "25", covered with retroreflecting tape. In the top and middle rows, the characters are mounted on a stick and moved rapidly in a top-to-bottom (↕) sweeping motion, with some additional forward-backward translation (⊙). In the bottom row, the characters are swung by hand using a string. For each row, the RGB image to the left shows the scene and visualizes the approximation motion trajectory. The images to the right show the depth frames at different points of the trajectory, captured at 180 fps (top and middle row) or 200 fps (bottom row). The supplement includes a video of continuous depth scanning as the targets move.

this data, we observe that at low scan rates, the Galvo mirror device projects most of the laser's light output to the two lines of interest, but as the scan rate increases, it projects most of the light output incorrectly to the region between the two lines. By contrast, the AO device always projects all light at only the two lines of interest.

We note that we could not scan the Galvo mirror device beyond a 1 kHz rate, as that is the physical limit for these devices. For our AO device, we could theoretically go up to 1 MHz to scan two lines (given that the transducer operates at 2 MHz). However, as the frame rate increases, the amount of light decreases, and we noticed empirically that at 20 kHz the amount of light reaching the event sensor is below its detection threshold. Therefore, though theoretically we can scan up to 1 MHz adaptively, practically, we are limited by our laser power to a 10 kHz scan rate, which is still 10× higher than the event camera's full-frame frame rate.

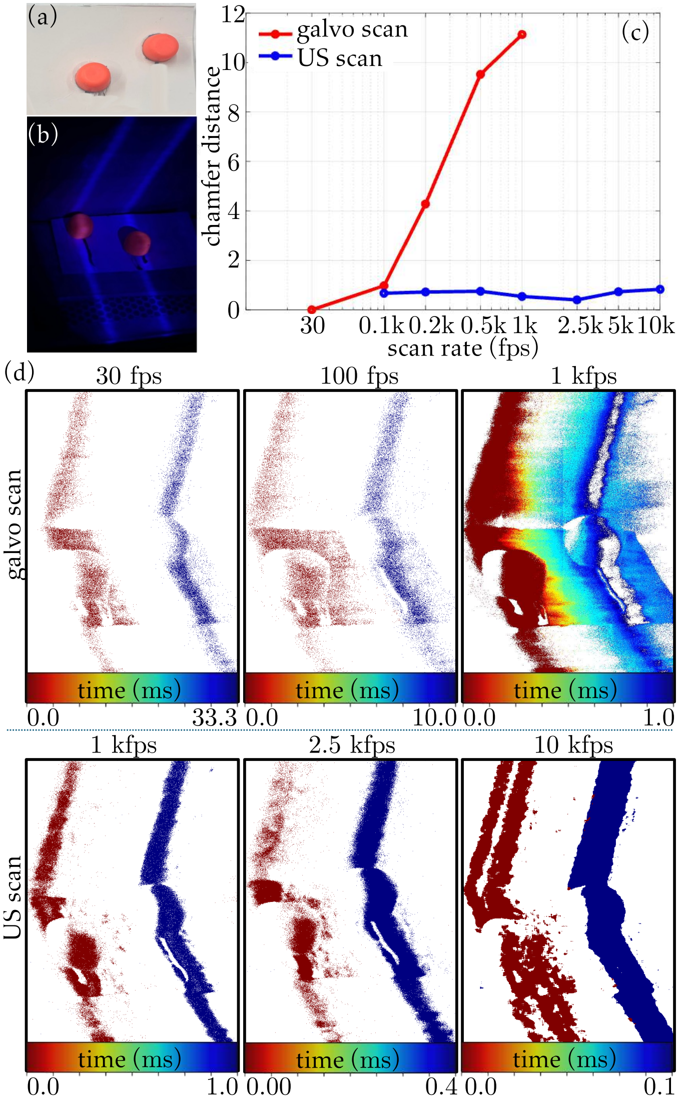


Fig. 10. Adaptive scanning of two regions of interest (orange knobs) in a scene shown in (a). We aim to illuminate and 3D scan only the knobs (b), using both the Galvo mirror and our AO device. In (c), we show the depth error measured for these knobs when we scan with the two devices. At low scan rates, the Galvo mirror is effective at adaptively scanning only the knobs, but its performance deteriorates as the scan rate increases: Due to the need to mechanically rotate the beam from one location to another, the Galvo mirror must also illuminate the region between the knobs. At high scanning speeds, the fraction of total scanning time spent illuminating this region increases, decreasing performance. By contrast, our AO device can illuminate only the knobs, resulting in effective adaptive scanning even at 10 kfps.

6 DISCUSSION AND CONCLUSION

We designed an acousto-optic device to scan light planes at unprecedented rates of up to 2 MHz, and combined it with an event camera to perform swept-plane structured light. This combination allowed us to achieve full-frame 3D scanning at 1000 fps, 3–4× faster than the state of the art [36]. However, our method has a few limitations, some of which can be addressed by modifications to the prototype, whereas others point towards future research directions.

On the first front, similar to Pediredla et al. [1], our prototype has a small aperture that, due to diffraction, results in large linewidths and thus reduced spatial and depth resolution. Using a transmission medium where the speed of sound is faster than in water would increase the ultrasound wavelength and hence the aperture size, addressing this issue. Additionally, our prototype uses a laser with low emission power ($\approx 1\,\mathrm{mW}$), necessitating light accumulation over multiple scans at higher frame rates, as well as the use of a large aperture lens that reduces depth of field. Using a better laser would mitigate both issues. Lastly, our current prototype has a limited field of view of $38\,\mathrm{mm}$, constraining the size of objects that can be scanned. This issue can be mitigated using better optics (for example, an F-theta lens), at the cost of increased linewidth. Lastly, our prototype works by modulating a large water volume,

In terms of more fundamental limitations, our prototype requires modulating a large volume of an optically transparent medium (e.g., water), resulting in a large form factor that is impractical for miniaturization. Enabling miniaturization requires research on realizing the light-focusing behavior of our prototype using standard acousto-optic modulators. Additionally, the limited bandwidth of the event camera is the bottleneck that prevents faster 3D scanning rates, creating a gap of three orders of magnitude with the light scanning rate. Closing this gap requires research into sensors that not only have drastically larger readout bandwidths, but also can operate at drastically shorter exposure times. Single-photon avalanche diodes are a promising option to this end, as they are suitable for ultra-wideband operation [75] and structured light under minuscule exposures [5]. Additionally, our system creates opportunities for combinations with other imaging modalities that use additional dimensions of light (e.g., polarization, wavelength) to bring robustness against effects such as specularity, glare, and scattering that are challenging for structured light systems—including ours. Lastly, incorporating learning-based depth priors and reconstruction methods (e.g., diffusion, neural rendering) into our pipeline could further enhance reconstruction fidelity and robustness, especially under noisy or sparse event conditions.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under awards 2047341, 2107454, 2326904, and 2403122, as well as a Sloan Research Fellowship for Ioannis Gkioulekas. We thank Aniket Dashpute (Rice University) and Manasi Muglikar (University of Zurich) for their valuable insights on event camera parameters. We also thank Ziyuan (Quinton) Qu and Sarah K. Friday (Dartmouth College), for their assistance in setting up the experimental scenes.

REFERENCES

[1] A. Pediredla, S. G. Narasimhan, M. Chamanzar, and I. Gkioulekas, "Megahertz Light Steering Without Moving Parts," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 1–12.

[2] M. Muglikar, D. P. Moeys, and D. Scaramuzza, "Event guided depth sensing," in *2021 International Conference on 3D Vision (3DV)*. IEEE, 2021, pp. 385–393.

[3] D. Yang, Y. Liu, Q. Chen, M. Chen, S. Zhan, N.-k. Cheung, H.-Y. Chan, Z. Wang, and W. J. Li, "Development of the high angular resolution 360° lidar based on scanning mems mirror," *Scientific Reports*, vol. 13, no. 1, p. 1540, 2023.

[4] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "Structured light 3d scanning in the presence of global illumination," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2011, pp. 713–720.

[5] V. Sundar, S. Ma, A. C. Sankaranarayanan, and M. Gupta, "Single-photon structured light," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 17 865–17 875.

[6] K. Ichimaru, D. Thomas, T. Iwaguchi, and H. Kawasaki, "Neural sdf for shadow-aware unsupervised structured light," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. IEEE, 2025, pp. 287–296.

[7] R. Qiao, H. Kawasaki, and H. Zha, "Depth reconstruction with neural signed distance fields in structured light systems," in *2024 International Conference on 3D Vision (3DV)*. IEEE, 2024, pp. 770–779.

[8] M. O'Toole, R. Raskar, and K. N. Kutulakos, "Primal-dual coding to probe light transport." *ACM Trans. Graph.*, vol. 31, no. 4, pp. 39–1, 2012.

[9] M. O'Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos, "Homogeneous codes for energy-efficient illumination and imaging," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 1–13, 2015.

[10] S. Achar, J R Bartels, W. L Whittaker, K. N Kutulakos, and S. G Narasimhan, "Epipolar time-of-flight imaging," *ACM Trans. Graph.*, vol. 36, no. 4, p. 37, 2017.

[11] N. Matsuda, O. Cossairt, and M. Gupta, "Mc3d: Motion contrast 3d scanning," in *2015 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2015, pp. 1–10.

[12] D. Chan, S. G. Narasimhan, and M. O'Toole, "Holocurtains: Programming light curtains via binary holography," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 17 886–17 895.

[13] J. Wang, J. Bartels, W. Whittaker, A. C. Sankaranarayanan, and S. G. Narasimhan, "Programmable triangulation light curtains," in *Proceedings of the European Conference on Computer Vision*, 2018, pp. 19–34.

[14] T. Ueda, H. Kubo, S. Jayasuriya, T. Funatomi, and Y. Mukaigawa, "Slope disparity gating using a synchronized projector-camera system," in *2019 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2019, pp. 1–9.

[15] H. Kubo, S. Jayasuriya, T. Iwaguchi, T. Funatomi, Y. Mukaigawa, and S. G. Narasimhan, "Programmable non-epipolar indirect light transport: Capture and analysis," *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 4, pp. 2421–2436, 2019.

[16] S. Chandran, H. Kubo, T. Ueda, T. Funatomi, Y. Mukaigawa, and S. Jayasuriya, "Slope disparity gating: System and applications," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 317–332, 2022.

[17] D. B. Lindell, G. Wetzstein, and M. O'Toole, "Wave-based non-line-of-sight imaging using fast fk migration," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 1–13, 2019.

[18] X. Liu, I. Guillén, M. La Manna, J. H. Nam, S. A. Reza, T. Huu Le, A. Jarabo, D. Gutierrez, and A. Velten, "Non-line-of-sight imaging using phasor-field virtual wave

optics," *Nature*, vol. 572, no. 7771, pp. 620–623, 2019.

[19] M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," *Nature*, vol. 555, no. 7696, pp. 338–341, 2018.

[20] A. Pediredla, A. Dave, and A. Veeraraghavan, "Snlos: Non-line-of-sight scanning through temporal focusing," in *2019 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2019, pp. 1–13.

[21] S. Xin, S. Nousias, K. N. Kutulakos, A. C. Sankaranarayanan, S. G. Narasimhan, and I. Gkioulekas, "A theory of fermat paths for non-line-of-sight shape reconstruction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 6800–6809.

[22] M. Abdo, V. Badilita, and J. Korvink, "Spatial scanning hyperspectral imaging combining a rotating slit with a dove prism," *Optics Express*, vol. 27, no. 15, pp. 20 290–20 304, 2019.

[23] Z. Wang, J. Cao, Q. Hao, F. Zhang, Y. Cheng, and X. Kong, "Super-resolution imaging and field of view extension using a single camera with risley prisms," *Review of Scientific Instruments*, vol. 90, no. 3, p. 033701, 2019.

[24] B.-S. Kim, S. Gibson, and T.-C. Tsao, "Adaptive control of a tilt mirror for laser beam steering," in *Proceedings of the 2004 American Control Conference*, vol. 4. IEEE, 2004, pp. 3417–3421.

[25] A. McCarthy, N. J. Krichel, N. R. Gemmell, X. Ren, M. G. Tanner, S. N. Dorenbos, V. Zwiller, R. H. Hadfield, and G. S. Buller, "Kilometer-range, high resolution depth imaging via 1560 nm wavelength single-photon detection," *Optics express*, vol. 21, no. 7, pp. 8904–8915, 2013.

[26] J. Rodriguez, B. Smith, B. Hellman, A. Gin, A. Espinoza, and Y. Takashima, "Multi-beam and single-chip lidar with discrete beam steering by digital micromirror device," in *Physics and Simulation of Optoelectronic Devices XXVI*, vol. 10526. SPIE, 2018, pp. 89–94.

[27] B. Smith, B. Hellman, A. Gin, A. Espinoza, and Y. Takashima, "Single chip lidar with discrete beam steering by digital micromirror device," *Optics Express*, vol. 25, no. 13, pp. 14 732–14 745, 2017.

[28] V. Gavryusev, G. Sancataldo, P. Ricci, A. Montalbano, C. Fornetto, L. Turrini, A. Laurino, L. Pesce, G. de Vito, N. Tiso *et al.*, "Dual-beam confocal light-sheet microscopy via flexible acousto-optic deflector," *Journal of biomedical optics*, vol. 24, no. 10, p. 106504, 2019.

[29] J. Xu and J. Tang, "Tunable prism based on piezoelectric metamaterial for acoustic beam steering," *Applied Physics Letters*, vol. 110, no. 18, p. 181902, 2017.

[30] L. Sun, J.-h. Kim, C.-h. Jang, D. An, X. Lu, Q. Zhou, J. M. Taboada, R. T. Chen, J. J. Maki, S. Tang *et al.*, "Polymeric waveguide prism-based electro-optic beam deflector," *Optical Engineering*, vol. 40, no. 7, pp. 1217–1222, 2001.

[31] X. Shang, J.-Y. Tan, O. Willekens, J. De Smet, P. Joshi, D. Cuypers, E. Islamaj, J. Beeckman, K. Neyts, M. Vervaeke *et al.*, "Electrically controllable liquid crystal component for efficient light steering," *IEEE Photonics Journal*, vol. 7, no. 2, pp. 1–13, 2015.

[32] X. Wang, L. Wu, C. Xiong, M. Li, Q. Tan, J. Shang, S. Wu, and Q. Qiu, "Agile laser beam deflection with high steering precision and angular resolution using liquid crystal optical phased array," *IEEE Transactions on Nanotechnology*, vol. 17, no. 1, pp. 26–28, 2016.

[33] C.-P. Hsu, B. Li, B. Solano-Rivas, A. R. Gohil, P. H. Chan, A. D. Moore, and V. Donzella, "A review and perspective on optical phased array for automotive lidar," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 1, pp. 1–16, 2020.

[34] C. V. Poulton, M. J. Byrd, B. Moss, E. Timurdogan, R. Millman, and M. R. Watts, "8192-element optical phased array with 100° steering range and flip-chip cmos," in *CLEO: QELS_Fundamental Science*. Optical Society of America, 2020, pp. JTh4A–3.

[35] M. Muglikar, G. Gallego, and D. Scaramuzza, "ESL: Event-based structured light," in *2021 International Conference on 3D Vision (3DV)*. IEEE, 2021, pp. 1165–1174.

[36] A. Dashpute, J. Wang, J. Taylor, O. Cossairt, A. Veeraraghavan, and F. Willomitzer, "Event-based Motion-Robust Accurate Shape Estimation for Mixed Reflectance Scenes," *arXiv preprint arXiv:2311.09652*, 2023.

[37] M. S. Gottlieb, "Acousto-optic tunable filters," in *Design and fabrication of acousto-optic devices*. CRC Press, 2021, pp. 197–283.

[38] J. Beller and L. Shao, "Acousto-optic modulators integrated on-chip," *Light: Science & Applications*, vol. 11, no. 1, pp. 1–2, 2022.

[39] Z. Yu and X. Sun, "Gigahertz acousto-optic modulation and frequency shifting on etchless lithium niobate integrated platform," *ACS Photonics*, vol. 8, no. 3, pp. 798–803, 2021.

[40] D. Jeong, H. Jang, M. U. Jung, and C.-S. Kim, "Angular resolution variable fmcw lidar with acousto-optic deflector," in *Imaging Systems and Applications*. Optica Publishing Group, 2022, pp. ITh4D–3.

[41] C. S. Tsai, *Guided-wave acousto-optics: interactions, devices, and applications*. Springer Science & Business Media, 2013, vol. 23.

[42] Q. Liu, H. Li, and M. Li, "Electromechanical brillouin scattering in integrated optomechanical waveguides," *Optica*, vol. 6, no. 6, pp. 778–785, 2019.

[43] Q. W. Song, X.-M. Wang, R. Bussjager, and J. Osman, "Electro-optic beam-steering device based on a lanthanum-modified lead zirconate titanate ceramic wafer," *Applied optics*, vol. 35, no. 17, pp. 3155–3162, 1996.

[44] S. Kang, M. Duocastella, and C. B. Arnold, "Variable optical elements for fast focus control," *Nature Photonics*, vol. 14, no. 9, pp. 533–542, 2020.

[45] M. N. Cherkashin, C. Brenner, G. Schmitz, and M. R. Hofmann, "Transversally travelling ultrasound for light guiding deep into scattering media," *Communications Physics*, vol. 3, no. 1, pp. 1–11, 2020.

[46] M. Chamanzar, M. G. Scopelliti, J. Bloch, N. Do, M. Huh, D. Seo, J. Iafrati, V. S. Sohal, M.-R. Alam, and M. M. Maharbiz, "Ultrasonic sculpting of virtual optical waveguides in tissue," *Nature communications*, vol. 10, no. 1, pp. 1–10, 2019.

[47] Y. Karimi, M. G. Scopelliti, N. Do, M.-R. Alam, and M. Chamanzar, "In situ 3d reconfigurable ultrasonically sculpted optical beam paths," *Optics express*, vol. 27,

no. 5, pp. 7249–7265, 2019.

[48] M. B. Hullin, H. P. Lensch, R. Raskar, H.-P. Seidel, and I. Ihrke, “Dynamic display of brdfs,” in *Comput. Graph. Forum*, vol. 30, no. 2. Wiley Online Library, 2011, pp. 475–483.

[49] J. Han, C. Zhou, P. Duan, Y. Tang, C. Xu, C. Xu, T. Huang, and B. Shi, “Neuromorphic camera guided high dynamic range imaging,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 1730–1739.

[50] S. Tulyakov, D. Gehrig, S. Georgoulis, J. Erbach, M. Gehrig, Y. Li, and D. Scaramuzza, “Time lens: Event-based video frame interpolation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 16 155–16 164.

[51] S. Tulyakov, A. Bochicchio, D. Gehrig, S. Georgoulis, Y. Li, and D. Scaramuzza, “Time lens++: Event-based frame interpolation with parametric non-linear flow and multi-scale fusion,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 17 755–17 764.

[52] Z. Qu, Z. Zou, V. Boominathan, P. Chakravarthula, and A. Pediredla, “Event fields: Capturing light fields at high speed, resolution, and dynamic range,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025, pp. 26 910–26 920.

[53] H. Kim, A. Handa, R. Benosman, S.-H. Ieng, and A. J. Davison, “Simultaneous mosaicing and tracking with an event camera,” *J. Solid State Circ*, vol. 43, pp. 566–576, 2008.

[54] H. Rebecq, T. Horstschäfer, G. Gallego, and D. Scaramuzza, “Evo: A geometric approach to event-based 6-dof parallel tracking and mapping in real time,” *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 593–600, 2016.

[55] C. Scheerlinck, N. Barnes, and R. Mahony, “Continuous-time intensity estimation using event cameras,” in *Asian Conference on Computer Vision*. Springer, 2018, pp. 308–324.

[56] L. Pan, R. Hartley, C. Scheerlinck, M. Liu, X. Yu, and Y. Dai, “High frame rate video reconstruction based on an event camera,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 5, pp. 2519–2533, 2020.

[57] P. Shedligeri and K. Mitra, “Photorealistic image reconstruction from hybrid intensity and event-based sensor,” *Journal of Electronic Imaging*, vol. 28, no. 6, pp. 063 012–063 012, 2019.

[58] H.-C. Liu, F.-L. Zhang, D. Marshall, L. Shi, and S.-M. Hu, “High-speed video generation with an event camera,” *The Vis. Comput.*, vol. 33, pp. 749–759, 2017.

[59] C. Haoyu, T. Minggui, S. Boxin, W. YIzhou, and H. Tiejun, “Learning to deblur and generate high frame rate video with an event camera,” *arXiv preprint arXiv:2003.00847*, 2020.

[60] L. Sun, C. Sakaridis, J. Liang, Q. Jiang, K. Yang, P. Sun, Y. Ye, K. Wang, and L. V. Gool, “Event-based fusion for motion deblurring with cross-modal attention,” in *Proceedings of the European Conference on Computer Vision*. Springer, 2022, pp. 412–428.

[61] L. Sun, C. Sakaridis, J. Liang, P. Sun, J. Cao, K. Zhang, Q. Jiang, K. Wang, and L. Van Gool, “Event-based frame interpolation with ad-hoc deblurring,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 18 043–18 052.

[62] F. Xu, L. Yu, B. Wang, W. Yang, G.-S. Xia, X. Jia, Z. Qiao, and J. Liu, “Motion deblurring with real events,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 2583–2592.

[63] L. Zhang, H. Zhang, J. Chen, and L. Wang, “Hybrid deblur net: Deep non-uniform deblurring with event camera,” *IEEE Access*, vol. 8, pp. 148 075–148 083, 2020.

[64] L. Pan, C. Scheerlinck, X. Yu, R. Hartley, M. Liu, and Y. Dai, “Bringing a blurry frame alive at high frame-rate with an event camera,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 6820–6829.

[65] C. Brandli, T. A. Mantel, M. Hutter, M. A. Höpflinger, R. Berner, R. Siegwart, and T. Delbruck, “Adaptive pulsed laser line extraction for terrain reconstruction using a dynamic vision sensor,” *Frontiers in neuroscience*, vol. 7, p. 275, 2014.

[66] J. N. Martel, J. Müller, J. Conradt, and Y. Sandamirskaya, “An active approach to solving the stereo matching problem using event-based sensors,” in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2018, pp. 1–5.

[67] T. Leroux, S.-H. Ieng, and R. Benosman, “Event-based structured light for depth reconstruction using frequency tagged light patterns,” *arXiv preprint arXiv:1811.10771*, 2018.

[68] A. R. Mangalore, C. S. Seelamantula, and C. S. Thakur, “Neuromorphic fringe projection profilometry,” *IEEE Signal Processing Letters*, vol. 27, pp. 1510–1514, 2020.

[69] J. L. Posdamer and M. D. Altschuler, “Surface measurement by space-encoded projected beam systems,” *Computer graphics and image processing*, vol. 18, no. 1, pp. 1–17, 1982.

[70] D. Caspi, N. Kiryati, and J. Shamir, “Range imaging with adaptive color structured light,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 5, pp. 470–480, 1998.

[71] P. S. Huang, C. Zhang, and F.-P. Chiang, “High-speed 3-d shape measurement based on digital fringe projection,” *Optical engineering*, vol. 42, no. 1, pp. 163–168, 2003.

[72] D. Lanman and G. Taubin, “Build your own 3d scanner: 3d photography for beginners,” in *ACM siggraph 2009 courses*. ACM, 2009, pp. 1–94.

[73] J.-Y. Bouguet and P. Perona, “3D photography using shadows in dual-space geometry,” *Int. J. Comput. Vis.*, vol. 35, pp. 129–149, 1999.

[74] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, and K. Daniilidis, “Event-based vision: A survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2020.

[75] M. Wei, S. Nousias, R. Gulve, D. B. Lindell, and K. N. Kutulakos, “Passive ultra-wideband single-photon imaging,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 8135–8146.

Dhawal Sirikonda is a Ph.D. student at Dartmouth College and a member of the Rendering and Imaging Science (RISc) Lab, advised by Prof. Adithya Pediredla. He works on problems in computational imaging. He earned his Master's degree from IIIT Hyderabad, where he worked with Prof. P. J. Narayanan at the intersection of 3D vision and real-time graphics.

Praneeth Chakravarthula is an Assistant Professor at the University of North Carolina at Chapel Hill. His research focuses on novel computational imaging and display systems, at the intersection of optics, perception, computer graphics, optimization, and machine learning. Dr. Chakravarthula is a Senior Member of IEEE and Optica, recipient of the IEEE VR Best Dissertation Award, as well as multiple Best Paper and Demo Awards at premier venues such as ACM SIGGRAPH, IEEE VR, and ISMAR. He completed his postdoctoral research at Princeton University, received his Ph.D. from UNC Chapel Hill, and holds B.Tech and M.Tech degrees in Electrical Engineering with a specialization in Signal Processing from the Indian Institute of Technology (IIT) Madras.

Ioannis Gkioulekas is an Associate Professor in the Robotics Institute at Carnegie Mellon University. He received his Ph.D. and M.S. in Engineering Sciences from Harvard University, and a Diploma in Electrical and Computer Engineering from the National Technical University of Athens. His research focuses on computational imaging and physics-based rendering. He is a recipient of the NSF CAREER Award, a Sloan Research Fellowship, and Best Paper Awards at CVPR and SIGGRAPH. He has served on numerous program committees and editorial boards in computer vision and graphics, and co-organizes the ICCP Summer School on Computational Imaging.

Adithya Pediredla is an Assistant Professor in the Department of Computer Science at Dartmouth College, where he leads the Rendering and Imaging Science (RISC) lab, focusing on the intersection of computer graphics, computational imaging, and computer vision. Before this role, he is a postdoctoral fellow at the Robotics Institute, Carnegie Mellon University. He has a Ph.D. from Rice University and a master's degree from the Indian Institute of Science. He won the Ralph Budd Best Engineering Thesis Award for his Ph.D. thesis, the Prof. K. R. Kambati Memorial gold medal, and an innovative student project award from the Indian National Academy of Engineering for his Master's thesis.

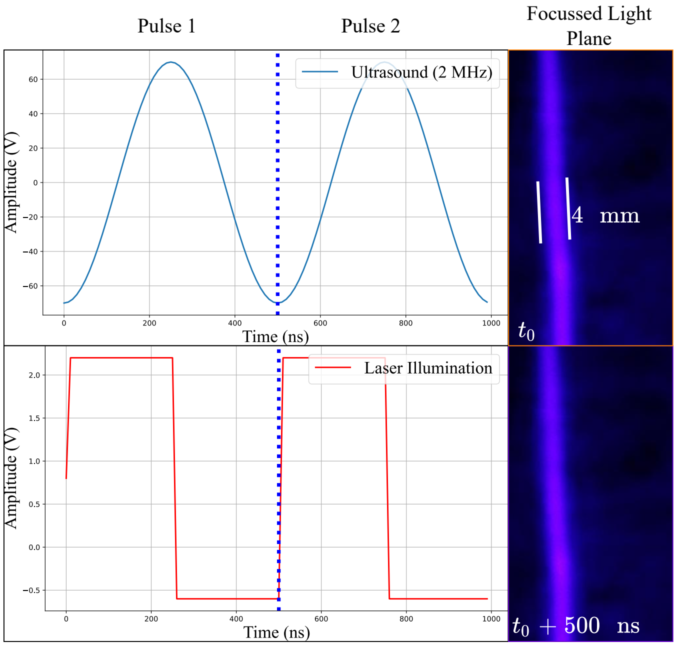


Fig. 11. Stationary light plane creation. The system toggles illumination on and off every $500\,\mathrm{ns}$, keeping the light plane stationary instead of moving with the speed of sound. The thickness of the focused line is $4\,\mathrm{mm}$

7 How to Keep the Light Plane Stationary Instead of Moving at the Speed of Sound in Water?

To keep the light plane stationary, rather than moving at the speed of sound along with the sculpted lenses, we synchronize the laser pulses with the frequency of the ultrasound wave. The ultrasound wave oscillates sinusoidally at $2\,\mathrm{MHz}$, and we pulse the laser at the same frequency ($2\,\mathrm{MHz}$), as we show in the duty cycle diagram in Fig. 11. This synchronization ensures the light plane remains stationary and unaffected by the movement of the sculpted lenses.

The method activates the laser each time a new sound wave reaches a specific point x in the medium, and deactivates it as the wave moves past x. By timing the laser pulses perfectly with the propagation of the sound wave, the system keeps the light plane stationary. Fig. 11 visualizes this method. The laser pulses every $500\,\mathrm{ns}$, corresponding to the $2\,\mathrm{MHz}$ pulsing frequency. This sequence stabilizes the light plane at a fixed position, preventing it from traveling through the medium.

8 How is the Light Plane Moved at a Certain Speed?

In Sec. 7, we demonstrated how to position a light plane at a stationary location. In this section, we examine how the pulsing frequency and delay influence the light plane's position and enable its controlled movement.

Assuming the laser continues pulsing at a frequency of $2\,\mathrm{MHz}$, we introduce a time delay in the pulse, which shifts the illuminated spot of the light plane. We can precisely control this time delay by adjusting the phase of the laser pulse relative to the ultrasound wave. Fig. 12 visualizes how varying the pulsing delay affects the position of the light plane. It demonstrates the direct relationship between phase adjustments and the spatial shift in illumination. We can express the time delay in illumination mathematically as:

$$\phi = \omega_{\mathrm{us}} t \implies t_{\mathrm{delay}} = \frac{\phi}{\omega_{\mathrm{us}}}, \tag{7}$$

where ϕ is the phase delay, ω_{us} is the angular frequency of the ultrasound, and t_{delay} is the resulting time delay. By adjusting this delay, we can control the location at which the light plane is illuminated.

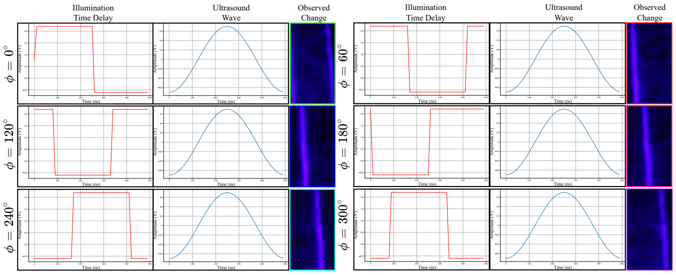


Fig. 12. Effect of time delay in illumination pulses for achieving various light plane positions.

If we define the phase delay ϕ as a time-varying function, such as $\phi(t) = \omega_{\mathrm{spf}} t$, where ω_{spf} represents the frequency of the swept plane, we can dynamically move the light plane at a desired speed. Using this formulation, we can sweep the light plane across space at a controlled speed determined by ω_{spf}.

9 How to Position and Oscillate Between Arbitrary Locations at Desired Speeds?

In Sec. 7, we explained how to position a light plane at a specific location by adjusting the delay in illuminations, and in Sec. 8, we described how to move these lines at desirable speeds. In this section, we detail how to illuminate two specific light planes at distinct locations of interest (Fig. 13(a)).

We assume we want to position two light planes, LP_1 and LP_2, at delays of $0\,\mathrm{ns}$ and $250\,\mathrm{ns}$, respectively, when pulsing the laser. To determine these time delay locations, we adjust the phase of the laser pulse as $\phi = \omega_{\mathrm{us}} t_{\mathrm{delay}}$ from Eq. (7). Creating a suitable waveform is straightforward, as we show in Fig. 13(b, c). These waveforms position the light planes at the desired locations, as we show in Fig. 13(d, e).

9.1 How do we control the oscillation speed between LP_1 and LP_2?

To control the speed of oscillation between these planes, we repeat the pulse with delay d_{i} multiple times before transitioning to $d_{\mathrm{i+1}}$. This repetition adjusts the oscillation speed as needed.

For a general scenario, where n_{l} represents the number of lines of interest and f_{d} represents the desired oscillation speed, we calculate the number of pulse repetitions P_{r} as follows:

$$P_{\mathrm{r}} = \frac{f_{\mathrm{us}}}{n_{\mathrm{l}} \times f_{\mathrm{d}}}. \tag{8}$$

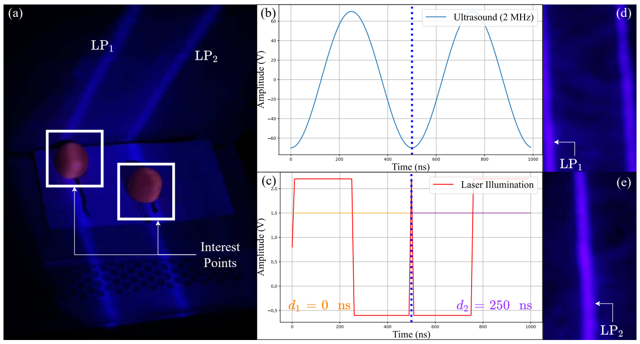


Fig. 13. Arbitrary line placement. This example demonstrates positioning light planes at two locations, LP$_1$ and LP$_2$ (a). We achieve this by delaying the pulsing of the laser between consecutive pulses by $d_1 = 0^\circ \times \frac{180^\circ}{2\pi}/\omega_{\text{us}} = 0\,\text{ns}$ and $d_2 = 180^\circ \times \frac{180^\circ}{2\pi}/\omega_{\text{us}} = 250\,\text{ns}$, where $\omega_{\text{us}} = 2\pi f_{\text{us}}$. With this configuration, the system illuminates the light planes LP$_1$ and LP$_2$ at a rate of $2\,\text{Mfps}$.

From this equation, using $f_{\text{us}} = 2\,\text{MHz}$, desired oscillation frequency $10\,\text{kHz}$, and $n_l = 2$ lines of interest, we compute:

$$P_r = \frac{2 \times 10^6}{2 \times 10 \times 10^3} = 100. \tag{9}$$

This calculation indicates that the pulse with $d_1 = 0\,\text{ns}$ should repeat 100 times, followed by 100 repetitions of the pulse with $d_2 = 250\,\text{ns}$, to achieve an oscillation speed of $10\,\text{kHz}$ ($10\,\text{kfps}$).

The laser pulse waveform is given by the ordered sequence:

$$\{\underbrace{d_1, d_1, \dots d_1}_{P_r}\ \underbrace{d_2, d_2, \dots d_2}_{P_r} \cdots, \underbrace{d_{n_l}, d_{n_l}, \dots d_{n_l}}_{P_r}\}, \tag{10}$$

For a desired oscillation speed of $10\,\text{kfps}$ with 2 light planes and $2\,\text{MHz}$ transducer, the waveform is given by:

$$\{\underbrace{d_1, d_1, \dots d_1}_{\text{100 repeats}}\ \underbrace{d_2, d_2, \dots d_2}_{\text{100 repeats}}\}, \tag{11}$$

10 REFERENCE PLANES FOR STRUCTURED LIGHT

Our structured light scanner uses the classical swept-plane procedure [72, 73], which we visualize in Fig. 14(a). Briefly, this procedure triangulates a 3D point for each event on the event camera, by backprojecting a ray from the pixel of the event camera and intersecting the ray with the light plane P that triggered this event.

The swept-plane method for structured light requires calibrating the light planes relative to the camera. To this end, we follow the procedure by Bouguet and Perona [73], which we visualize in Fig. 14(b): We place two fixed approximately orthogonal reference planes in the scene (planes V and H), and calibrate their location and orientation with respect to the event camera. Then, we calibrate the scanned light planes using their linear intersections (lines λ_V and λ_H) with the two reference planes. To create reference planes that are easy to calibrate with respect to the event camera (which does not take conventional intensity measurements), we use two displays as we show in Fig. 14(c). During the calibration, the displays project a blinking checkerboard pattern that creates events for the camera. During the calibration and scanning of the swept planes, we keep the displays off. As the displays are black, we apply translucent tape to make them reflective, ensuring that the camera captures events when the swept light plane hits the surface.

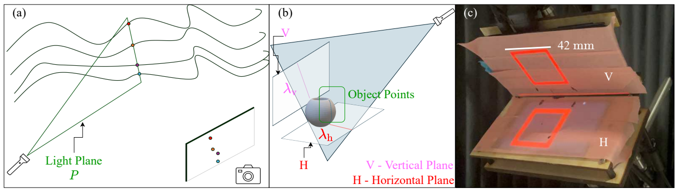


Fig. 14. (a-b)Structured light by light plane sweeping. The image shows a light plane illuminating the scene. A camera observes the illuminated points, and the system retrieves depth information by triangulating the light plane. (c) Reference planes for calibration. We create the reference planes using an approximately orthogonal screen arrangement. We display blinking square patterns of size $42\,\text{mm}$ on the screen to calibrate the reference planes with respect ot the event camera.

TABLE 3. List of components for the hardware prototype.

component	manufacturer
water tank	B01K8WDE4I, Amazon.com, Inc.
transducers	P 12.7 mm-12.7 mm-1 mm-855-WFB, APC International, Ltd (Item No: 1976)
amplifiers	Generic 1 MHz–3 MHz Amplifier with 50 W output (available on Ebay)
signal generators(AWG)	SDG1032X, Siglent Tech.
Q-switched laser	Thorlabs NPL45B
40 mm convex lens	Thorlabs LA1274
galvo	GVS-211, Thorlabs Inc.
event camera	Prophesee, IMX636 sensor

11 HARDWARE DETAILS

Sec. 11 shows the circuit diagram of our prototype: A dual-channel arbitrary waveform generator (AWG) drives both the transducer (T) and the pulsed laser (PL). Both channels operate at $2\,\text{MHz}$ to maintain a stationary light plane. To move the light plane at a desired speed, we introduce a time-varying pulse delay $\phi(t)$, as we explained in Eq. (7).

Tab. 3 lists the components we used to realize the hardware setup. The core components of the acousto-optic light scanning device are a water tank, transducers, and amplifiers. Additional components, such as the galvo-scanning hardware, are only for comparisons. The system requires a pulsed laser, the most expensive component, costing approximately USD 1750 apart from the signal generator. This is similar to Pediredla et al. [1]. Unlike their work, our design eliminates the need for a costly, high-power, high-bandwidth amplifier by shifting the steering control to the illumination side.

12 VOLTAGE VS LIGHT PLANE INTENSITY

In Fig. 16, we show the focusing behavior of the line as the pressure profile evolves with increasing input voltage. The

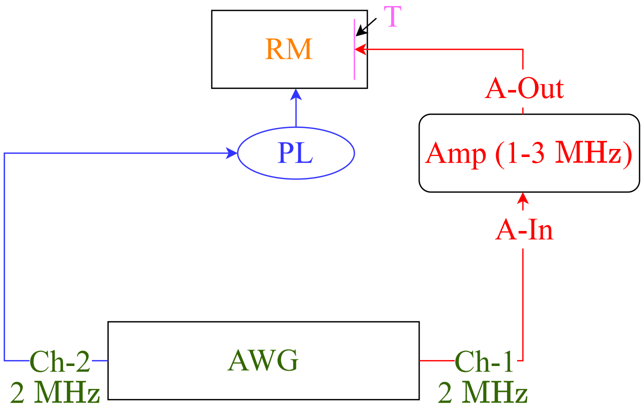
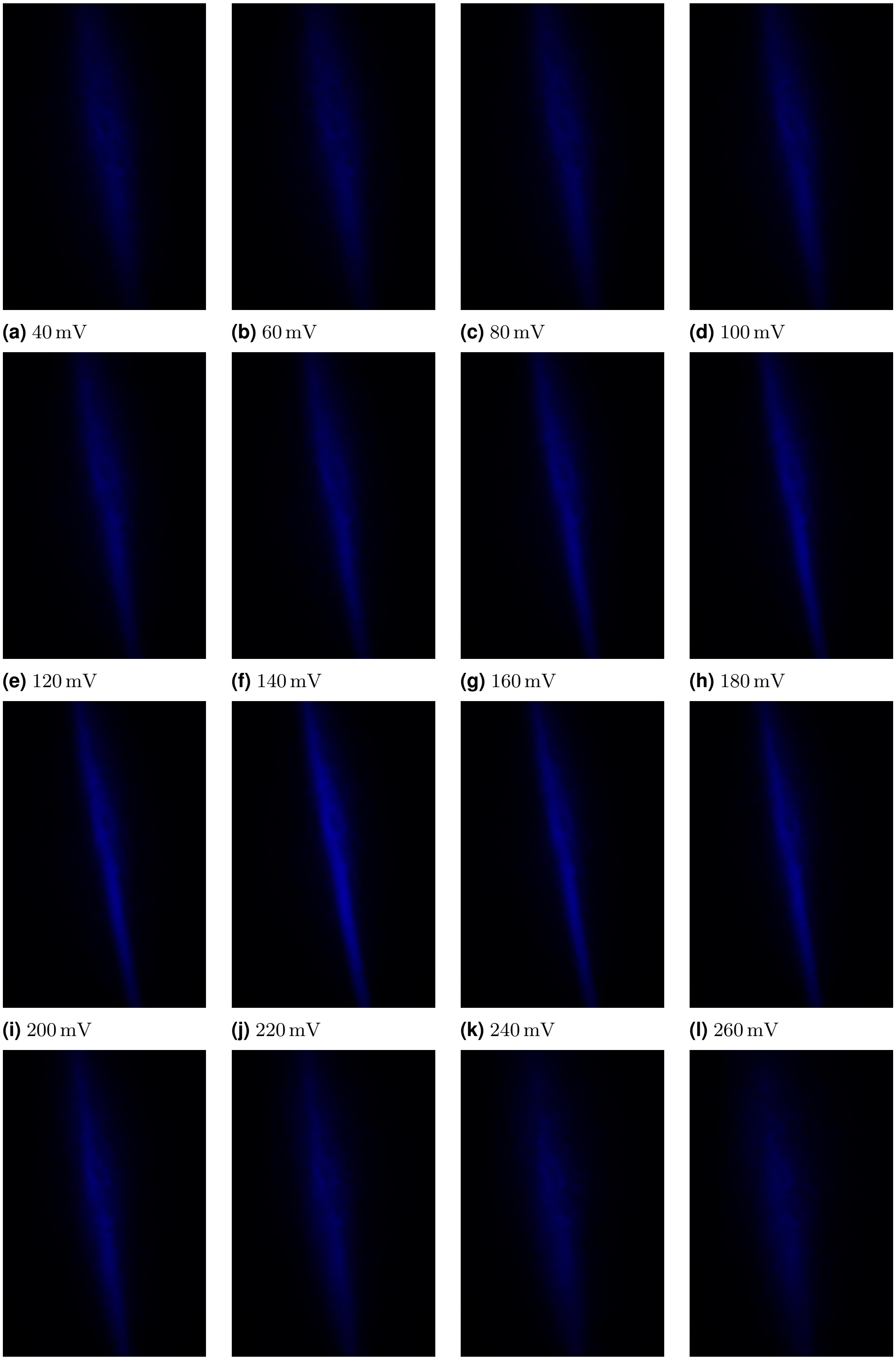


Fig. 15. Circuit diagram of the hardware prototype. A two-channel arbitrary waveform generator (AWG) drives both the laser and the amplifier (Amp) whereby providing a PLL between two channels. Channel-1 (Ch-1) of the AWG connects to the Amp via the amplifier input (A-In) and then to the transducer (T) through the amplifier output (A-Out). The transducer (T) operates within the transparent medium (RM), sculpting the cylindrical lens. Channel-2 (Ch-2) of the AWG drives the pulsed laser (PL), which generates the light planes based on the mechanisms described in Sec. 7 and Sec. 8. The setup includes the following components: AWG - Arbitrary Waveform Generator, Ch-1 - Channel 1, Ch-2 - Channel 2, Amp - Amplifier ($1\,\mathrm{MHz}$–$3\,\mathrm{MHz}$), A-In - Amplifier Input, A-Out - Amplifier Output, T - Transducer, RM - Refractive Media, PL - Pulsed Laser.

line begins to focus around $220\,\mathrm{mV}$ and starts to degrade beyond $240\,\mathrm{mV}$. Unlike the results reported by [1], we observe a slight shift in the optimal voltage range, which we attribute to differences in the amplifier used. Nevertheless, our setup exhibits similar behavior to the more expensive alternative employed in their work.



(a) $40\,\mathrm{mV}$ **(b)** $60\,\mathrm{mV}$ **(c)** $80\,\mathrm{mV}$ **(d)** $100\,\mathrm{mV}$

(e) $120\,\mathrm{mV}$ **(f)** $140\,\mathrm{mV}$ **(g)** $160\,\mathrm{mV}$ **(h)** $180\,\mathrm{mV}$

(i) $200\,\mathrm{mV}$ **(j)** $220\,\mathrm{mV}$ **(k)** $240\,\mathrm{mV}$ **(l)** $260\,\mathrm{mV}$

(m) $280\,\mathrm{mV}$ **(n)** $300\,\mathrm{mV}$ **(o)** $320\,\mathrm{mV}$ **(p)** $340\,\mathrm{mV}$

Fig. 16. The images illustrate the change in line width with increasing input voltage, starting from $40\,\mathrm{mV}$ in $20\,\mathrm{mV}$ increments. We found the optimal voltage range to be between $200\,\mathrm{mV}$ and $240\,\mathrm{mV}$.