

Dense Lissajous Sampling and Interpolation for Dynamic Light-Transport

XIAOMENG LIU,^{1,*} KRISTOFER HENDERSON,¹ JOSHUA REGO,²
SUREN JAYASURIYA,^{2,3} AND SANJEEV KOPPAL¹

¹Department of Electrical and Computer Engineering, University of Florida

²School of Electrical, Computer and Energy Engineering, Arizona State University

³School of Arts, Media and Engineering, Arizona State University

*liuxiaomeng51@ufl.edu

Abstract: Light-transport represents the complex interactions of light in a scene. Fast, compressed, and accurate light-transport capture for dynamic scenes is an open challenge in vision and graphics. In this paper, we integrate the classical idea of Lissajous sampling with novel control strategies for *dynamic light-transport applications* such as relighting water drops and seeing around corners. In particular, this paper introduces an improved Lissajous projector hardware design and discusses calibration and capture for a microelectromechanical (MEMS) mirror-based projector. Further, we show progress towards speeding up the hardware-based Lissajous subsampling for dual light transport frames, and investigate interpolation algorithms for recovering back the missing data. Our captured dynamic light transport results show complex light scattering effects for dense angular sampling, and we also show dual non-line-of-sight (NLoS) capture of dynamic scenes. This work is the first step towards adaptive Lissajous control for dynamic light-transport. **Please see accompanying video for all the results.**

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Light transport effects include direct bounces of light to the detector, global light effects due to multiple reflections, refraction, and scattering in the scene. Capturing the full light transport enables numerous applications in computer vision and graphics including post-capture visualization using image-based relighting techniques.

The linear relationship between light source and camera sensor is typically modeled as the light transport matrix \mathbf{T} [1–4]. This matrix maps the illuminating pattern \mathbf{p} to the camera image \mathbf{c} , governed by the equation $\mathbf{c} = \mathbf{T} \mathbf{p}$. There has been much research in acquiring \mathbf{T} for static scenes, primarily focused on improved capture efficiency [5–10]. Applications include relighting static scenes post-capture [11], separating direct and global components of light in the scene [12], and creating dual imagery [7] using Helmholtz’s reciprocity.

However, capturing *dynamic* light-transport is challenging due to the divergent requirements of fast capture. Here, the light transport matrix is a function of time $\mathbf{T}(t)$, and usual static light transport acquisition face additional challenges.

1.1. Dynamic light transport capture challenges

To capture $\mathbf{T}(t)$, researchers have developed high-speed acquisition setups with fast projectors and cameras. There are however fundamental trade-offs with high-speed capture for dynamic light transport. For example, high-speed capture, which happens in short bursts (i.e. low exposure), suggests a focused, bright single spot that is scanned over the scene, such as a “flying spot” [13]. Unfortunately, sequentially scanning the flying spot takes time. Abandoning the flying spot in favor of flood-lit parallel projected illumination patterns can allow smart, fast sampling [7]. However, we then lose the flying-spot advantage and the spread of energy over the scene reduces the signal-to-noise ratio (SNR).

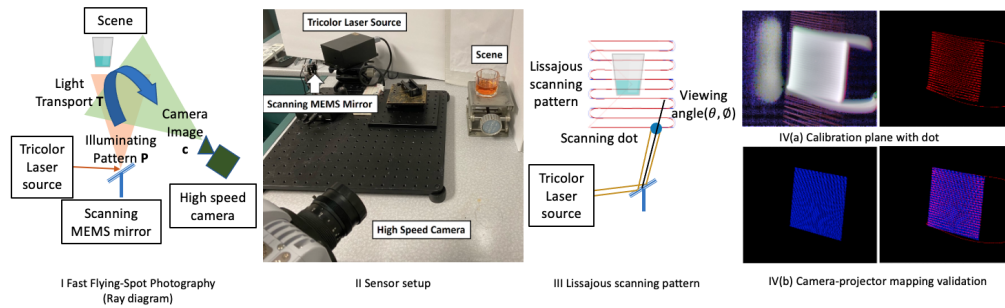


Fig. 1. Fast Flying-Spot Photography setup and mapping from slide coordinates to Lissajous pattern. In (I), we show our tri-color laser which emits optically combined red, green and blue lasers. This is reflected off a small, scanning mirror, creating a narrow pencil or rays, or spot, that is incident on the scene. A color, high-speed camera allows the capture of the moving "flying-spot", as it is scanned over the scene by the mirror. II shows the sensor setup of the ray diagram. III shows the Lissajous scanning pattern and the variables for each scanning dot. IV(a) left, we show a Lambertian plane we use for calibration. The light-transport of this calibration plane has been captured, and each column consists of a single bright dot. The centers of each dot are shown in red in IV(a). Note that these dots form a sinusoidal, Lissajous pattern, instead of the conventional uniform grid sampling in a slide image. In IV(b) we show mapping of these uniform projector pixels in blue, and compare them to the dot centers in red in IV(b) right. To map between the real Lissajous pattern and the desired slide coordinates, the relative dot distances on this plane are used in weighted nearest neighbors.

Most previous work in dynamic light-transport has avoided these trade-offs by giving up illumination resolution and focusing instead on complex, omnidirectional, environment maps. These "light-stage" methods that use gantry-based setups have been extremely successful for relighting human faces or human motion [3, 14–16]. Additionally, robust methods exist to capture slices of the dynamic light transport [17–20].

1.2. Dense dynamic light transport capture with Lissajous sampling

In this paper, we focus on capturing densely sampled full light-transport, with high-resolution both in the camera and in the illumination, using Lissajous sampling for MEMS-mirror control. This type of sampling is widely used for conventional imaging applications, such as endoscopy [21], remote sensing [22] and imaging [23], where MEMS-scanning has enabled a variety of commercially available portable projectors. *In this paper, we investigate the impact of Lissajous sampling on light-transport, showing results such as video-rate "seeing around corners" and relighting of dynamic water droplets and caustics.*

Recently, Henderson et al. [24] introduced a flying-spot projector that captured dynamic light transport with moderate illumination resolution, and showed applications such as video-based relighting and dual videography. Our work builds upon their work [24], but we introduce an upgraded hardware design as well as new sampling and interpolation algorithms for dynamic light transport. This paper presents the first steps towards control and sampling of a bright flying-spot, followed by light-transport interpolation. **Our contributions are:**

1. We present a simple, new optical setup by combining a microelectromechanical (MEMS) mirror modulated tri-color laser with a high-speed camera in Sect. 3. This setup provides higher resolution than recent MEMS-mirror designs [24], and can demonstrate dual-NLoS imagery of dynamic scenes (Fig. 4). We also show new calibration that provides similar

quality to previous work [25], extended to dynamic scenes (Fig. 2). We show dynamic relighting and dual imaging for moving objects, glass/liquid scenes and fog effects (Sect. 6).

2. We show the theory of how to adaptively change the scanning pattern, based on a desired illumination for light-transport capture, by finding a hardware-based Lissajous pattern that controls the MEMS mirror motion. This Lissajous pattern is fitted to a target pattern in the dual light transport frame. In Sect. 4, we show that the physical constraints of the second order mass-spring-damper system allows the existence of an optimal Lissajous pattern that satisfies these constraints. We validate this theory with simulations.
3. We investigate heuristics for fine-tuning a state-of-the-art video interpolation neural network [26] for light-transport, showing global effects such as caustics (Sect. 5). We present real results and simulations showing potential speedups in capture time of up to 8 times without affecting quality, including a real scene at 24 FPS.
4. Finally, we provide evaluations of the approach, such as Table 1 and Fig. 9, which validate these early steps towards adaptive scanning for light-transport capture.

Organization: For readers desiring high level understanding, this introduction along with Section 3, particularly Sections 3.1 and 3.2 discuss the design considerations for a flying-spot projector. For readers focused on the hardware setup and calibration, we refer them to Section 3.3. Section 4 discusses Lissajous sampling of the scene due to the MEMS mirror including a proof about an optimal fit to a desired target control pattern. Section 5 then proceeds to describe interpolation using a deep neural network for recovering missing light transport samples. Finally, Section 6 shows applications in light transport using our prototype setup.

2. Related work

Efficient light-transport capture. Capturing light-transport has been an active research area [1–4]. Efficient capture has been shown with compressive sensing [5, 6], adaptive illumination [7], symmetry priors [8], and low-rank approximations [9, 10]. However, most of these methods have not impacted light-transport for *dynamic scenes*. Our strategy of scanning a bright flying-spot improves SNR and enables dual-NLoS video, which has not been shown with optically parallel, projected patterns such as those in compressive light-transport [5, 6].

Dynamic light-transport with light stages. Dynamic light transport capture has been shown in interactive or video frame rates [15, 16, 27]. These *light-stages* typically contain programmable sources arranged in a dome to acquire reflectance fields. These stages trade-off dense angular resolution over the dome for high-quality capture in a large field-of-view, allowing applications such as relighting human faces [3, 14], capturing moving actors [15], high-speed photometric stereo [27], and even 7D information for relighting walking/running humans [16]. For these applications, omnidirectional ambient illumination (i.e. environment map) is sufficient. In contrast, we show dense dynamic light transport capture, for relighting caustics and fog (without requiring its 3D reconstruction [28]).

The paper most similar to ours introduced a MEMS-based flying-spot projector for capturing dynamic light transport [24]. This paper used a LED source and MEMS-mirror to steer the flying spot and capture dynamic light transport for a variety of scenes. These results were enabled by novel calibration and denoising algorithms. Our work extends on these ideas but features a novel laser source in the hardware design eliminating the need for denoising algorithms, the demonstrated capture of dual-NLoS videography, and Lissajous sampling and interpolation strategies for light transport acquisition.

Lissajous sampling: Lissajous patterns are an excellent model for MEMS mirror modulation, since these are run by actuators in resonance, in silicon. Many MEMS mirror enabled devices,

such as OCT scanning for endoscopy and imaging [21] as well real-time projection [23], use these ideas. MEMS mirror-based projectors are used in a variety of commercial devices, such as Microvision ShowWWX, as well as in Mirrorcle LIDARs [29]. In contrast to all these methods, we investigate Lissajous control for dynamic light-transport capture, with applications such as seeing around corners and relighting dynamic caustics and water drops.

Dense light-transport slices. Most dense light-transport methods focus on static scenes [7, 10, 30]. There has been recent work on capturing dense light-transport components using fast optical modulation, such as coded exposure [31, 32], digital micromirror devices [18, 33, 34], MEMS projectors [35, 36], and synced camera-projectors [20, 37–39]. Our work seeks to match the speed of such approaches for *full* light transport capture. To achieve this, we extend ideas from laser scanning dual light-stages for static scenes [25] with MEMS-mirror based flying-spot projectors that suffer from noisy capture [24].

Deep image-based relighting Recent advances in neural relighting and machine learning have achieved state-of-the-art in image-based relighting. Facial relighting performs realistic portrait relighting and handles complex shadows, including some involving light stages [40–42]. Neural networks have been used to estimate light transport matrices [43] with much less measurements, or perform tasks such as direct/global separation [44]. Our work shares, with these other methods, the goal of deep interpolation of missing data in the light-transport. Our method is able to demonstrate this for dynamic light transport, including the global component of the light transport.

NLoS imaging: Recent non-line-of-sight (NLoS) imaging research [45–49] includes using ultra-fast lasers [50], low-cost sensors [51], special optics [52], elliptic tomography [53], the light cone transform [54, 55], and Fermat paths [56]. High-speed corner imaging has also enabled NLoS imaging [57]. Recently NLoS video reconstructions have been shown for retroreflective scenes [58, 59].

In this paper, we focus on the *relatively simpler problem of dual-NLoS*, where the light-source directly illuminates the scene, allowing for dynamic dual videography. We exploit Helmholtz reciprocity [7, 60–62], which has been shown for video using epipolar imaging [20]. Our application of dual-NLoS imaging is closest to efforts that reconstruct dual views from shadows [63] and steady-state-based inversion [64]. Our work is also similar to Musarra et al. and Nam et al. [65, 66]. Musarra et al. achieved 0.8 second NLOS reconstruction with a single-pixel GameraG [65]. Nam et al. used two 16×1 SPAD arrays to reconstruct live real time videos of NLOS scenes at 5 fps [66]. In contrast, we also show dual-NLoS color videos of diffuse scenes and relighting of dynamic caustic effects, but at a faster speed of 7 fps and higher resolution 328×768 as compared to these previous works [65, 66].

3. Capturing dense dynamic transport

In this section, we introduce a novel laser dot scanning optical design that builds upon two previous light-transport capture designs [24, 25]. Both these designs, and ours, build on Baird’s flying-spot scanner [67] which allowed live television [13] before the advent of cathode ray tubes. In contrast to large environmental lighting from light stages [15, 16, 27], flying-spot scanners can achieve dense angular resolution that enables capturing scene lighting effects such as caustics of glass objects, light scattering in fog, and subsurface scattering at high visual fidelity. The hardware setup and calibration described here is later used for dynamic relighting and dual videography for both line-of-sight (LoS) and non-line-of-sight (NLoS) scenes in Section 6.

3.1. Design of flying-spot projectors

Consider a scene being illuminated by a flying spot scanner. We denote the light transport matrix \mathbf{T} [1–4] as mapping the illuminating pattern \mathbf{p} to the camera image \mathbf{c} , as $\mathbf{c} = \mathbf{T} \mathbf{p}$. Flying spot data are impulse responses of the light transport matrix \mathbf{T} [7, 68]: $\mathbf{T}^i = \mathbf{T} \delta_i$, where δ_i is the impulse

response and the i^{th} column of the matrix is T^i . Now consider a dynamic scene, where the light transport matrix varies with time t :

$$\mathbf{c}(t) = \mathbf{T}(t) \mathbf{p}(t). \quad (1)$$

Flying-spot projectors capture $\mathbf{T}(t)$ within a time interval Δt by scanning a flying spot, as an optical analog to the impulse response δ_i . Flying-spot scanning only works if (1) the spot brightness is detectable within the time interval Δt , and (2) the spot modulation is fast compared to scene motion within the Δt time interval. We now formally outline these two constraints, and we assume they are held going forward in all our experiments.

We first model exposure, building on [69], for the capture interval Δt and sensor gain g , where N_T are the number of light transport columns, then the requirement for the light source power Φ should satisfy:

$$\frac{\Phi B_{min} \frac{\Delta t}{N_T}}{g} + n > I_{min} \quad (2)$$

where n is a sensor noise term, I_{min} is the lowest discernable sensor irradiance, and B_{min} is a reflectance term denoting the loss due to scene transport, and is scene dependent.

Secondly, within the Δt scanning interval, there should be no apparent motion or, equivalently, the change in the light transport matrix should be less than an acceptable error ϵ :

$$\|T(t) - T(t + \Delta t)\| < \epsilon. \quad (3)$$

Eqns. 2 and 3 are the two constraints necessary for fast flying spot capture of light transport matrices. In particular for dual-NLoS scenes, B_{min} can be quite small due to indirect light reflections, and thus satisfying both constraints can be hard.

3.2. Flying-spot vs. conventional projectors.

Our flying-spot projector concentrates light in a region, whereas traditional light transport capture uses parallel light projection [6, 7, 70]. O’Toole et al. [20] showed that such conventional parallel projectors, using spatial light modulators, are not as efficient as concentrating the light into a small region. Formally, flying spot scanning focuses the available illumination energy into a smaller cone ω_{small} compared to over the entire FOV ω_{FOV} of a conventional projector. This multiplexing of the energy results in a k times increase in scene radiance, given by how much bigger the FOV is compared to the smaller cone, $k = \frac{\omega_{FOV}}{\omega_{small}}$.

In Henderson et al.’s work [24], these ideas were extended for time-efficiency, noting that since parallelism gives a conventional projector a speed advantage (say m times faster), it can use this extra time to increase exposure and reduce noise. This suggests a simple ratio to determine when flying spot scanning, known to be energy efficient [20], is also time-efficient,

$$\frac{m}{k} < 1. \quad (4)$$

In our hardware setup, we increase k by using a tri-color laser light source combined with a pentaprism, blending 3 colors of light (RGB) to generate a bright white dot. This reduces the illumination solid angle ω_{small} more than the LEDs of Henderson et al.’s work [24] while allowing dynamic light transport capture, unlike previous work [25]. While this is a simple hardware enhancement, it results in never-before-seen results such as dynamic dual NLOS photography as shown in Section 6.

3.3. Fast flying laser spot system

Our system is shown in Fig 1, with the tri-color laser, high-speed camera, and MEMS mirror. We designed a customized, collimated 5V DC RGB (red-green-blue) combined Class 2 laser module. The maximum laser output power is $1mW$ for the three lasers at $450nm$, $520nm$, and $635nm$. The beam size is $1mm$ with a divergence of $0.8\mu rad$, which favors reasonably when coupled with a Mirrorcle MEMS mirror of $3.6mm$ -diameter. For high-speed capture, we use a Photron model SA-X2 color high-speed camera with a $50mm$ $f/0.95$ Navitar lens set at $f/1.2$, with a focal plane set to $0.73m$, exposure time $0.018ms$, frame rate $50k$ FPS, resolution 328×768 .

Calibration for interpolation. We first perform standard high-speed sensing calibration steps including dark calibration, shading correction, and hot pixel correction. We then replicate the batching process of previous work [24], using the reset curve of the MEMS mirror path as shown in Fig 1IV(a). This batches the high-speed camera data into individual flying-spot images. The i^{th} flying spot is the corresponding column in the dynamic light-transport $\mathbf{T}^i(t)$ at time t . Every video frame consists of $8k$ scanning dots, resulting in a 50×160 flying spot scanning resolution.

For dual-NLoS results, we require an additional calibration step to find the mapping between projector and camera coordinates. Conventional dual view imaging [7] uses the transpose of the light transport $\mathbf{T}'(t)$. In other words, each pixel of the dual view is the sum of the corresponding column of the light-transport matrix at time t . However, this mapping only works for conventional projectors where rectangular pixels are arranged in a grid, whereas our scanning path is determined by the sinusoidal-like Lissajous pattern of the MEMS mirror.

The indices required for interpolation are found by using a Lambertian backplane, explicitly placed in the scene at the beginning of each dual-NLoS experiment, to avoid synchronization complications between MEMS mirror and camera across experiments. We manually identify six point correspondences between projector pattern and camera image to apply standard homography estimation techniques as shown in Fig 1IV(b). Due to this homography, every flying dot location (and matching light-transport column) is mapped to a location on the calibration plane, as a series of correspondences.

This mapping also helps sub-sampling, explained in Sect. 6, which creates gaps in the light transport. Consider a flying spot position *not* captured, due to subsampling, which corresponds to a missing light-transport column. We map the missing flying spot position to the calibration plane and compute weights to nearby measured flying spot positions along the calibration plane. We apply either weighted nearest neighbors or deep interpolation (described in Sect. 5) to combine the light transport columns corresponding to these positions to obtain the missing data.

Limitations of scanning. Based on the limitation of the current hardware, the MEMS mirror we use can only scan a relative small angle of the dynamic scene ($\pm 5.5^\circ$ horizontally and $\pm 1.375^\circ$ vertically). This limits our current maximum field-of-view of the system. However, additional optics could be applied to expand the scanning range. An ultra wide-angle fish eye lens would help extend the angular range of the system, but additional calibration algorithm would be needed.

4. Control of light-transport capture

In the previous section, we discussed how we could increase the scene radiance multiplier, k , via a brighter laser source to increase the time-efficiency $\frac{m}{k}$ of the system. In this section, we drive the MEMS mirror faster, reducing the quantity m in the system.

Our idea is to sample only a subset of light-transport T columns by skipping certain areas of the scene with the flying spot. This sparse sampling would leave gaps, which we would interpolate back in post-processing as explained in Sect. 5.

Our method assumes a target control pattern M that correspond to how important they are to be sampled. This target control pattern could come from another sensor (thermal or event cameras) or from a vision algorithm. We now explain how our hardware setup offers the unique ability to have control over the MEMS mirror to design the scanning pattern of the flying-spot,

Optimal Lissajous sampling for light-transport. Fast flying-spot scanning requires running the MEMS mirror at or near resonance. Unlike a conventional projector, the flying dot is controlled by the angle of the MEMS mirror. It has been shown that this angle can be represented as different Lissajous patterns [23], parameterized on a plane perpendicular to the optical axis at unit distance away from the mirror. Our system captures the light-transport at time t over the period $t_L \in (t, t + \Delta t)$ in the Lissajous pattern given by

$$x(t_L) = A \sin(\omega_h t_L + \phi), \quad y(t_L) = B \sin(\omega_v t_L), \quad (5)$$

where A and B are the amplitudes of the driving sinusoids of the Lissajous pattern, $\omega_{h,v}$ are the driving frequency of the MEMS mirror in the horizontal and vertical direction respectively, and ϕ is the phase difference between these. Thus our goal is to find a set of parameters $\mathbf{\Pi} = (\omega_h, \omega_v, \phi, A, B)$ to realize a Lissajous pattern $L(\mathbf{\Pi}) = (x(t_L), y(t_L))$ to realize sampling patterns for the scene. In Section 5, we then discuss how to use either nearest-neighbor interpolation or neural networks to recover the missing light transport columns that are not sampled by this pattern.

Suppose we are given a target control pattern $M(x, y)$, defined on the virtual plane that maps $\mathbb{R}^2 \rightarrow [0, 1]$, which represents the probability of that (x, y) location being sampled. We then define the overlap measure $E_{overlap}$ for a given Lissajous pattern $\mathbf{\Pi}$ as follows:

$$E(\mathbf{\Pi})_{overlap} = \int_{L(\mathbf{\Pi})=(x(t_L), y(t_L))} M(x(t_L), y(t_L)) dx dy \quad (6)$$

We now present the following proposition, given an assumption on the continuity of the target control pattern, that there exists an optimal Lissajous pattern which is possible in our MEMS mirror scanning scenario.

Proposition 1. *Assume that M is a continuous function. Then there exists at least one Lissajous pattern $\mathbf{\Pi}^*$ such that the overlap measure $E_{overlap}$ is maximized.*

Proof sketch. First we show that the overlap measure $E(\mathbf{\Pi})_{overlap}$ is continuous. Then we show that certain physical qualities of MEMS mirror-based scanning imply the domain for this measure, $\mathbf{\Pi}$, is bounded. Since we show a continuous measure $E(\mathbf{\Pi})_{overlap}$ with a bounded domain $\mathbf{\Pi}$, the proof follows from the extreme value theorem.

Proof. Part 1: E is continuous. We assume M is a continuous function, which is to say, it is Riemann integrable on a bounded domain. Now consider $M(x(t_L), y(t_L))$. If the domain $(x(t_L), y(t_L))$ is bounded, then $M(x(t_L), y(t_L))$ is Riemann integrable too. The domain $(x(t_L), y(t_L))$ comes from the Lissajous function Eq. 5 which, due to sine functions, is bounded $[-1, 1]$. Therefore $M(x(t_L), y(t_L))$ is Riemann integrable, and the integral of a Riemann integrable function is continuous. Therefore the overlap measure $E(\mathbf{\Pi})_{overlap}$ is continuous.

Part 2: $\mathbf{\Pi}$ is bounded. We show that $\mathbf{\Pi}$ degenerates into three parameters with bounded domain, for the MEMS mirror setup. Most 2D MEMS mirror resonant scanning can be simplified as a 2nd order mass-spring-damper system. This means that each axis of the mirror has a resonant frequency, ω_{h0} for the horizontal axis and ω_{v0} for the vertical axis.

Note that we do not have to drive the MEMS at these frequencies, but at any driving frequencies in the horizontal and vertical, ω_h and ω_v respectively. However, this impacts the amplitude of the scan. Therefore, instead of being two free parameters, the amplitudes A and B of the MEMS mirror's Lissajous pattern are determined by the system transfer function of the mass-spring-damper system at these resonant frequencies,

$$A = \|H_A(j\omega_h)\| = \left\| \frac{\omega_{h0}^2}{\omega_{h0}^2 - \omega_h^2 + \frac{\omega_{h0}}{Q_h} j\omega_h} \right\| \quad (7)$$

$$B = \|H_B(j\omega_v)\| = \left\| \frac{\omega_{v0}^2}{\omega_{v0}^2 - \omega_v^2 + \frac{\omega_{v0}}{Q_v} j\omega_v} \right\| \quad (8)$$

where Q_h and Q_v are physical factors relating to the fabrication quality of the MEMS mirror. Our key observation is that the only non-resonant driving frequencies that we will consider are *greater* than the resonant frequency, and we might take the hit on the loss in amplitude if the overlap on the desired pattern M worked in our favor. Therefore, rewriting,

$$A = \|H_A(j\omega_h)\| = \left\| \frac{\left(\frac{\omega_{h0}}{\omega_h}\right)^2}{\left(\frac{\omega_{h0}}{\omega_h}\right)^2 - 1 + \left(\frac{\omega_{h0}}{Q_h}\right)j} \right\|, \quad (9)$$

$$B = \|H_B(j\omega_v)\| = \left\| \frac{\left(\frac{\omega_{v0}}{\omega_v}\right)^2}{\left(\frac{\omega_{v0}}{\omega_v}\right)^2 - 1 + \left(\frac{\omega_{v0}}{Q_v}\right)j} \right\|. \quad (10)$$

Note that if we specify $k_v = \frac{\omega_0}{\omega_v}$ and $k_h = \frac{\omega_0}{\omega_h}$ then $k_v, k_h \in [0, 1]$ because we always want faster MEMS motion to end within the time period Δt . Thus the numerators of A, B are ≤ 1 . For the denominator, we note that there is no choice of $\omega_{h,v}$ which causes a pole in the transfer function (definition of quality $Q_{h,v}$), and thus A, B are finite quantities and thus bounded. Also, due to sinusoidal periodicity, the phase $\phi \in [0, 2\pi]$. Finally, $\omega_{h,v}$ are bounded due to hardware parameters to some max frequency possible by the mirror. Therefore the domain $(x, y) = L(\mathbf{\Pi})$ is bounded as $\mathbf{\Pi}$ is bounded.

Conclusion: We thus note from Parts 1 and 2 of the proof that the overlap function $E_{overlap}(\mathbf{\Pi})$ is continuous on a bounded domain. Therefore by the extreme value theorem, there exists at least one optimal Lissajous pattern $\mathbf{\Pi}^*$ such that $E_{overlap}(\mathbf{\Pi}^*)$ is maximized. \square

Scope: The above proof tells us that for a given target pattern (represented in continuous probability), there is a best fit Lissajous pattern that exists. It does not tell us how good this overlap is, nor the discrete approximation error when $M(x, y)$ becomes binary labels. In practice, in Sect. 6, we have noticed relaxing these assumptions does not affect performance and we utilize these optimal Lissajous patterns to perform light transport subsampling.

5. Light transport interpolation

In the previous subsections, we discussed improving the ratio $\frac{m}{k}$ in our favor by increasing k with a tri-color laser and reducing m with a custom sampling pattern. We now discuss a subsequent interpolation step to recover those missing column information from the custom pattern. This has two advantages: it enables the faster scanning of the scene discussed previously as well as allows fewer images stored at capture time.

There are several existing methods for interpolating between the columns (or equivalently dot images) including nearest-neighbor and classical optical flow. In fact, we show later that the nearest-neighbor method is particularly effective for dual-NLoS sampling where global light is so weak that any further processing on the images can distort the dual result. However, complex lighting effects such as global interreflections or caustics and/or geometry boundaries such as non-planar surfaces, occlusions, and shadows will deform the flying spot to change shape in that region. Thus interpolation becomes more challenging for these methods.

To overcome these challenges, we utilize a modern state-of-the-art neural network for video interpolation from previous work [26] to interpolate missing columns of the light transport matrix. We fine-tune this network on a previous frame's light transport matrix and then deploy the network at test time for future frames. We show how this network can outperform both traditional Farneback optical flow [71] as well as FlowNet2.0 [72] in estimating light transport columns, particularly for global light for line-of-sight (LoS) scenes. The disadvantage of fine-tuning is that

it is required for each new scene. However fine-tuning only takes a few epochs when initialized with pretrained weights from previous work [26].

Network architecture. We adopt the network from [26] to use as our network architecture for frame interpolation (which we entitled ZSM). This network interpolates between any two frames by extracting visual features, then uses deformable convolutions to help temporally interpolate these features, before feeding both the interpolated features and the original frames into a bidirectional LSTM to output the final interpolated frame Xiang et al. [26]. We do not use the final upsampling layer from the original network in our implementation.

To interpolate between two column images with indices i and j , we utilize a simple binary search tree procedure to fill in the missing indices between i and j . The network first interpolates the midpoint $k = \text{mid}(i, j)$, then $k_1 = \text{mid}(i, k)$, $k_2 = \text{mid}(k, j)$ in the second pass, and so on until all interpolations are complete. Since the flying spot travels in a linear fashion, typically i and j lie on an approximately horizontal line and can be interpolated using our method. Depending on the pattern, as in Figure 8, we perform vertical interpolation to fill the gaps.

This ZSM network was trained on everyday scenes, and thus we investigated whether the network could generalize to flying spot images taken from our setup. In Section 6.2, we show the results of these experiments which demonstrate the need for fine-tuning the network, particularly on interpolating global light effects such as caustics. To fine-tune the network, we require additional losses to help improve performance. We utilize both an MSE loss as well as a perceptual loss [73] using a pretrained VGG-19 [74] network to extract features $\phi(I)$ of a flying-spot image: $\mathcal{L}_{\text{percep}}(I, \hat{I}) = \|\phi(I) - \phi(\hat{I})\|_2^2$, where \hat{I} is the network output and I is the ground truth flying spot image. In training, we assign loss weights of 1.0 for the MSE loss and 0.01 for the perceptual loss. Section 6.2 shows that our fine-tuned network trained on these losses can interpolate complex light transport effects.

6. Experimental results

Capture details. All scenes were captured with the mirror scanning $\pm 5.5^\circ$ in the horizontal and $\pm 1.375^\circ$ in the vertical at 50,000 samples/sec, and with the Photron SA-X2 color camera at 50,000 FPS. For line-of-sight (LoS) scenes, the target was placed in direct view of light-source and camera. For dual-NLoS scenes, we used a diffuse v-groove where the projector targeted one side of the v-groove where the object was located, and the camera captured the other side. Thus the camera did not have the object visible in its image, but the illumination directly hit the object.

LoS results. In Fig. 2, we show four LoS scenes obtained by our system. These scenes include objects showing sub-surface scattering (such as a candle), caustics (water droplets and glass), and vapor (due to dry ice in liquid). We have captured the full light transport of each dynamic scene, shown by the dual view and relighting results.

We obtained similar illumination data to Henderson et al. [24] through their authors’ website, allowing qualitative comparison. Other than a simple thresholding scalar, we do not perform any of the complex denoising of previous paper [24], despite having brighter results. In Fig. 2I(a-b) the dual and relighting are consistent through thick glass refraction. In Fig. 2II(a) the shadow of the water droplets reveals a large gap that is reproduced in the dual in Fig. 2(b). Other effects include radial caustics in Fig. 2III and dense fog relighting in Fig. 2IV produced by dry ice in colored liquid.

The quality of our dual images also shows the utility of weighting light transport columns rather than hole filling in the image domain. The dual quality shows the dense capture of projector resolution, but we also encourage the reader to zoom in onto the water droplet and glass caustics, where colorful effects show the dense resolution of the illumination. This is in contrast to results by Debevec [75], where such caustics are not shown for dynamic scenes. In Fig. 2I-II(b) in particular, each streak seems to be illuminated by a flying spot, resulting in visually appealing results.



Fig. 2. **Real Line-of-sight results.** Here we show floodlit, dual and relighting results for four dynamic scenes. Please see the supplementary video. In (I) we show a wax candle placed in a liquid-filled cup. In (II) we show water droplets falling off fingers. In (III) we show a toy being rotated, and in (IV) we show fog-like vapors from dry-ice placed in a cup containing colored water. Note the caustics in the relighting, and the shadows/specularities maintained in the dual views.

Comparisons to Henderson et al. [24]: Our hardware design has improved the brightness in the light source as well as better overall SNR in the capture of the dot images. The improved design uses a brighter laser which achieves 625 Lux, while Henderson et al. used 14.51 Lux [24]. Henderson et al. only had a dot size of 3.6mm, while our laser generates a much more parallel light ray to create smaller dot (0.9mm in diameter @ 1m), making the angular resolution as small as 0.103° . Computing the SNR for a single dot image for comparison, our system achieves 51dB, which is more than twice of Henderson et al.’s 21dB. This also explains why extensive denoising algorithms were necessary in their system to achieve suitable visual results.

Fig. 3 shows qualitative comparisons of our approach to that in Henderson et al. [24]. Note that these are not exact comparisons, since the scenes are different and the experiment parameters could be different (such as glass thickness). We encourage the user to watch our supplementary video and compare the results over the entire scene. Given these caveats, we note that the use of a bright, tri-color laser provides natural caustics that do not suffer from the hard boundaries, created by complex denoising in their work [24]. Further, when comparing the fog scenes, our method provides denser sampling without dark/bright stripes (due to smaller angular extent of the laser) and a brighter, more realistic vapor cloud image, showing dense and wispy regions of fog.

NLoS results. Unlike most NLoS efforts [45–51], we work on the simpler dual-NLoS problem,

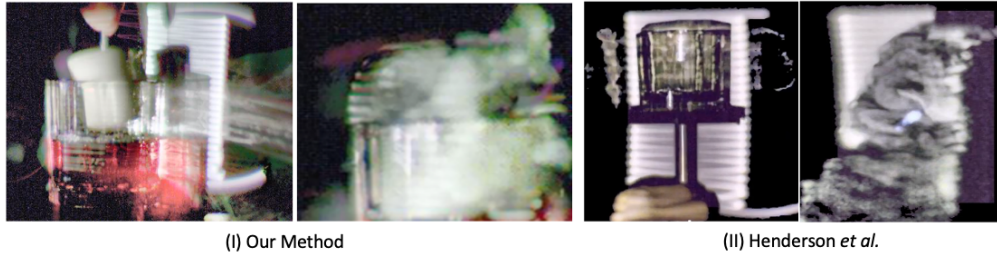


Fig. 3. **Qualitative Comparison to [24].** Here we show qualitative comparisons between our method, and that of [24], for scenes with caustics and specular interreflections. Note that our results occurs after a simple thresholding (three numbers for RGB), whereas the [24] is the output after a complex denoising process. Note that our caustics look bright and natural due to the use of a laser source. The choice of LED and denoising in [24] results in masked caustics.

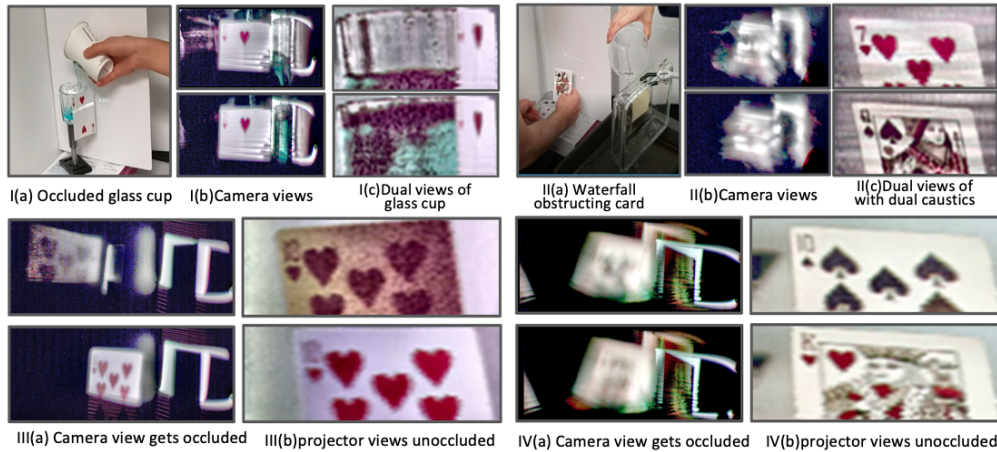


Fig. 4. **Real Dual NLoS results.** (I) Pouring blue liquid in I(a) is shown with camera and dual views I(a-b). In II the camera views the scene through falling water II(a-b), and the card textures are unscrambled with caustics caused by a virtual source behind the waterfall II(c). III and IV show more dual scenes, and IV shows a view through thick refractive glass that causes distortions. Please see supplementary video.

where the light-source directly illuminates the scene, but the camera is occluded. While previous work has enabled video reconstructions [58, 59] for retroreflective scenes, solving dual-NLoS allows color, video, high-resolution NLoS imaging for the first time. In Fig. 4I-I(a), we show one frame in ambient illumination, for the reader’s benefit. The rest of the experiment occurs in darkness to maximize the SNR of the measured data. Every scene in the experiments is placed between two non-glossy Lambertian planes (i.e. a diffuse corner).

Fig. 4(I) shows an experiment where a blue liquid is poured into an occluded glass. Note that in the camera view, the glass is occluded by the playing card. Refractive effects due the glass and liquid are visible, but the glass itself is only visible in the dual view, where the edge of the glass and the bubbles in the poured liquid are clearly visible.

Fig. 4(II) shows a stack of playing cards being revealed, one-by-one, by dropping the topmost card. The entire scene is viewed through a transparent glass with water flowing over it. The flowing water distorts the camera view, but since the projector directly illuminates the scene, the light-transport captures all these effects, which can be inverted to produce the cards. Note that

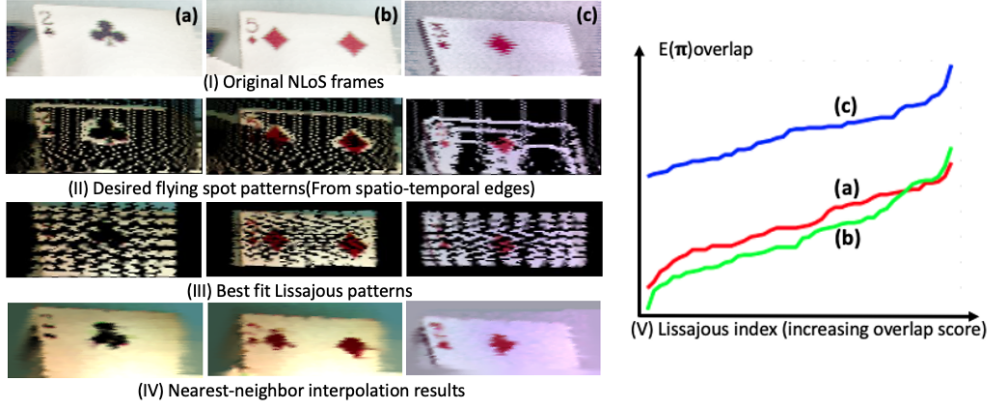


Fig. 5. **Simulated Lissajous subsampling on real NLoS data.** Here we show Lissajous simulations validating our Proposition 1. Using the ground-truth transport (I), Lissajous patterns with fit to spatio-temporal edges (II). The best fit (III) was used in nearest neighbor interpolation (IV) and was found by grid-search over the Lissajous parameters, visualized in increasing order of overlap (V).

cards appear illuminated through water, creating dual caustics.

Fig. 4(III) shows an experiment where a playing card is being rotated in and out of view. When the playing card is edgewise to the camera, only diffuse reflections from its surface are visible. However, since the projector directly illuminates the camera, the dual view always sees the playing card clearly. Finally, Fig. 4(IV) shows cards viewed through thick, refractive glass. The geometry, refractive index, and location of the glass pieces are unknown, but the projector-view image can also be obtained by inverting the effects that the light-transport captured.

6.1. Light transport Lissajous simulations

In Sect. 4 we discussed the selection of a Lissajous pattern, given a desired sampling pattern. Here we show simulations on the NLoS scenes that we discussed earlier. We use a target illumination pattern M that prefers intensity edges in the projector space, i.e. either spatial or spatio-temporal edges of the *dual image*. Intensity edges in the dual space tell us how differential changes in the flying spot at some location influenced the camera image. Large changes (i.e. strong dual edges) are places where we need dense illumination sampling, and a lack of edges (i.e. uniform intensities in the dual image) are places which could be approximated by, say, nearest-neighbor interpolation.

Dual edge detection involves thresholding the gradients of spatio-temporal dual video. This creates a target pattern $M(x, y)$ where the edge strength is over some threshold e_{dual} . We note that the desired edges can only be calculated knowing the full light-transport T to form the dual image. We solve this problem by using previous frames as proxies for the current light transport and diffusing edges to compensate for motion.

To perform this algorithm, we first find 3 consecutive dual images (I_1, I_2, I_3) , and store them in grayscale as a 3D matrix $G(x, y, t)$. We then compute directional gradients of this 3D matrix, denoted G_x, G_y, G_t . We compute the maximum of G_x, G_y, G_t along the time dimension, and then compute the magnitude of the 3D gradient image $(I = \sqrt{G_x^2 + G_y^2 + G_t^2})$. We perform edge detection on this gradient image, and then optimize different Lissajous patterns to this image until selecting the one with the best overlap score. To determine the subsampling which corresponds to this pattern, we utilize our homography between projector and camera to map each non-zero

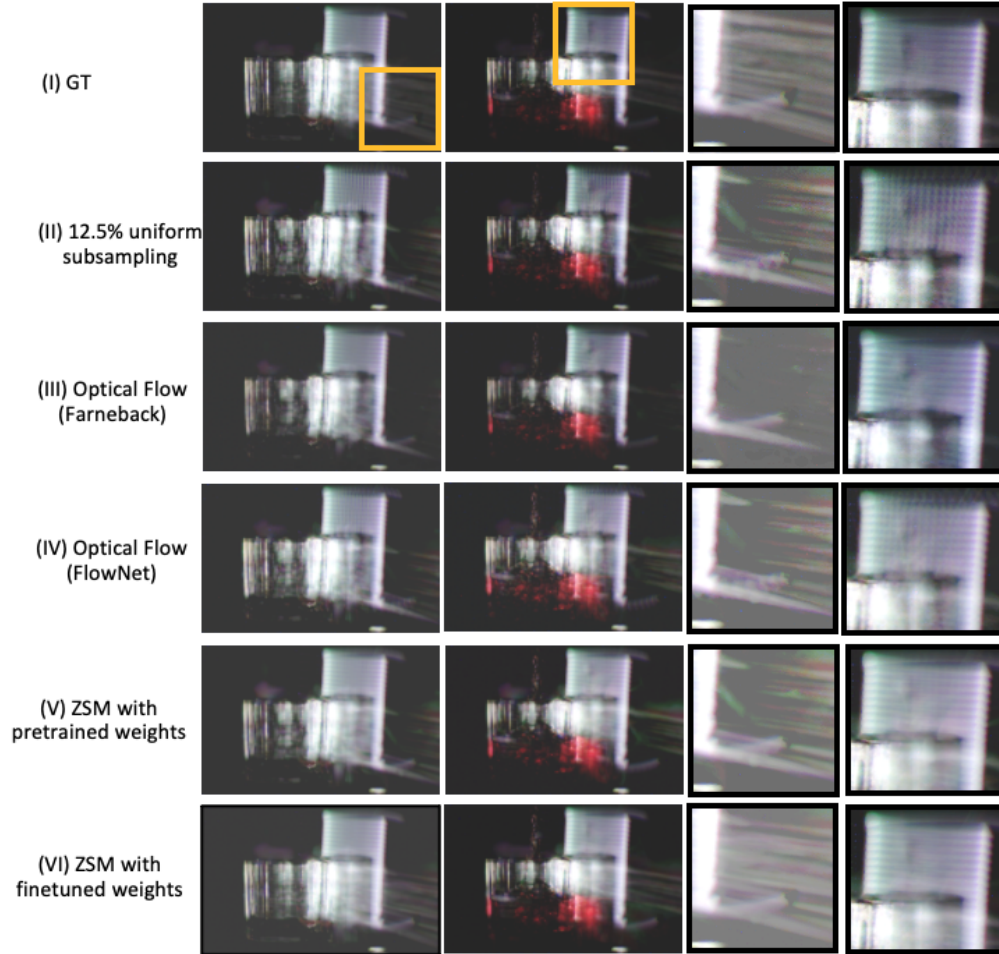


Fig. 6. **Simulated uniform subsampling for real LoS data.** Here we show results on fine-tuning a state-of-the art video interpolation method ZSM [26] for flying spot images. The ground-truth (GT) in (I) is sub-sampled by 12.5% (II). Compared to optical flow methods (III-IV) and the pretrained model, fine-tuning with our loss heuristics produces better results in the caustic insets. We used a third frame, not shown, for fine-tuning.

Lissajous pixel in that optimized pattern to their corresponding dot image indices (or equivalently columns of the light transport matrix).

We show our Lissajous fit simulations in Fig. 5. To find the resonant frequencies ω_{h0}, ω_{v0} we fit a dense Lissajous pattern to the 100×160 projector resolution. Since Proposition 1 guarantees a good fit, multiple optimization techniques are possible. The fact that the ratios k_h, k_v are bounded $[0, 1]$ and the phase difference $\phi \in [0, 2\pi]$ allowed us to perform a grid search to find the best fit. Warping the edges detected using light-field-based optical flow could give better results [76] and remains an avenue of future work.

Once we compute the Lissajous patterns in Fig. 5(III), we then use our homography to determine what columns from the light transport should be subsampled. Our subsampling performs well with simple nearest-neighbor interpolation as shown in (IV) with most of the details of the dual-frame being recovered.

6.2. Light transport interpolation results

Implementation details. Our network architecture has the same layers and parameters as in Xiang et al. [26] but without the last upsampling layer. We train on a 2080Ti GPU with a learning rate $\lambda = 5e - 4$ using the ADAM optimizer. The network takes approximately 12 hours to converge, and at test time, interpolation takes around 1-2 hours per light transport matrix T for a given frame, utilizing the binary search tree procedure described earlier.

Main results. We conduct two types of experiments for interpolation: (1) simulated subsampling post-capture for real data to enable comparison to ground truth at a higher sampling rate (Figs. 6 and 7), and (2) real Lissajous subsampling performed with our hardware setup without ground truth (Fig. 8). The former experiments allow us to compare different subsampling strategies and quantitatively evaluate performance, while the latter shows real experimental proof-of-concept on our current hardware setup. Since MEMS mirrors need to operate at resonance for high speed operation, Lissajous subsampling is used throughout the paper as it is most physically feasible for high-speed patterns with the current hardware setup. However, we do show a qualitative comparison of simulated subsampling methods such as uniform and random subsampling in the Supplemental Material (Fig. S5).

In Fig. 6(I) we show ground-truth (GT) of two frames in a LoS floodlit video, and in Figure 6(II) we show the results for simulated subsampling of the light transport columns uniformly by a factor of 16. Caustic effects in the inset (third row) are affected and the Lambertian plane shows Moire effects. In Fig. 6(II-IV) we show interpolation with conventional optical flow methods [71, 72]. The third row insets show that Farneback and FlowNet2.0 perform badly on caustics, probably due to generalization error since flying spot images are qualitatively different from what FlowNet2.0 is trained on.

Using the pretrained network from Xiang et al. [26] as an off-the-shelf frame interpolation method, we show the results of subsampling the light transport columns uniformly by a factor of 8 in Fig. 6(V). Note how at 12.5% subsampling, the interpolation recovers the direct light transport columns particularly on the back Lambertian plane, but is missing several global light effects such as the caustics of the glass. As described in Sect. 5, we performed finetuning with an MSE and perceptual loss on a prior frame of a scene’s dynamic light transport to enhance performance. The prior frame was immediately preceding the frame in the first row of Fig. 6, but more than 10 frames behind the second row frame. As shown in Fig. 6(VI), the network interpolates global lighting effects such as interreflections and caustics at high visual quality.

In Figure 7, we simulate a hardware-based Lissajous sampling for our captured data and subsequent interpolation using the pre-trained network. As we can see, the Lambertian plane and glass caustics are recovered by the network. This is an important example because the scanning pattern is physically realizable in the hardware setup.

Finally, in Fig. 8, we show a real scene captured with a real Lissajous pattern. Due to the

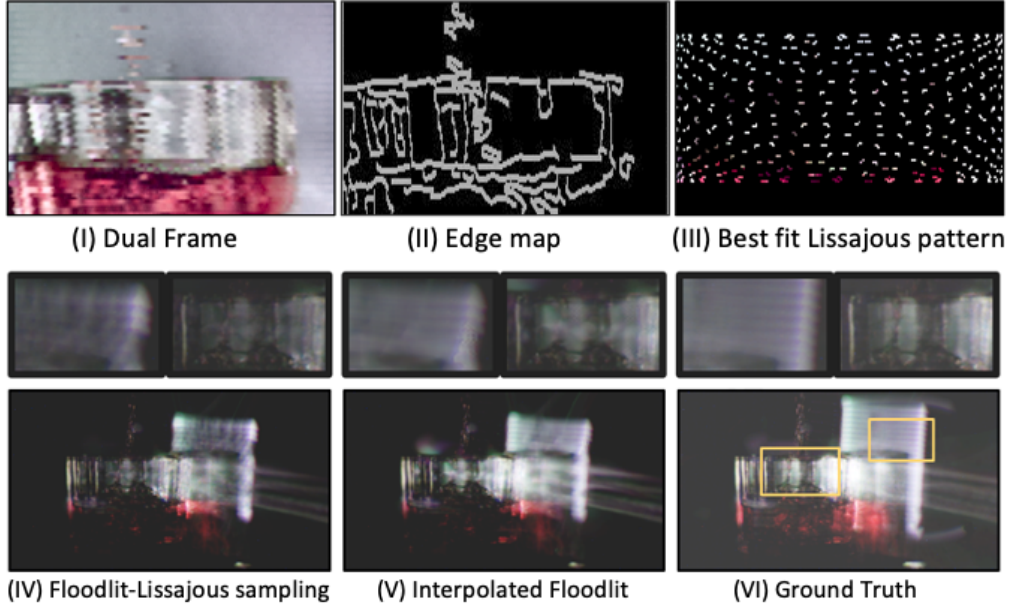


Fig. 7. **Simulated Lissajous sampling on real LoS data.** Here we show results for LOS interpolation for Lissajous pattern sampling. The dual image (I) of the previous frame is used to detect the edge map in (II) used to simulate the best fit indexes for a lissajous pattern simulation (III). The sampled indexes in the floodlit image (IV) show lost dots along the Lambertian plane and reflections in the glass. The network interpolation (V) is able to recover most of the Lambertian plane and partially recover reflections in the glass compared to the ground truth floodlit image (VI)

hardware limitations of our setup, we only used a single pattern for the entire capture. The gaps in the light-transport are never captured, and we enable **24 FPS** (i.e. near real-time) capture of the scene. This is an improvement over the 5-7 FPS results we get for all other real scenes in this paper as well as the 6 FPS of [24]. In the figure, we show interpolation using the pretrained network from [26], where the missing light transport information is interpolated. Please see the supplementary video for more of this interpolation result.

6.3. Quantitative evaluation of transport interpolation

In Table 1, we evaluate the interpolation using the light-transport for the scenes described qualitatively in the previous section. For each method, we randomly selected 100 pairs of columns in the light transport matrix, and interpolated a column in between this pair. We compared the interpolated light transport column to the ground truth using both PSNR and SSIM metrics.

Note that the fine tuning heuristics, represented by ZSM-FT in the table, outperform traditional optical flow methods by 3-4 dB in PSNR. Furthermore, the fine-tuning does better than both network-based optical flow (Flownet2) and the original pretrained weights of ZSM.

Finally, in Fig. 9, we show an ablation study that concretely supports our hypothesis that sparse flying-spot photography, improving the ratio in Eq. 3.2, is the right choice over parallel projection, given good interpolation.

In the figure, we conduct two experiments on the light-transport used in the qualitative results in the previous section. The first uses the PSNR metric and second uses the SSIM metric. In each experiment, we follow the change in quality (compared to ground truth) of the interpolation in two scenarios. The first scenario is the interval, or distance between the interpolated column and

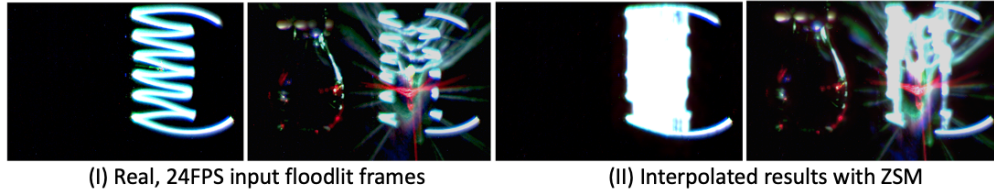


Fig. 8. **Real Subsampling Experiment.** In (I) we show a difficult scene captured with real Lissajous pattern sub-sampling interpolation, i.e. light-transport gaps are never captured, that has a frame rate of 24 FPS. In (II) we show recovery of both the lambertian back-plane and a frame with the glass with red liquid.

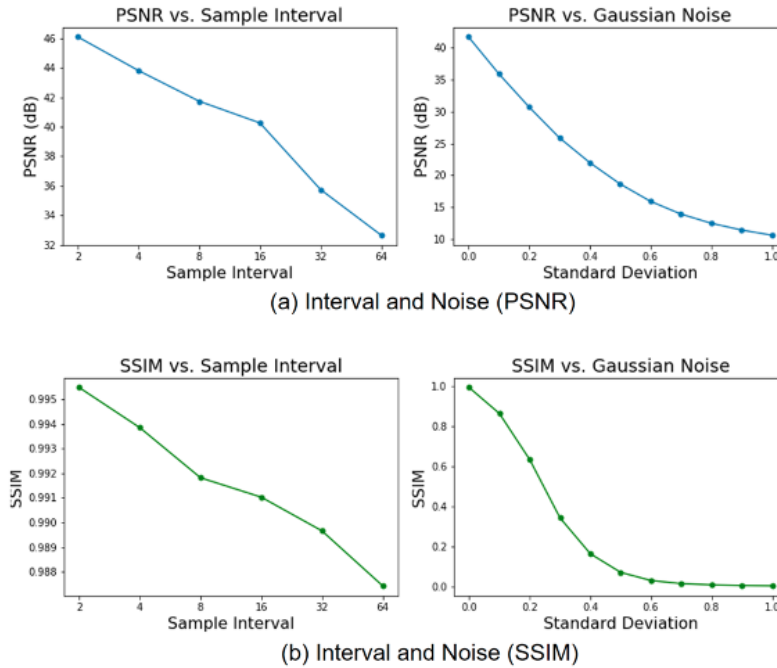


Fig. 9. **Sampling Interval and Noise Studies.** Here we show (a) PSNR and (b) SSIM evaluations of how the interpolation breaks down across two factors. The first is interval, or the angle between two flying spots (or equivalently distance between column indices on a single row) whose midpoint is interpolated. The second is noise, as a proxy for illumination brightness, where the intensity of the flying spot is reduced, decreasing SNR. *The key point here is that the neural network interpolation falls linearly with interval compared to noise which falls non-linearly.* In other words, for the illumination power budget, focusing energy into a flying-spot is preferred to parallel illumination, as long as neural network interpolation provides good accuracy.

its inputs. The second is image noise, which acts as proxy for illumination power, and increased noise means the light-source (i.e. our tricolor laser) is reduced in power.

Each experiment shows the same pattern. The fall-off in quality when the distance between the interpolated column and its input pair is near-linear, where the fall-off in quality due to reduced illumination energy (modeled as noise) falls off non-linearly.

In other words, flying-spot photography is successful on two fronts First, it improves the ratio (Eq. 3.2) in being fast and reducing the number of captured samples. Second, using the state-of-the-art neural network interpolation, the large gaps between sparse flying-spot patterns have less error than smaller gaps for a dimmer light source, such as a parallel DLP or LCD projector of the same wattage as the tri-color laser.

Table 1. Quantitative comparison of interpolation methods.

	Farneback	FlowNet2	ZSM	ZSM-FT
PSNR (dB)	36.68	37.48	40.17	41.73
SSIM	0.9883	0.8792	0.9910	0.9938

7. Conclusion and limitations

We present a new hardware setup for flying spot light transport capture which utilizes a tri-color laser. We describe calibration and implementation details and demonstrate numerous light transport applications, including color dual-NLoS videos for the first time in the literature. We then show how adaptive control of the MEMS mirror and scanning pattern can yield even more time speedups for scanning. Finally, we show missing light transport columns can be interpolated to achieve speedup without sacrificing visual quality.

For future work, we are building a new system allowing real-time Lissajous adaptive control and combining sampling/interpolation strategies in hardware for end-to-end learning. This paper is the first step in a promising new direction for adaptive flying-spot scanners in computer vision and graphics. The tricolor laser is bright and focused when compared to the LED light-source in [24]. While each laser individually is safe, the overall source is not eye safe. In the future, we hope to make it eye safer by pulse width modulation over wavelength. The strange color artifacts, as seen in the water droplet example, will need to be fixed as well. While we captured the sparse flying spot images in full resolution, the efficiency could be improved by applying adaptive camera capturing strategies.

Funding.

Acknowledgments. The authors would like to thank the anonymous reviewers and the editors for their helpful feedback. The authors would also like to thank members of the FOCUS lab at UFL and at the Imaging Lyceum at ASU for productive discussions about the work.

Disclosures. The authors declare no conflicts of interest.

Data Availability Statement. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile, "Modeling the interaction of light between diffuse surfaces," in *ACM SIGGRAPH Computer Graphics*, vol. 18 (ACM, 1984), pp. 213–222.
2. J. T. Kajiya, "The rendering equation," in *ACM Siggraph Computer Graphics*, vol. 20 (ACM, 1986), pp. 143–150.

3. P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *ACM SIGGRAPH*, (2000), pp. 145–156.
4. R. Ng, R. Ramamoorthi, and P. Hanrahan, "All-frequency shadows using non-linear wavelet lighting approximation," in *ACM Transactions on Graphics (TOG)*, vol. 22 (ACM, 2003), pp. 376–381.
5. P. Peers, D. K. Mahajan, B. Lamond, A. Ghosh, W. Matusik, R. Ramamoorthi, and P. Debevec, "Compressive light transport sensing," *ACM Transactions on Graph. (TOG)* **28**, 3 (2009).
6. P. Sen and S. Darabi, "Compressive dual photography," in *Computer Graphics Forum*, vol. 28 (Wiley Online Library, 2009), pp. 609–618.
7. P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. Lensch, "Dual photography," in *ACM Transactions on Graphics (TOG)*, vol. 24 (ACM, 2005), pp. 745–755.
8. G. Garg, E.-V. Talvala, M. Levoy, and H. P. Lensch, "Symmetric photography: exploiting data-sparseness in reflectance fields," in *Proceedings of the 17th Eurographics Conference on Rendering Techniques*, (Eurographics Association, 2006), pp. 251–262.
9. J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo, "Kernel nystrom method for light transport," *ACM Trans. Graph.* **28**, 29:1–29:10 (2009).
10. M. O'Toole and K. N. Kutulakos, "Optical computing for fast light transport analysis," in *ACM Transactions on Graphics (TOG)*, vol. 29 (ACM, 2010), p. 164.
11. S. M. Seitz, Y. Matsushita, and K. N. Kutulakos, "A theory of inverse light transport," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 2 (2005), pp. 1440–1447 Vol. 2.
12. D. Wu, A. Velten, M. O'Toole, B. Masia, A. Agrawal, Q. Dai, and R. Raskar, "Decomposing global light transport using time of flight imaging," *Int. journal computer vision* **107**, 123–138 (2014).
13. J. Haines and G. Tingley, "Live flying-spot color scanner," *Electr. Eng.* **75**, 528–533 (1956).
14. T. Hawkins, J. Cohen, C. Tchou, and P. Debevec, "Light Stage 2.0," in *SIGGRAPH Technical Sketches*, (2001), p. 217.
15. P. Debevec, A. Wenger, C. Tchou, A. Gardner, J. Waese, and T. Hawkins, "A Lighting Reproduction Approach to Live-Action Compositing," in *SIGGRAPH 2002*, (San Antonio, TX, 2002), pp. 547–556.
16. P. Einarsson, C.-F. Chabert, A. Jones, W.-C. Ma, B. Lamond, T. Hawkins, M. Bolas, S. Sylwan, and P. Debevec, "Relighting human locomotion with flowed reflectance fields," in *Proceedings of the 17th Eurographics conference on Rendering Techniques*, (Eurographics Association, 2006), pp. 183–194.
17. M. O'Toole, R. Raskar, and K. N. Kutulakos, "Primal-dual coding to probe light transport." *ACM Trans. Graph.* **31**, 39–1 (2012).
18. M. O'Toole, J. Mather, and K. N. Kutulakos, "3d shape and indirect appearance by structured light transport," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2014), pp. 3246–3253.
19. M. O'Toole, F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos, "Temporal frequency probing for 5d transient analysis of global light transport," *ACM Transactions on Graph. (ToG)* **33**, 87 (2014).
20. M. O'Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos, "Homogeneous codes for energy-efficient illumination and imaging," *ACM Transactions on Graph. (TOG)* **34**, 35 (2015).
21. H.-C. Park, Y.-H. Seo, and K.-H. Jeong, "Lissajous fiber scanning for forward viewing optical endomicroscopy using asymmetric stiffness modulation," *Opt. express* **22**, 5818–5825 (2014).
22. Z.-I. Song, S. Li, and T. F. George, "Remote sensing image registration approach based on a retrofitted sift algorithm and lissajous-curve trajectories," *Opt. Express* **18**, 513–522 (2010).
23. Q. A. Tanguy, O. GaiFFE, N. Passilly, J.-M. Cote, G. Cabodevila, S. Bargiel, P. Lutz, H. Xie, and C. Gorecki, "Real-time lissajous imaging with a low-voltage 2-axis mems scanner based on electrothermal actuation," *Opt. express* **28**, 8512–8527 (2020).
24. K. Henderson, X. Liu, J. Folden, B. Tilmon, S. Jayasuriya, and S. Koppal, "Design and calibration of a fast flying-dot projector for dynamic light transport acquisition," *IEEE Transactions on Comput. Imaging* **6**, 529–543 (2020).
25. T. Hawkins, P. Einarsson, and P. E. Debevec, "A dual light stage." *Render. Tech.* **5**, 91–98 (2005).
26. X. Xiang, Y. Tian, Y. Zhang, Y. Fu, J. P. Allebach, and C. Xu, "Zooming slow-mo: Fast and accurate one-stage space-time video super-resolution," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2020), pp. 3370–3379.
27. A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec, "Performance relighting and reflectance transformation with time-multiplexed illumination," in *ACM Transactions on Graphics (TOG)*, vol. 24 (ACM, 2005), pp. 756–764.
28. T. Hawkins, P. Einarsson, and P. Debevec, "Acquisition of time-varying participating media," *ACM Transactions on Graph. (ToG)* **24**, 812–815 (2005).
29. A. Kasturi, V. Milanović, D. Lovell, F. Hu, D. Ho, Y. Su, and L. Ristic, "Comparison of mems mirror lidar architectures," in *MOEMS and Miniaturized Systems XIX*, vol. 11293 (International Society for Optics and Photonics, 2020), p. 112930B.
30. P. Peers, T. Hawkins, and P. Debevec, "A Reflective Light Stage." ICT Technical Report ICT TR 04 2006, University of Southern California Institute for Creative Technologies (2006).
31. M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "Flexible voxels for motion-aware videography," in *European Conference on Computer Vision*, (Springer, 2010), pp. 100–114.
32. Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar, "Video from a single coded exposure photograph using a learned over-complete dictionary," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, (IEEE,

- 2011), pp. 287–294.
33. A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec, “Rendering for an interactive 360 light field display,” *ACM Transactions on Graph. (TOG)* **26**, 40 (2007).
 34. S. J. Koppal, S. Yamazaki, and S. G. Narasimhan, “Exploiting dlp illumination dithering for reconstruction and photography of high-speed scenes,” *Int. J. Comput. Vis.* **96**, 125–144 (2012).
 35. R. Hoskinson, B. Stoerber, W. Heidrich, and S. Fels, “Light reallocation for high contrast projection using an analog micromirror array,” *ACM Transactions on Graph. (TOG)* **29**, 165 (2010).
 36. R. Hoskinson and B. Stoerber, “High-dynamic range image projection using an auxiliary mems mirror array,” *Opt. Express* **16**, 7361–7368 (2008).
 37. H. Kubo, S. Jayasuriya, T. Iwaguchi, T. Funatomi, Y. Mukaigawa, and S. G. Narasimhan, “Acquiring and characterizing plane-to-ray indirect light transport,” in *Computational Photography (ICCP), 2018 IEEE International Conference on*, (IEEE, 2018), pp. 1–10.
 38. H. Kubo, S. Jayasuriya, T. Iwaguchi, T. Funatomi, Y. Mukaigawa, and S. G. Narasimhan, “Programmable non-epipolar indirect light transport: Capture and analysis,” *IEEE Transactions on Vis. Comput. Graph.* pp. 1–1 (2019).
 39. T. Ueda, H. Kubo, S. Jayasuriya, T. Funatomi, and Y. Mukaigawa, “Slope disparity gating using a synchronized projector-camera system,” *IEEE Int. Conf. on Comput. Photogr. (ICCP)* (2019).
 40. H. Zhou, S. Hadap, K. Sunkavalli, and D. W. Jacobs, “Deep single-image portrait relighting,” in *Proceedings of the IEEE International Conference on Computer Vision*, (2019), pp. 7194–7202.
 41. D. Shahlaei, M. Piotraschke, and V. Blanz, “Lighting design for portraits with a virtual light stage,” in *2016 IEEE International Conference on Image Processing (ICIP)*, (IEEE, 2016), pp. 1579–1583.
 42. T. Nestmeyer, J.-F. Lalonde, I. Matthews, and A. Lehrmann, “Learning physics-guided face relighting under directional light,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2020), pp. 5124–5133.
 43. B. Mildenhall, P. P. Srinivasan, R. Ortiz-Cayon, N. K. Kalantari, R. Ramamoorthi, R. Ng, and A. Kar, “Local light field fusion: Practical view synthesis with prescriptive sampling guidelines,” *ACM Transactions on Graph. (TOG)* **38**, 1–14 (2019).
 44. S. Nie, L. Gu, A. Subpa-Asa, I. Kacher, K. Nishino, and I. Sato, “A data-driven approach for direct and global component separation from a single image,” in *Asian Conference on Computer Vision*, (Springer, 2018), pp. 133–148.
 45. M. Buttafava, J. Zeman, A. Tosi, K. Eliceiri, and A. Velten, “Non-line-of-sight imaging using a time-gated single photon avalanche diode,” *Opt. Express* **23**, 20997–21011 (2015).
 46. G. Satat, M. Tancik, and R. Raskar, “Towards photography through realistic fog,” in *2018 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2018), pp. 1–10.
 47. M. O’Toole, D. B. Lindell, and G. Wetzstein, “Confocal non-line-of-sight imaging based on the light-cone transform,” *Nature* **555**, 338 (2018).
 48. F. Heide, M. O’Toole, K. Zang, D. B. Lindell, S. Diamond, and G. Wetzstein, “Non-line-of-sight imaging with partial occluders and surface normals,” *ACM Transactions on Graph. (ToG)* **38**, 1–10 (2019).
 49. M. O’Toole, D. B. Lindell, and G. Wetzstein, “Real-time non-line-of-sight imaging,” in *ACM SIGGRAPH 2018 Emerging Technologies*, (2018), pp. 1–2.
 50. A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, “Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging,” *Nat. Commun.* **3**, 1–8 (2012).
 51. F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich, “Low-budget transient imaging using photonic mixer devices,” *ACM Transactions on Graph. (ToG)* **32**, 1–10 (2013).
 52. X. Liu, I. Guillén, M. La Manna, J. H. Nam, S. A. Reza, T. H. Le, A. Jarabo, D. Gutierrez, and A. Velten, “Non-line-of-sight imaging using phasor-field virtual wave optics,” *Nature* **572**, 620–623 (2019).
 53. M. V. Klibanov, J. Su, N. Pantong, H. Shan, and H. Liu, “A globally convergent numerical method for an inverse elliptic problem of optical tomography,” *Appl. Analysis* **89**, 861–891 (2010).
 54. M. O’Toole, D. B. Lindell, and G. Wetzstein, “Confocal Non-Line-of-Sight Imaging Based on the Light-Cone Transform,” *Nature* (2018).
 55. S. I. Young, D. B. Lindell, B. Girod, D. Taubman, and G. Wetzstein, “Non-line-of-sight surface reconstruction using the directional light-cone transform,” in *Proc. CVPR*, (2020).
 56. 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 Conference on Computer Vision and Pattern Recognition*, (2019), pp. 6800–6809.
 57. K. L. Bouman, V. Ye, A. B. Yedidia, F. Durand, G. W. Wornell, A. Torralba, and W. T. Freeman, “Turning corners into cameras: Principles and methods,” in *Proceedings of the IEEE International Conference on Computer Vision*, (2017), pp. 2270–2278.
 58. M. Isogawa, Y. Yuan, M. O’Toole, and K. M. Kitani, “Optical non-line-of-sight physics-based 3d human pose estimation,” in *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2020), pp. 7013–7022.
 59. M. Isogawa, D. Chan, Y. Yuan, K. M. Kitani, and M. O’Toole, “Efficient non-line-of-sight imaging from transient sinograms,” in *16th European Conference on Computer Vision (ECCV)*, (2020).
 60. T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman, “Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction,” *Int. J. Comput. Vis.* **49**, 215–227 (2002).

61. M. Holroyd, J. Lawrence, and T. Zickler, "A coaxial optical scanner for synchronous acquisition of 3d geometry and surface reflectance," in *ACM Transactions on Graphics (TOG)*, vol. 29 (ACM, 2010), p. 99.
62. S. J. Koppal and S. G. Narasimhan, "Beyond perspective dual photography with illumination masks," *IEEE Transactions on Image Process.* **24**, 2083–2097 (2015).
63. M. Aittala, P. Sharma, L. Murmann, A. Yedidia, G. Wornell, B. Freeman, and F. Durand, "Computational mirrors: Blind inverse light transport by deep matrix factorization," in *Advances in Neural Information Processing Systems*, (2019), pp. 14311–14321.
64. W. Chen, S. Daneau, F. Mannan, and F. Heide, "Steady-state non-line-of-sight imaging," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2019), pp. 6790–6799.
65. G. Musarra, A. Lyons, E. Conca, Y. Altmann, F. Villa, F. Zappa, M. J. Padgett, and D. Faccio, "Non-line-of-sight three-dimensional imaging with a single-pixel camera," *Phys. Rev. Appl.* **12**, 011002 (2019).
66. J. H. Nam, E. Brandt, S. Bauer, X. Liu, E. Sifakis, and A. Velten, "Real-time non-line-of-sight imaging of dynamic scenes," arXiv preprint arXiv:2010.12737 (2020).
67. J. L. Baird, "Television apparatus. u.s. patent 2006124a," (1935).
68. S. M. Seitz, Y. Matsushita, and K. N. Kutulakos, "A theory of inverse light transport," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2 (IEEE, 2005), pp. 1440–1447.
69. S. W. Hasinoff, F. Durand, and W. T. Freeman, "Noise-optimal capture for high dynamic range photography," *CVPR* (2010).
70. Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Transactions on Pattern Analysis Mach. Intell.* **29**, 1339–1354 (2007).
71. G. Farneback, "Two-frame motion estimation based on polynomial expansion," in *Scandinavian Conference on Image analysis*, (Springer, 2003), pp. 363–370.
72. E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "FlowNet 2.0: Evolution of optical flow estimation with deep networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2017), pp. 2462–2470.
73. J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European Conference on Computer Vision*, (Springer, 2016), pp. 694–711.
74. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556 (2014).
75. P. Debevec, "The light stages and their applications to photoreal digital actors," *SIGGRAPH Asia* **2** (2012).
76. S. Ma, B. M. Smith, and M. Gupta, "Differential scene flow from light field gradients," *Int. J. Comput. Vis.* **128**, 679–697 (2020).