

Design and Calibration of a Fast Flying-Dot Projector for Dynamic Light Transport Acquisition

Kristofer Henderson, Xiaomeng Liu, Justin Folden, Brevin Tilmon, Suren Jayasuriya , *Member, IEEE*,
and Sanjeev Koppal , *Senior Member, IEEE*

Abstract—The light transport captures a scene’s visual complexity. Acquiring light transport for dynamic scenes is difficult, since any change in viewpoint, materials, illumination or geometry also varies the transport. One strategy to capture dynamic light transport is to use a fast “flying-dot” projector; i.e., where an impulse light-probe is quickly scanned across the scene. We have built a novel fast flying-dot projector prototype using a high speed camera and a scanning MEMS (Micro-electro-mechanical system) mirror. Our contributions are calibration strategies that enable dynamic light transport acquisition at near video rates with such a system. We develop new methods for overcoming the effects of MEMS mirror resonance. We utilize new algorithms for denoising impulse scanning at high frame rates and compare the trade-offs in visual quality between frame rate and illumination power. Finally, we show the utility of our calibrated setup by demonstrating graphics applications such as video relighting, direct/global separation, and dual videography for dynamic scenes such as fog, water, and glass. Please see our accompanying video for dynamic scene results.

Index Terms—Computer vision, robotics and automation, smart cameras.

I. INTRODUCTION

IMAGE-BASED rendering aims to directly synthesize new images from real scene measurements. To enable realistic results, these scene measurements must capture the subtle, rich visual back-story that occurs as light goes through several interactions between the source and the viewer. For example, effects due to complex geometries and materials, including

diffuse/specular reflection and subsurface scattering must be captured. Interactions between objects and immersion media can cause optical scattering. Additionally, inter-reflections between nearby objects can cause color-bleeding, often in the form of dynamic caustics.

The primary vehicle in which to study and exploit these effects is the scene’s light transport. Light transport gives the linear relationship between input illumination and captured images in a scene. Researchers have investigated many aspects of transport over the past several decades. For static scenes, these have resulted in full light-transport capture [1], [2], including a galvanometer-modulated laser scan for complex, static scenes [3]. Other work has enabled specific decompositions such as direct/global [4] and bounce separation [5].

Unfortunately, for dynamic moving scenes, light transport must be captured at each moment in time. This challenge has been tackled by using large, expensive light stages with high speed cameras and synchronized flash illumination [6]–[9] for relighting of performance capture. Additionally, cameras that exploit epipolar geometry and MEMS (Micro-electro-mechanical systems) resonance scanning have allowed for slices of the transport to be directly captured [10]–[12].

Our Approach: We consider a light-weight setup for desktop-based light-transport capture. We wish to acquire the capabilities of the light-stage while exploiting the compactness of MEMS mirror based approaches. Our approach involves illuminating a small solid angle (i.e., a “flying dot” [13]) and quickly scanning this dot over the whole scene. This paper describes our prototype leveraging recent advances in MEMS scanning mirrors and high-speed cameras along with calibration strategies for such a system. We describe the trade-offs between this approach and spatial light modulators in terms of energy and time efficiency. We develop several algorithms for this prototype including calibration and denoising. Finally, we show applications for computer graphics using this setup including visualizing physically-realistic relighting for moving complex scenes including specular, transparent materials (glass) and scattering media (fog).

Our Specific Contributions Are:

- A system implementation of the fast flying-dot idea using a scanning MEMS mirror and high speed camera.
- A characterization of the trade-off between energy and time efficiency for flying dot projectors.
- A denoising algorithm for the low-light challenges that are fundamental to fast flying-dot designs.

Manuscript received April 25, 2019; revised September 30, 2019 and December 16, 2019; accepted December 22, 2019. Date of publication January 6, 2020; date of current version January 21, 2020. The University of Florida authors have been partially supported by the National Science Foundation through NSF IIS: 1514154 and NSF IIS: 1909729, and the Office of Naval Research through ONR N00014-18-1-2663. S. Jayasuriya has been supported by the National Science Foundation through NSF IIS: 1909192, as well as joint support from both the Herberger Research Initiative in the Herberger Institute for Design and the Arts (HIDA) and the Fulton Schools of Engineering (FSE) at Arizona State University. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. D. Gutierrez. (*Corresponding author: Sanjeev Koppal.*)

K. Henderson, X. Liu, J. Folden, B. Tilmon, and S. Koppal are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: khenderson5@ufl.edu; liuxiaomeng51@ufl.edu; jfolden@ufl.edu; brevinjt@gmail.com; sjkoppal@ece.ufl.edu).

S. Jayasuriya is with the School of Arts, Media and Engineering and the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85281 USA (e-mail: sjayasur@asu.edu).

This article has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the authors.

Digital Object Identifier 10.1109/TCI.2020.2964246

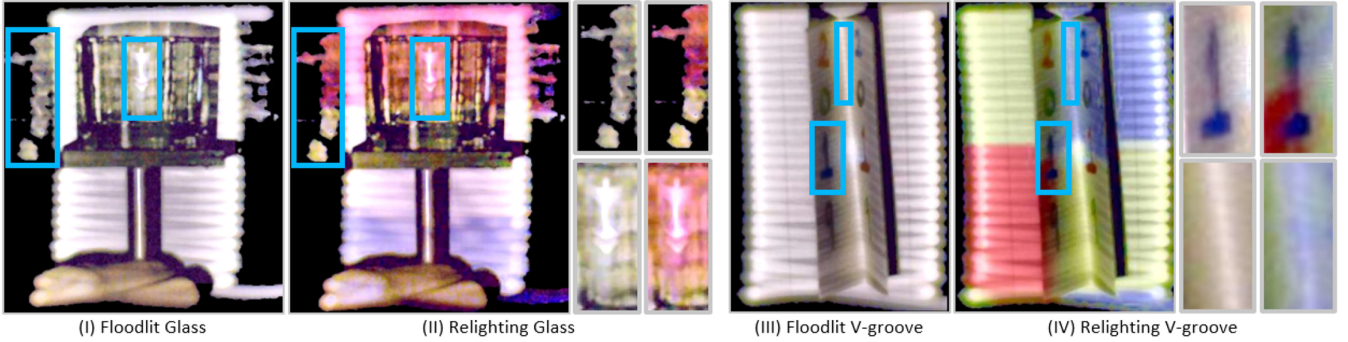


Fig. 1. Relighting dynamic scenes with a fast flying-dot projector: We present calibration and denoising algorithms that enable a fast flying-dot projector-camera system to capture light-transport. Our approach enables graphics applications such as relighting of dynamic scenes with a new, desired illumination at near video-rate. In (I–II) we show a glass scene, floodlit (I) and also illuminated with a tricolor flag (II). Complex effects, such as caustics and specularities are correctly relit (see close-ups). In (III–IV) we show a v-groove with text, floodlit (III) and also illuminated with a color checker pattern (IV). Effects such as interreflections and color-bleeding across the v-groove vertex are correctly reproduced (see close-ups).

- New calibration strategies for overcoming the effects of resonance, inherent to fast MEMS mirror scanning, particularly the issue of non-uniform angular sampling.
- Applications including video relighting, direct/global separation, and dual videography using our setup.

Organization: For readers desiring high level understanding, this introduction followed by the the first few paragraphs of Section III will describe the system broadly. For readers focused on the calibration steps, Section III analyzes the tradeoffs in time and energy efficiency for fast optical modulators in general, Section IV introduces our optical setup and calibration and Section V presents our denoising algorithm and comparisons of visual quality for frame rate and illumination power. Finally, Section VI shows applications in light transport using our prototype setup.

II. BACKGROUND AND RELATED WORK

This section provides the fundamental light-transport equations and describes where our design and calibration of a fast flying-dot project fits into the larger effort on light-transport acquisition and analysis. For a scene illuminated by a projector or other programmable light sources, the light transport between captured image $\mathbf{i}$ and illumination $\mathbf{I}$, is given by the light transport matrix $\mathbf{T}$ [6], [14]–[16]:

$$\mathbf{i} = \mathbf{T} \mathbf{I}. \quad (1)$$

Capturing light-transport matrices require large measurements and storage. To alleviate these issues, researchers have used compressive sensing to reduce acquisition time [2], [17], adaptive schemes to multiplex illumination [1], symmetry priors [18], and low rank approximations using the Nyström method [19] or optical computing [20]. Other methods have decomposed light transport into bounces of light [5] and its sub-components [21].

Light Transport Components: Nayar *et al.* [4] showed separation of direct/global illumination using high frequency patterns, which has been extended to defocused and global illumination [22], video via motion compensation [23], multiple sources [24], short/long-range indirect light [25] and

primal-dual [10]. Exploiting the projector-camera epipolar geometry can enable video-rate separation [11], [12], [26].

Dual Photography/Videography: Using Helmholtz reciprocity of light, the transposed light-transport matrix describes an optical setup when the projector and camera have been switched [1], [27]. Reciprocity has been used to reconstruct complex surface geometry [28]. Hawkins *et al.* [3] used reciprocity using a similar setup to ours, but required a static scene and a diffuse spherical shell enclosing the object. Finally, using projector-camera system aligned in a disparity-gating configuration, live dual videography was demonstrated [12].

A. Dynamic Light Transport Matrices

Now consider a dynamic scene, where either objects in the scene are moving and/or the camera and light-source are in motion and/or the illumination patterns are changing. In this scenario, the light transport matrix varies with time t :

$$\mathbf{i}(t) = \mathbf{T}(t) \mathbf{I}(t). \quad (2)$$

It is the goal of this paper to acquire this **dynamic light transport** $\mathbf{T}(t)$ for desktop scenes by proposing calibration and denoising algorithms that improve the ease and cost of capture. Our work builds on previous efforts that capture components of the light-transport using fast optical modulation, such as coded exposure [29], [30], digital micromirror devices [11], [31], [32], MEMS projectors [33], [34] and synced camera-projectors [12], [26], [35], [36].

Light Stages: Closest to our work are the light stages which capture reflectance fields for a given scene. These stages trade-off dense angular resolution over (say) a light-dome for high-quality capture in a large field-of-view. Our desktop scenes are illuminated by denser angular samples in a setup without a special dome, but require calibration strategies that we provide and are directional in their illumination as opposed to the environmental lighting of the light stage.

Various versions of the light stage have allowed relighting human faces [6], [37], and have been extended to moving actors in a scene using a LED dome [7], high-speed photometric stereo [8] and even 7D information for relighting

walking/running humans [9]. Other light stages include the dual light-stage which uses a scanning laser, similar to our setup, but is limited to static scenes in a diffuse spherical shell [3]. The reflective light stage [38] uses mirrors and reflected illumination to capture reflectance fields for highly specular objects. Acquisition by the DISCO system allows for the capture of translucent materials [39]. Our work differs from all these light stages by capturing a denser dynamic light transport for a smaller desktop scene, enabling local effects for dynamic scenes such as specular interreflections, caustics, and multiple scattering of light in media.

III. THE FLYING-DOT PROJECTOR

This section introduces the flying-dot projector, with discussion of our imaging setup and inherent scene assumptions. It also provides comparative analysis of its advantages and disadvantages to fast parallel optical modulators. To capture dynamic light transport, we implement an optical setup inspired by the “flying-spot” scanner first demonstrated by John Baird [40]. This scanner operated by shining an arc lamp through the holes of a spinning punctured disk (called a Nipkow disk), creating a bright light spot in the scene. This was measured as video by photoelectric cell arrays at 15–20 frames per second. This technique enabled early live television [13] until being replaced by cathode-ray tubes.

In computer graphics, the flying dot images are equivalent to light-transport impulse responses [1], [5]: $\mathbf{T}^i = \mathbf{T} \delta_i$, where δ_i is a vector with 1 at the i -th location, 0 elsewhere, and T^i is correspondingly the i th column. Now let us consider a dynamic scene with a time-varying light transport $\mathbf{T}(t)$. We assume that impulse images are captured by a high-speed camera without motion blur. Then we require our impulse sampling of dynamic light transport to occur within time interval Δt such that $\|\mathbf{T}(t) - \mathbf{T}(t + \Delta t)\| < \epsilon$. Thus we can safely assume that all the columns captured are from the same light-transport matrix at time t , given by $\mathbf{T}^i(t)$ for the i th column created by the i th flying dot image.

The flying-dot projector inherently trades off temporal and spatial resolution. The impulse response at the i th location, corresponding to a column of the light-transport matrix $\mathbf{T}^i(t)$ at time t is also the impulse response of the scene at time $t + \delta t$, where δt is time taken to capture the first i frames of the light-transport, where, $\delta t \leq \Delta t$. The assumption in the previous paragraph translates to an assumption that the mirror in our setup moves much faster than scene motion.

MEMS-Mirror Modulated Fast Flying-Dot Projectors: Recent advances in MEMS mirrors have enabled optical modulation in high KHz range used for mobile laser projection [41]. These have been effectively used to capture slices of the light-transport [12], [20]. We have designed and built an optical setup that combines such fast MEMS mirror modulation (>10 KHz) with a high-speed camera (>10000 FPS) which consists of a high-speed CMOS sensor with high-bit parallel A/D conversion. The optical hardware design for the MEMS-modulated light source and high-speed camera, as well as an associated suite of calibration algorithms, is detailed in Section IV. However, a variety of optical modulation technologies exist to enable

such flying-dot imaging, and we discuss the energy and time efficiency for our design choices in the following subsection.

A. Energy and Time Efficiency for the Flying-Dot

1) *Energy Efficiency:* Optical implementations of the flying-dot fall between two extremes of energy efficiency [12]. In one class of designs, a spatial light modulator (SLM) (e.g. DMDs, LCoS, etc.) masks the entire sensor field-of-view ω_{FOV} and creates the flying-dot with modulation patterns; for e.g. sequentially unmasking light-rays with a smaller illumination cone ω_{small} .

Our MEMS-mirror setup lies in the second class of designs, where a fast optical modulator (FOM) controls only a single, smaller illumination cone ω_{small} and sequentially steers this pencil of rays over the FOV ω_{FOV} . O’Toole *et al.* [12] showed that SLMs (the first class) are not as energy-efficient as FOMs (the second class). Intuitively, FOMs focus the available illumination energy into a smaller cone ω_{small} compared to over the entire FOV ω_{FOV} . This multiplexing of the energy by FOMs results in a k times increase in scene radiance, given by how much bigger the FOV is compared to the smaller cone,

$$k = \frac{\omega_{FOV}}{\omega_{small}}. \quad (3)$$

Despite this, for dynamic scenes, exposure times for flying-dot imaging for dynamic light-transport are low enough that images still contain significant amounts of shot noise. One of our contributions deals with the low-light present in impulse dot images, which we overcome with a simple and novel denoising algorithm and additional optics in Section V.

2) *Time Efficiency:* While fast optical modulators (FOMs) such as the MEMS mirror used in our setup have been proven to be energy efficient [12], it is important to also consider the time efficiency of the two design classes. For most commercially available FOM technology (including MEMS steering mirrors, phase-based MEMS mirror arrays, etc.) the modulation can be still slower than SLMs which can illuminate parallel patterns simultaneously (for e.g. hierarchical patterns [1] implemented on a DMD array). If SLMs effectively scan m times faster, they can use this extra time to increase exposure and reduce noise. This suggests a simple ratio to determine when impulse responses from FOMs, known to be energy efficient [12], are also time efficient,

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

3) *Example Use Case:* Consider our scenario with a 3.6 mm Mirrocle MEMS mirror, detailed in Section VI, capable of 6 FPS light-transport capture. The FOV of the device is 5.25° in the horizontal and 9.27° in the vertical, and for ease of discussion we use the average FOV $\omega_{FOV} \propto 7.3^\circ$.¹ We measured the angular support of the dot in our system to be $\omega_{small} \propto 0.42^\circ$. We compare to a commercial DMD array, the TI DLP LightCrafter, such as that used in [12]. For simplicity, we assume that no scene point is illuminated by more than one dot. In this situation, the

¹Proportionality due to solid angles.

energy advantage of our system is

$$k = \frac{\omega_{FOV}}{\omega_{small}} \approx \frac{2\pi(1 - \cos(7.3^\circ))}{2\pi(1 - \cos(0.42^\circ))} \approx 301, \quad (5)$$

assuming a conical solid angle. Note that since there are no overlapping dots, that means our system can cover the full FOV in 301 unique directions. However, since the TIDL can perform 10 KHz binary patterns, this translates to 1667 patterns at our sensor's 6 fps rate. Using the adaptive algorithm from [1], one can analyze the difference between complex scenes (such as glass and water) vs. a Lambertian scene with no inter-reflections. For a complex scene, the adaptive algorithm has worst-case complexity $O(n)$, and therefore the time-multiplexing advantage of the DLP is $m = \frac{1667}{301} \approx 6$ and so the ratio $\frac{m}{k} = \frac{6}{301} < 1$ favors our system. However, for Lambertian scenes, [1] provides a complexity of $O(\log_4(n))$, which translates to $m = \frac{1667}{\log_4(301)} \approx 416$, and the ratio $\frac{m}{k} = \frac{416}{301} > 1$ favors the TIDL. While most scenes lie in-between these extremes, the ratio we have provided allows for sensor design decisions, given different use cases.

4) *Impact on Flying-Dot Photography*: In other words, given a desired FOV ω_{FOV} , it is recommended to run any FOM regardless of beam-steering technology, including our sensor, as fast as possible, to reduce the parallelizable advantage m of SLMs. For most device technology, including the scanning MEMS mirror that we use, this means running the modulator at resonance.

Resonance scanning also helps us drive the systems at the highest spatial resolution possible, since it reduces the time interval Δt discussed above. There are additional consequences to requiring the MEMS mirror modulator to run at resonance. First, the MEMS mirror must be placed in air/vacuum, requiring a cover-glass that induces optical effects. Second, the MEMS mirror modulation at resonance is sinusoidal, meaning that graphics algorithms for light-transport, such as relighting, separation and dual photography discussed in Section VI, must succeed with non-grid sampling. In the next sub-section, introduce the first of such algorithms for overlap compensation to deal with scanning fast at resonance.

B. Scanning Dot Compensation for Flying-Dot Projectors

A fast flying dot projector solves the problem of measuring columns of the dynamic light-transport $T(t)$ before it changes significantly. However, it introduces a new challenge due to significant overlaps between adjacent scanning dots. This is due to the generation of our circular "dots" by reflection off a circular MEMS mirror. Commercially available projector pixels are rectangular, with high fill-factor when the illumination is in focus. In contrast, our projector cannot illuminate every scene point without causing illumination overlaps for some regions in the scene.

Formally, this non-uniform sampling is caused by the fast scanning of a single, small illumination solid angle over the entire sensor's FOV. Illumination overlap means that two projector pixels i and j with directions (θ_i, ϕ_i) and (θ_j, ϕ_j) , have solid angles such that $\omega_i < \omega_{diff}$ and/or $\omega_j < \omega_{diff}$, where ω_{diff} is the inter-probe solid angle $\omega_{diff} = \|\theta_i - \theta_j\| \|\phi_i - \phi_j\| \sin \theta_i$. Such illumination overlaps are common for systems whose point

spread function (PSF) is larger than a pixel including defocused conventional projectors. Therefore a valid imaging strategy is to simply ignore the overlap region, if the MEMS mirror scanning allows for densely-packed projector pixels. Many of our results in Section VI follow this policy, and the relighting is not badly affected.

However, as we reduce the number of scanning dots (for instance, to accelerate capture of even faster scenes), the effect of overlap cannot be ignored. In Fig. 2(I-II) we show a simulation using PBRT [42] where the gap between the scanning dot is half the angular extent of each dot. The floodlit scene illuminated by the flying-dot project or then contains edge artifacts due to this gap.

Our strategy here is to produce a set of compensation weights for the direct component of each pixel that normalizes the pixel value by how many times it was illuminated by a single dot. Since our system is radiometrically calibrated, all sensor measurements are linear and simply dividing by these weights is enough to remove the effects of overlap compensation. We simulate multiple overlap of dots in the scene per pixel, and calculate the compensation weights by binarizing these images and counting how many times a single pixel was covered by a dot. By combining dot patterns into checkerboard-like sets, similar to [4], we can separate this compensation mask only for the direct component, ignoring it for the global component. In Figure 2(III) we show the compensated image.

For relighting, the masks can be computed for each color channel, for particular patterns. As we describe in later sections, the color value given to each pixel in the projector depends on the illumination direction (θ, ϕ) . Given a desired illumination pattern, recovering the color for a given direction requires interpolation. However, the resolution of the projector is limited both by the camera and the spatially varying angular extent of the pixel's illumination. In Figure 2(IV) we use bicubic interpolation and our overlap compensation algorithm to generate a relit image. For dynamic scenes, this computation is done per frame, and we discuss the details of the compensation algorithm on real experimental data in Section VI-B.

IV. FAST FLYING-DOT PROJECTOR OPTICS

In this section we describe the optical properties of the fast flying-dot projector, including issues regarding mirror resonance and noise. In the next section, we discuss how to compensate or calibrate for these effects, which can impact visual quality of our system. Our approach is to capture the light-transport for dynamic scenes using fast sensing and optical modulation. The optical setup we use is depicted in Fig. 3, where a high-speed camera images a scene illuminated from a broadband light emitting diode (LED). The LED illumination is reflected off a circular MEMS mirror, creating a small illumination cone that is modulated over the sensor FOV. The scene appears floodlit when viewed directly by the naked eye or with a low-frame rate camera.

In our setup in Fig. 3, we use a Photron SA-X2 color camera that can capture 1024×1024 resolution at 13.5k FPS. The broadband LED was generated from a 460 mW Prizmatix light engine, stopped down via a physical aperture, and reflected off a

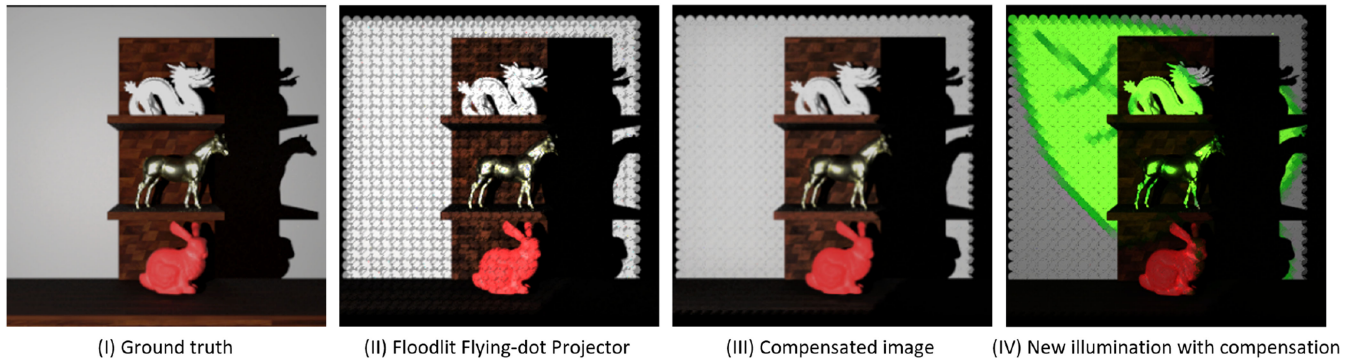


Fig. 2. Compensating for circular dot pixels in the flying-dot projector: Scanning an illuminated solid angle over the sensor's FOV creates overlap artifacts (II) when compared to ground truth floodlit illumination (I). We propose a method to extract checkerboard patterns to allow for robust compensation, allowing removal of artifacts in floodlit (III) and relit images (IV). By using the checkerboard pattern, similar to [4], this compensation can be targeted only at the direct component.

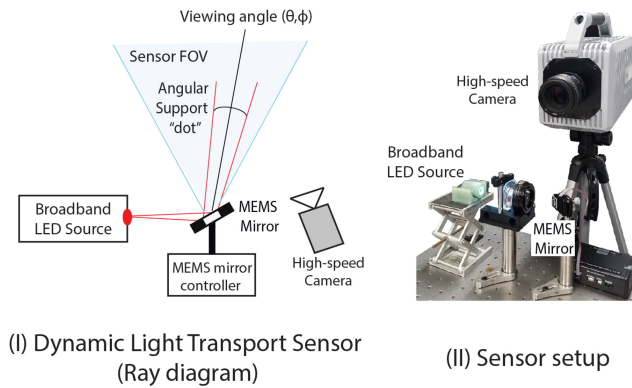


Fig. 3. Setup: We use a high-speed camera to image a scene with floodlit illumination. The source is modulated by a fast moving MEMS mirror. Effectively, we collect “dot” images at a fast enough rate that allows capturing light-transport for dynamic scenes.

Mirroracle 3.6 mm-diameter MEMS mirror. Whether the design is electrostatic or electrothermal, MEMS mirrors trade-off size versus resonance speed. While multiple tens of KHz are possible with tiny mirrors, these result in low SNR and a small illumination solid angle. We found that 3.6 mm allowed for a large enough dot size while still allowing scanning speeds in the low KHz range for capturing dynamic scenes.

Commercial laser projectors, light-source selection and lighting optics: The novelty in our flying-dot projector design is a MEMS mirror that is run at a balanced speed in *both* angular dimensions (θ, ϕ), which is different from many commercial laser projectors, such as [41], that run at least one axis at 60 Hz, i.e., at video frame rate, and the other at multiple KHz. This imbalance, while potentially faster than our solution, creates large amounts of redundant imagery, reducing efficiency in storage and pre-processing. Further, our flying-dot projector system can be implemented with different light-sources. We use an incoherent LED light-source, since we target light-transport applications and broad-band spectral response is attractive. Since cheap, broadband LEDs have large beam divergence, we use a simple two-lens telescopic optics to focus the light on the MEMS mirror, with an iris to prevent reflection onto the MEMS packaging.

Laser sources, such as those used in [3] would also work well with our system, and in Section V a discussion is provided on the impact of power on visual quality, which would be the main advantage of lasers. Commercial laser projectors, such as [41], combine multi-wavelength lasers to provide metamer coverage for human perception of color. For light-transport applications however, this can be a disadvantage; e.g. in the degenerate case of relighting a prism, the spectrum caustic caused by dispersion would collapse to three stripes for a tri-wavelength laser projector. Finally, eye-safety issues for scanning laser displays [43], are reduced for incoherent scanning devices such as ours.

i) Fast Operation via MEMS Resonance: Finding the fastest mirror speed $L_{\max}$ for some task is crucial. Let the frame rate of the camera be F in frames/sec, and the mirror speed be L in steradians/sec. Let $F_{\max}$ be the maximum camera frame rate possible. The maximum optical modulation provided depends on the amplitude (i.e., total solid angle) over which modulation is occurring, as well as line sample rate (i.e., resolution) and the control points (i.e., specific angular directions in (θ, ϕ) space that must be sampled). Further, MEMS mirrors exhibit *resonance*, i.e., a combination of these parameters, for a particular amplitude, that produces optimal energy transfer and fast mirror motion.

To simplify the goal of finding the fastest dot movement, we only consider mirror parameters that produce a desired maximum FOV, or amplitude. We plot *iso-amplitude* parameter curves, and select the fastest parameter combination along any such curve. As an example of how to use iso-amplitude curves to find the fastest mirror speed $L_{\max}$, consider Fig. 4(I), which shows three such curves measured with our setup. The z axis is a proxy for speed (i.e., lower is faster) and the x and y axes are the sample rate and control point parameters mentioned above. The top and bottom curves, for 5 cm and 8 cm respectively, are near monotonic and the fastest setting is close to the bottom of the curve. For the 6.5 cm curve, a resonance configuration exists and can be exploited.

ii) Mirror Speed, Camera Rate, and Projector Resolution: We present a simple flowchart shown in Fig. 4(II) to determine camera frame-rate and mirror-speed to ensure a single non-motion blurred dot per frame. We first image a vertical fronto-parallel Lambertian plane inside the depth-of-field of the

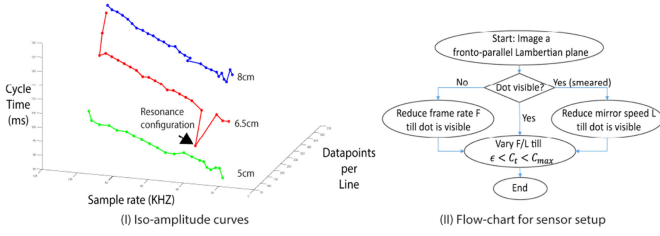


Fig. 4. Procedure for discovering sensor parameters: In (I) we show an example of collecting iso-amplitude curves while varying resolution and control points. The z-axis depicts speed of the dot, collected on a lambertian plane placed in the working volume, and the lowest value shows the highest speed. Once the fastest MEMS speed has been selected, in (II) we provide a flowchart to select the remaining sensor parameters.

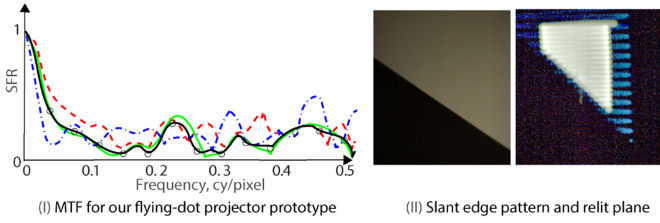


Fig. 5. MTF of our optical system is measured in (I) for the slanting calibration image shown in (II) and obtained as an image measurement from one mirror scan of the sensor FOV.

camera. We then analyze whether the dot is under-exposed (i.e., below noise threshold) or over-exposed (i.e., smeared due to motion) using thresholds on dot-pixels. We also require discernible motion detection between frame-to-frame, which we denote as $C_t = \|I_t - I_{t-1}\|$, and vary frame rate over mirror speed F/L until some criteria is met $\epsilon < C_t < C_{\max}$. If the projector FOV is ω steradians, then the time taken for 1 scan is $t = \frac{\omega}{L}$. The resolution of the projector is then the number of frames captured in this time, $N = Ft$.

iii) *Modulation Transfer Function*: The Modulation Transfer Function (MTF) measures the spatial frequencies preserved by an optical system (similar to the optical transfer function but neglecting phase effects in the system [44]). We measure MTF using the slanted edge method [45] with the software sfrmat3.² We project a binary slanted edge image using our virtual flying-dot projector as our target, so that we can measure the loss of spatial frequencies in both the virtual projector as well as the camera.

In Figure 5, we show the system MTF, measured as modulation contrast as a function of spatial frequency in units cycles/pixel, and extends up to the Nyquist frequency of 0.5 cycles/pixel. Note that MTF50 metric, i.e., when the MTF curve loses 50% of its contrast, occurs at 0.05 cycles/pixel, which means we have a factor of 10x loss in frequency resolution compared to the ideal optical system given our pixel pitch. This validates with our intuition, as the dot size of the pixel prevents the projection of high spatial frequencies in the scene, and is a limitation of our hardware.

²[Online]. Available: <http://losburns.com/imaging/software/SFRedge/index.htm>

A. Fast Flying-Dot Projector Calibration

Our calibration utilizes the Lambertian plane scene as shown in Fig. 6 (II), and imaging the scanning dot over the entire FOV in one batch as determined in (I). Synchronization is done by binarizing the dot on a “trailing dot,” circled in red, that lies on the Lambertian plane and is directly viewed by the camera. This easily gives a leading edge-based temporal signal to extract each scanned batch and synchronize batches together over the duration of the dynamic scene capture.

Additionally, we compute a confusion matrix, by binarizing each frame to obtain the overlap between scanned dots from different viewing directions. We do this by forming a $N \times N$ matrix M where M_{ij} is that dot product between spotlight images i and j . This confusion matrix calibrates the angular overlap for different MEMS-mirror viewing directions (θ, ϕ) . The Lambertian scene has almost zero global illumination with the direct bounce dot clearly visible, allowing reliable and robust binarization and calibration. This matrix, which we term the *calibration confusion matrix*, completely describes our projector, and is calculated per scene to characterize any drifting or shift in position due to non-repeatability of the mirror.

Scanning Pattern for Every Scene: Our scanning pattern captures the entire light transport by scanning the dot within a brief instance of time, ($\frac{1}{5}$ sec in most of our experiments). To synchronize each batch of dot images, we place a static Lambertian plane outside the imaging working volume, so that the scanning dot illuminates it at the same trailing dot location at the end of every batch. Binarizing this dot helps synchronize every batch. Additionally, we require that every mirror pattern that is used shines light out of the working volume, once every batch at the same time index, onto a distant, optically-absorbing material. This image, which we call $I_{\text{background}}$, contains the ambient light in the scene, and is used to background subtract ambient light from the other frames. Finally, the first few frames of every experiment require at least one batch of the Lambertian plane as a backdrop, to allow the computation of the calibration confusion matrix for light-transport applications.

V. IMPROVING VISUAL QUALITY FOR THE FAST FLYING-DOT PROJECTOR

Fast light-transport capture is challenging due to noise issues that affect visual quality. In this section we discuss the system strategies that have proven useful in increasing the quality of the captured light-transport. Any speed-up of the flying-dot projector will reduce the exposure time for each dot image. Since each dot image is created by illuminating a small solid angle of light-rays, this exacerbates an existing low-light scenario. While previous light-transport capture with static scenes [1], [2] obtained high dynamic range (HDR) images for each column T^i of the light transport, we cannot acquire HDR images for fast-moving scenes and must extract signal from shot-noise dominated images. Figure 7 (I) shows raw flood-lit data from our sensor, demonstrating shot noise and fixed-pattern noise present in our data.

In addition, the raw image in (I) also exhibits a bluish tinge. MEMS mirrors are sensitive to dust particles and vibration, and

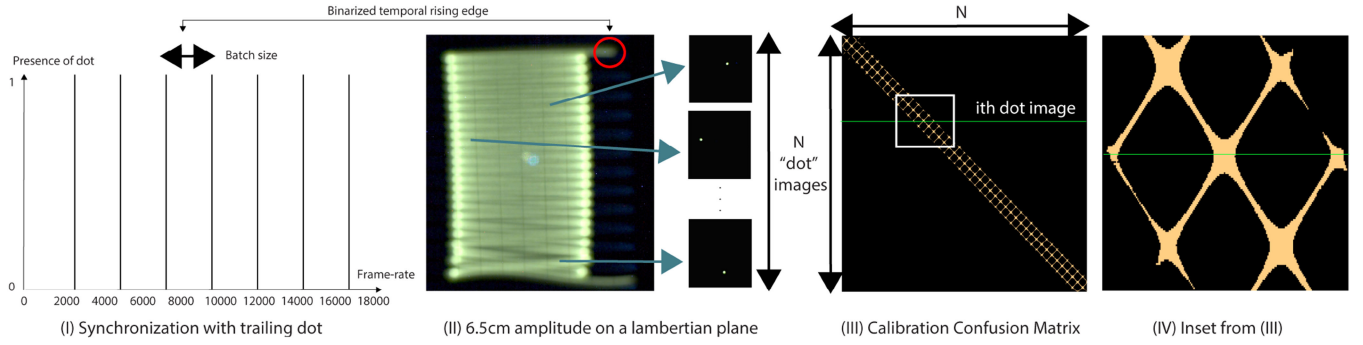


Fig. 6. We depict our synchronization and calibration process. We require a static lambertian plane placed in the background of our experiments. A trailing dot, circled in red, is required to be visible throughout the experiment, and serves as a synchronization leading edge, as shown in (I). In (II) we show dot images from the lambertian plane, which allow for easy detection and binarization. Dot-products between binarized dot images form the calibration confusion matrix, shown in (III). The close-up in (IV) shows the partially repeated structure has gaps, which make this matrix critical for light-transport applications.

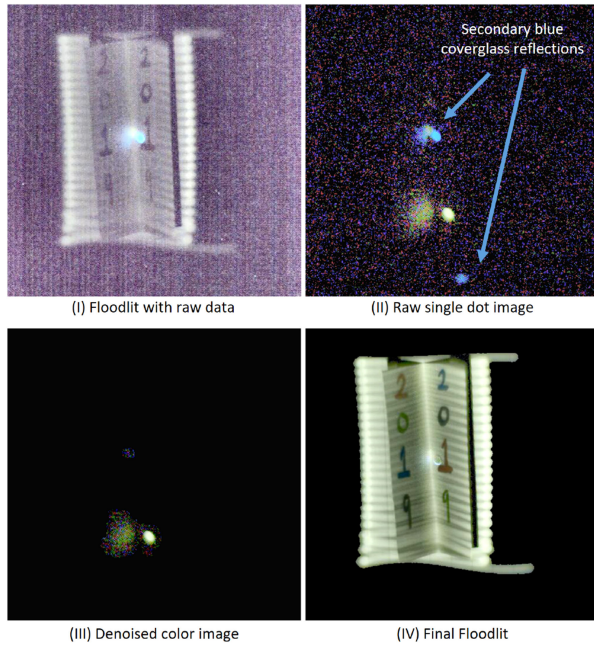


Fig. 7. Our system captured data at high-frame rates, resulting in low-light characteristics. Here we depict our denoising algorithm to remove these and enable light-transport based relighting.

therefore are protected by a coverglass. The optics of the coverglass contain anti-reflectance coatings that induce secondary, dynamic dots in the scene, shown in Figure 7 (II). To alleviate these issues, we perform HSV thresholding and a median filter to clean up the final videos we synthesize, which are described in Section VI.

We propose a simple but novel denoising algorithm tailored for our system to extract signal from noise, which is outlined in Algorithm 1. We convert all our dot images to grayscale images for this denoising algorithm.

Let a single dot image be denoted as I . First we perform background subtraction $I = I - I_{background}$, using the background image calculated in Section IV-A. We then extract a region of pixels in I that is only a small non-illuminated region of the image composed of solely background noise, and

Algorithm 1: Denoising Algorithm.

Input: Light-transport matrix from a single scan (i.e., whose columns are dot images)

Output: Denoised scanning dot image O

```

1: for each scanning dot image  $I$  do
2:   Initialize output  $O$  to all zeros.
3:   Background subtraction:  $I = I - I_{background}$ .
4:   Calculate  $(\mu_b, \sigma_b)$  of background patch in  $I$ .
5:   Calculate  $(\mu_I, \sigma_I)$  of the entire image.
6:   repeat
7:      $(x^*, y^*) = \operatorname{argmax}_{(x,y)} I(x, y)$ .
8:      $N_r(x^*, y^*) = \{(x, y) : \|(x, y) - (x^*, y^*)\| < r\}$ .
9:      $I = I - N_r(x^*, y^*)$ .
10:     $O = O + N_r(x^*, y^*)$ .
11:    Recalculate  $(\mu_I, \sigma_I)$ .
12:   until  $(\mu_I, \sigma_I) = (\mu_b, \sigma_b)$ ;
13: end
    
```

calculate the mean and variance (μ_b, σ_b) for those pixels. We also calculate the mean and variance (μ_I, σ_I) for the entire image. We then perform the following iterative algorithm: (0) Initialize O to be an empty output matrix that is the same size as I . Then (Step 1) find the index of the brightest pixel in the image $(x^*, y^*) = \operatorname{argmax}_{(x,y)} I(x, y)$, (Step 2) extract a neighborhood of pixels $N_r(x^*, y^*)$ with user-defined radius r , (Step 3) set $I = I - N_r(x^*, y^*)$, and $O = O + N_r(x^*, y^*)$, and re-calculate (μ_I, σ_I) .

Steps 1–3 are repeated until $(\mu_I, \sigma_I) = (\mu_b, \sigma_b)$. This termination criterion is equivalent to when the noise statistics of the whole image matches the noise statistics of that background patch. The extracted neighborhoods collectively form our denoised image O . In Figure 7(III) we show the denoised image and the resulting floodlit image when we use this denoised impulses.

A. Visual Quality Trade-Offs Between Frame Rate and Illumination Strength

In the design of our system, there are several key parameters that affect performance. This includes primarily the frame rate of the camera and the illumination strength of the source, which

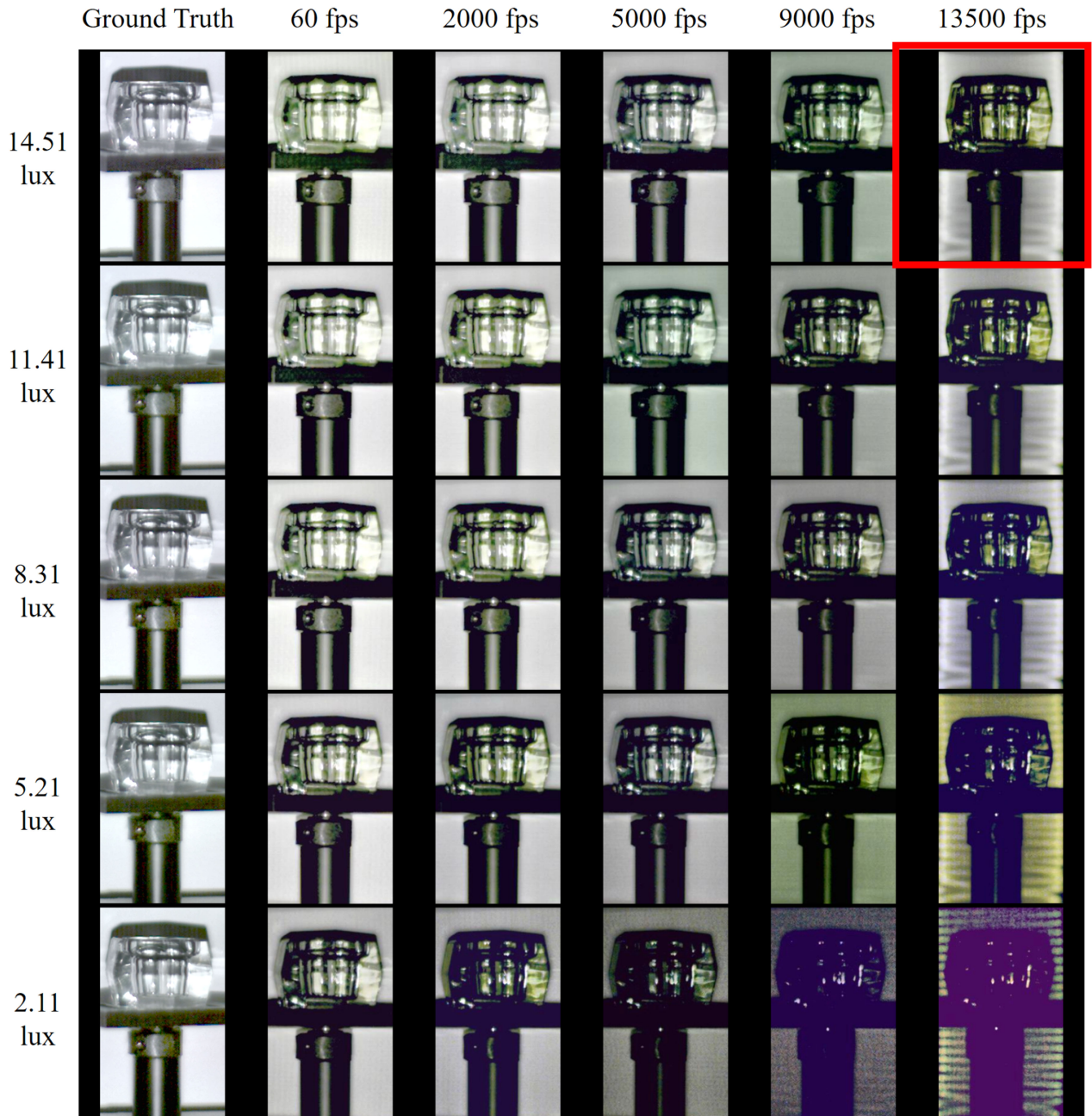


Fig. 8. A sweep of illumination (measured in lux) and frame rate of the high-speed camera for the proposed flying-dot projector, visualizing the synthesized floodlit images from the captured light transport. On the left column, ground truth light transport was captured using a co-located regular projector using Hadamard codes (i.e. a fast, parallel high-SNR capture strategy). Note that higher illumination power or slower frame rate results in better visual quality, and illustrates the trade-offs of designing fast flying-dot projectors for light transport acquisition. All the results in this paper are from the settings given by the red rectangle, i.e., fast frame rate of the high-speed camera and high-power, allowing 6 FPS light-transport acquisition. **Note that for static scenes our method provides high quality at a faster rate of capture. For example, a video-rate capture of this glass static scene at 14.51 lux, using Hadamard codes, took about 300 seconds. Using our setup, at 2000 FPS, a similar result was obtained in about 1 second.**

determines the visual quality of the resulting light transport captured. To benchmark our performance, we co-located a regular color projector (Epson Powerlite Home Cinema 8350) with our MEMS-based projector. This was achieved by using a 50/50 beam splitter where we aligned shadows projected from uniform illumination of the regular projector to match the shadows cast by the MEMS projector. We then perform Hadamard coding to

multiplex the light transport acquisition [46], where the minimum pixel size of the projector was set to roughly correspond to the size of the dot from the MEMS mirror (32×32 pixels).

In Figure 8, we show the floodlit (e.g. adding all dot images together) results of sweeping frame rate versus illumination for our flying-dot projector. We used a EXTECH LT300 light meter to measure the luminous flux (lux) of the projected dot by the

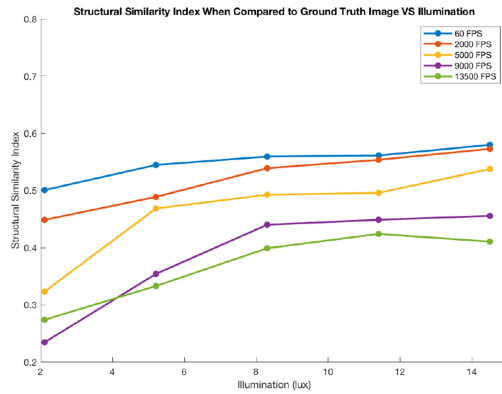


Fig. 9. Quantitative analysis of the SSIM as a function of illumination and frame rate. All SSIM calculations used the Ground Truth, 14.51 lux setting as the reference. We see clearly the trend of stronger illumination or longer exposure time results in higher visual quality.

MEMS mirror. On the left column, we show the equivalent ground truth light transport measurement at the same illumination lux using our co-located regular projector, and the high speed camera at 60 fps. Note that except for a slight change in color, our flying-spot projector at 60 fps has similar visual quality to the ground truth. As one can see from the graph, visual quality follows our intuitions in terms of frame rate and illumination strength, with higher frame rates and lower lux resulting in a degradation in visual quality. For our system, we operate at 14.51 lux and 13,500 fps for our experimental results, as we wanted to achieve light transport at interactive/near-video rates of 6 fps while sacrificing some visual quality. However, one can see that a brighter source would help alleviate some artifacts in our light transport.

In Figure 9, we quantitatively evaluate the visual quality of the images shown in Figure 8 using the Structural Similarity Index (SSIM) [47] between the reference image (Hadamard ground truth light transport captured at 14.51 lux, 60 fps) and every other image. Note that the trends in this graph conform to our intuition: lower frame rates or higher illumination results in higher visual quality. This gives us a good design point for trade-offs between these powers.

Note that all the results are for floodlit imagery. For completeness, we also include here a rendering of the same scene with a black-and-white horizontal stripes pattern, created for the 2000 FPS case. Note that the quality of the image is high, noise is low and the different specular properties can be clearly seen.



VI. LIGHT TRANSPORT APPLICATIONS USING OUR FLYING-DOT PROJECTOR

We now present light-transport results using our approach. We refer the reader to the supplemental material for video results of several of the scenes presented in this section.

Implementation Details: All scenes were captured with a 3.6 mm Mirrocle MEMS mirror scanning 5.25° in the horizontal and 9.27° in the vertical with sample rate of 39,000 Hz. The Photron camera is set to a frame rate of 13,500 FPS. For

every scene, we perform dark calibration, lens shading correction, and hot pixel correction using the Photron software. We output linear raw measurements from the camera, and then perform our batching synchronization described in Section IV-A. Batching takes approximately 3–5 hours per scene. We then calculate the calibration confusion matrix which takes an additional 3–5 hours. We then perform our denoising algorithm and store the dynamic light transport which takes about 1–2 days per video. We generate videos using our algorithms for relighting (1–2 hours per video), direct/global separation (30–40 minutes per frame, 1–2 days per video), and dual videography (1–2 hours per video). The final frame rate of our captured videos is 6 FPS. After all videos have been generated, we perform white balancing, median filtering, color thresholding in HSV space to remove blue artifacts caused by the coverglass, tone-mapping, and then mask the background of the scene to generate the final videos displayed in the supplemental material and presented as still frames.

A. Video Relighting

Light transport capture enables scene relighting after images have been captured. In traditional light-transport relighting, a desired illumination $\mathbf{l}$ is simply a matrix-multiplied by the light-transport to obtain the new synthesized image $\mathbf{i} = \mathbf{T}\mathbf{l}$. However, due to MEMS resonance in our flying-dot projector, we cannot simply perform a matrix-vector product, as the indices change in time, and thus we must perform re-indexing of the illumination vector for every frame.

We do this by working directly in the dot image space, since we require a Lambertian plane placed behind the working volume for our calibration. We first compute the centroid of the dot in the i th image, which we define as a mapping $M_i = (a_i, b_i)$. We then take our lighting vector $\mathbf{l}$ and resize it to the image space of M_i , and normalize each color between 0 and 1. We then use the mapping M_i to sample the illumination and recolor the corresponding spotlight image which the illumination maps to. This is equivalent to mapping rays from the MEMS mirror through the synthetic illumination image to the final scene locations. We then sum up all the spotlight images to form the final relight image. This is done for every frame of the video scene.

Experimental Results. V-Groove: Here we relight a moving V-groove with two illuminations: a moving pastel-colored checkerboard in the teaser Fig. 1, and a waving tri-color flag in Fig. 10(I). Notice in Fig. 1 how the groove center shows color-bleeding due to strong diffuse interreflections near the vertex. This demonstrates that our light-transport capture obtains subtle multi-bounce effects.

Candle: We relight a wax candle in Fig. 10(II) with a tricolor flag. The specularities change color, and note that the hues on the hands and the background change spatially as objects in the scene move.

Whiskey Glass: In this scene, we rotated a whiskey glass to demonstrate complex caustics and specular reflections, as shown in both the teaser Fig. 1 and Fig. 10(III). In Fig. 1 the specularities change color, and the caustics on the background show mixing

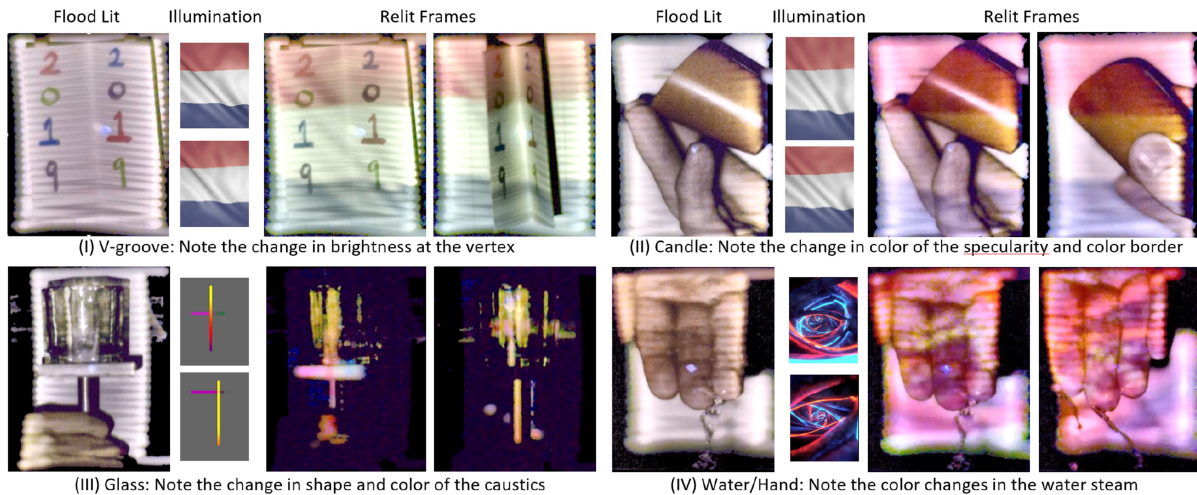


Fig. 10. Relighting results using our new flying-dot projector and high speed camera. Note that the floodlit images on the left are also synthetically generated by our method. Several frames of the input illumination and video relighting frames are shown for four scenes: (I) V-Groove, (II) Whiskey glass, (III) Candle, and (IV) Pouring water. Various global light transport effects such as specularities, caustics, and diffuse interreflections are captured in time. We encourage the reader to view the supplemental videos for each of these scenes.

between illuminant colors. In Fig. 10(III), we relight with synthetically created color bars. This high-frequency illumination probes individual regions of the caustic generating fascinating color/light combinations.

Pouring Water: We show a water being poured across fingers in Fig. 10 (IV). While the relighting result matches the input illumination, this scene also demonstrates some of the limitations of our light transport acquisition. Since the effective frame rate is only 6 fps, we do not capture the water droplets, but rather experience motion blurring along the water. Further, since our illumination pattern has a high spatial frequency, it gets washed out (loss of contrast given by the MTF function described earlier) and loses its sharpness. However, the motion blurred water droplets behave as small transparent artifacts in the light-transport, and our relighting is able to capture the complex internal reflections inside these. By pausing the video in the supplementary material, a reader can view relighting the reflections and specularities in water.

Fog: Finally in Fig. 11, we show that we can even relight scattering media such as fog moving in the scene. Previous work [48] achieved relighting of fog by capturing shape and scattering properties using a swept laser sheet and high-speed camera. Such relighting requires post-capture Monte Carlo rendering. In contrast, our method directly captures the light transport to enable image-based relighting of fog.

To conduct the experiment, we utilized dry ice mixed with water to generate the fog, which we then poured from above down into our imaging volume. Note that our static relighting of the letter “A” softens and conforms to the shape of the moving fog, showing the effects of scattered light within the volumetric media.

B. Overlap Compensation

Since our optical setup allows for dense motion of the MEMS modulated illuminated solid angle in the scene, the dots

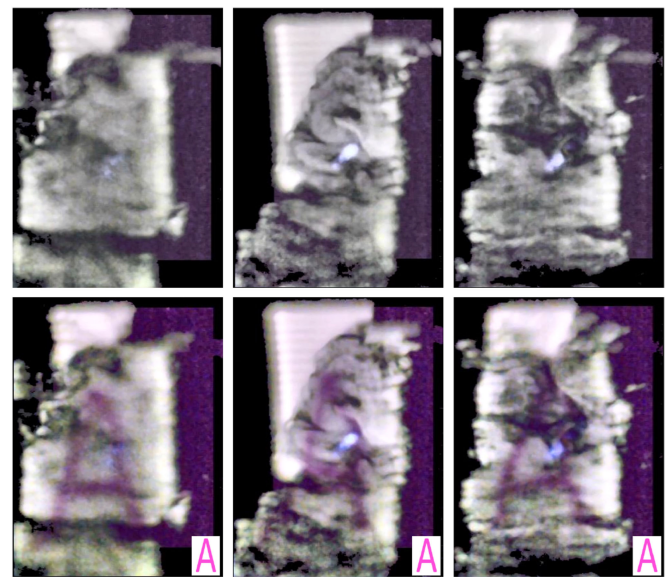


Fig. 11. We show floodlight (top row) and relighting (bottom row) of fog, generated by dry ice, moving dynamically in the scene. Note how the relight “A” in the scene conforms to the warping of the fog and has edges that are blurred due to the scattering of light.

captured by the projector are heavily overlapped. As explained in Section III, this reduces visual artifacts, compared to sparse sampling, and many of the results in this section show no significant artifacts.

Another option is to apply the overlap compensation algorithm described in Section III-B. We do so for the V-groove scene, as shown in Fig. 12, where the artifacts of alternating light and dark rows are removed in the compensated image. The algorithm we use is identical to that described in Section III-B with one major difference - the dot projected onto the scene cannot be assumed to be uniform. In other words, the PSF of our system is not simply the pillbox function, but has some

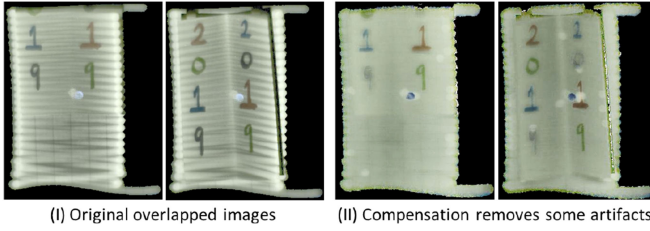


Fig. 12. We apply our overlap compensation algorithm for the circular dots in the real data of the V-groove scene. Note that the artifacts are diminished while the texture remains intact for the scene.

non-uniform radiance distribution across the solid angle reflected off of the MEMS mirror.

We estimate this distribution for each dot in the scene and we have found that assuming a continuous, unimodal form works well. We normalize the peak value at 1, and accumulate the PSFs. If the PSF is non-unimodal (e.g. if there are texture boundaries in the dot), we default back to uniform distribution (i.e., pill-box) over the dot's solid angle. We accumulate these radiance distribution estimates, to compute a compensation mask for each frame. Note that in Fig. 12, most of the visual artifacts are removed, while much of the texture (colored digits in this example) are preserved.

C. Direct/Global Separation

Since we have the full light-transport, we show direct-global separation in dynamic scenes similar to other works [4], [12], [23], [49]. We do require post-capture processing (rather than live results) yet we also allow for other results such as relighting with data captured.

For scenes which conform to the frequency assumptions of global illumination from [4], we present a similar “shadow dot” algorithm. We use our relighting method to generate images that are floodlit except for a single unique dark spotlight per frame in our videos. This “shadow dot” is then swept across the entire frame, and we apply the min-max algorithm from [4] to extract the direct and global from the scene. We show direct/global separation on the V-groove scene in the top row of Figure 13.

We also implement a more general approach for separation, based on the epipolar constraint, similar to that proposed by [11]. This allows us to also demonstrate separation for scenes with high-frequency caustics, which may not conform to the assumptions in [4]. Our approach involves, first, computing the fundamental matrix between the high-speed camera and the flying-dot projector. This is done by utilizing the calibration stage of our data capture, where a fronto-normal background Lambertian plane is scanned by the flying-dot, and was used in the calibration confusion matrix described earlier.

For each mirror angle, in time t , given by (θ_t, ϕ_t) , we compute the centroid of the corresponding dot image on the Lambertian plane given by (c_x^t, c_y^t) in pixel coordinates in the high-speed camera. We then drop a virtual plane Π perpendicular to the optical axis of our flying dot projector (the axis is along the direction given by the mirror resting position, i.e., $(\theta = 0, \phi = 0)$), and we require that this virtual plane is at unit distance

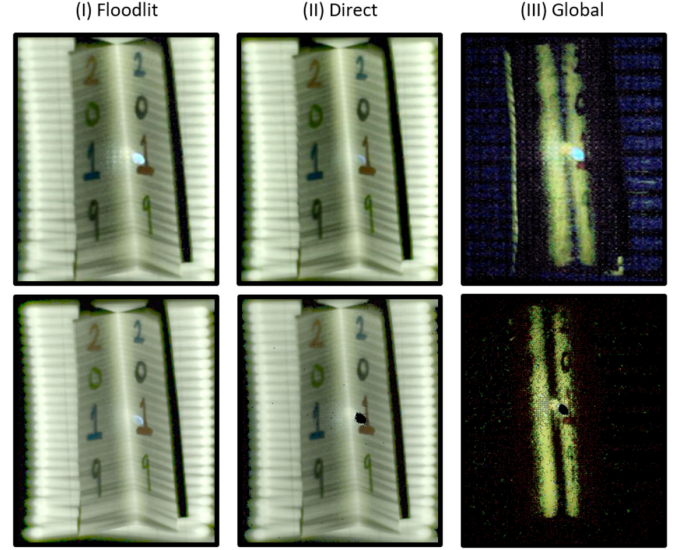


Fig. 13. Direct/global separation algorithms for the V-groove (top row) using [4]'s algorithm. Please see the supplemental video for video results of this direct/global separation. In the bottom row we show similar results using the more general approach proposed by [11].

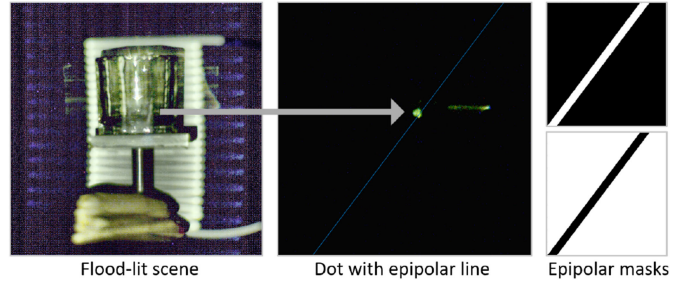


Fig. 14. We utilize a variant of the epipolar constraint-based algorithm from [11] to generate direct and global separation for dynamic scenes.

from the mirror. This allows a transformation from mirror angles to 2D location on this arbitrary plane, given by $m_x^t = (\cos(\theta_t) \cdot \sin(\phi_t))$, $m_y^t = (\cos(\theta_t) \cdot \cos(\phi_t))$. In practice, these can be directly obtained by scaling and transforming the centroid locations (c_x, c_y) , since the mapping between these and the virtual plane Π is given by a homography. Given correspondences between the flying-dot projector and the high-speed camera, and setting $\mathbf{c} = (c_x^t, c_y^t, 1)$, $\mathbf{m} = (m_x^t, m_y^t, 1)$, we can find the fundamental matrix $\mathbf{F}$ such that

$$\mathbf{m}^T \mathbf{F} \mathbf{c} = 0. \quad (6)$$

Once the fundamental matrix is obtained for particular experiment, the location of the flying-dot at time t and at angle (θ_t, ϕ_t) can be constrained to be having a pixel value c along the epipolar line. In Fig. 14, we show, for an arbitrary dot on one frame of the glass video, the epipolar line that passes through it. Note that caustics and other global effects lie outside this line. To generate the global and direct images, we create a mask M and its complement $M' = 1 - M$, where M matches the resolution

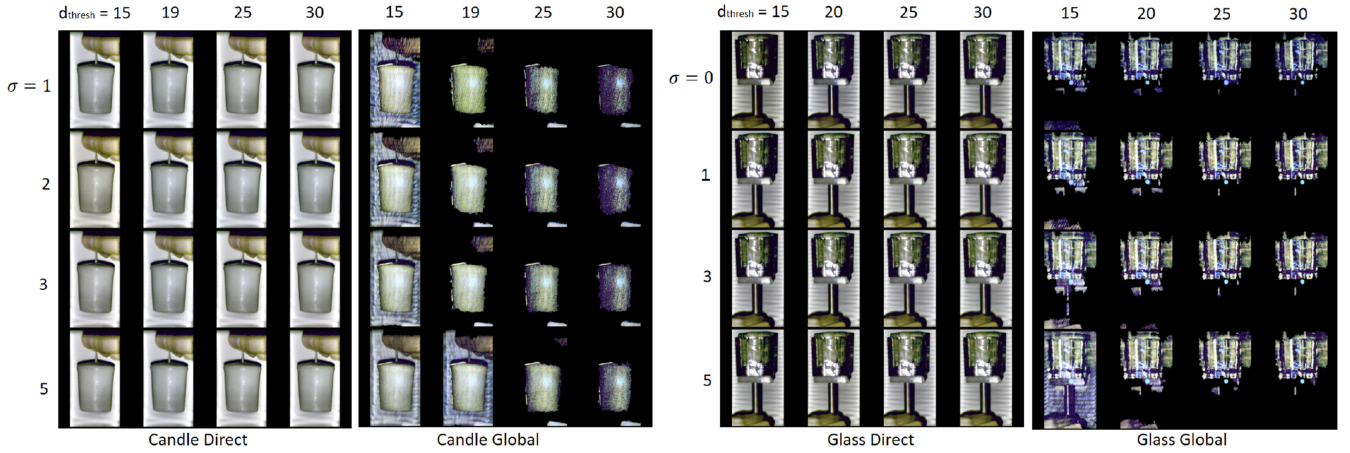


Fig. 15. Analysis of the d_{thresh} vs. σ parameters in our direct-global separation algorithm. The parameters affect the values in mask M in the direct-global separation Eq. VI-C. Here, for two particular frames from our candle and glass experiments, we vary the line-width d_{thresh} and the soft threshold parameter σ . Different types of global illumination (sub-surface scattering vs. caustics above) lend themselves to different separation masks for proper separation.

of the high-speed camera, according to the following equation,

$$M = \begin{cases} 0 & f_{thresh}(\mathbf{m}^T \mathbf{F} \mathbf{c}) > d_{thresh} \\ \mathcal{N}(f_{thresh}(\mathbf{m}^T \mathbf{F} \mathbf{c}) | f_{thresh}(\mathbf{m}^T \mathbf{F} \mathbf{c}) \leq d_{thresh}, \mu = 0, \sigma^2) & f_{thresh}(\mathbf{m}^T \mathbf{F} \mathbf{c}) \leq d_{thresh} \end{cases} \quad (7)$$

where f_{thresh} is a function that returns the perpendicular distance to the epipolar line and d_{thresh} and σ are user defined thresholds. In other words, if the pixels lie within a threshold distance of the epipolar line, their weights are assigned according to a 1D Gaussian with zero mean and sigma given by the user. To generate direct and global images from N dot images I_t , we perform the following masking and additions in post-capture,

$$I_{direct} = \sum_{t=0}^N M \cdot I_t \quad (8)$$

$$I_{global} = \sum_{t=0}^N M' \cdot I_t. \quad (9)$$

Analysis and experimental results: In Fig 15, we show analysis of the two user parameters, the line width d_{thresh} and the fall-off σ , showing their influence on separation. For a fixed σ , increase in line-width produces incorrect I_{direct} images that contain global components, since more and more of the “dot” in the flying-dot images gets claimed as part of the epipolar line. Conversely, as the line width decreases, the global image I_{global} obtains direct components, most visibly being the lambertian background plane. For a fixed line-width, at small values of σ , the masks are near-binary, and the line intersects the dot at arbitrary angles depending on scene depth creating high-frequency effects in the I_{global} image due to hard boundaries in the dot mask. As σ becomes larger, the boundaries become soft and these artifacts disappear. Fig. 16 shows separation videos for specific values of the mask ($d_{thresh} = 19, \sigma = 2$) for the wax candle and ($d_{thresh} = 30, \sigma = 0$) for glass. Note that the specularities of the wax candle are successfully removed to the direct, while the caustics of the glass and only found in the global images. Also note that there is significant noise in the sub-surface scattering component of the wax candle, at the frame-rate and LED power

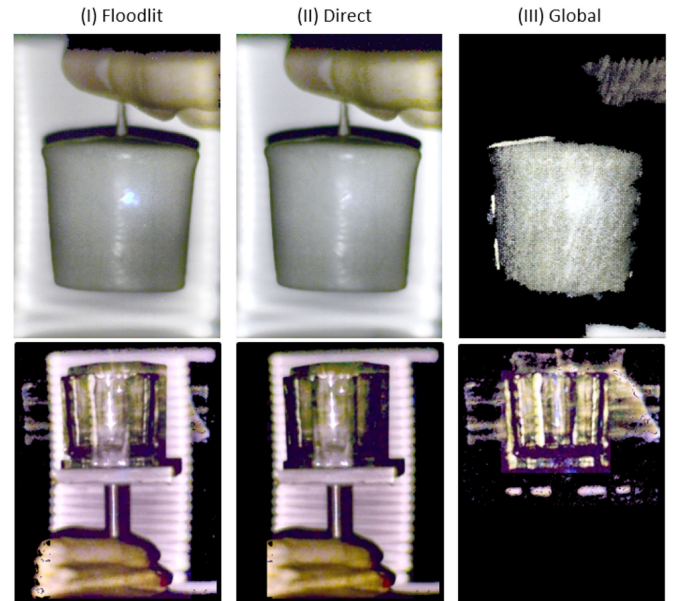


Fig. 16. Direct/global separation algorithms for the wax candle and glass. The specularity in the candle remains in the direct, and although there is significant noise in the sub-surface scattering component of the wax candle, it demonstrates the soft glow (about $\frac{1}{10}$ the power of the direct). This can be improved by choosing a different frame-rate/power pair, as in Fig 8. For the specular interreflections of the glass, this frame-rate/power pair is sufficient. Please see the supplemental video for video results of both these direct/global separation scenes.

(highlighted by the red rectangle in Fig 8.) that is necessary to capture moving scenes.

D. Dual Videography

Helmholtz reciprocity states that light rays traveling in a scene can have their directions reversed while preserving the same radiance. This fact has been exploited by several researchers in computer graphics and vision [1], [50], [51]. For dual videography, [12] were able to show live imaging for only the direct

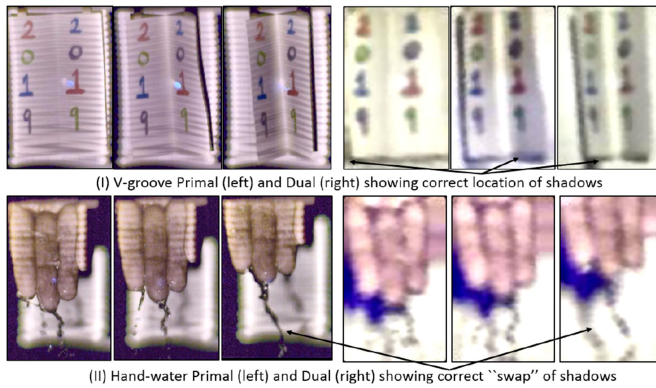


Fig. 17. Dual videography switching the viewpoints for the projector and camera are shown for the V-groove and pouring water scene. Note the correct swapping of objects and shadows between primal and dual domains. Please refer to the supplemental videos of these dual videography scenes.

component of light by utilizing the epipolar configuration of a synchronized projector-camera system.

In our setup, we can also perform dual videography, but can capture the full dual image since we are capturing the full light transport at each time instance. However, since our projector does not have a fixed resolution as described earlier, we need to perform our dual videography algorithm directly on the acquired dot images rather than taking a full matrix transpose as [1].

To perform dual videography without a matrix transpose, we again use our mapping $M_i = (a_i, b_i)$ of centroid locations per spotlight image. For each spotlight image i , we compute the average color spatially, and store that value in the (a_i, b_i) coordinate of the dual image. This results in some holes in our dual image, which we in-paint by applying horizontal and vertical averaging of non-zero neighbors repeatedly.

In Figure 17, we show the results of our dual videography for both the V-groove and the hand-water scene. Notice that the dual video is at low resolution, but still shows correct effects - most notably, the shadows in the primal are the water-stream in the dual and vice-versa. In addition, foreshortening effects on the V-groove text show the view is from the left, i.e., the correct location of our MEMS mirror.

VII. DISCUSSION

In this paper, we present an optical setup to capture dynamic light transport using a fast camera and a MEMS mirror flying dot projector. We analyzed the trade-off between energy and time efficiency for these systems, presented calibration strategies for our particular setup, and a denoising algorithm to boost its performance. Finally, we show several compelling computer graphics and vision applications including relighting of dynamic scenes including specular objects, subsurface scattering, and participating media, direct/global separation, and dual videography.

There are several limitations to our proposed approach. While we do mitigate several visual quality issues, we are limited to only scanning a single dot using our MEMS mirror system, and thus cannot use multiplexed illumination to increase SNR as

other works [2], [46]. We still have a maximum scene motion we can tolerate in a scene, and so still will suffer from motion artifacts for very fast moving scenes; our effective frame-rate is 6 fps, and we demonstrate this is still enough to see complex visual effects such as those present in glass, fog and water. In addition, our direct/global separation has limited effectiveness for subsurface scattering due to SNR limitations. Finally, our system is optimized for directional or projected illumination, and is not designed for environmental lighting commonly used for light stages [6] due to its reduced field-of-view.

We believe that controllable MEMS mirrors can provide a new avenue for adaptive light transport sampling. There is an interesting spatio-temporal interpolation that can be done, where smart choices of what components of the light transport to keep (specularities, high global, etc) versus what can be recovered in post-processing using optical flow and image interpolation algorithms. This can determine which light transport impulses are key to sample, and can perform energy-efficient scanning patterns that adapt to the moving scene with applications to 3D scanning. All these reasons point to dynamic light transport opening new avenues for computer graphics and vision research.

REFERENCES

- [1] P. Sen *et al.*, "Dual photography," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 745–755, 2005.
- [2] P. Peers *et al.*, "Compressive light transport sensing," *ACM Trans. Graph.*, vol. 28, no. 1, p. 3, 2009.
- [3] T. Hawkins, P. Einarsson, and P. E. Debevec, "A dual light stage," *Rendering Techn.*, vol. 5, pp. 91–98, 2005.
- [4] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 935–944, 2006.
- [5] S. M. Seitz, Y. Matsushita, and K. N. Kutulakos, "A theory of inverse light transport," in *Proc. IEEE Comput. Vision, Tenth Int. Conf.*, 2005, vol. 2, pp. 1440–1447.
- [6] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proc. ACM SIGGRAPH*, 2000, pp. 145–156.
- [7] P. Debevec, A. Wenger, C. Tchou, A. Gardner, J. Waese, and T. Hawkins, "A lighting reproduction approach to live-action compositing," in *Proc. SIGGRAPH*, San Antonio, TX, Jul. 2002, pp. 547–556.
- [8] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec, "Performance relighting and reflectance transformation with time-multiplexed illumination," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 756–764, 2005.
- [9] P. Einarsson *et al.*, "Relighting human locomotion with flowed reflectance fields," in *Proc. 17th Eurographics Conf. Rendering Techn.* Eurographics Association, 2006, pp. 183–194.
- [10] 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.
- [11] M. O'Toole, J. Mather, and K. N. Kutulakos, "3d shape and indirect appearance by structured light transport," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2014, pp. 3246–3253.
- [12] 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, 2015, Art. no. 35.
- [13] J. Haines and G. Tingley, "Live flying-spot color scanner," *Elect. Eng.*, vol. 75, no. 6, pp. 528–533, 1956.
- [14] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile, "Modeling the interaction of light between diffuse surfaces," *ACM SIGGRAPH Comput. Graph.*, vol. 18, no. 3, pp. 213–222, 1984.
- [15] J. T. Kajiya, "The rendering equation," *ACM Siggraph Comput. Graph.*, vol. 20, no. 4, pp. 143–150, 1986.
- [16] R. Ng, R. Ramamoorthi, and P. Hanrahan, "All-frequency shadows using non-linear wavelet lighting approximation," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 376–381, 2003.

- [17] P. Sen and S. Darabi, "Compressive dual photography," *Comput. Graph. Forum*, vol. 28, no. 2, pp. 609–618, Wiley Online Library, 2009.
- [18] G. Garg, E.-V. Talvala, M. Levoy, and H. P. Lensch, "Symmetric photography: Exploiting data-sparseness in reflectance fields," in *Proc. 17th Eurographics Conf. Rendering Techn.* Eurographics Association, 2006, pp. 251–262.
- [19] J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo, "Kernel nystrom method for light transport," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 29:1–29:10, Jul. 2009.
- [20] M. O'Toole and K. N. Kutulakos, "Optical computing for fast light transport analysis," *ACM Trans. Graph.*, vol. 29, no. 6, 2010, Art. no. 164.
- [21] J. Bai, M. Chandraker, T.-T. Ng, and R. Ramamoorthi, "A dual theory of inverse and forward light transport," *Comput. Vision–ECCV*, vol. 6312, pp. 294–307, 2010.
- [22] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "A practical approach to 3D scanning in the presence of interreflections, subsurface scattering and defocus," *Int. J. Comput. Vision*, vol. 102, no. 1–3, pp. 33–55, 2013.
- [23] S. Achar, S. T. Nuske, and S. G. Narasimhan, "Compensating for motion during direct-global separation," in *Proc. IEEE Int. Conf. Comput. Vision*, 2013, pp. 1481–1488.
- [24] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar, "Multiplexed illumination for scene recovery in the presence of global illumination," in *Proc. IEEE Comput. Vision Int. Conf.*, 2011, pp. 691–698.
- [25] D. Reddy, R. Ramamoorthi, and B. Curless, "Frequency-space decomposition and acquisition of light transport under spatially varying illumination," in *Proc. Eur. Conf. Comput. Vision*. Springer, 2012, pp. 596–610.
- [26] H. Kubo, S. Jayasuriya, T. Iwaguchi, T. Funatomi, Y. Mukaigawa, and S. G. Narasimhan, "Acquiring and characterizing plane-to-ray indirect light transport," in *Proc. IEEE Comput. Photography Int. Conf.*, 2018, pp. 1–10.
- [27] S. J. Koppal and S. G. Narasimhan, "Beyond perspective dual photography with illumination masks," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2083–2097, Jul. 2015.
- [28] M. Holroyd, J. Lawrence, and T. Zickler, "A coaxial optical scanner for synchronous acquisition of 3d geometry and surface reflectance," *ACM Trans. Graph.*, vol. 29, no. 4, 2010, Art. no. 99.
- [29] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "Flexible voxels for motion-aware videography," in *Proc. Eur. Conf. Comput. Vision*. Springer, 2010, pp. 100–114.
- [30] 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 *Proc. IEEE Comput. Vision Int. Conf.*, 2011, pp. 287–294.
- [31] A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec, "Rendering for an interactive 360 light field display," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. no. 40.
- [32] S. J. Koppal, S. Yamazaki, and S. G. Narasimhan, "Exploiting DLP illumination dithering for reconstruction and photography of high-speed scenes," *Int. J. Comput. Vision*, vol. 96, no. 1, pp. 125–144, 2012.
- [33] R. Hoskinson, B. Stoeber, W. Heidrich, and S. Fels, "Light reallocation for high contrast projection using an analog micromirror array," *ACM Trans. Graph.*, vol. 29, no. 6, 2010, Art. no. 165.
- [34] R. Hoskinson and B. Stoeber, "High-dynamic range image projection using an auxiliary mems mirror array," *Opt. Express*, vol. 16, no. 10, pp. 7361–7368, 2008.
- [35] 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. Visualization Comput. Graph.*, to be published.
- [36] T. Ueda, H. Kubo, S. Jayasuriya, T. Funatomi, and Y. Mukaigawa, "Slope disparity gating using a synchronized projector-camera system," *IEEE Int. Conf. Comput. Photography*, 2019, pp. 1–9.
- [37] T. Hawkins, J. Cohen, C. Tchou, and P. Debevec, "Light Stage 2.0," in *Proc. SIGGRAPH Tech. Sketches*, 2001, p. 217.
- [38] P. Peers, T. Hawkins, and P. Debevec, "A Reflective Light Stage," Univ. Southern California Inst. Creative Technologies, ICT Tech. Rep. ICT TR 04 2006, 2006.
- [39] M. Goesele, H. P. A. Lensch, J. Lang, C. Fuchs, and H.-P. Seidel, "Disco: Acquisition of translucent objects," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 835–844, Aug. 2004.
- [40] J. L. Baird, "Television apparatus," U.S. Patent 2 006 124a, Jun. 1935.
- [41] W. O. Davis, R. Sprague, and J. Miller, "Mems-based pico projector display," in *Proc. IEEE Opt. MEMs Nanophoton.,/LEOS Int. Conf.*, 2008, pp. 31–32.
- [42] M. Pharr, W. Jakob, and G. Humphreys, *Physically Based Rendering: From Theory to Implementation*. San Mateo, CA, USA: Morgan Kaufmann, 2016.
- [43] A. Frederiksen, R. Fieß, W. Stork, S. Bogatscher, and N. Heußner, "Eye safety for scanning laser projection systems," *Biomedizinische Technik/Biomed. Eng.*, vol. 57, no. 3, pp. 175–184, 2012.
- [44] J. W. Goodman, *Introduction to Fourier Optics*. Greenwood Village, CO, United States: Roberts and Company Publishers, 2005.
- [45] P. D. Burns, "Slanted-edge mtf for digital camera and scanner analysis," in *Proc. IS&T PICS Conf.*, 2000, pp. 135–138.
- [46] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1339–1354, Aug. 2007.
- [47] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [48] T. Hawkins, P. Einarsson, and P. Debevec, "Acquisition of time-varying participating media," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 812–815, Jul. 2005.
- [49] 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 Trans. Graph.*, vol. 33, no. 4, 2014, Art. no. 87.
- [50] S. J. Koppal *et al.*, "Toward wide-angle microvision sensors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2982–2996, Dec. 2013.
- [51] T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *Int. J. Comput. Vision*, vol. 49, no. 2, pp. 215–227, 2002.



Kristofer Henderson is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. He is also a Research Assistant with FOCUS Lab. His research interests include computational photography, computer vision, and sensor fusion.



Xiaomeng Liu received the M.S. degree in computer engineering from the Electrical and Computer Engineering Department, University of Florida, Gainesville, FL, USA, where she is currently working toward the Ph.D. degree. She is also a Research Assistant with FOCUS Lab. Her research interests include thermal sensors and novel cameras.



Justin Folden received the B.S. degree in electrical engineering from Florida Polytechnic University, Lakeland, FL, USA. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. He is also a Research Assistant with FOCUS Lab. After receiving the B.S. degree he joined the Gator Nation. His research interests include material science, image processing, electromagnetics, and computational photography.



Brevin Tilmon received the B.S.E. degree in engineering physics from Murray State University, Murray, KY, USA, in 2019. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. He is also a Research Assistant with FOCUS Lab. His research interests include computer vision, computational photography, and machine learning.



Suren Jayasuriya received the B.S. degree in mathematics and the B.A. degree in philosophy from the University of Pittsburgh, Pittsburgh, PA, USA, in 2012, and the Ph.D. degree in 2017 from the Electrical and Computer Engineering Department, Cornell University, Ithaca, NY, USA. He is currently an Assistant Professor jointly between the Schools of Arts, Media and Engineering and Electrical, Computer and Energy Engineering. He was a Postdoctoral Fellow with the Robotics Institute, Carnegie Mellon University.

His research interests include computational imaging and photography, computer vision and graphics, signal processing, and sensors. He was the recipient of the Best Paper Award for IEEE International Conference on Computational Photography (ICCP) in 2014 and the best demo award in ICCP 2019.



Sanjeev Koppal received the B.S. degree from the University of Southern California, Los Angeles, CA, USA, in 2003, and the master's and Ph.D. degrees from the Robotics Institute, Carnegie Mellon University (CMU), Pittsburgh, PA, USA. He is currently an Assistant Professor with the Electrical and Computer Engineering Department, University of Florida (UF), Gainesville, FL, USA, and a PI of FOCUS Lab. Prior to joining UF, he was a Researcher with the Texas Instruments Imaging R&D Lab, Dallas, TX, USA.

After getting the degrees from CMU, he joined the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA, as a Postdoctoral Research Associate. He is a Coauthor on best student paper awards for ECCV 2016 and NEMS 2018, and work from his FOCUS lab was a CVPR 2019 best-paper finalist. His interests span computer vision, computational photography and optics, novel cameras and sensors, 3D reconstruction, physics-based vision and active illumination.