PS^{2}F: Polarized Spiral Point Spread Function for Single-Shot 3D Sensing

Bhargav Ghanekar, Vishwanath Saragadam, Dushyant Mehra, Anna-Karin Gustavsson,
Aswin C. Sankaranarayanan, Ashok Veeraraghavan

Abstract—We propose a compact snapshot monocular depth estimation technique that relies on an engineered point spread function (PSF). Traditional approaches used in microscopic super-resolution imaging such as the Double-Helix PSF (DHPSF) are ill-suited for scenes that are more complex than a sparse set of point light sources. We show, using the Cramér-Rao lower bound, that separating the two lobes of the DHPSF and thereby capturing two separate images leads to a dramatic increase in depth accuracy. A special property of the phase mask used for generating the DHPSF is that a separation of the phase mask into two halves leads to a spatial separation of the two lobes. We leverage this property to build a compact polarization-based optical setup, where we place two orthogonal linear polarizers on each half of the DHPSF phase mask and then capture the resulting image with a polarization-sensitive camera. Results from simulations and a lab prototype demonstrate that our technique achieves up to 50% lower depth error compared to state-of-the-art designs including the DHPSF and the Tetrapod PSF, with little to no loss in spatial resolution.

Index Terms—Computational Photography, 3D Sensing, Microscopy, Phase Mask Design, Polarization-encoded PSFs

✦

1 INTRODUCTION

3D scanning is crucial to a wide range of applications, including microscopy [1], autonomous driving [2], and robot-assisted surgeries [3]. Among the multitude of approaches to measure depth, perhaps the hardest are those that involve passive, monocular, and snapshot measurements—a scenario that prioritizes compactness and time resolution. Estimating 3D information in this context relies on cues such as shading [4] or defocus [5, 6]; however, the underlying inverse problem is challenging and ill-posed.

A key property of real cameras is that the defocus blur changes with the depth of scene points, which has been leveraged by prior work to obtain depth. However, the blur produced by conventional pupils is not conducive to robust depth estimation; to mitigate this, there has been significant interest in the design of engineered pupil plane masks in the form of an amplitude [5], or phase mask [7, 8, 9, 10, 11, 12, 13]. In particular, the seminal work of Pavani *et al.* [9] has demonstrated that the so-called Double-Helix PSF (DHPSF), consisting of two lobes rotating about a center, can provide high spatial and depth resolutions, at least in the context of super-resolution localization microscopy.

The majority of prior work on engineered PSFs focus on isolated point light sources; this is a consequence of their intended application—namely super-resolved localization of fluorescent particles—and the simplification that this assumption provides. However, real-world scenes often consist of more complex geometric primitives such as lines, edges, and curves [14, 15]; a PSF optimized for point sources is inadequate at recovering such complex geometries due

to ambiguities in depth estimation. This is all the more important in applications involving linear structures including blood vasculature, neuronal networks (of the biological kind), and microtubules.

We propose and evaluate polarized spiral PSF (PS^{2}F), a novel engineered PSF that is well-suited for linear structures. Our key enabling observation is that asymmetric PSFs achieve higher depth accuracy for scenes comprising of linear structures. We achieve this asymmetry by observing that the DHPSF is comprised of two lobes that rotate about a common center; by capturing images with the individual lobes, we can reliably estimate depth. We propose a compact realization of our approach with a novel polarization-based imaging system. PS^{2}F is generated by combining a DHPSF mask with orthogonal linear polarizers on the two halves of the mask. We then use a polarization sensor, capable of measuring images along four polarization angles in a snapshot by using a Bayer-like tiling. Finally, the 3D scene is estimated by solving a depth-dependent deconvolution problem. An overview of depth estimation with PS^{2}F is illustrated in Fig. 1.

Contributions. We propose a new engineered PSF for imaging linear structures and make the following contributions.

- We theoretically demonstrate that the DHPSF is ill-suited for 3D imaging of linear structures by analysing its Cramér-Rao lower bound (CRLB).
- We propose a compact polarization-based setup, using a novel polarizer-phase mask encoding that is capable of better 3D estimation in a snapshot manner.
- We demonstrate the advantages of PS^{2}F through several experiments with our lab prototype over a wide range of scene geometries.

Limitations. At its core, PS^{2}F assumes that the scene intensity is unpolarized; if the scene is polarized, the resultant

- *B.G., V.S., A.V. are with the Department of Electrical and Computer Engineering, Rice University, Houston, TX.*
  E-mail: bhargav.ghanekar@rice.edu
- *D.M., A.-K.G. are with the Department of Chemistry, Rice University, Houston TX.*
- *A.C.S. is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh PA.*
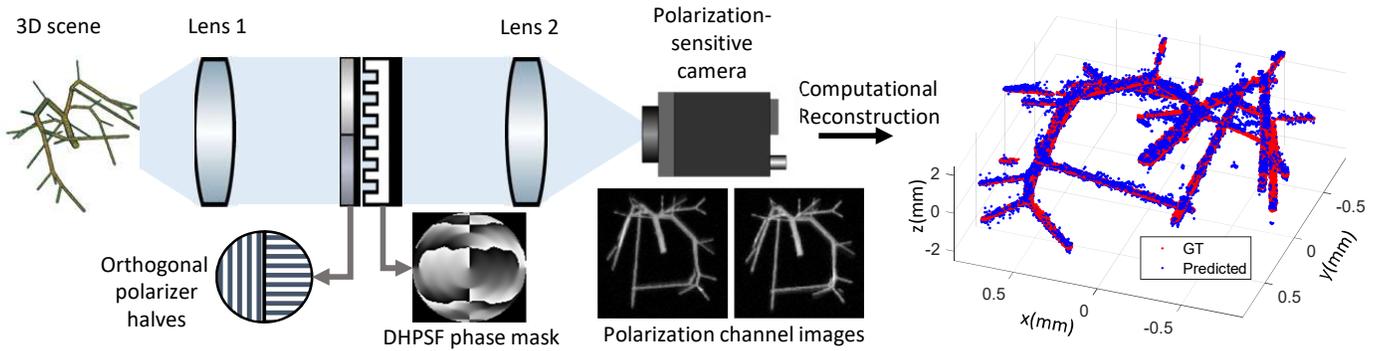
Fig. 1: **Polarized Spiral PSF (PS²F) for single-shot 3D scene estimation.** We propose a novel monocular depth imager based on an engineered PSF. Numerous real-world scenes can be modeled as linear structures. We observe that asymmetric PSFs, such as a single lobe of the DHPSF, enable us to accurately estimate the 3D geometry of such linear structures. We leverage this observation and propose a compact polarization-based setup that produces accurate depth maps in a snapshot manner.

depth estimate may not be accurate. Our technique, like most techniques that rely on defocus blur, performs poorly for regions with little or no textures or structures.

2 RELATED WORK

PS²F draws motivation from the prior work on engineered PSFs. We discuss the relevant ones in this section.

2.1 Mask-based PSF encoding

Encoding PSFs by designing masks has been used in several works, including extended depth-of-field imaging [16, 17, 18], 3D sensing in the macroscopic [6] and microscopic [7, 10] regimes, and in a lensless imaging setting [19]. The designed masks can be placed in front of the imaging lens or in the pupil plane of the imaging system. Earlier works include designing amplitude masks or coded apertures [5, 6] – such amplitude masks however lead to poor light efficiency and very coarse depth information. Several follow up works utilized phase-only modulation masks which include the rotating PSFs [7, 20, 21], and PSFs obtained by optimizing the Fisher information [10]. Phase masks have also been designed in an end-to-end optimization setting, both for macroscopic [22, 23] and microscopy [12, 13] tasks.

2.2 3D super-resolution microscopy using phase masks

Coupled with STORM [24] and PALM [25, 26], masks that result in rotating PSFs with depth have been able to perform super-resolution microscopy [7, 20] of individual fluorophore emitters to nanometer-scale accuracy. The most successful of such designs is the DHPSF [7], a popular choice for 3D localization of point sources [11, 24, 27]. 3D localization performance have been further improved by designing Tetrapod PSFs [10], which were generated by directly optimizing the Fisher information of the phase mask for (x, y, z) localization of point sources. Recently, deep learning techniques have enabled an end-to-end design framework that simultaneously learn phase masks as well as neural networks for obtaining 3D localizations from raw captures [12, 13].

Of particular interest is the work by Nehme *et al.* [13], where a pair of masks were designed in an end-to-end

learning process. The two masks are placed in two parallel $4f$ imaging channel. The mask pair, coupled with a learnt deep network, allows for high accuracy 3D localization of single fluorophores over a large depth range and for a high density of point-like emitters. Their work focuses on super-resolution microscopy and on 3D localization of *point-like* emitters, while our result tackles extended linear structures.

2.3 3D sensing of linear and extended structures

Human/animal bodies have a rich, dense 3D network of blood vessels running through them. Capturing the 3D structure of vasculature is usually done using variants of light-sheet microscopy (LSM) [28, 29, 30], confocal microscope (CM) [31, 32], optical projection tomography [33], and optical coherence tomographic angiography [34]. However, confocal and LSM methods involves scanning individual points, lines, or planes which makes the systems bulky and decreases the achievable imaging rate, making them non-ideal for fast dynamic 3D imaging.

2.4 Depth sensing using rotating PSFs

There is a rich body of work devoted to PSFs with lobes that rotate with defocus. These can be clustered into two groups based on their use of Gauss-Laguerre (GL) modes [9] or Fresnel rings [21].

GL-based rotating PSFs [7, 20] have been used for super-resolving and localizing single molecules in 3D [7, 35, 36, 37]. Both GL-based [38] and Fresnel-ring-based [39, 40] rotating PSFs have also been used for single-shot depth estimation of scenes. However, the first work employs a two phase-mask solution. The latter two employ a patch-based technique involving processing in the cepstrum domain. They only recover very simple scenes at low spatial and depth resolutions. Such recovery approaches do not extend to more complex geometry such as linear structures.

2.5 Using polarization channels for PSF encoding

There have been a few works that leverage polarization to help in 3D super-resolution imaging [13, 41, 42]. These works typically use a beam-splitter to split the incoming light into two separate $4f$ optical system channels which

enabled individual channels to be encoded with a different mask. The images were then captured over two non-overlapping regions of a single sensor, or on two separate sensors. Such solutions however call for bulkier optics, and more importantly, a need for sub-pixel alignment to achieve high spatial resolution. We instead leverage snapshot polarimetric cameras that are naturally well-aligned, and hence easier to work with, and more compact.

3 POLARIZED SPIRAL POINT SPREAD FUNCTION

We begin with a brief background on the DHPSF, an analysis of its properties, and its performance for linear/extended structures. Subsequently, we introduce our proposed PSF and its associated reconstruction technique.

3.1 Background: DHPSF and mask design

This subsection outlines the background on the generation of the DHPSF [7] and other GL-based rotating PSFs [20].

3.1.1 Phase mask for GL-based rotating PSFs

Any propagating paraxial wave/beam can be expressed in terms of the orthogonal basis of GL modes, which are characterized by two integers (n, m). Different combinations of the GL modes lead to varying properties of paraxial beams. In [43], it was shown that beams that constituted of GL modes that are in arithmetic progression continuously rotated and scaled upon propagation. Using a beam corresponding to GL modes of (1,1), (5,3), (9,5), (13,7), it was shown in [8] that the amplitude and phase profile of the beam at $z = 0$ can be used as a mask in front of a lens (or in the Fourier plane of a $4f$ lens system) to generate a PSF that rotates with defocus. Subsequently, this mask was implemented in a phase-only manner using a variant of Gerchberg-Saxton optimization procedure with constraints included for the GL modal domain [9]. The resulting phase function creates a PSF that rotates with changing point source depth, and is popularly known as the DHPSF due to its double helical structure. This PSF has been successfully used in 3D localization of fluorescent particles [7][35].

3.1.2 Properties of GL-based rotating PSFs

Piestun *et al.* [43] present a theory on paraxial rotating beams using GL modes, and characterizing the resulting rotation range, rotation rates, and beam scaling rates. Specifically, given a $4f$ system with lenses of focal length f, a GL-based mask designed with beam width w_0 for wavelength λ, and having GL modes lying on a single line with slope V_1, the following can be established:

- The total rotation in one direction of defocus is $(V_1 \pi / 2)$
- The angle of the rotating PSF

$$\phi(z) = \phi_0 + V_1 \tan^{-1} \left(\frac{z}{\lambda f^2 / \pi w_0^2} \right)$$

- The PSF will rotate by $(V_1 \pi / 4)$ over a depth of $\frac{\lambda f^2}{\pi w_0^2}$.

Even though phase-only masks for rotating PSFs are generated after an optimization procedure [9], the above rates and ranges are still a good approximation to understand the properties of any GL-based rotating PSF.
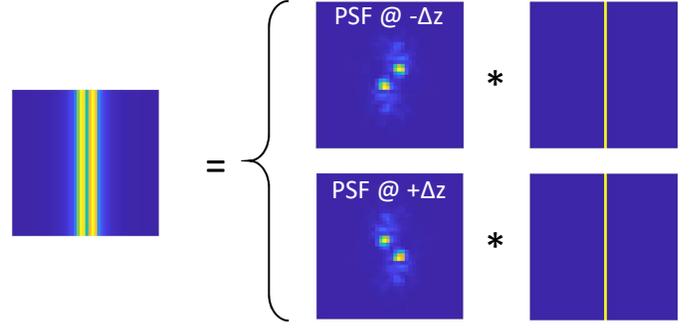
Fig. 2: **Ambiguities caused by DHPSF for lines**. Estimation of depth and orientation of a line using the DHPSF is an ill-conditioned problem. There is an inherent global ambiguity in depth estimation, since for every line image, two possible candidate depths are possible.

3.2 Challenges in imaging linear/extended structures

Phase masks such as the DHPSF [7], or the Tetrapod PSF [10], were designed for 3D estimation of *point sources*. To image scenes with extreme fluorophore density, the concept of a scene made up of point-like emitters is not accurate. With increasing fluorophore/light source density, scenes consist of edge-like and linear structures. Using a DHPSF for recovering depth of a line then leads to an ambiguity between depth and orientation of the line. This is illustrated in Fig. 2. Given a line image formed using the DHPSF, there will always be a global depth ambiguity—two different depths can give rise to the same line image. This makes the DHPSF ill-suited for imaging edge/linear structures.

3.3 Designing a PSF for 3D lines

A line with constant intensity blurred with an arbitrary PSF produces another line – thereby leading to loss of information along the line direction. In case of the DHPSF, the loss of information implies that two lines at two different depths ($-\Delta z$ and $+\Delta z$) produce the same measurement. This ambiguity can however be resolved if we measure two separate images with each lobe of the PSF. Imaging with the individual DHPSF lobes separately will remove the global degeneracy issues (up to a 2π lobe rotation range). Next, we provide a theoretical analysis of the advantages to be obtained by such a separation. For estimation from data (image(s) in our case), we can use theoretical tools to understand what ultimate precision is possible. The Fisher information and Cramér-Rao Lower bound (CRLB) give the best possible precision that can exist by any estimator. In the case of an image of an isolated point source, the Fisher information matrix for estimating parameters $\theta = (x, y, z)$ is given by [10]:

$$[FI(\theta)]_{ij} = \sum_{k=1}^{N_p} \frac{1}{\mu_\theta(k) + \beta} \left(\frac{\partial \mu_\theta}{\partial \theta} \right)_i \left(\frac{\partial \mu_\theta}{\partial \theta} \right)_j^T, \quad (1)$$

where N_p is the number of pixels, k is a variable for summing over all pixels, T corresponds to the transpose operator, β is the background Poisson noise level, and μ_θ is the scaled PSF image for a point source at $\theta = (x, y, z)$ (scaled by the number of photons N). The diagonal elements

of the inverse of the Fisher information matrix provides the CRLB for estimation of each of the parameters $\theta = (x, y, z)$.

$$[\text{CRLB}(\theta)]_i \equiv [FI(\theta)]_{ii}^{-1}$$

We next extend the Cramér-Rao analysis to imaging of lines. Given a line image (or a one-sided edge image), the parameters to be estimated are the associated depth z of the line/edge (at center of the patch), the associated orientation ϕ, i.e. $\theta = (z, \phi)$. Thus, given a line image ψ, the CRLB estimates for depth z is given by

$$[FI(\theta)]_{ij} = \sum_{k=1}^{N_p} \frac{1}{\psi_\theta(k) + \beta} \left(\frac{\partial \psi_\theta}{\partial \theta} \right)_i \left(\frac{\partial \psi_\theta}{\partial \theta} \right)_j^T \quad (2)$$

$$[\text{CRLB}(\theta)]_z \equiv [FI(\theta)]_{zz}^{-1} \quad (3)$$

The estimation of line orientation from a line path is easier than depth estimation, as orientation can be computed from the global structure of the line image. Hence, for simplicity of exposition, we focus on the analysis of the CRLB_z values here. Analysis regarding the CRLB_ϕ values is shown in the Supplementary. Using Eqns 2, 3 we can calculate the $\sqrt{\text{CRLB}_z}$ values for a given line image for all (z, ϕ). This calculation can be readily extended to line images captured using two separate PSFs. In such a case, the final Fisher information matrix is the sum of the individual Fisher matrices. Assuming a $4f$ system with 50 mm focal length lenses, and $N = 100,000$ photons, $\beta = 5$ photons/pixel, we compute the $\sqrt{\text{CRLB}_z}$ plots (as a function of z, ϕ) for the DHPSF [9], the Tetrapod PSF [10], and the individual DHPSF lobes with 25% photons in each lobe. Note that $\sqrt{\text{CRLB}_z}$ obtained with a mask pair should be compared with a single mask having 2x SNR (2x total photons), as shown in [13]. However, our polarizer-phase mask setup causes a 50% light loss (see further in Section 3). Hence, for an appropriate comparison, we compare the individual DHPSF lobe pair, with other PSFs having $4\times$ the SNR. There are possible ways to remove this loss, with a setup demonstrated in [13], which will further lower the estimated $\sqrt{\text{CRLB}_z}$ values.

Fig. 3 illustrates the $\sqrt{\text{CRLB}_z}$ values (log-scale) for a given line at depth z and having orientation ϕ. The mean and standard deviation of the $\sqrt{\text{CRLB}_z}$ values are also shown in the individual insets. The DHPSF plot in Fig. 3 shows several peaks with large magnitude. This occurs at specific (z, ϕ) where the line connecting the two lobes of the PSF and the scene line are either perpendicular or parallel to each other. The ability to discern lines at slightly differing depths is greatly reduced at these points due to symmetry. For the Tetrapod PSF [10] and the PhaseCam3D PSF [22], the plots in Fig. 3 also show more peaks and ridges, but with lower magnitude compared to the DHPSF. For the PSF pair of the individual DHPSF lobes, we see that most of the peaks are removed, except for a central peak in the plot. The peak corresponds to the case when the PSF is in focus and the the line orientation is perpendicular to the PSF lobes. This central peak does not affect the effectiveness of separating lobes, because it is a narrow peak over a only small area in the parameter space of (z, ϕ). The average $\sqrt{\text{CRLB}_z}$ values for the individual DHPSF lobe pair is similar to the DHPSF with 4x SNR. Thus, with 4x lower SNR, the PS^{2}F is able to achieve a comparable $\sqrt{\text{CRLB}_z}$ value with other PSFs,
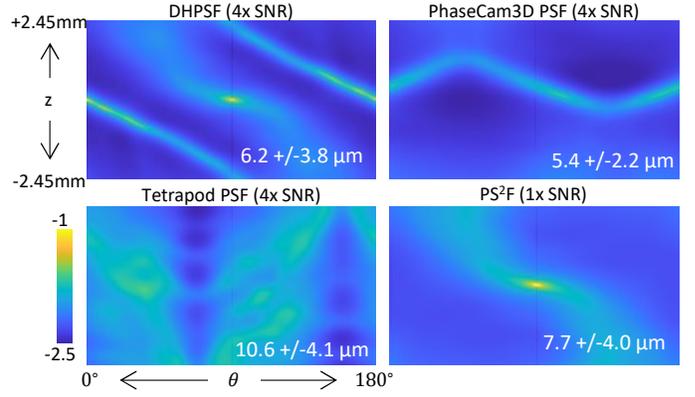
Fig. 3: **CRLB comparison for various phase masks.** The figure shows $\sqrt{\text{CRLB}_z}$ log-intensity (log-mm) plots for line images, as a function of line depth (z) and line orientation (θ). PS^{2}F achieves a comparable CRLB over the whole range of depth and orientations with fewer peaks, even with 4x lower SNR. In contrast, competing approaches have curves with large CRLB values, which introduce ambiguities in the depth estimate.

and does not show consistent peaks and ridges as seen in the plots. Note that CRLB calculations only provide insight about local ambiguities or precision levels. However, we also observe that a PSF pair created out of separating the DHPSF lobes removes global ambiguities as well. Thus overall, a PSF pair created from the individual DHPSF lobes is better for estimating depths of line/edge patches.

3.4 Using polarizers to separate out the DHPSF lobes in a single $4f$ optical system

Imaging with a PSF pair is possible by having two separate parallel $4f$ imaging systems with separate sensors, as demonstrated in [13]. However, the DHPSF mask (and the Fresnel ring-based two-lobe masks) enjoy the property of PSF lobe separability. Partitioning such a mask into two halves along a particular axis, and allowing light through only one half leads to the creation of only a single rotating PSF lobe, as depicted in Fig. 4. The primary reason is that the light falling on one half is modulated to form one lobe, while the light falling on the other half is modulated to form the other lobe. Such a partitioning leads to partitioning of the DHPSF into two separate, distinct lobes.

The separability of DHPSF lobes can be leveraged to build a compact imaging system. In PS^{2}F, we add two linear polarizers on each mask half, one oriented along s-polarization and the other oriented along p-polarization. This ensures that the s-polarization component is modulated using one mask half, and consequently the p-polarized component is modulated using the other mask half. The PSF created out of such a polarizer-phase mask encoding contains two lobes, but with each lobe being in a different polarization (either s or p). This addition of polarizers causes a 50% light loss, which was taken into consideration while estimating the $\sqrt{\text{CRLB}_z}$ values. Such a *polarized spiral PSF* can be captured efficiently using a polarized camera sensor. A polarized camera sensor contains a 2D array of pixels with a Bayer pattern that has 90, 45, 135, 0 degree polarizers. After addition of the polarizers in the mask-plane, the orthogonally polarized DHPSF lobes can be im-
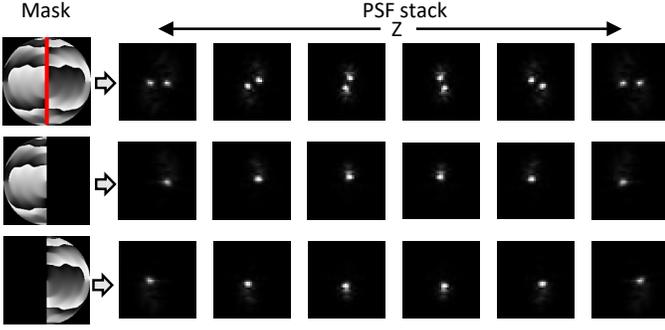
Fig. 4: **Effect of separating the phase mask.** The phase mask for generating the DHPSF has a special property that separating the phase profile along the highlighted red line yields a separation of the two lobes. We leverage this to implement a compact realization of Polarized Spiral PSF (PS^{2}F) using a polarization camera sensor and two polarizers in the pupil plane.

aged in their respective channels of a polarization sensor. This allows for a compact $4f$ system, allowing the imaging of a PSF pair using just a single $4f$ system channel, making the proposed method a single-shot, single-sensor method.

3.5 Imaging model and reconstruction procedure

We assume that the light from the scene is unpolarized in our imaging model. The imaging of a 3D scene with a depth-dependent PSF can be approximated as a sum of 2D convolutions between the depth-dependent PSF and the per-plane scene intensity:

$$I_c(x,y) = \sum_{z=z_{start}}^{z_{end}} h_c(x,y;z) * s(x,y,z) \quad (4)$$

where $I_c(x,y)$ is the image intensity at (x,y) in polarization channel c, $h_c(x,y;z)$ is the 2D PSF corresponding to depth z and polarization channel c, and $s(x,y,z)$ is the scene intensity at point (x,y,z).

The goal of 3D reconstruction is to estimate a 3D matrix $s(x,y,z)$ from the noisy measurements of $I(x,y)$. There are numerous ways to solve this problem including regularized least-squares [19, 44, 45], data learning-based [22, 23], and more recently, ones based on deep network-based regularizer [46]. To keep the exposition simple, we demonstrate recovery by modelling the 3D reconstruction problem as a regularized least squares optimization problem.

Specifically, we formulate the 3D scene estimation problem as:

$$\operatorname*{argmin}_{\mathbf{x}} \left\| \begin{bmatrix} I_0 \\ I_{90} \end{bmatrix} - S \begin{bmatrix} H_0 \\ H_{90} \end{bmatrix} \mathbf{x} \right\|_2^2 + \lambda_{TV} \|\Psi(\mathbf{x})\|_1 + \lambda_{L1} \|\mathbf{x}\|_1 \quad (5)$$

where $\mathbf{x}$ is a 3D matrix of scene intensities, H_0, H_{90} are the depth-specific PSF operators corresponding to the individual two lobes that are polarized to $0°$ and $90°$ states respectively, and I_0, I_{90} are captured images in the $0°$ and $90°$ polarization channels respectively. S is the summing operator that sums across the depth channels of $H_0\mathbf{x}$ and $H_{90}\mathbf{x}$ separately. We employ TV and L1 regularizers as scene priors, with $\lambda_{TV}, \lambda_{L1}$ as hyperparameters to control their regularization effects respectively. We solve the optimization problem using autograd functionality in PyTorch and the

Adam optimizer [47] to leverage the speed of graphical processing units (GPUs).

A key requirement for our technique is that the scene emits unpolarized light. This is required so as to ensure that both the lobes have produce PSFs with similar intensity levels that we can calibrate a priori. When the incident light is polarized, the blur kernels in Eq. 4 will have an unknown scaling, which leads to a model mismatch; in an extreme scenario, we lose all the information in one of the lobes if the polarization angle of the incident light is orthogonal to the corresponding polarizer in the pupil plane. While this is a limitation of our technique, we largely encountered unpolarized light in our intended application of fluorescence microscopy.

4 SIMULATIONS

To evaluate the performance of our proposed PS^{2}F, we perform extensive simulations on the VascuSynth 2013 dataset [48, 49], which provides 10 simulated 3D volumes of vascular trees for each of 12 levels of complexity. Scene complexity is set by the number of bifurcations in the vascular tree (ranging from 1 to 56 in steps of 5). Each 3D vasculature volume in the dataset consists of $101 \times 101 \times 101$ voxels. These volumes were converted into 3D scenes of size $256 \times 256 \times 256$ through interpolation methods. This new volume was assumed to cover a volume of $1.76 \times 1.76 \times 5.00$ mm^3 with each voxel occupying a $6.9 \times 6.9 \times 19.5$ μm^3.

We further assume a $4f$ system with both lenses having focal lengths $f = 50$ mm. We render images of the vascular structures using several different masks of diameter 3 mm in the Fourier plane. We assume that the light from the vasculature is monochromatic of wavelength $\lambda = 532$ nm. We render images assuming occlusion of vasculature and that light is received only from the surface of the vasculature. We assume that the center of the 3D scene is $f = 50$ mm away from the first lens, thus, the scene spans a defocus range of $[-2.5 \text{ mm}, +2.5 \text{ mm}]$. The masks (and their corresponding PSFs) used in rendering the simulated images were:

- DHPSF: We optimize for a GL-based DHPSF mask (following procedure in [9]), based on the rotating paraxial beam with $w_0 = 0.4$ mm and having GL modes corresponding to $(1,1), (3,5), (5,9), (7,13)$. Full 180 degree lobe rotation was achievable over $2\frac{\lambda f^2}{\pi w_0^2} = 5.3$ mm.
- PS^{2}F: Using the above mask, we construct an equivalent mask pair corresponding to PS^{2}F by partitioning the mask into two halves appropriately.
- Tetrapod PSF: We obtain a Tetrapod PSF mask designed for 550 nm wavelength from [36] and repurpose it for imaging over our specified 5 mm depth range.

We then added Poisson and Gaussian read-out noise to the final measurements. For an appropriate comparison, we perform 3D reconstructions with individual single lobes (PS^{2}F), two-shot capture with DHPSF, and a two-shot capture using the Tetrapod PSF. To accurately account for the 50% light loss in the PS^{2}F case, we also render out images with half the signal level as compared to the signal levels in DHPSF and Tetrapod PSF case. After obtaining the 3D volumetric estimate, we filter out points whose sum across the z-stack is lower than a fixed threshold. We then estimate depth using
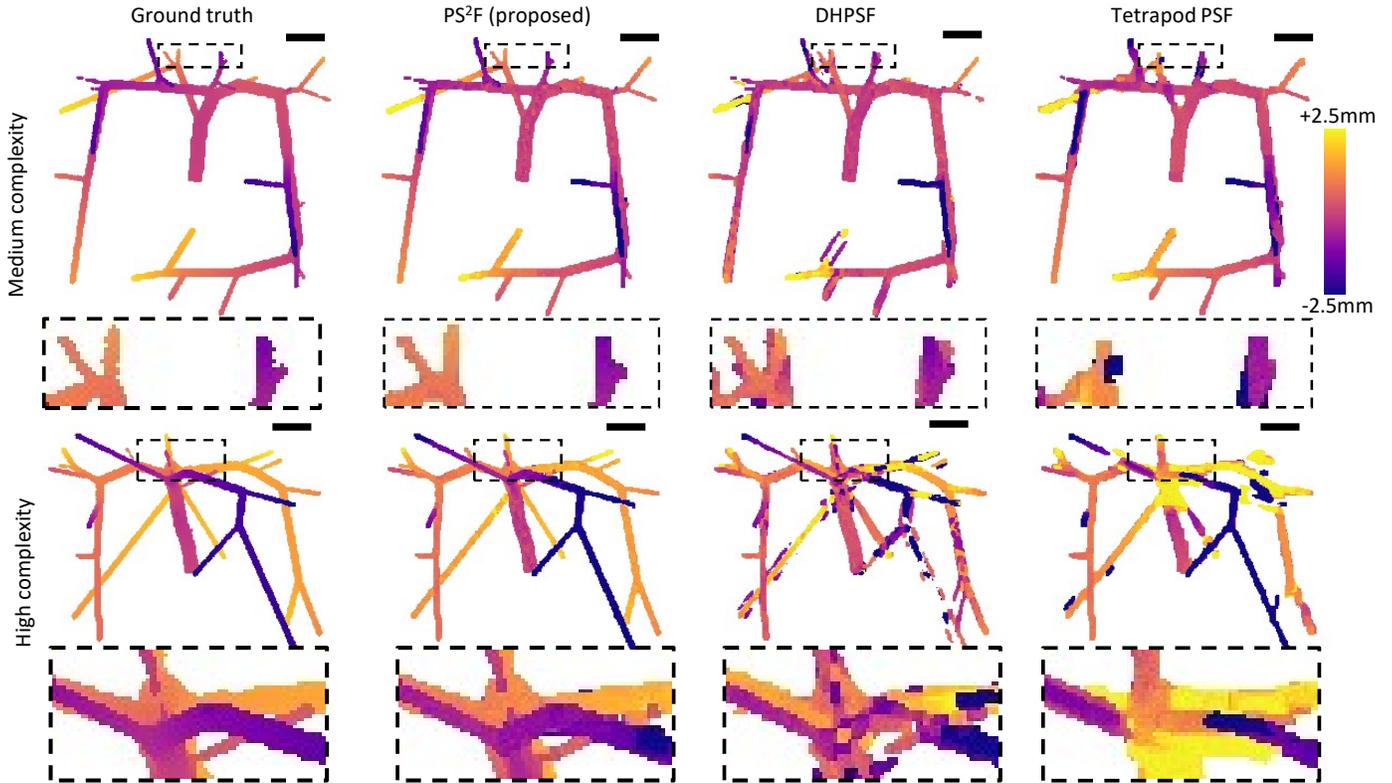
Fig. 5: **Simulations on vascusynth dataset with varying complexity.** The figure shows reconstructions with various PSF designs on scenes with medium complexity (top row) and high complexity (bottom row). PS²F is superior in depth estimation, especially in regions where there is higher vasculature (scene) complexity (see insets). Scale bars indicate 0.25 mm.
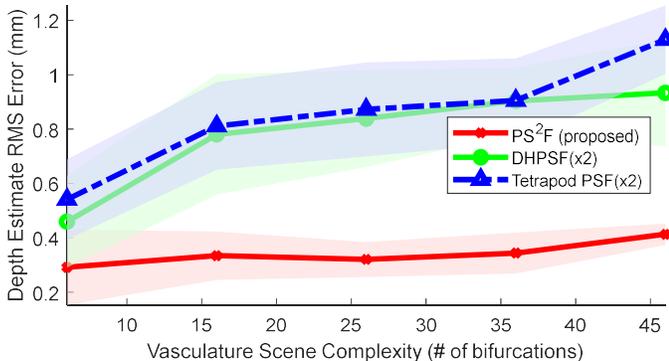
Fig. 6: **Accuracy vs. scene complexity**. Depth estimation performance, for the proposed PS²F and other comparing PSFs (averaged over 10 scenes for each complexity level, shaded regions show variance across the 10 scenes). This demonstrates the suitability of PS²F, and that the separation of the DHPSF lobes helps in 3D reconstruction.

the index corresponding to a maximum-intensity-projection (MIP) of the 3D estimate.

4.1 Comparison with DHPSF and Tetrapod PSF

Fig. 6 shows the depth estimation performance for the proposed PS²F, DHPSF, and the Tetrapod PSF. We compare the PSFs depth estimation using the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the multi-scale Structural Similarity Index Measure [50] (MS-SSIM) metrics. The entire set of results can be seen in the Supplementary.

The proposed PS²F method performs $\sim 2\times$ better in terms of the RMS error for depth estimation. The depth estimation performance improvement is even higher ($\sim 3\times$) for vasculature with more bifurcations, i.e. greater complexity. Depth estimation results for two example vasculature scenes can be seen in Fig. 5. In Fig. 5 (top row), the vasculature scene is not very complex. The depth map estimation results are fairly similar for all the three PSFs (PS²F, DHPSF, Tetrapod PSF), but the zoom-in insets highlight the errors seen in the DHPSF and Tetrapod PSF depth estimation. In Fig. 5 (bottom row), the scene has greater complexity, and the PS²F produces much better depth estimation than the other two PSFs. The zoom-in insets especially highlight the same.

4.2 Resolution performance

Furthermore, we analyse the reconstruction resolution performance by reconstructing a simulated skew USAF target. We assume a USAF target places at the focus plane, but tilted with respect to the optical axis so that its depth changing horizontally across the target; we render using the same three masks in the same $4f$ optical system as above. The results are shown in Fig. 7. The proposed PS²F generates a better reconstruction than the DHPSF and Tetrapod PSF reconstructions, both in lateral and axial dimensions. As seen from the XZ MIP plots in Fig. 7, the PS²F and Tetrapod PSF reconstructions are able to obtain the skew angle of the target, with the former being much better. The DHPSF fails to get the skew angle, due to the global depth ambiguity issues explained in Section 3. In the lateral dimensions, the Tetrapod PSF is unable to reconstruct Group 4 and 5. This
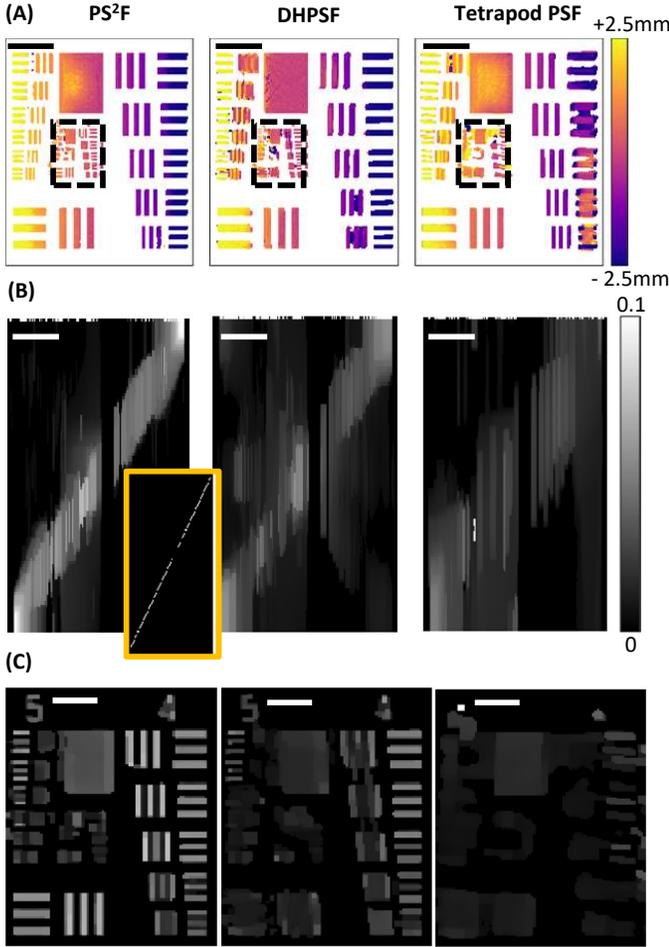
Fig. 7: **Simulated reconstruction results of a USAF target at a skew angle.** (A) Shows the depth map estimates for PS²F, DHPSF, and Tetrapod PSF respectively. The XZ maximum intensity projections are shown in (B) in their respective columns, with ground truth shown in the inset. (C) shows zoom-in insets of depth maps (after thresholding to remove small artifacts) of highlighted regions in (A). PS²F has a better reconstruction in lateral and axial dimensions as compared to the DHPSF and the Tetrapod PSF. Scale bars for (A), (B), (C) indicate 0.5 mm, 0.5 mm, and 0.1 mm respectively.

could be attributed to the fact that the PSF has a larger support. On the other hand, the PS²F and DHPSF have concentrated lobes, and thus are able to reconstruct elements of Group (5,3) (12.40 μm linewidth) and (5,2) (13.92 μm linewidth) respectively.

The obtained xy-resolution can be theoretically justified. The separation of lobes in the PS²F method allows for resolution of finer scene features due to the compactness of the PSF. The resolution will be roughly determined by the width of the individual PSF lobes. Fitting a 2D Gaussian to each individual lobe, the average 2σ-value for the Gaussian fit is ~ 1.74 pixels $= 12\ \mu$m in the object space. This readily matches with the fact that the PS²F can resolve elements of Group (5,3), which has line elements that are spaced at $12.40\ \mu$m distance. The diffraction-limited resolution of an Airy disk PSF for the same optical system configuration is given by $1.22\lambda f/D = 10.81\ \mu$m, indicating that the xy-resolution obtained by the PS²F is close to the diffraction-
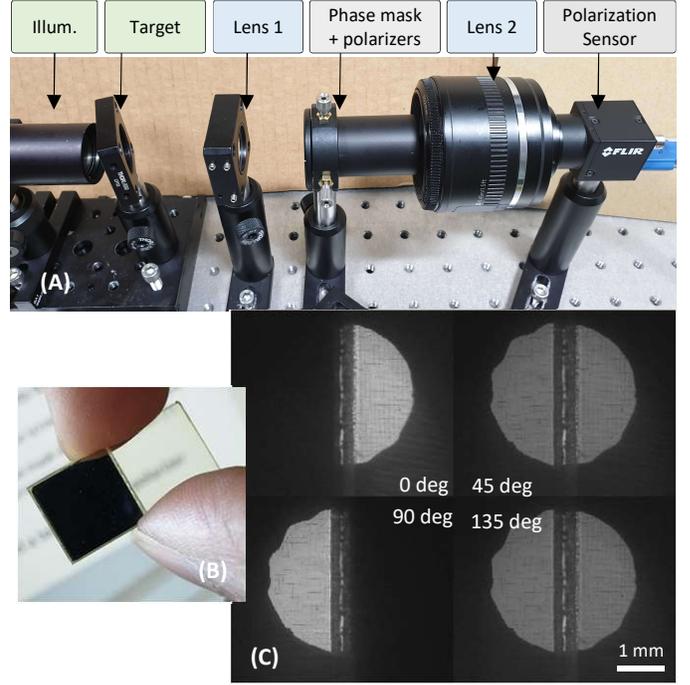
Fig. 8: **Lab prototype.** (A) Our lab prototype is a $4f$ system consists of a polarization sensor (FLIR BFS-U3-51S5P-C) equipped with a 50 mm F-mount lens. The phase mask and polarizers are on an XY stage. An achromat (AC254-050-A) forms the other half of the $4f$ system. We used an additional z-axis motorized translation stage to capture the PSF stack. (B) Two orthogonal polarizer pieces. (C) Images of the polarizers and phase mask with a camera focused on the pupil-plane.

limited resolution ($1.15\times$) as well.

For PS²F, we also analysed the spread of reconstruction signal across the z-channel. For pixels with a significant signal level, we fit a 1D Gaussian curve to estimate this spread. The median 2σ-value across z-channel obtained was $2 \times 304 = 608\ \mu$m, which is an estimate of the axial resolution performance of the proposed PS²F. Note that the diffraction-limited axial performance of an Airy disk PSF for the same optical system configuration is given by $4\lambda(f/D)^2 = 591\ \mu$m. We are able to achieve $\sim 1.03\times$ the diffraction-limited axial resolution with the proposed PS²F.

5 EXPERIMENTAL RESULTS

We perform real-world imaging with PS²F to verify our simulations. In this section we detail out the experiments and compare the PS²F with the corresponding DHPSF.

5.1 Imaging setup description

We built a $4f$ imaging setup with the first lens being an achromatic lens with focal length 50 mm, and the second lens being a Canon camera lens with effective focal length $= 50$ mm. The polarizer-phase mask is added to the pupil plane of the $4f$ system. Fig. 8 shows the prototype.

5.1.1 Polarizer-Mask design

We fabricated a 3 mm diameter DHPSF mask using a 3-D printer employing two-photon lithography. See Supplementary for more details on fabrication. The fabricated mask was

Distance from in-focus plane
-2.5mm ← → +2.5mm

0 deg

45 deg

90 deg

135 deg
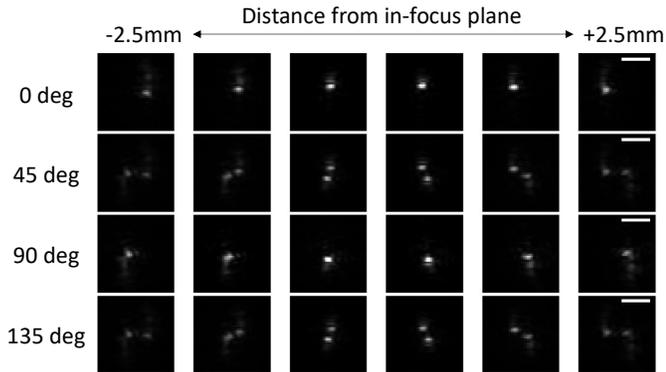
Fig. 9: **PSF stacks captured with a lab prototype.** We achieved high quality separation between the two lobes in 0 and 90 degree images with our setup. We noticed no significant effect of any misalignment between the phase mask and the two polarizer halves. Scale bars indicate $100\ \mu$m.

the same mask used in the Section 4 to perform simulations of the VascuSynth dataset. To create a single-mask, single-sensor design for the PS^{2}F as outlined in Section 3, two rectangular pieces of thin-film polarizers were laser cut and put side-by-side in a 3D printed holder. Both the polarizer pieces were laser cut such that their polarization orientation was $0\deg$ and $90\deg$ respectively. The fabricated DHPSF mask was attached to the back side of the 3D printed holder, carefully aligning the two mask halves w.r.t. to the polarizer edges under a bright-field 4x microscope. This polarizer-mask aperture is illustrated in Fig. 8(B),(C). Further details can be found in the Supplementary figures.

5.1.2 PSF capture and calibration

To reconstruct 3D scenes, we use an experimentally captured PSF stack for the PS^{2}F. Using the polarization sensor, we image a $5\ \mu$m pinhole over a range of $[-2.5\ \text{mm}, +2.5\ \text{mm}]$ on either side of the focus plane. The experimentally-captured PSFs can be seen in Fig. 9.

5.2 Imaging experiments

With the fabricated mask and optical setup, we perform a series of experiments with 3D skew planar targets (using transmissive illumination) and with 3D fluorescence targets. We image using a polarization camera sensor (FLIR BFS-U3-51S5P-C) to obtain the individual PS^{2}F lobes in two polarization channels. By averaging across the Bayer pattern, we can readily obtain the image of the scene as imaged by a DHPSF, which we use for comparison purposes.

5.2.1 3D skew planar targets

We imaged a USAF planar target placed at a skew angle, and then performed single-shot 3D reconstruction of the target. Furthermore, we obtain a focal stack of the target so as to obtain ground truth depth map. The ground truth depth map was obtained from the focal stack using the Shape-from-focus method [51]. For the FOV being imaged, the relative intensity of each of the PS^{2}F lobes changed across the FOV, causing a spatial variance in the PSF. This was readily accounted for by adding a per-pixel, per-channel weight

term to the 3D scene variable **x** in Eq. 5, and jointly optimizing for the 3D scene **x** and the weights. The results are depicted in Fig. 10. After modelling of the relative strengths of the PS^{2}F lobes as variables to be estimated during the optimization, we achieved an accurate reconstruction result matching the simulation results in terms of resolving Group (5,3) elements for the PS^{2}F. The YZ maximum intensity projections demonstrate that the skew plane of the USAF has been correctly estimated using the PS^{2}F, as compared to using the DHPSF. When judged against the ground truth, we see a $\sim 50\%$ lower RMS error achieved with the PS^{2}F as compared to the DHPSF.

5.2.2 3D fluorescence imaging

We, next, imaged a set of prepared fluorescent samples, consisting of $10\ \mu$m fluorescent beads in a PDMS substrate (for more information, see Supplementary). The bead concentration was varied from approximately 3 to 8 beads/mm^3. The resulting samples were ~ 2 mm in thickness. We imaged these samples using the same mask as before, but in a $4f$ optical system using lenses of focal length 30 mm and 50 mm resulting in a $1.6\times$ magnification. We reconstructed using the same optimization algorithm as specified in Eq. 5, and compared against ground truth data obtained with a confocal microscope. The results are shown in Fig. 11. Bead localization performance using the DHPSF seems to get relatively worse as the bead concentration increases. In contrast, PS^{2}F enabled accurate results with increasing bead concentration. It is important to note that in this experiment, we use an optimization objective to perform bead reconstruction. More sophisticated and accurate methods (such as likelihood-based methods) exist for reconstruction of beads or point-like objects. Moreover, the ground truth (GT) data was captured using confocal imaging method. However, due to the complexity of matching/registering imaging volumes, we only obtain an approximated alignment or approximate registration between the GT and the imaging volume captured.

Using the same optical setup as above, we also image fluorescent cotton strands suspended in a PDMS sample (see Supplementary for more information about sample preparation). Fig. 12 shows the results for the same. PS^{2}F method is better able to capture the 3D structure of the fluorescent strands - especially in parts with highly complex geometry. Imaging with PS^{2}F results in sharper reconstructions, which helps in reconstructing the fine, complex structure of the strands. The estimated depth map from the PS^{2}F is also more accurate as compared to the DHPSF case.

6 CONCLUSION

We presented PS^{2}F for single shot, monocular depth estimation of extended (linear) structures where we separated the two lobes of the DHPSF into the two orthogonal states of polarization. This separation enabled us to remove ambiguities when estimating depth of line segments by breaking the inherent symmetry of DHPSF with respect to depth. The snapshot capabilities of PS^{2}F will enable faster microscopic imaging, including high resolution light sheet microscopy at real-time rates with fewer captures. We designed and demonstrated a compact physical realization of PS^{2}F with a
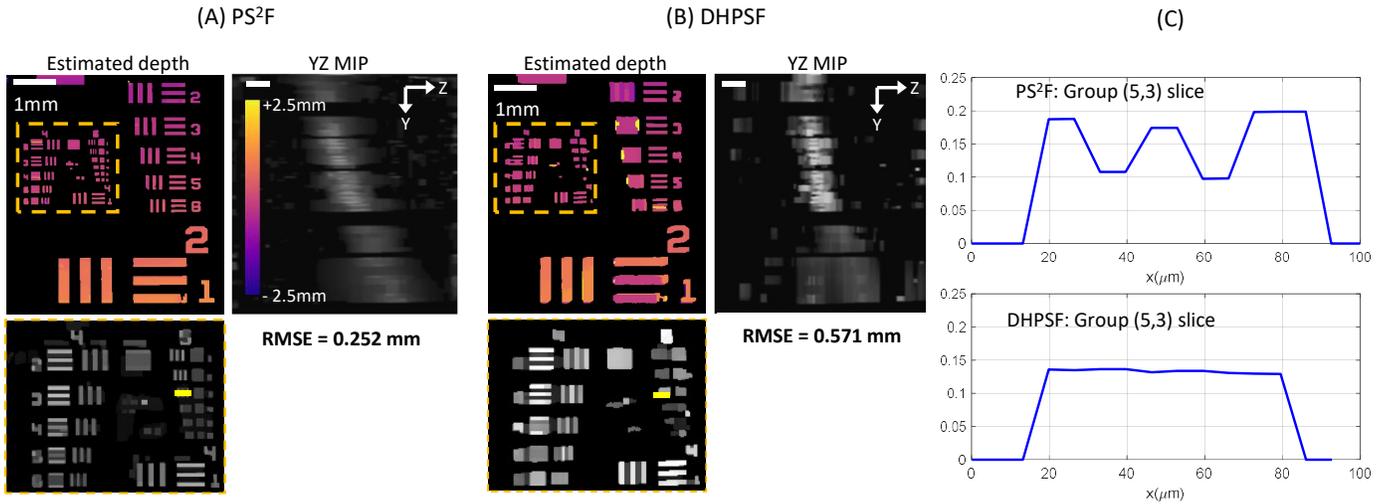
Fig. 10: **Skew USAF target reconstruction.** Showing reconstructed depth map, YZ maximum intensity projection, zoom-in inset of highlighted ROI for (A) PS^{2}F, and for (B) DHPSF. Column (C) shows slice plots for marked yellow lines in the zoom-in insets, corresponding to Group (5,3) elements for PS^{2}F (top plot) and DHPSF (bottom plot). We were able to resolve up to Group (5,3) elements (12.40 μm linewidth) using the PS^{2}F, whereas only up to Group (4,5) elements (19.69 μm linewidth) using the DHPSF. Text below YZ MIP plots indicate the depth RMS error obtained for reconstruction when compared to ground truth, showing that the PS^{2}F performs $\sim 2\times$ better. All scale bars indicate 1 mm.
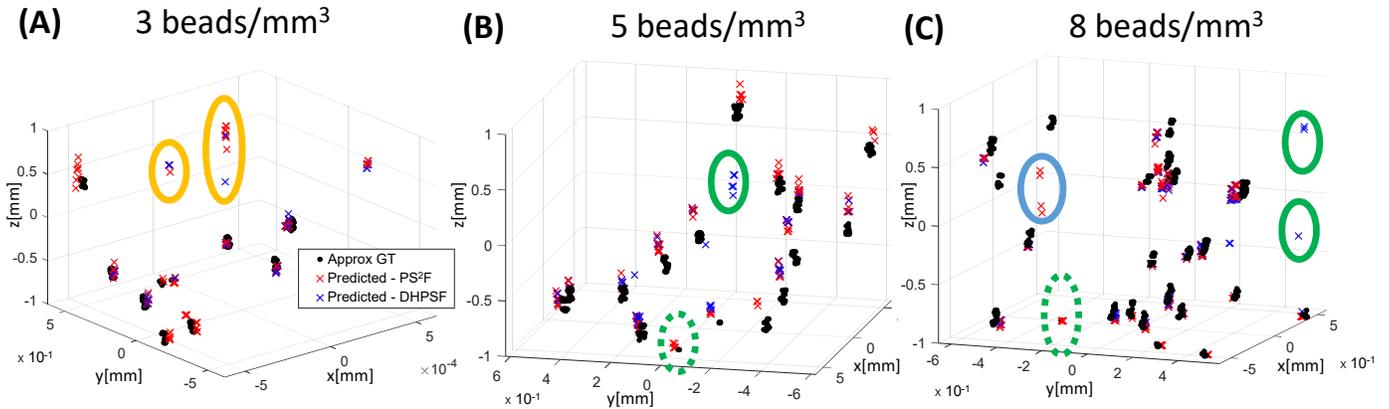
Fig. 11: **3D reconstructed fluorescent bead volumes.** The scatter plots show 3D reconstructed views with PS^{2}F and the DHPSF for bead densities approximately corresponding to (A) 3 beads/mm^3, (B) 5 beads/mm^3, and (C) 8 beads/mm^3. We captured ground truth information with a semi-aligned confocal ground truth (hence an approximate ground truth). Highlighted regions indicate various errors: incorrect estimation by DHPSF (solid green), missing estimation by DHPSF (dashed green), incorrect estimation by PS^{2}F (solid blue), and incorrect estimation by both PSFs (solid yellow). Axes values are in units of mm.

single, polarization-sensitive camera, and showed that PS^{2}F results in $2\times$ or higher accuracy compared to current state-of-the-art phase masks such as DHPSF and the Tetrapod PSF. We believe that our approach of leveraging polarization multiplexed phase masks combined with polarization-sensitive cameras opens new avenues for high resolution snapshot depth imaging at micro and macro scales.

PS^{2}F is inherently designed for extended/linear structures, and imaging such structures will involve high photon counts than when imaging single molecules or point-like emitters. Hence, we do not specifically deal with a low-photon count scenario in this work. Nevertheless, PS^{2}F can be adapted to low-photon count scenarios by using advanced recovery techniques including deep network-based approaches. Improvements such as modeling spatially varying PSF, as well as accounting for polarized input light will further improve PS^{2}F results.

REFERENCES

[1] R. S. Fischer, Y. Wu, P. Kanchanawong, H. Shroff, and C. M. Waterman, "Microscopy in 3d: a biologist's tool-box," *Trends in Cell Biology*, vol. 21, no. 12, pp. 682–691, 2011.
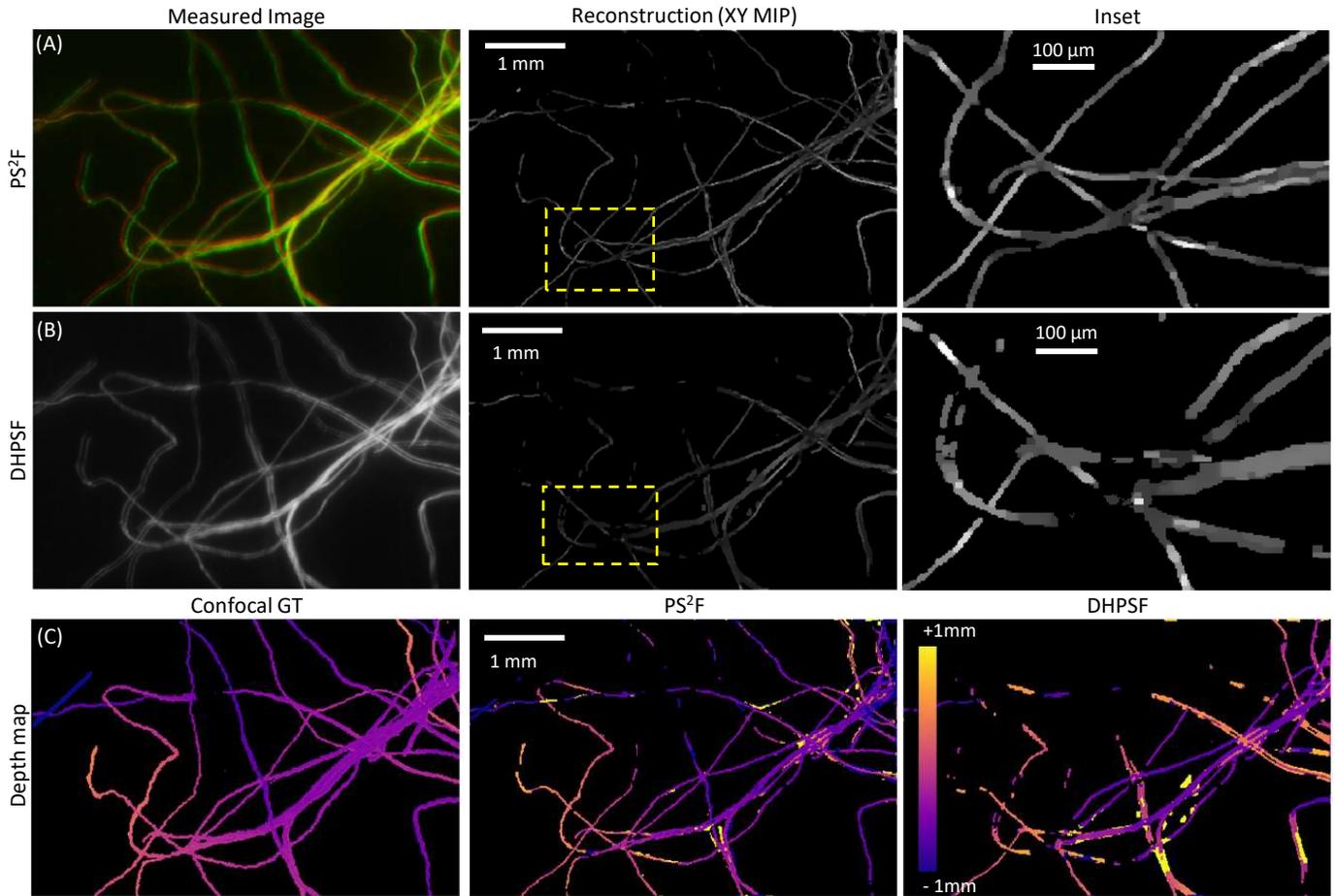
Fig. 12: **Experimental results on fluorescent cotton strands**. (A) PS^{2}F and (B) DHPSF. First column in (A), (B) show the captured scene images (polarization channels in the PS^{2}F image are shown in red and green colors). Second column in (A), (B) shows the reconstructions. Highlight ROIs (in yellow) are shown in the last column of (A), (B) – depicting the inaccuracy of DHPSF, as well as the ability of PS^{2}F to reconstruct scenes with complex linear structures. In the last row (C), columns (from left-to-right) show color-coded depth maps obtained from confocal imaging (treated as ground truth), PS^{2}F reconstructions, and DHPSF reconstructions respectively. Note the confocal GT depth map is an approximate depth map, due to imperfect registration between GT and the imaged volumes.

[2] E. Arnold, O. Y. Al-Jarrah, M. Dianati, S. Fallah, D. Oxtoby, and A. Mouzakitis, "A survey on 3d object detection methods for autonomous driving applications," *IEEE Trans. Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3782–3795, 2019.

[3] A. Reiter, A. Sigaras, D. Fowler, and P. K. Allen, "Surgical structured light for 3d minimally invasive surgical imaging," in *IEEE Intl. Conf. Intelligent Robots and Systems*, 2014.

[4] T. Ping-Sing and M. Shah, "Shape from shading using linear approximation," *Image and Vision Computing*, vol. 12, no. 8, pp. 487–498, 1994.

[5] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Trans. Graphics*, vol. 26, no. 3, pp. 70–es, 2007.

[6] C. Zhou, S. Lin, and S. Nayar, "Coded aperture pairs for depth from defocus," in *IEEE Intl. Conf. Computer Vision (ICCV)*, 2009.

[7] S. R. P. Pavani, A. Greengard, and R. Piestun, "Three-dimensional localization with nanometer accuracy using a detector-limited double-helix point spread function system," *Appl. Physics Letters*, vol. 95, no. 2, p. 021103, 2009.

[8] A. Greengard, Y. Y. Schechner, and R. Piestun, "Depth from diffracted rotation," *Optics Letters*, vol. 31, no. 2, pp. 181–183, 2006.

[9] S. R. P. Pavani and R. Piestun, "High-efficiency rotating point spread functions," *Optics Express*, vol. 16, no. 5, pp. 3484–3489, 2008.

[10] Y. Shechtman, S. J. Sahl, A. S. Backer, and W. E. Moerner, "Optimal point spread function design for 3d imaging," *Physical Review Letters*, vol. 113, no. 13, p. 133902, 2014.

[11] M. Badieirostami, M. D. Lew, M. A. Thompson, and W. Moerner, "Three-dimensional localization precision of the double-helix point spread function versus astigmatism and biplane," *Appl. Physics Letters*, vol. 97, no. 16, p. 161103, 2010.

[12] E. Nehme, D. Freedman, R. Gordon, B. Ferdman, L. E. Weiss, O. Alalouf, T. Naor, R. Orange, T. Michaeli, and Y. Shechtman, "Deepstorm3d: dense 3d localization

microscopy and psf design by deep learning," *Nature Methods*, vol. 17, no. 7, pp. 734–740, 2020.

[13] E. Nehme, B. Ferdman, L. E. Weiss, T. Naor, D. Freedman, T. Michaeli, and Y. Shechtman, "Learning an optimal psf-pair for ultra-dense 3d localization microscopy," *arXiv preprint arXiv:2009.14303*, 2020.

[14] T. S. Cho, S. Paris, B. K. Horn, and W. T. Freeman, "Blur kernel estimation using the radon transform," in *IEEE Comp. Vision and Pattern Recognition (CVPR)*, 2011.

[15] N. Joshi, R. Szeliski, and D. J. Kriegman, "Psf estimation using sharp edge prediction," in *IEEE Comp. Vision and Pattern Recognition (CVPR)*, 2008.

[16] A. Castro and J. Ojeda-Castañeda, "Asymmetric phase masks for extended depth of field," *Appl. Optics*, vol. 43, no. 17, pp. 3474–3479, 2004.

[17] Q. Yang, L. Liu, and J. Sun, "Optimized phase pupil masks for extended depth of field," *Optics Communications*, vol. 272, no. 1, pp. 56–66, 2007.

[18] P. Zammit, A. R. Harvey, and G. Carles, "Extended depth-of-field imaging and ranging in a snapshot," *Optica*, vol. 1, no. 4, pp. 209–216, 2014.

[19] V. Boominathan, J. K. Adams, J. T. Robinson, and A. Veeraraghavan, "Phlatcam: Designed phase-mask based thin lensless camera," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 42, no. 7, pp. 1618–1629, 2020.

[20] M. D. Lew, S. F. Lee, M. Badieirostami, and W. Moerner, "Corkscrew point spread function for far-field three-dimensional nanoscale localization of pointlike objects," *Optics Letters*, vol. 36, no. 2, pp. 202–204, 2011.

[21] S. Prasad, "Rotating point spread function via pupil-phase engineering," *Optics Letters*, vol. 38, no. 4, pp. 585–587, 2013.

[22] Y. Wu, V. Boominathan, H. Chen, A. Sankaranarayanan, and A. Veeraraghavan, "Phasecam3d—Learning phase masks for passive single view depth estimation," in *IEEE Intl. Conf. Computational Photography (ICCP)*, 2019.

[23] J. Chang and G. Wetzstein, "Deep optics for monocular depth estimation and 3d object detection," in *IEEE Intl. Conf. Computer Vision (ICCV)*, 2019.

[24] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science*, vol. 319, no. 5864, pp. 810–813, 2008.

[25] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[26] S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophysical J.*, vol. 91, no. 11, pp. 4258–4272, 2006.

[27] M. F. Juette, T. J. Gould, M. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Bewersdorf, "Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples," *Nature Methods*, vol. 5, no. 6, pp. 527–529, 2008.

[28] J. M. Girkin and M. T. Carvalho, "The light-sheet microscopy revolution," *J. Optics*, vol. 20, no. 5, p. 053002, 2018.

[29] A. P. Di Giovanna, A. Tibo, L. Silvestri, M. C. Müllenbroich, I. Costantini, A. L. A. Mascaro, L. Sacconi, P. Frasconi, and F. S. Pavone, "Whole-brain vasculature reconstruction at the single capillary level," *Scientific Reports*, vol. 8, no. 1, pp. 1–11, 2018.

[30] E. Lugo-Hernandez, A. Squire, N. Hagemann, A. Brenzel, M. Sardari, J. Schlechter, E. H. Sanchez-Mendoza, M. Gunzer, A. Faissner, and D. M. Hermann, "3d visualization and quantification of microvessels in the whole ischemic mouse brain using solvent-based clearing and light sheet microscopy," *J. Cerebral Blood Flow & Metabolism*, vol. 37, no. 10, pp. 3355–3367, 2017.

[31] C. M. St. Croix, S. H. Shand, and S. C. Watkins, "Confocal microscopy: comparisons, applications, and problems," *Biotechniques*, vol. 39, no. 6, pp. S2–S5, 2005.

[32] I. D. Kelch, G. Bogle, G. B. Sands, A. R. Phillips, I. J. LeGrice, and P. R. Dunbar, "Organ-wide 3d-imaging and topological analysis of the continuous microvascular network in a murine lymph node," *Scientific Reports*, vol. 5, no. 1, pp. 1–19, 2015.

[33] A. Bassi, L. Fieramonti, C. D'Andrea, G. Valentini, and M. Mione, "In vivo label-free three-dimensional imaging of zebrafish vasculature with optical projection tomography," *J. Biomedical Optics*, vol. 16, no. 10, p. 100502, 2011.

[34] S. Makita, Y. Hong, M. Yamanari, T. Yatagai, and Y. Yasuno, "Optical coherence angiography," *Opt. Express*, vol. 14, no. 17, pp. 7821–7840, 2006.

[35] G. Grover, S. Quirin, C. Fiedler, and R. Piestun, "Photon efficient double-helix psf microscopy with application to 3d photo-activation localization imaging," *Biomedical Optics Express*, vol. 2, no. 11, pp. 3010–3020, 2011.

[36] A.-K. Gustavsson, P. N. Petrov, M. Y. Lee, Y. Shechtman, and W. Moerner, "3d single-molecule super-resolution microscopy with a tilted light sheet," *Nature Communications*, vol. 9, no. 1, pp. 1–8, 2018.

[37] H. W. Bennett, A.-K. Gustavsson, C. A. Bayas, P. N. Petrov, N. Mooney, W. Moerner, and P. K. Jackson, "Novel fibrillar structure in the inversin compartment of primary cilia revealed by 3d single-molecule super-resolution microscopy," *Molecular Biology of the Cell*, vol. 31, no. 7, pp. 619–639, 2020.

[38] S. Quirin and R. Piestun, "Depth estimation and image recovery using broadband, incoherent illumination with engineered point spread functions," *Appl. Optics*, vol. 52, no. 1, pp. A367–A376, 2013.

[39] R. Berlich, A. Bräuer, and S. Stallinga, "Single shot three-dimensional imaging using an engineered point spread function," *Optics Express*, vol. 24, no. 6, pp. 5946–5960, 2016.

[40] Z. Wang, Y. Cai, Y. Liang, X. Zhou, S. Yan, D. Dan, P. R. Bianco, M. Lei, and B. Yao, "Single shot, three-dimensional fluorescence microscopy with a spatially rotating point spread function," *Biomedical Optics Express*, vol. 8, no. 12, pp. 5493–5506, 2017.

[41] C. Roider, A. Jesacher, S. Bernet, and M. Ritsch-Marte, "Axial super-localisation using rotating point spread functions shaped by polarisation-dependent phase modulation," *Optics Express*, vol. 22, no. 4, pp. 4029–4037, 2014.

[42] H. Ikoma, T. Kudo, Y. Peng, M. Broxton, and G. Wet-

zstein, "Deep learning multi-shot 3d localization microscopy using hybrid optical–electronic computing," *Optics Letters*, vol. 46, no. 24, pp. 6023–6026, 2021.

[43] R. Piestun, Y. Y. Schechner, and J. Shamir, "Propagation-invariant wave fields with finite energy," *JOSA A*, vol. 17, no. 2, pp. 294–303, 2000.

[44] K. Yanny, N. Antipa, W. Liberti, S. Dehaeck, K. Monakhova, F. L. Liu, K. Shen, R. Ng, and L. Waller, "Miniscope3d: optimized single-shot miniature 3d fluorescence microscopy," *Light: Science & Applications*, vol. 9, no. 1, pp. 1–13, 2020.

[45] Y. Xue, I. G. Davison, D. A. Boas, and L. Tian, "Single-shot 3d wide-field fluorescence imaging with a computational miniature mesoscope," *Science Advances*, vol. 6, no. 43, p. eabb7508, 2020.

[46] K. Zhang, M. R. Kellman, E. Bostan, and L. Waller, "3d fluorescence deconvolution with deep priors (conference presentation)," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXVII*, 2020.

[47] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[48] G. Hamarneh and P. Jassi, "Vascusynth: Simulating vascular trees for generating volumetric image data with ground-truth segmentation and tree analysis," *Computerized Medical Imaging and Graphics*, vol. 34, no. 8, pp. 605–616, 2010.

[49] P. Jassi and G. Hamarneh, "Vascusynth: Vascular tree synthesis software," *Insight J.*, 2011.

[50] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Asilomar Conf. Signals, Systems & Computers*, 2003.

[51] S. Peruz, "Shape from focus," https://www.mathworks.com/matlabcentral/fileexchange/55103-shape-from-focus, 2022.

Bhargav Ghanekar received his B.Tech. degree in Engineering Physics from the IIT Madras, Chennai, India in 2018, and the M.S. degree from the Electrical and Computer Engineering (ECE) Department, Carnegie Mellon University, Pittsburgh, PA, USA. He is currently a Ph.D. student in the ECE Department at Rice University, Houston, TX, USA. His research interests include computational imaging and vision. He is the recipient of Sri Jandhyala Lakshmi Kantam and Srimati Sitamahalakshmi Prize for Second Best Academic Record in B.Tech. Engg. Physics in 2018.

Vishwanath Saragadam received his Ph.D. from Carnegie Mellon University, Pittsburgh and is currently a postdoctoral researcher at the DSP group at Rice University. His research interests include hyperspectral imaging, thermal imaging, computational photography, compressive sensing and computer vision. He is the recipient of the Future Faculty Fellowship at Rice in 2020, A.G. Jordan outstanding thesis award in 2020, the Prabhu and Poonam Goel graduate fellowship in 2019, the outstanding TA award from the ECE Department at CMU in 2018, and Siemens award for highest GPA in the EE department in IIT Madras in 2014.

Dushyant Mehra received his bachelor's degree in biomedical engineering from the University of Minnesota, Twin Cities and his Ph.D. degree in biomedical engineering and physiology from the Mayo Clinic Graduate School of Biomedical Sciences. He is currently a postdoctoral associate in the department of chemistry at Rice University. His research interests include using computational imaging techniques to improve single molecule localization microscopy.

Anna-Karin Gustavsson received her BS, MS, and PhD in Physics from the University of Gothenburg, Sweden, in 2008, 2010, and 2015, respectively. After receiving her PhD, she worked as a Postdoctoral Fellow in the Department of Chemistry at Stanford University, Stanford, CA, USA. In 2020, she joined Rice University, Houston, TX, USA, where she currently is an Assistant Professor and a CPRIT Scholar in Cancer Research in the Department of Chemistry, the Department of BioSciences, the Smalley-Curl Institute, and the Institute of Biosciences and Bioengineering. Her research group strives to gain detailed information about cellular nanoscale structure, dynamics, and molecular mechanisms by designing and applying innovative and versatile microscopy approaches primarily for 3D single-molecule tracking and super-resolution imaging. She is the recipient of the FEBS Journal Richard Perham Prize for Young Scientists in 2012, the 3-year Swedish Research Council International Postdoc Fellowship in 2016, the PicoQuant Young Investigator Award in 2018, the NIH K99/R00 Pathway to Independence Award in 2019, the CPRIT Recruitment of First-Time Tenure-Track Faculty Members Award in 2020, and the Scialog: Advancing Bioimaging Fellowship in 2021.

Aswin C. Sankaranarayanan is a professor in the ECE department at CMU, where he is the PI of the Image Science Lab. His research interests are broadly in compressive sensing, computational photography, signal processing and machine vision. His doctoral research was in the University of Maryland where his dissertation won the distinguished dissertation award from the ECE department in 2009. Aswin is the recipient of best paper awards at CVPR 2019 and ICCP 2021, the CIT Dean's Early Career Fellowship, the Spira Teaching award, the NSF CAREER award, the Eta Kappa Nu (CMU Chapter) Excellence in Teaching award, and the Herschel Rich Invention award from Rice University.

Ashok Veeraraghavan received his bachelor's degree in electrical engineering from the Indian Institute of Technology, Madras, Chennai, India, in 2002, and the MS and PhD degrees from the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA, in 2004 and 2008, respectively. He is currently a professor in the Electrical and Computer Engineering, with Rice University, Houston, TX, USA. Before joining Rice University, he spent three years as a research scientist at Mitsubishi Electric Research Labs, Cambridge, MA, USA. His research interests are broadly in the areas of computational imaging, computer vision, machine learning, and robotics. His thesis received the Doctoral Dissertation Award from the Department of Electrical and Computer Engineering at the University of Maryland. He is the recipient of the National Science Foundation CAREER Award in 2017. At Rice University, he directs the Computational Imaging and Vision Lab.