High Resolution Diffuse Optical Tomography using Short Range Indirect Subsurface Imaging

Chao Liu, *Student Member, IEEE,* Akash K. Maity, *Student Member, IEEE,*
Artur W. Dubrawski, *Member, IEEE,* Ashutosh Sabharwal, *Fellow, IEEE,*
and Srinivasa G. Narasimhan *Member, IEEE*

Abstract—Diffuse optical tomography (DOT) is an approach to recover subsurface structures beneath the skin by measuring light propagation beneath the surface. The method is based on optimizing the difference between the images collected and a forward model that accurately represents diffuse photon propagation within a heterogeneous scattering medium. However, to date, most works have used a few source-detector pairs and recover the medium at only a very low resolution. And increasing the resolution requires prohibitive computations/storage. In this work, we present a fast imaging and algorithm for high resolution diffuse optical tomography with a line imaging and illumination system. Key to our approach is a convolution approximation of the forward heterogeneous scattering model that can be inverted to produce deeper than ever before structured beneath the surface.We show that our proposed method can detect reasonably accurate boundaries and relative depth of heterogeneous structures up to a depth of $8\ mm$ below highly scattering medium such as milk. This work can extend the potential of DOT to recover more intricate structures (vessels,tissue, tumors, etc.) beneath the skin for diagnosing many dermatological and cardio-vascular conditions.

Index Terms—Computational Photography; Radiative Transfer Equation; Light scattering; Diffuse Optical Tomography.

1 Introduction

Imaging below the skin and through tissue is important for diagnosis of several dermatological and cardiovascular conditions. MRI remains the best current approach to obtain a 3D dimensional visualization below the skin. But MRI is expensive, requires visits to a hospital or imaging center, and the patients are highly inconvenienced. Non-invasive imaging using visible or near-infrared light has the potential to make devices portable, safe, and convenient to use at home or at point-of-care centers.

While light penetrates deep through tissue, it scatters continuously resulting in poor image contrast. This makes it challenging to recover useful properties about the anatomical structures below the skin. Further, the anatomical structures include a complex heterogeneous distribution of tissue, vasculature, tumors (benign or malignant) that vary in optical properties (density, scattering, absorption) and depths below the skin. This makes the modeling of light propagation below skin challenging.

Fortunately, under the highly scattering regime, the photons can be assumed to be traveling diffusely in the medium and can be described as a random walk. This has enabled accurate forward models under diffuse photon propagation. In order to improve contrast, imaging detectors and sources are placed at different locations on the skin. This arrangement captures only indirectly scattered light while eliminating the dominant direct reflection and backscatter [1]. The approaches that attempt to invert the diffusion model with such indirect light imaging systems are commonly classified as "Diffuse Optical Tomography" (DOT). Due to their portability and ease of use, DOT is becoming an attractive choice over traditional modalities like MRI for cerebral as well as hemodynamic imaging [26], [35]. More recently, DOT has been shown to be a promising tool in detecting strokes [9],

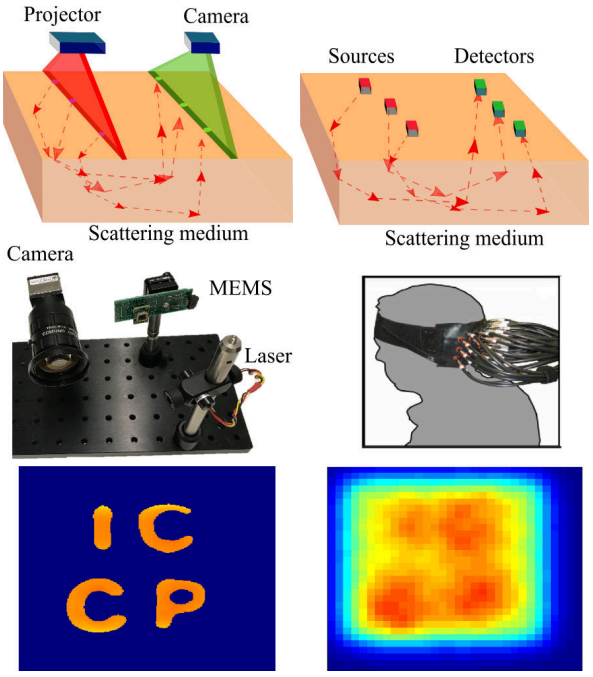


Fig. 1. Diffuse optical tomography (DOT) with line-scanned camera and line scanning MEMS projector (left) compared with traditional DOT [1] with 25 source-detector pairs (right). Both arrangements capture short range indirect subsurface scattered light but our approach is more efficient and recovers the medium (bottom row) at much higher resolution. The results shown here are from simulated data. Please refer to Section 6 for more details and real results.

- *C.Liu, A.W.Dubrawski and S.G.Narasimhan are with the Robotics Institute, Carnegie Mellon University, Pennsylvania, PA, 15213. E-mail: {chaoliu1, awd, srinivas}@cs.cmu.edu*
- *A.K.Maity and A.Sabharwal are with the Department of Electrical and Computer Engineering, Rice University, Texas, 77005. E-mail: {akm8, ashu}@rice.edu*

1. Analogously, in vision and graphics, works measure the Bi-directional Sub-surface Scattering Reflectance Distribution Function (BSSRDF) [12], [13], [16], [22], [34]

breast cancer [19], and thyroid imaging [17].

But there are two important drawbacks to existing DOT approaches. First, the number of source-detector pairs limits the form-factor of devices built so far. Even with multiple source-detector pairs, applying traditional inverse methods for DOT results in poor resolution, as shown in the second column of Figure 1. Second, as the number of source-detector pairs increases, the computational complexity of the algorithms that recover the volumetric medium increases prohibitively [7]. In this work, we present an imaging and algorithmic approach to resolve these fundamental drawbacks in DOT systems. Instead of separate detector-source pairs, we use a high resolution 2D camera and a MEMS projector to obtain a large number of effective source-detector pairs, as is commonly done in vision and graphics. This makes the DOT system much more compact and programmable. Second, to increase the speed of acquisition, we illuminate a line on the skin and capture a different line in the sensor, as described in [24]. This arrangement captures short-range indirect light transport much faster than point-to-point illumination and sensing. But [24] uses a rectified configuration where the projector and camera planes are parallel [28], leading to a low spatial resolution over a large overlapping stereo volume. We develop a new design with a verged configuration that enables high spatial resolution imaging within a small region on the skin (approximately 8 cm x 8 cm).

Using this verged design, we develop an efficient algorithm that is based on the convolution approximation to the forward model for light propagation in a heterogeneous subsurface medium. We show that the convolution kernel is independent of the heterogeneous structure and only depends on the imaging configuration and the scattering properties of the homogeneous scattering medium. This allows us to recover the heterogeneous structures at much higher spatial resolution compared with the traditional DOT, as shown in the last row of Figure 1.

We evaluate our imaging and algorithmic approaches on simulated data by borrowing optical parameters of skin, tissue, blood and vasculature from bio-optical literature [25]. We then demonstrate our approach with an imaging setup that consists of a high resolution 2D camera and a custom-built MEMS based laser projector that are verged to capture near-microscopic spatial resolution beneath a small area of the surface. This imaging system is used to recover heterogeneous structures immersed within a highly scattering medium, such as milk. We show that with the proposed hardware and algorithm, we can detect reasonably accurate boundaries of structures up to a depth of $8mm$ below the surface of milk. We believe this work represents strong progress in achieving high-spatial resolution diffuse optical tomography for the first time at subsurface depths of several millimeters in highly scattering media.

2 Related Works

Over the past few years, there have been developments mainly on two aspects of DOT - improving the instrument system design [32], [33], [39], and secondly, on theoretical aspects that involves accurate forward modeling and rendering [2], [4]. The traditional DOT setup consists of illumination sources and detectors which are placed on the tissue or skin surface. To improve the reconstruction of optical parameters of the tissue volume, multiple configurations of DOT have been explored. In general, it is not possible to obtain depth information from a single source-detector system and therefore multiple source-detector configurations are needed [31].

More recent systems utilize the time-domain (TD) information of photon propagation. TD-DOT systems consist of a source emitting a pulse of light and fast-gated detectors that capture the time-profile of photon arrival. The detectors and the sources are located on a probe with fixed strategic distances between them, so that the photons have traversed a certain depth and rejecting early arriving photons [29]. The most important drawback in these systems is that the limited number of detectors constrains the spatial resolution of the reconstructed optical parameters. More recent DOT systems use structured illumination, which involves projection of patterns instead of discrete sources [3], [18], [23], [30]. The light after interaction with the tissue is captured by either a single-pixel detector or a 2D CCD camera. The use of structured illumination addresses the issues of low speed and sparse spatial sampling, which are associated with traditional DOT systems [30].

The reconstruction of optical parameters involves fitting a forward model to the acquired measurements. The forward model can be obtained either from mesh-based Monte-Carlo simulations [15], [37] or from a diffusion approximation [5], [6], [14] derived from radiative-transfer equations (RTE). While the Monte-Carlo based forward model is more accurate, it requires hours of computing time. Under Born or Rytov approximations, the forward model relates the optical parameters and the measurements by a linear set of equations [27]. Generally, the number of optical parameters to be reconstructed per voxel is very large compared to the number of measurements, and therefore the inverse problem is severely ill-posed. Tikhonov regularizer or sparsity-inducing regularizers are commonly applied for solving the inverse problem [10], [20]. However, with dense sampling the computational load increases as the reconstruction process involves inversion of a large-scale Jacobian matrix [36].

A notable alternative approach [38] is to use the Monte Carlo estimator to differentiate the RTE with respect to any arbitrary differentiable changes of a scene, such as volumetric scattering property, anisotropic phase function or location of heterogeneity. This approach shares the same generality as RTE. However, the performance of the differentiation-based method is highly dependent on the initial estimation of the variables and tend to be trapped in local minimal. Our conventional formulation simplifies the RTE such that the inverse problem is convex. As a result, the result of our method can be used as a good initial guess for the full RTE method [38].

3 Forward Model

In this section, we will review the derivation of the basic theory in DOT for dense scattering tissues. First, we will derive the Diffusion Equation for the surrounding homogeneous medium from the Radiative Transfer Equation [8], [21], assuming that the homogeneous scattering medium surrounding the heterogeneities is dense enough such that the light propagation in the medium is dominated by the high-order light scattering events and the angular distribution of light propagation direction is isotropic. Then we will derive the Diffusion Equation for the heterogeneities, assuming that the light absorption coefficient discrepancy dominates the scattering property difference between the heterogeneous embedding and the surrounding medium. Although the assumptions do not always hold perfectly, we find that our proposed method is

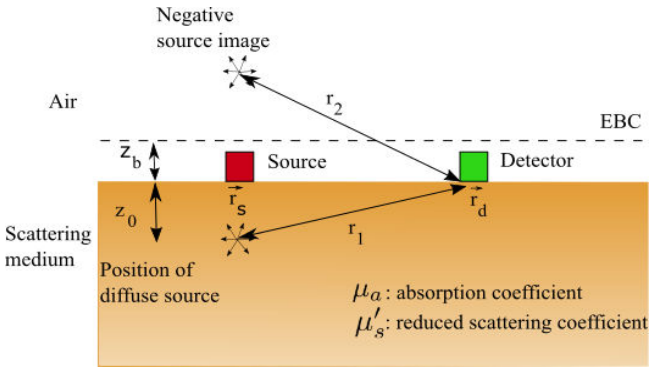


Fig. 2. Source-detector configuration in typical DOT system. The fluence rate at the detector is given by superposition of the real diffuse source located z_0 below the surface, and a negative image source around the zero flux line denoted by EBC.

robust to the cases where the assumptions fail through evaluations in Section 6.1.

The Radiative Transfer Equation (RTE) describes the light radiance, which models light propagation, at a particular position in the scattering medium at a specific time instant. It is generally difficult to solve the RTE in closed form. When the medium is highly scattering, as in the case of biological tissue, the diffusion approximation is commonly applied to obtain the diffusion equation [14], [22]. The photon diffusion equation models the fluence rate Φ, that is defined as the total light radiance integrated over all directions, at a position $\vec{r}$ in the scattering medium for a continuous intensity light source S, given as,

$$(-D(\vec{r})\nabla^2 + \mu_a(\vec{r}))\Phi(\vec{r}) = S(\vec{r}), \tag{1}$$

where $\mu_a(\vec{r}), \mu_s'(\vec{r})$ are the absorption coefficient and the reduced scattering coefficient of the medium respectively, and $D(\vec{r}) = 1/(3(\mu_s'(\vec{r}) + \mu_a(\vec{r})))$ is known as the diffusion coefficient of the scattering medium. The tissue is commonly modeled as a semi-infinite scattering homogeneous medium, with the source and the detector positioned at the air-medium interface. When the light source is treated as a constant pencil beam source, i.e. $S(\vec{r}) = S\delta(\vec{r}_s)$, the solution for fluence rate in (1) for the configuration in Figure 2 can be written in a analytical form using extrapolated zero boundary conditions (EBC):

$$\Phi_0(\vec{r}_d, \vec{r}_s) = \frac{S}{4\pi D_0}\left[\frac{e^{-\beta r_1}}{r_1} - \frac{e^{-\beta r_2}}{r_2}\right], \tag{2}$$

where, $\Phi_0(\vec{r}_d, \vec{r}_s)$ is the fluence rate at detector kept at a position $\vec{r}_d$ with a constant source at $\vec{r}_s$ [5]. The diffusion coefficient of the homogeneous medium is denoted by D_0 and the term $\beta = \sqrt{3\mu_s'\mu_a}$ depend on the optical properties of the homogeneous scattering medium. The extrapolated boundary condition (EBC) accounts for the refractive index mismatch of air and the medium. Solving for the boundary condition defines a zero fluence rate line located z_b above the air-medium interface. This boundary line is imposed by placing a negative image of the source around the zero-crossing line [5]. The terms r_1 and r_2 are the distances from the detector to the real and the negative image source respectively, and they are defined as:

$$\begin{aligned} r_1 &= |\vec{r}_s + z_0 - \vec{r}_d|, \\ r_2 &= |\vec{r}_s - z_0 - 2z_b - \vec{r}_d|, \end{aligned} \tag{3}$$

where, $z_0 = 3D$ is the location of diffused source in the medium. The term z_b is the distance of the zero fluence rate boundary from the air-medium interface.

Often, we are interested in reconstructing objects like veins or tumors embedded within human tissue. Typically these objects have different optical parameters than the background medium. In the presence of heterogeneity, the change in absorption coefficient of the medium can be defined as,

$$\mu_a(\vec{r}) = \mu_{a_0} + \delta\mu_a(\vec{r}) \tag{4}$$

where $\delta\mu_a(\vec{r})$ is the difference in absorption coefficient of the heterogeneous object over the background medium. We assume that the change in the reduced scattering coefficient $\mu_s'(\vec{r})$ is negligible and can be ignored. The resultant fluence rate at the detector position $\vec{r}_d$ for a source at $\vec{r}_s$ is written as a linear addition of fluence rate from homogeneous component $\Phi_0(\vec{r}_d, \vec{r}_s)$ and the change in fluence rate $\Delta\Phi(\vec{r}_d, \vec{r}_s)$ due to the heterogeneous object,

$$\Phi(\vec{r}_d, \vec{r}_s) = \Phi_0(\vec{r}_d, \vec{r}_s) + \Delta\Phi(\vec{r}_d, \vec{r}_s). \tag{5}$$

The change in fluence rate is due to the absorption coefficient change across the volume around the point of interest [5]:

$$\Delta\Phi(\vec{r}_d, \vec{r}_s) = -\int \Phi_0(\vec{r}_s, \vec{r}_j)\frac{\delta\mu_a(\vec{r}_j)}{D_0}G_0(\vec{r}_j, \vec{r}_d)d\vec{r}_j, \tag{6}$$

where G_0 represents the Green's function for a homogeneous slab and is related to Φ_0 in (2) as $G_0 = D_0\Phi_0/S$.

We acquire images using a CCD camera, which records the radiant exitance at the surface. The radiant exitance is proportional to the diffuse reflectance R, which is the projection of current density along the surface normal of the detector for a unit-power source,

$$R(\vec{r}_d, \vec{r}_s) = D_0\left[\frac{\delta\Phi}{\delta z_d}\right]_{z_d=0}, \tag{7}$$

where z_d is the z component of the detector location $\vec{r}_d$.

Applying a derivative to (5) with respect to z_d and multiplying by D_0, we obtain,

$$R(\vec{r}_d, \vec{r}_s) = R_0(\vec{r}_d, \vec{r}_s) + \Delta R(\vec{r}_d, \vec{r}_s), \tag{8}$$

where $R_0 = D_0\left[\delta\Phi_0/\delta z_d\right]_{z_d=0}$ is the diffuse reflectance due to the homogeneous background medium and is obtained by taking a derivative of Φ_0 in (2) with respect to z_d, given by [22],

$$R_0 = \frac{1}{4\pi}\left[\frac{z_0(1+\beta r_1)e^{-\beta r_1}}{r_1^3} + \frac{(z_0+2z_b)(1+\beta r_2)e^{-\beta r_2}}{r_2^3}\right] \tag{9}$$

Similarly, ΔR represents the change in diffuse reflectance for the heterogeneous object. The expression for ΔR is obtained by taking a derivative of (6) with respect to z_d and multiplying by D_0, resulting in the following expression,

$$\begin{aligned} \Delta R(\vec{r}_d, \vec{r}_s) &= -\int \Phi_0(\vec{r}_s, \vec{r}_j)\delta\mu_a(\vec{r}_j)\left[\frac{\delta G_0(\vec{r}_j, \vec{r}_d)}{\delta z_d}\right]_{z_d=0} d\vec{r}_j, \\ &= -\int \Phi_0(\vec{r}_s, \vec{r}_j)\delta\mu_a(\vec{r}_j)R_0(\vec{r}_j, \vec{r}_d)d\vec{r}_j. \end{aligned} \tag{10}$$

We discretize the integral above by dividing the medium into N voxels, and the optical properties are defined for each voxel. If the medium is discretized into N voxels with volume of each

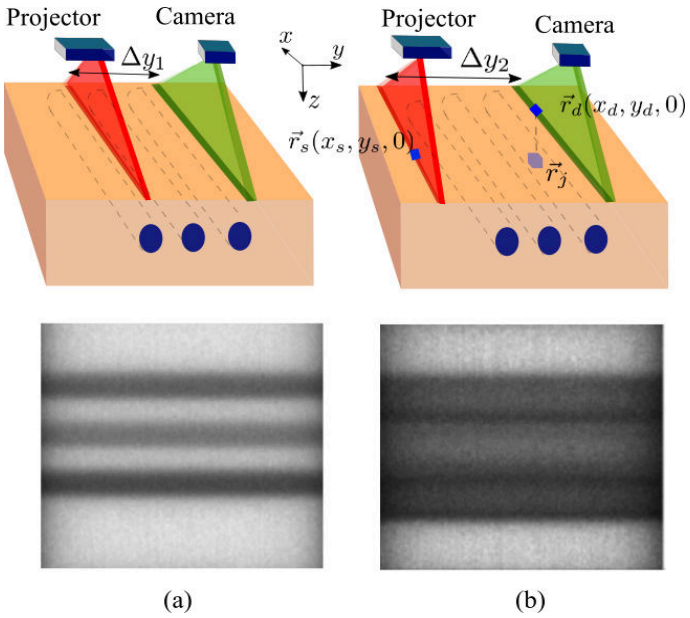


Fig. 3. Generation of short range indirect images for a small (a) and a large (b) pixel to illumination distance Δy. The simulated scene consists of three cylinders embedded in a scattering medium. The offset Δy is kept constant while scanning the entire scene to obtain an image. For a shorter Δy as in (a), the rods are distinctly visible, whereas for longer Δy, the blurring increases with reduction of signal-to-noise ratio.

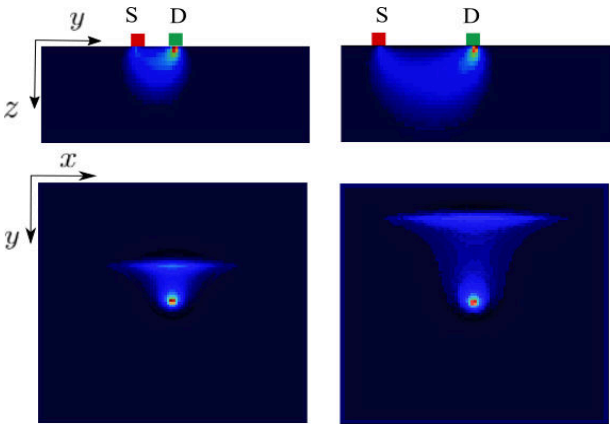


Fig. 4. Visualization of phase function for different pixel to illumination line distance in y-z plane (top row), and x-y plane (bottom row). S and D represents the illumination line and pixel location respectively. As the pixel to illumination line distance increases, the photons tend to travel deeper into the scattering medium but leads to reduced number of photons reaching the pixel, thereby reducing the signal-to-noise ratio.

voxel as h^3, then (10) can be written in the discrete summation form given by

$$\Delta R(\vec{r}_d, \vec{r}_s) = -\sum_{j=1}^{N} P(\vec{r}_s, \vec{r}_j, \vec{r}_d)\delta\mu_a(\vec{r}_j), \tag{11}$$

with

$$P(\vec{r}_d, \vec{r}_j, \vec{r}_s) = \Phi_0(\vec{r}_s, \vec{r}_j)R_0(\vec{r}_j, \vec{r}_d)h^3. \tag{12}$$

The term $P(\vec{r}_s, \vec{r}_j, \vec{r}_d)$ is commonly known as the phase function defined at each voxel position $\vec{r}_j$ in the medium. The values of the phase function depend on the optical properties of the background homogeneous medium as well as the location of the source $\vec{r}_s$ and the detector $\vec{r}_d$. Note that the phase function is independent from the structure of the heterogeneous object.

4 Convolution Approximation of Heterogeneous Model

In this section, we describe how the diffuse forward model can be adapted to our experimental setup. We project a line illumination on the scene using a laser projector as in [24]. So the light source is now considered as a slit source instead of a point source. By using a slit source we reduce the acquisition time since line scanning is significantly faster than point scanning. We incorporate a slit source in the forward model by using the linear superposition principle. The quantities described in the previous section which are functions of the source location $\vec{r}_s$ are now obtained by adding up the contributions corresponding to all the point sources on the illumination line.

On the detector side, we use a rolling shutter camera synchronised with the laser projector. The advantage of using a camera is that each pixel in the camera sensor can now be considered as an individual detector, and hence we have a detector array with millions of detectors. Secondly, since the camera sensor can be considered as a grid array of detectors, we can derive a convolution form of the forward model, significantly speeding up the computation time. We acquire several images by varying the pixel to illumination line distance shown in Figure 3. These images are referred to as short-range indirect images. The boundaries of the heterogeneous structures become more blurry in the short range indirect image as the pixel to illumination line distance Δy increases. The blurring effect is related to Δy and the depth of the structures. This is similar to the depth from (de)focus methods, where the blurring effect is related to the focal setting of the camera and the scene depth.

The values of phase function at each voxel for the short-range indirect images can be interpreted as the number of photons that have traversed through the corresponding voxel for a given pixel to illumination line distance. Typically, the most probable path between a pixel and the source illumination line follows a well-known "banana shape" [11] and is shown for different pixel to illumination line distances in the Figure 4.

We note two important properties of the phase function P. Firstly, in case of simple geometry like the semi-infinite homogeneous background medium under consideration, the expression for the Green's function G_0 as well as Φ_0 can be written in terms of relative voxel location rather than the absolute location, i.e,

$$\begin{aligned} P(\vec{r}_d, \vec{r}_j, \vec{r}_s) &= \Phi_0(\vec{r}_s - \vec{r}_j)R_0(\vec{r}_j - \vec{r}_d), \\ &= P(\vec{r}_j - \vec{r}_d, \vec{r}_s - \vec{r}_d). \end{aligned} \tag{13}$$

Secondly, we note that the values of the phase function decays fast spatially as the distance between a voxel and source or detector position increases. Hence, we can neglect the contribution of voxels that are far away from both the illumination line and the pixel. Since we are using a projected line illumination as our light source, we approximate the phase function in (13) by the summation of truncated phase function for each source point along the illumination line. Additionally, as evident from the figure, the contribution of light from the illumination line to a center pixel is dominant only near the center of the illumination line, and hence we can use a spatially-invariant phase kernel κ. We define the pixel to illumination line distance $\Delta y = y_s - y_d$, where y_s and y_d are the y component of illumination row $\vec{r}_s$ and the pixel location $\vec{r}_d$

respectively. The phase kernel for a line illumination can then be written as,

$$\kappa(\vec{r}_j - \vec{r}_d; \Delta y) = \sum_{\vec{r}_s} P(\vec{r}_j - \vec{r}_d, \vec{r}_s - \vec{r}_d), \tag{14}$$

where the summation over $\vec{r}_s$ is for all the point sources lying on the illumination line. In the following, we will denote the phase kernel as $\kappa(\Delta y)$ for denotation simplicity unless the spatial dependency needs to be emphasized.

Similarly, the diffuse reflectance $R(\vec{r}_d, \vec{r}_s)$, the change in diffuse reflectance $\Delta R(\vec{r}_d, \vec{r}_s)$ and the homogeneous diffuse reflectance $R_0(\vec{r}_d, \vec{r}_s)$ in (8) are modified for line illumination as the sum of contribution from all point sources lying on the illumination line, and are defined as $R(\vec{r}_d; \Delta y)$, $R(\vec{r}_d; \Delta y)$ and $R_0(\vec{r}_d; \Delta y)$ respectively. We denote (x_d, y_d) as the surface coordinates of the pixel location $\vec{r}_d$ as shown in Figure 3. If the change in absorption coefficient $\delta\mu_a(\vec{r}_j)$ in (11) is represented by a 3D volume Q, then the change in diffuse reflectance ΔR in (11) can now be expressed in a convolution notation as

$$\begin{aligned} \Delta R(x_d, y_d; \Delta y) &= -\sum_{\vec{r}_s} \sum_{j=1}^{N} P(\vec{r}_j - \vec{r}_d, \vec{r}_s - \vec{r}_d) \delta\mu_a(\vec{r}_j) \\ &= -\sum_{j=1}^{N} \kappa(\vec{r}_j - \vec{r}_d; \Delta y) \delta\mu_a(\vec{r}_j) \end{aligned} \tag{15}$$

where $\Delta R \in \mathbb{R}^{M \times N}$ is defined over a sensor array of dimension $M \times N$ and corresponds to each pixel to illumination line distance Δy as shown in Figure 3. By representing the change of absorption coefficient $\delta\mu_a$ by a 3D volume Q, we can rewrite (15) as the sum of a 3D convolution results:

$$\Delta R(x_d, y_d; \Delta y) = -\sum_{z} \kappa(\Delta y) * Q(x_d, y_d, z) \tag{16}$$

The change in absorption coefficient in the 3D volume is denoted by $Q \in \mathbb{R}^{M \times N \times D}$, where D is the depth resolution. The 3D truncated kernel $\kappa \in \mathbb{R}^{m \times n \times D}$ is the defined for each Δy, and has the same depth resolution as that of the 3D volume Q. Using (8), the resultant diffuse reflectance R acquired at each pixel to illumination line distance Δy can be written as a linear summation of the contribution from homogeneous background medium R_0 and the perturbation term due to presence of heterogeneity ΔR,

$$R(x_d, y_d; \Delta y) = R_0(x_d, y_d; \Delta y) - \sum_{z} \kappa(\Delta y) * Q(x_d, y_d, z)$$

where $R \in \mathbb{R}^{M \times N}$ is the diffuse reflectance on an $M \times N$ grid.

4.1 Reconstruction of heterogeneous structure

For the set of captured images which correspond to different pixel to illumination line Δy, we capture a set of short range indirect images $I(\Delta y)$. For the given set of images, we reconstruct the volume Q of unknown optical parameters by solving the following optimization problem,

$$\min_{Q} . \sum_{\Delta y = T_{d_{min}}}^{T_{d_{max}}} ||I(\Delta y) - l(R_0(\Delta y) - \kappa(\Delta y) * Q)||_F^2 + \lambda ||Q||_1, \tag{17}$$

where $||.||_F$ denotes the Frobenius norm, and l is an unknown scaling factor which depends on the intensity and width of the laser profile and the sensitivity of the camera. The procedure for determining this factor l is highlighted in more detail in Section 6.2. We also assume the reconstructed volume to be sparse, which essentially implies that the heterogeneous object only occupies a fraction of the total reconstructed volume.

The optimization is done over a range of Δy values. For smaller Δy values, the diffusion approximation breaks down, as the photon propagation is largely governed by single or very few scattering events. For very large Δy, not enough photons reach the camera pixels, and therefore the measurement images have a poor signal-to-noise ratio. Therefore, the range of Δy values needs to be chosen appropriately.

If we know the exact optical parameters μ_s' and μ_a of the homogeneous background medium, then we can construct the kernel $\kappa(\Delta y)$ as in (14). However in some cases, the background optical parameters of the material are not known. In those cases, we select a homogeneous patch inside the field of view, and fit the pixel intensity measurements with lR_0 with respect to the unknown optical coefficients as in (9). We then use the estimated values of the coefficients to construct the phase kernel $\kappa(\Delta y)$ for solving the optimization in (17).

We use PyTorch for implementation given it is highly optimized for convolution operations. The running time on a workstation with TianX GPU is around 5 minutes for 300 iterations for Q with a depth resolution of 64. The λ value in (17) is heuristically chosen to be 0.0001. We start the optimization with an initial value of all zeros for Q, and the reconstruction accuracy can be further improved if a better initialization is provided based on prior knowledge of the scene.

5 Hardware

In this section, we describe our imaging setup for capturing short-range indirect images. In [24], a rectified configuration where the projector and camera are parallel is used for capturing the short-range images. That setup leads to a low spatial resolution over a large overlapping stereo volume. To capture high resolution images for small area of interest, we need a high spatial resolution over a smaller overlapping stereo volume. One way to achieve smaller overlapping stereo volume is to verge the projector and camera. This motivates us to design a verged setup for capturing high resolution short-range indirect images.

Our setup consists of a pair of synchronized rolling shutter camera and a laser projector implemented with Micro-Electro-Mechanical-Systems (MEMS). Our imaging setup is shown in Figuer 1 and Figure 6 (a). We use IDS-3950CPv2 industrial camera and Mirrorcle MEMS development kit. The central wavelength for the laser light is 680 nm. The MEMS mirror reflects the laser beam from the laser diode and the orientation of the MEMS mirror can be controlled in terms of two rotation axes (vertical and horizontal). The size of the imaged area on the sample is 8 cm by 8 cm. We model the laser diode and MEMS mirror pair as a pinhole projector whose center of projection is the center of rotation of the MEMS.

During the imaging process, the projector is scanned through the epipolar planes of the projector-camera pair. The camera is synchronized such that the pixels having a pre-defined offset from the corresponding epipolar line on the camera image plane are exposed. Each offset corresponds to one pixel to illumination line distance Δy as discussed in Section 4. For the physically rectified projector-camera pair as in [24], the epipolar lines on the projector and camera image are horizontal. This simply corresponds to

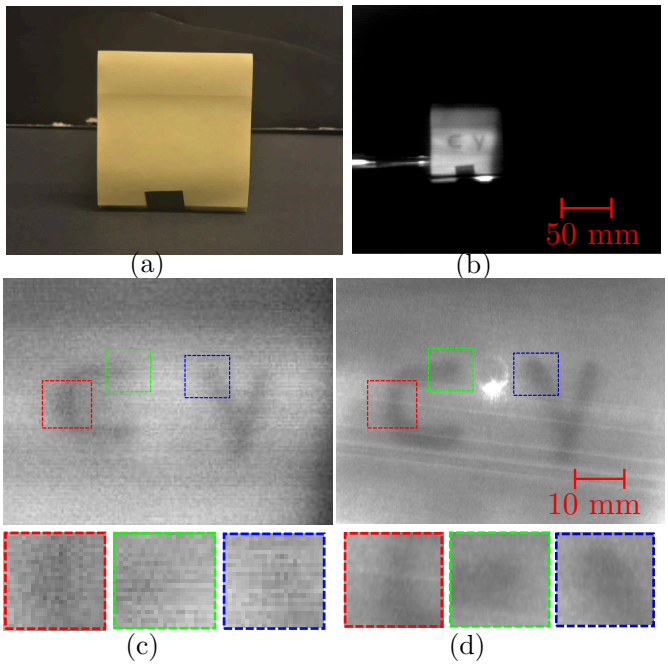
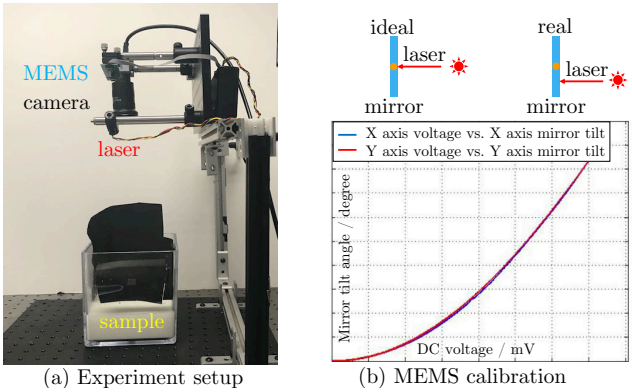


Fig. 5. Images of a paper sticker captured using different devices. The sticker page with letters is occluded by several other pages so no letters can be seen under regular lighting. (a) Image captured with cellphone camera under regular lighting. ; (b) Short-range indirect image captured with the device in [24]; (c) Enlarged image for the sticker region in (b); (d) Short-range indirect image captured with our device. Our device has smaller FOV due to non-zero vergence angle. The images captured with our device has higher resolution, SNR and contrast as shown in the insects in (c) and (d). The bright spot in the center and the lines in (d) is due to the reflection and inter-reflections from the protective glass in front of the MEMS mirror.

illuminating and exposing the corresponding rows of projector and camera. In contrast, in our setup, the epipolar lines in the projector and camera are not horizontal due to the verged setup. So we cannot capture the short range indirect images by illuminating and exposing corresponding rows. Instead, on the projector side, we control the angle of the MEMS mirror to scan the light laser beam across a pencil of epipolar lines with different 2D slopes in the projector image plane. On the camera side, we interpolate over offset epipolar lines to get the short range indirect images. As a special case, for $\Delta y = 0$, the interpolated line overlaps with the epipolar. The resultant image is the direct light image.

Our image setup has smaller FOV than the rectified system in [24] due to the non-zero vergence angle between the project and camera. As a result, we can place the sample closer to the camera while the sample can still be illuminated by the projector. This enables higher image resolution for smaller area of interest so that more fine-grained (sub)surface details can be captured. In Figure 5, we show the images of a paper sticker captured with different devices. The sticker page with letters is occluded by several other pages so no letters can be seen under regular lighting. The occluded letters are visible in the short range indirect images from both [24] and our setup. Our device has smaller FOV and higher spatial resolution over the region of interest due to the verged configuration. In addition, we have better contrast and higher SNR because the laser light source used in our setup is of higher intensity compared to the pico-projector in [24]. The bright spot in the center and the lines in(d) is due to the reflection and inter-reflections from the protective glass in front of the MEMS.



(a) Experiment setup (b) MEMS calibration

Fig. 6. Experiment setup and calibration to compensate the laser-mirror misalignment and non-linearity of MEMS. (a) The device is mounted vertically above the sample container, with no cover above the scattering medium. (b) During MEMS calibration, we consider the misalignment between the laser and mirror (above) and non-linearity of MEMS mechanics. Due to misalignment, the incident laser beam onto the MEMS mirror will not be perpendicular to the mirror surface and align with the MEMS rotation center; Due to non-linearity of MEMS mechanics, the input control signal and degrees of rotation are not linearly related.

5.1 Calibration

The device is mounted vertically above the liquid container as shown in Figure 6 (a), with no cover above the scattering medium. We model the laser-MEMS as a pinhole projector whose center of projection is the rotation center of the MEMS mirror. During the calibration process, we estimate the relative pose between the MEMS mirror and the camera. For MEMS, we compensate for the non-linear mapping between the input voltage for the MEMS and the mirror tilt angle, and account for the mis-alignment of the MEMS mirror and the laser, as shown in Figure 6 (b).

More specifically, we illuminate planes with given poses relative to the camera with a set of dot patterns. As shown in Figure 7, given the laser dot images for different plane positions, we can fit the laser rays in 3D and triangulate the rays to get the origin of the rays, *i.e.* the rotation center of the MEMS mirror. Due to the laser-MEMS misalignment and fitting error for the rays, the rays will not intersect at one 3D point. We solve a least square problem for the intersection point where the point to ray distances are minimized. The fitted rays are also used to account for the non-linear relation between the MEMS input voltage and the rotation angle. In calibration, we build a lookup table relating the input voltage for the MEMS and the rotation angle for the mirror to account for their non-linear relation. During imaging, given the target laser ray direction, we can estimate the required input voltage by interpolating over lookup table.

6 Experiment Results

6.1 Simulation

We test the proposed algorithm using Monte Carlo rendered images. For the homogeneous medium, we use the scattering coefficients of human skin measured in [22]. The heterogeneous inclusions are located up to $4\ mm$ below the surface. For the imaging setup, the width of the laser illumination line is $1mm$. The distance between the illumination line and the camera pixel ranges from 0 to $15\ mm$. To make the diffusion approximation valid for the algorithm, we only use the images with the illumination to pixel distance Δy larger than $2\ mm$.

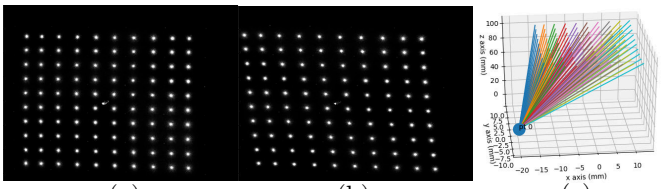


Fig. 7. The MEMS-camera pose calibration. We illuminate a plane with known pose with 2D array of beams as shown in (a) and (b). Given the plane orientations, we can get the 3D parameters for the rays from multiple such images. Then we triangulate all the fitted rays to determine the center of projection for the projector, in our case, the rotation center for the MEMS mirror.

The simulated direct and global light images are shown in the first two rows in Figure 8. The global light image is the sum of the images captured with different Δy's except for $\Delta y = 0$. The inclusions can not be seen in the direct only image. For the global light image, because highly scattering property of skin, the contrast is low for some of the deeper inclusions, such as the solid circle in the right column. This makes the detection and localization for such objects (e.g. tumor beneath the skin surface) difficult. For each short-range indirect image, the image intensity is contributed in part by the indirect light that travels from the illumination line with a preset distance to the imaged scene point. As a result, compared with the global light image, the contrast of the inclusions are much higher for the short-range indirect images shown in the third and fourth rows of Figure 8. On the other hand, for larger pixel to illumination line distance, the SNR is low because there are less photons reaching the imaged scene point due to multiple scattering of light. In addition, because the non-zero support of the diffuse scattering kernel increases with the pixel to line illumination distance, the boundaries of the inclusions in the image becomes more blurry for larger distance. Despite low SNR and blurring in the short-range indirect images, using the proposed method, we are able to localize the 2D boundaries and estimate the relative depth for the inclusions.

For the input short range indirect images, the contrast of the image decreases with the inclusion depth, as shown in Figure 9 (a). This is because as the inclusions become deeper, most light reaching the pixel is contributed by the scattered light from the homogeneous medium without traveling through the inclusions. Another intuition is that the diffuse scattering phase function diminishes with the depth increase, as shown in Figure 4.

One key factor for the proposed method is the size of the diffuse scattering kernel. Smaller kernel enables faster optimization process, but it will lead to more errors in the convolutional approximation, hence less accurate results; while larger kernel leads to better performance, it induces more processing time. The choice for the size of diffuse scattering kernel is also related to the pixel to illumination line distance. In addition, as shown in Figure 4, the non-zero support region for the kernel varies with the pixel to illumination line distance. For large pixel to illumination line distance, the magnitude of the kernel would be small due to multiple scatterings, so the performance will saturate at certain distance. In Figure 9 (b), we show how the performance changes with the kernel size when the medium is human skin. The performance is evaluated using the IoU score of segmentation results for the inclusions. As we can see, for highly scattering medium like human skin, the performance saturates when the kernel size approaches 20 mm in diameter.

6.1.1 Performance for different scene settings

The derivation of the forward model in Section. 3 is based on two assumptions about the scattering mediums: (1) the scattering coefficient of the surrounding homogeneous medium is large such that the light propagation direction distribution is isotropic; (2) the absorption coefficient discrepancy dominates the scattering property difference between the heterogeneous embedding and the surrounding medium. We evaluate the robustness of our method against the failure of those assumptions.

To this end, we perform four simulation experiments and evaluate the performance with different scattering property and scene settings. For all experiments, the measured images are rendered using Monte Carlo simulations. The performance is evaluated in terms of the IoU scores. We use the same denotation as in Section 3: μ_{s_0}, μ_{a_0} are the scattering and absorption coefficients for the surrounding homogeneous medium; μ_{s_1}, μ_{a_1} are the scattering and absorption coefficients for the embedded heterogeneous material. The performance is shown in Figure 10.

Performance vs. heterogeneity depth As seen, the performance of our method decreases with the depth of the inclusion since the image contrast reduces with depth. For small depth, although single-scattering events can dominate, the large image contrast of the heterogeneous medium makes the reconstruction task easier for the proposed method.

Performance vs. scattering coefficients of the homogeneous background An interesting observation is the parabolic-type performance curve. For lower scattering coefficient μ_{s_0}, the diffusion approximation starts to become less valid, resulting in modeling error. If the scattering coefficient μ_{s_0} is larger, multiple scattering events govern the photon propagation inside tissue, increasing the accuracy of our forward model. However, it becomes progressively difficult to recover the position of the heterogeneous object since now few photons actually sample the heterogeneity embedded at a particular depth and get detected by the detector.

Performance vs. scattering coefficients of the heterogeneity This set of experiments addresses the robustness with different scattering properties of the heterogeneous medium μ_{s_1}. The scattering coefficient of the homogeneous medium μ_{s_0} was kept constant. We noted that even though we vary μ_{s_1}, the performance of our method does not change much. The invariance of the performance is due to the sparse nature of the heterogeneous object inside the medium.

Performance vs. absorption coefficients of the heterogeneity For smaller μ_{a_1} values, the performance is lower due to the fact that the contrast of the heterogeneous object compared to the background medium is lower in the short-range indirect images. Though we assume that the change in absorption coefficient of the object is small compared to the background medium, the increase in contrast helps our algorithm to recover the location of the object.

6.1.2 Comparison with traditional DOT method

We compared our method with the traditional DOT method [5] quantitatively in terms of computational time and performance through simulation. The scene setup is a homogeneous medium with 3 rods embedded at certain depth with different absorption coefficient than the background medium.

For traditional DOT, We considered a fixed number of source-detectors (80 sources and 80 detectors) to reconstruct a volume of $64 \times 64 \times 8$ resolution. The reconstruction process took 15

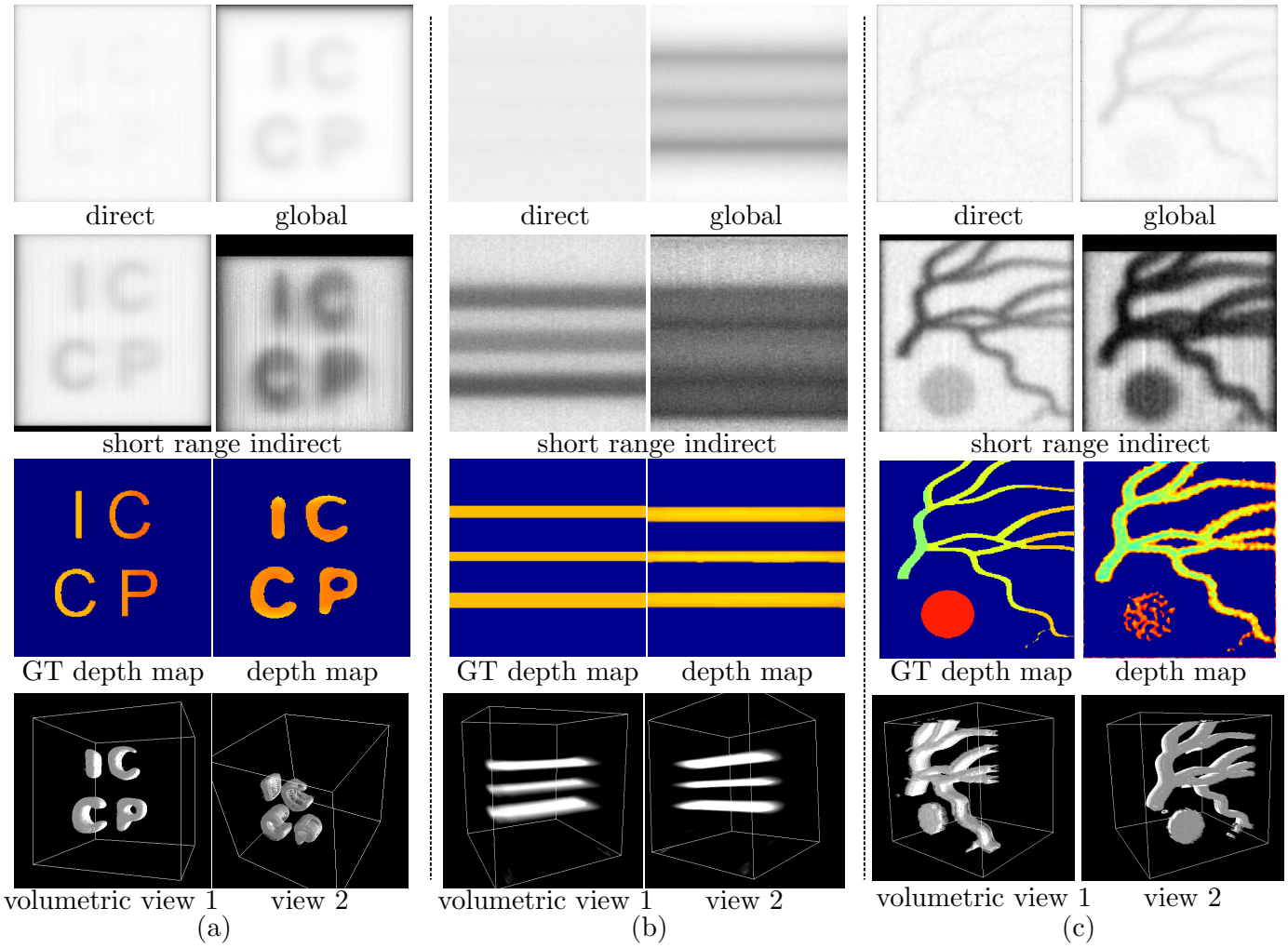


Fig. 8. Simulated direct/global light image, the short range images with different Δy settings, and the DOT results. The homogeneous medium is skin in (a) and (c), skim milk in (b), with the scattering coefficients measured in [22]. Photon and read-out noise are added to the input images. The depths of inclusions in (a) and (b) are $4mm$ and $3mm$ respectively. In (c), the depth of the vein structure is $5mm$ while the solid circle is $10mm$ deep, The inclusion boundaries in the global and short-range indirect images are either blurred or shifted due to light scattering. The signal-to-noise ration decreases as the pixel to illumination line distance increases since less photons are received by the detectors. Our methods recovered the inclusion boundaries and their relative depths despite the blurring and noises.

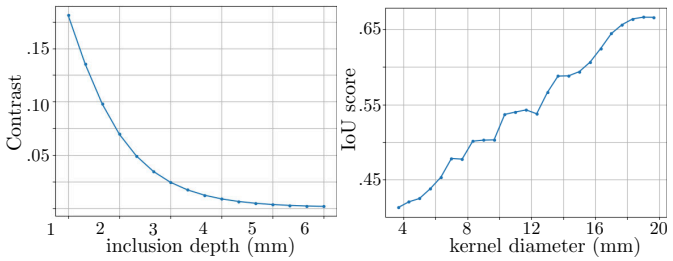


Fig. 9. (a) The average contrast of the short range indirect image varies with the inclusion depth. The inclusion depth is the distance between the inclusion and the embedding medium surface. (b) IoU vs. kernel sizes. For human skin, the performance saturates when the kernel size approaches 20 mm in diameter For all simulations, the images are synthesized using Monte Carlo method with scattering properties for human skin measured in [22].

minutes. The reconstructed volume was upsampled to the scene resolution and the IoU was computed to be only 0.18. For our method, we reconstructed the same scene using our setup with a resolution of $256 \times 256 \times 64$. The IoU from our method was 0.71 and it took 4 minutes. The results show that we are able to perform reconstruction of much higher resolution and accuracy using our method compared to traditional DOT.

6.2 Real Data

We test the proposed method on images captured using the calibrated imaging setup shown in Figure 6 (a). We choose the embedding medium to be milk with little or no water dilution because its scattering property is close to human skin and it can be well described using the diffusion approximation. Because the small FOV and high camera resolution, the laser line illumination spans multiple pixels in the image. We calibrate the laser light source for the its intensity and the width of laser beam. More specifically, we use a ideal white diffuser with albedo close to one, and illuminate the diffuser with the laser line illumination. Then we extract the 1D profile for the laser line illumination, by averaging the intensity along the reflected illumination line. During the optimization in Equation 17, the rendered image is

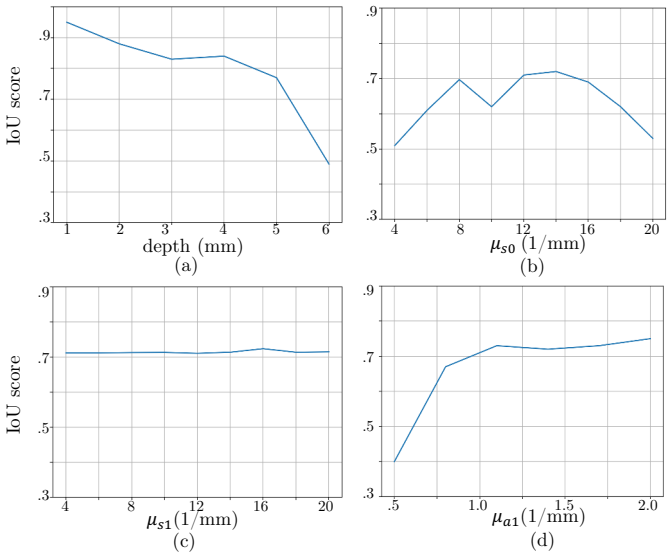


Fig. 10. Performance with different (a) heterogeneity depth; (b) scattering coefficient μ_{s_0} of homogeneous background; (c) scattering coefficient of the heterogeneity μ_{s_1}; (d) absorption coefficient of the heterogeneity μ_{a_1}.

convolved with the measured 1D laser profile to account for the width of the line illumination.

During imaging, we use the short range indirect images with pixel to illumination distance Δy ranges from $20mm$ to $40mm$, such that the light source to sensor distance is large enough for the diffusion approximation. This configuration is different from simulation because the laser beam spans more pixels in the real data. We capture the HDR images to include the large range of image intensity under laser illumination. During optimization, for efficiency, we set the size of the diffuse scattering kernel to be $30mm$. The initialization of the reconstructed volume for all experiments is set to zeros. The measured 1D laser profile is convolved with the rendered images to account for the laser span of multiple pixels. For each scene, we manually select a homogeneous region and fit the scattering properties using the dipole model in Equation 2. For all the results, we use 300 iterations and it takes around 5 minutes for optimization on a workstation with TitanV GPU.

We test on scenes with single and multiple inclusions within the scattering medium. In Figure 11, we show the captured images and reconstructions for single inclusion. Note that in Figure 11 (a) and (c), the inclusion boundaries in the global and short-range indirect images are blurred due to multiple light scattering. Compared with the short range indirect images, the contrast of the inclusions is lower in the global image. In addition, as shown in Figure 11 (b), the inclusion is barely visible if it is deep below the surface. Our method is able to localize the boundary and reconstruct the 3D structures (e.g. the wire structure) despite low visibility and lack of contrast in the input images. Similarly, as shown in Figure 12, for multiple inclusions, the boundary of the inclusions and their relative depths can be recovered although the contrast and visibility of the inclusions in the input short-range indirect images are low due to highly light scattering. The inclusions are up to 8 mm beneath the whole milk surface and no water dilution is added. The dark dots in the images are mask of the light reflection from the protection glass surface for the MEMS mirror, which cannot be controlled and can only be removed in the clean room to prevent the mirror from being contaminated.

7 Conclusion and Future Work

Our work addresses two fundamental limitations of existing diffuse optical tomography methods: low resolution reconstruction and high computational complexity. We overcome these limitations by (1) extending the design for short-range indirect subsurface imaging to a verged scanning projector and camera configuration and (2) a novel convolution based model and efficient computational algorithm for estimating the subsurface medium with heterogeneous structures. This allowed us to recover detailed heterogeneous structures immersed up to $8mm$ deep in a highly scattering medium, such as milk, for the first time. Avenues of future work include using other source spectra (near-infra red) to recover structures deeper within tissue, and using resonant MEMS scanning for capturing subsurface videos of dynamic structures, such as blood flow in microscopic capillary veins.

Acknowledgment This research was funded in part by NSF Expeditions "See Below the Skin" Grants #1730147 and #1730574, US Army Medical Research and Materiel Command "Autonomous Trauma Care in the Field" #W81XWH-19-C-0083 and a gift from the Sybiel Berkman Foundation. We thank Ioannis Gkioulekas for the insightful discussions and reviewers for their helpful feedback and comments.

References

[1] https://www.ese.wustl.edu/~nehorai/research/dot/dot_overview.html. 1
[2] Simon R Arridge and John C Schotland. Optical tomography: forward and inverse problems. *Inverse Problems*, 25(12):123010, dec 2009. 2
[3] Samuel Belanger, Maxime Abran, Xavier Intes, Christian Casanova, and Frederic Lesage. Real-time diffuse optical tomography based on structured illumination. *Journal of Biomedical Optics*, 15(1):016006, 2010. 2
[4] Bo Bi, Bo Han, Weimin Han, Jinping Tang, and Li Li. Image reconstruction for diffuse optical tomography based on radiative transfer equation. *Computational and Mathematical Methods in Medicine*, 2015:1–23, 2015. 2
[5] D. A. Boas, J. P. Culver, J. J. Stott, and A. K. Dunn. Three dimensional monte carlo code for photon migration through complex heterogeneous media including the adult human head. *Optics Express*, 10(3):159, feb 2002. 2, 3, 7
[6] David A Boas. *Diffuse photon probes of structural and dynamical properties of turbid media: theory and biomedical applications*. PhD thesis, University of Pennsylvania, 1996. 2
[7] Samuel Bélanger, Maxime Abran, Xavier Intes, Christian Casanova, and Frederic Lesage. Real-time diffuse optical tomography based on structured illumination. *Journal of Biomedical Optics*, 15(1):1 – 7, 2010. 2
[8] Kenneth M. Case, Paul F. Zweifel, and G. C. Pomraning. Linear transport theory. *Physics Today*, 21(10):72–73, oct 1968. 2
[9] Maria Chalia, Laura A. Dempsey, Robert J. Cooper, Chuen-Wai Lee, Adam P. Gibson, Jeremy C. Hebden, and Topun Austin. Diffuse optical tomography for the detection of perinatal stroke at the cot side: a pilot study. *Pediatric Research*, 85(7):1001–1007, feb 2019. 1
[10] Chen Chen, Fenghua Tian, Hanli Liu, and Junzhou Huang. Diffuse optical tomography enhanced by clustered sparsity for functional brain imaging. *IEEE Transactions on Medical Imaging*, 33(12):2323–2331, dec 2014. 2
[11] S. B. Colak, D. G. Papaioannou, G. W. 't Hooft, M. B. van der Mark, H. Schomberg, J. C. J. Paasschens, J. B. M. Melissen, and N. A. A. J. van Asten. Tomographic image reconstruction from optical projections in light-diffusing media. *Applied Optics*, 36(1):180, jan 1997. 4
[12] Craig Steven Donner. *Towards Realistic Image Synthesis of Scattering Materials*. PhD thesis, La Jolla, CA, USA, 2006. AAI3226771. 1
[13] Julie Dorsey, Alan Edelman, Henrik Wann Jensen, Justin Legakis, and Hans Køhling Pedersen. Modeling and rendering of weathered stone. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '99, pages 225–234, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co. 1

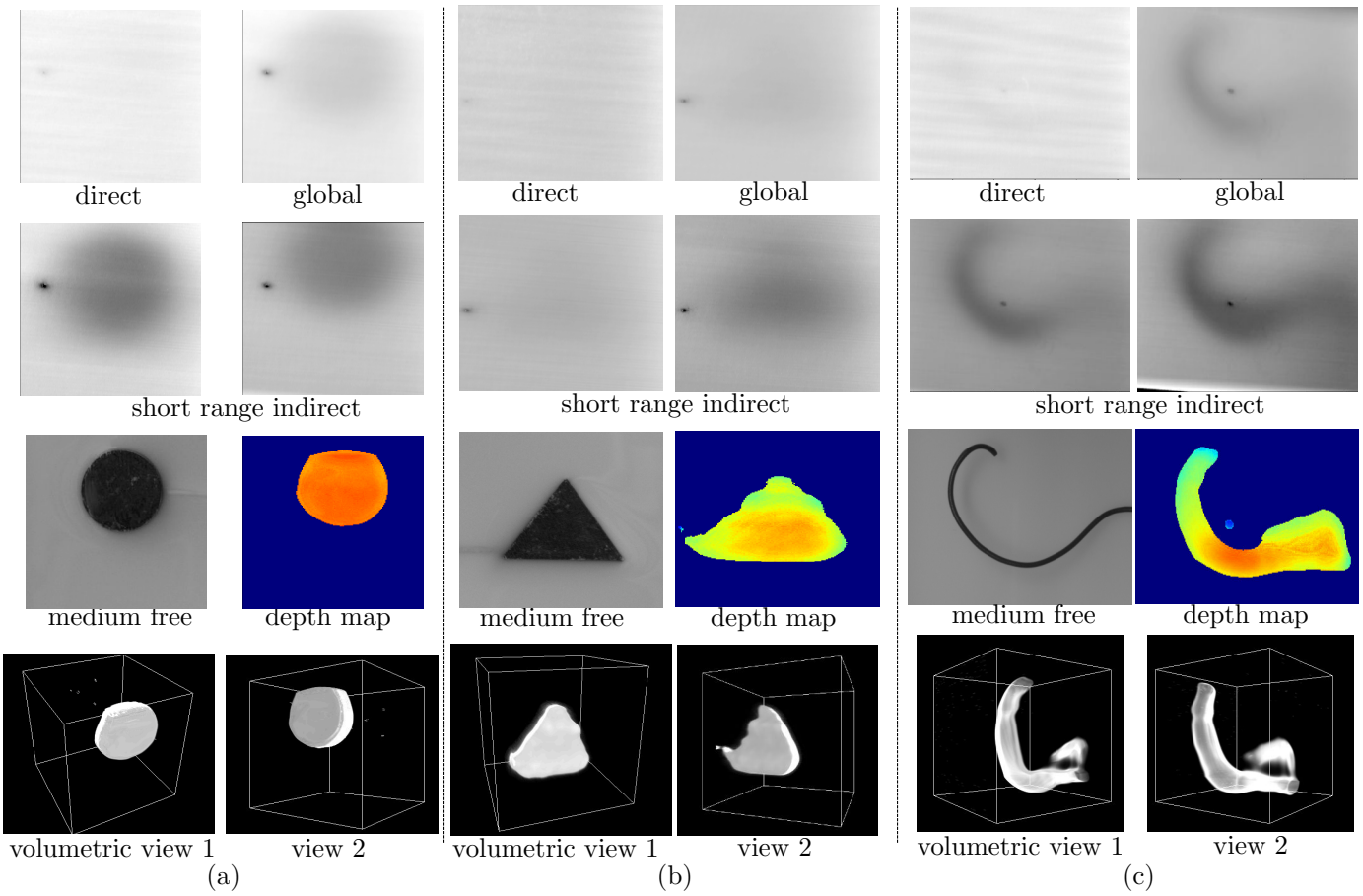


Fig. 11. Real data images and results for single inclusion. The scattering medium is skim milk with no or little water dilution. Rows from top to bottom: the direct and global light images, the short-range indirect images with different Δy settings, the medium free image and masked depth map, two views of the 3D tomography results. (a) Solid circle plate inclusion. The inclusion boundary in the global and short-range indirect images are either blurred or shifted due to light scattering. The boundary is faithfully recovered from our method. (b) Solid triangle plate inclusion. The inclusion is either blurry or barely visible in the global light and short-range indirect images since the inclusion is relatively deep in the scattering medium. Our method is able to reconstruct the triangle structure despite highly scattering effects. (c) Curved thin black wire. The 3D wire structure is recovered in the 3D tomography results, even though it is not obvious at all in the short range indirect images.

[14] T Durduran, R Choe, W B Baker, and A G Yodh. Diffuse optics for tissue monitoring and tomography. *Reports on Progress in Physics*, 73(7):076701, jun 2010. 2, 3

[15] Andrew Edmans and Xavier Intes. Mesh optimization for monte carlo-based optical tomography. *Photonics*, 2(2):375–391, apr 2015. 2

[16] Jeppe Revall Frisvad, Toshiya Hachisuka, and Thomas Kim Kjeldsen. Directional dipole model for subsurface scattering. *ACM Trans. Graph.*, 34(1):5:1–5:12, December 2014. 1

[17] H. Fujii, Y. Yamada, K. Kobayashi, M. Watanabe, and Y. Hoshi. Modeling of light propagation in the human neck for diagnoses of thyroid cancers by diffuse optical tomography. *International Journal for Numerical Methods in Biomedical Engineering*, 33(5):e2826, oct 2016. 2

[18] Sylvain Gioux, Amaan Mazhar, and David J. Cuccia. Spatial frequency domain imaging in 2019: principles, applications, and perspectives. *Journal of Biomedical Optics*, 24(07):1, jun 2019. 2

[19] Anuradha Godavarty, Suset Rodriguez, Young-Jin Jung, and Stephanie Gonzalez. Optical imaging for breast cancer prescreening. *Breast Cancer: Targets and Therapy*, page 193, jul 2015. 2

[20] Murat Guven, Birsen Yazici, Xavier Intes, and Britton Chance. Diffuse optical tomography with a priori anatomical information. *Physics in Medicine and Biology*, 50(12):2837–2858, jun 2005. 2

[21] L. C. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *The Astrophysical Journal*, 93:70, jan 1941. 2

[22] Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. A practical model for subsurface light transport. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 511–518, New York, NY, USA, 2001. ACM. 1, 3, 6, 8

[23] Soren D. Konecky, Amaan Mazhar, David Cuccia, Anthony J. Durkin, John C. Schotland, and Bruce J. Tromberg. Quantitative optical tomography of sub-surface heterogeneities using spatially modulated structured light. *Optics Express*, 17(17):14780, aug 2009. 2

[24] H. Kubo, S. Jayasuriya, T. Iwaguchi, T. Funatomi, Y. Mukaigawa, and S. G. Narasimhan. Acquiring and characterizing plane-to-ray indirect light transport. In *2018 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10, May 2018. 2, 4, 5, 6

[25] Tom Lister, Philip A. Wright, and Paul H. Chappell. Optical properties of human skin. *Journal of Biomedical Optics*, 17(9):1 – 15, 2012. 2

[26] Hellmuth Obrig and Arno Villringer. Beyond the visible—imaging the human brain with light. *Journal of Cerebral Blood Flow & Metabolism*, 23(1):1–18, jan 2003. 1

[27] Maureen A. O'Leary. *Imaging with diffuse photon density waves*. PhD thesis, University of Pennsylvania, 1996. 2

[28] Matthew O'Toole, Supreeth Achar, Srinivasa G. Narasimhan, and Kiriakos N. Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM Trans. Graph.*, 34(4):35:1–35:13, July 2015. 2

[29] Agathe Puszka, Laura Di Sieno, Alberto Dalla Mora, Antonio Pifferi, Davide Contini, Anne Planat-Chrétien, Anne Koenig, Gianluca Boso, Alberto Tosi, Lionel Hervé, and Jean-Marc Dinten. Spatial resolution in depth for time-resolved diffuse optical tomography using short source-detector separations. *Biomedical Optics Express*, 6(1):1, dec 2014. 2

[30] Matthew D. Reisman, Zachary E. Markow, Adam Q. Bauer, and Joseph P. Culver. Structured illumination diffuse optical tomography for noninvasive functional neuroimaging in mice. *Neurophotonics*, 4(2):021102, apr 2017. 2

[31] Juliette Selb, Anders M. Dale, and David A. Boas. Linear 3d reconstruction of time-domain diffuse optical imaging differential data: improved

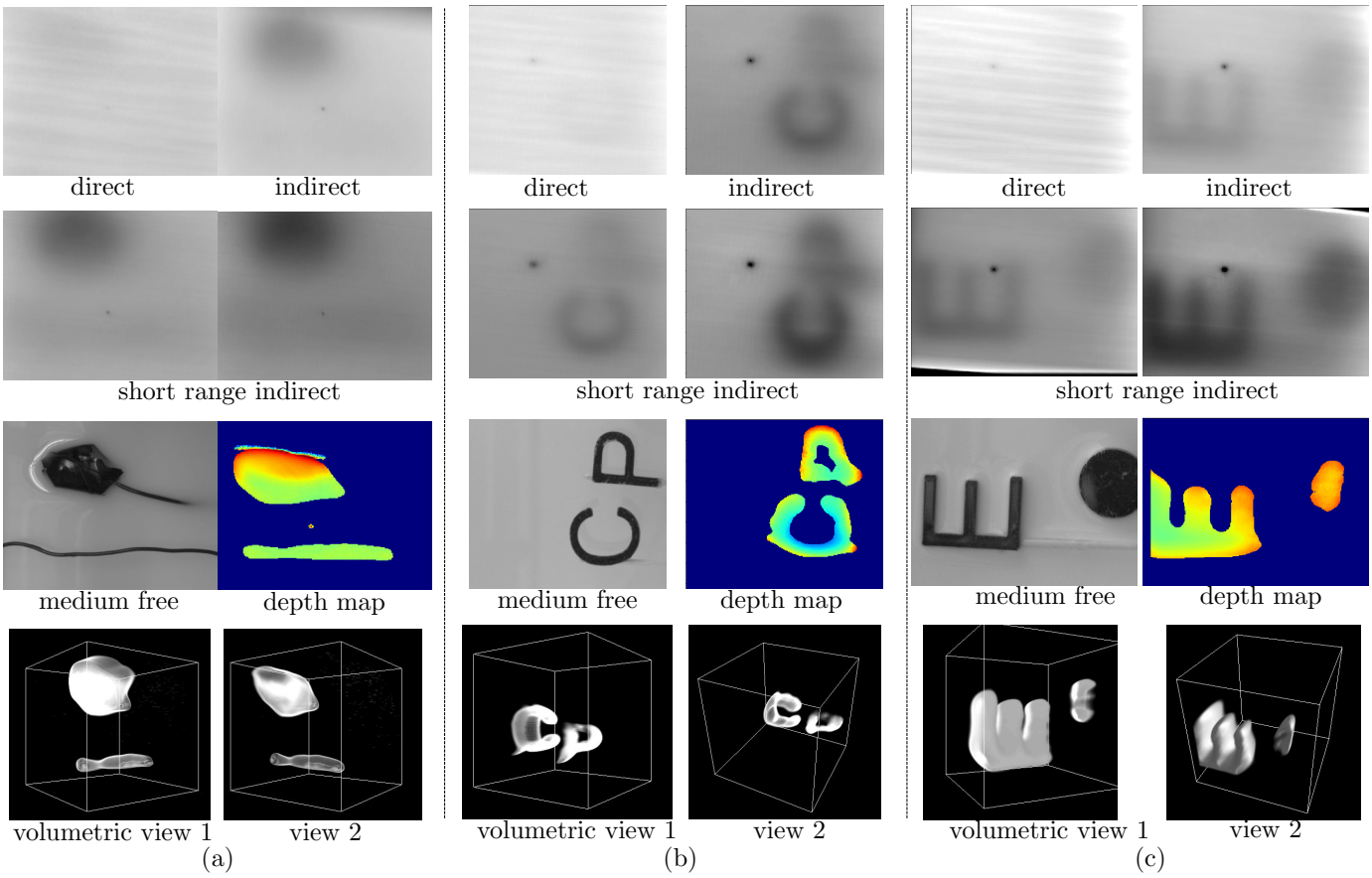


Fig. 12. Real data images and results for multiple objects inclusions. The scattering medium is skim milk with no or little water dilution. Real data results. Rows from top to bottom: the direct and global light only images, the short-range indirect images with different Δy settings, the medium free image and masked depth map, two views of the 3D tomography results. (a) Thin wire and black tape blob. The thin wire is barely visible in the input images due to its small width and light scattering of the medium. However, the location and boundary can be recovered using our method. In (b) and (c), the 3D structures are recovered even though the letters are blurred in the input short range indirect images because of light scattering. In (c), the inclusions are up to 8 mm beneath the whole milk surface and no water dilution is added.

depth localization and lateral resolution. *Optics Express*, 15(25):16400, 2007. 2

[32] Takeaki Shimokawa, Toshihiro Ishii, Yoichiro Takahashi, Satoru Sugawara, Masa aki Sato, and Okito Yamashita. Diffuse optical tomography using multi-directional sources and detectors. *Biomedical Optics Express*, 7(7):2623, jun 2016. 2

[33] A. Siegel, J. J. Marota, and David Boas. Design and evaluation of a continuous-wave diffuse optical tomography system. *Optics Express*, 4(8):287, apr 1999. 2

[34] Xin Tong, Jiaping Wang, Jiaping Wang, Stephen Lin, Baining Guo, Heung-Yeung Shum, and Heung-Yeung Shum. Modeling and rendering of quasi-homogeneous materials. *ACM Trans. Graph.*, 24(3):1054–1061, July 2005. 1

[35] A Villringer. Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neurosciences*, 20(10):435–442, oct 1997. 1

[36] Fang Yang, Feng Gao, Pingqiao Ruan, and Huijuan Zhao. Combined domain-decomposition and matrix-decomposition scheme for large-scale diffuse optical tomography. *Applied Optics*, 49(16):3111, may 2010. 2

[37] Xi Yi, Weiting Chen, Linhui Wu, Wei Zhang, Jiao Li, Xin Wang, Limin Zhang, Huijuan Zhao, and Feng Gao. Towards diffuse optical tomography of arbitrarily heterogeneous turbid medium using GPU-accelerated monte-carlo forward calculation. In Fred S. Azar and Xavier Intes, editors, *Multimodal Biomedical Imaging VIII*. SPIE, mar 2013. 2

[38] Cheng Zhang, Lifan Wu, Changxi Zheng, Ioannis Gkioulekas, Ravi Ramamoorthi, and Shuang Zhao. A differential theory of radiative transfer. *ACM Trans. Graph.*, 38(6), 2019. 2

[39] Xiaofeng Zhang. Instrumentation in diffuse optical imaging. *Photonics*, 1(1):9–32, mar 2014. 2

Chao Liu is a PhD student in Robotics Institute at Carnegie Mellon University. Before coming to CMU, he received Master degree from Imaging Science Department at Rochester Institute of Technology working with Jinwei Gu and Bachelor degree from University of Electronic Science and Technology of China. He is working with Srinivasa Narasimhan and Artur Dubrawski. His research interest is in computational photography, computer vision and medical imaging.

Akash K. Maity is a PhD student in Electrical and Computer Engineering, Rice University, He is working with Dr. Ashok Veeraraghavan and Dr. Ashutosh Sabharwal in Scalable Health Labs. He works in computational imaging with focus on developing non-invasive sensors to obtain clinical information embedded inside human tissue, by using light. He He received his Masters' degree from Rice University in 2018, and Bachelor's degree from Jadavpur University, India, in 2016.

Artur Dubrawski is a Research Professor at Carnegie Mellon University Robotics Institute where he directs the Auton Lab, a research group of 90 focusing on fundamental research and applications of Artificial Intelligence. He received a Ph.D. in robotics and automation from the Institute of Fundamental Technological Research, Polish Academy of Sciences, and a M.Sc. in aircraft engineering from Warsaw University of Technology. Dr. Dubrawski combines academic and real world experience having served in executive and research leadership roles in high tech industry. He had started up a small company which turned out successful in integration and deployment of advanced computerized control systems and novel technological devices. He had also been affiliated with startups incorporated by others: Schenley Park Research, a data mining consultancy and a CMU spin-off, where he was a scientist; and (more recently) with Aethon, a company building robots to automate transportation in hospitals, where he served as a Chief Technical Officer.

Ashutosh Sabharwal is the Department Chair and Professor in the Department of Electrical and Computer Engineering, Rice University. He currently works in two research areas. His first area of research focuses on health, at the intersection of engineering, behavioral sciences and medicine, and established Scalable Health Labs (sh.rice.edu). Scalable Health Labs' mission to develop methods to uncover behavior-biology causal pathways, with a specific focus in three areas: bio-behavioral sensing, mobile bio-imaging, and data science methodologies. Sabharwal is leading several NSF-funded center-scale proposals; most notably, he is the PI of "See below the skin" (seebelowtheskin.org) for non-invasive bio-imaging multi-university effort, and co-investigator on NSF Engineering Research Center, PATHS-UP (pathsup.org) that is developing cost-effective personalized technologies for under-served populations. His second area of research is wireless. He is the founder of WARP project (warp.rice.edu), an open-source project which is now in use at more than 125 research groups worldwide, and have been used by more than 500 research articles. He is currently leading RENEW (renew-wireless.org), an NSF PAWR project, that is developing open-source massive MIMO platform for a deployed city-scale national wireless testbed. He is also co-leading the new ARL-Rice Center for advanced materials and networks. His research has led to four commercial spinoffs (one in wireless and three in healthcare). He is an IEEE Fellow, and was awarded the 2017 IEEE Jack Neubauer Memorial Award, 2018 IEEE Advances in Communications Award, 2019 ACM Sigmobile Test-of-time Award and 2019 ACM MobiCom Community Contributions Award.

Srinivasa G. Narasimhan is the Interim Director and Professor of the Robotics Institute at Carnegie Mellon University. He obtained his PhD from Columbia University in Dec 2003. His group focuses on novel techniques for imaging, illumination and light transport to enable applications in vision, graphics, robotics, agriculture and medical imaging. His works have received several awards: Best Paper Award (IEEE CVPR 2019), Best Demo Award (IEEE ICCP 2015), A9 Best Demo Award (IEEE CVPR 2015), Marr Prize Honorable Mention Award (ICCV 2013), FORD URP Award (2013), Best Paper Runner up Prize (ACM I3D 2013), Best Paper Honorable Mention Award (IEEE ICCP 2012), Best Paper Award (IEEE PROCAMS 2009), the Okawa Research Grant (2009), the NSF CAREER Award (2007), Adobe Best Paper Award (IEEE Workshop on Physics based methods in computer vision, ICCV 2007) and IEEE Best Paper Honorable Mention Award (IEEE CVPR 2000). His research has been covered in popular press including NY Times, BBC, PC magazine and IEEE Spectrum and is highlighted by NSF and NAE. He is the co-inventor of programmable headlights, Aqualux 3D display, Assorted-pixels, Motion-aware cameras, Episcan360, Episcan3D, EpiToF3D, and programmable triangulation light curtains.