Double Your Corners, Double Your Fun: The Doorway Camera

William Krska, *Student Member, IEEE,* Sheila W. Seidel, *Student Member, IEEE,* Charles Saunders, *Member, IEEE,* Robinson Czajkowski, Christopher Yu, *Member, IEEE,* John Murray-Bruce, *Member, IEEE,* and Vivek Goyal, *Fellow, IEEE*

Abstract—In a built environment, wanting to see without direct line of sight is often due to being outside a doorway. The two vertical edges of the doorway provide occlusions that can be exploited for non-line-of-sight imaging by forming *corner cameras*. While each corner camera can separately yield a robust 1D reconstruction, joint processing suggests novelties in both forward modeling and inversion. The resulting *doorway camera* provides accurate and robust 2D reconstructions of the hidden scene. This work provides a novel inversion algorithm to jointly estimate two views of change in the hidden scene, using the temporal difference between photographs acquired on the visible side of the doorway. Successful reconstruction is demonstrated in a variety of real and rendered scenarios, including different hidden scenes and lighting conditions. A Cramér–Rao bound analysis is used to demonstrate the 2D resolving power of the doorway camera over other passive acquisition strategies and to motivate the novel biangular reconstruction grid.

Index Terms—biangular coordinates, computational photography, computer vision, corner camera, non-line-of-sight imaging, remote sensing, stereo vision

1 Introduction

Over the past decade, the possibility of forming images of objects hidden from line-of-sight view has emerged as an intriguing and potentially important expansion of computational imaging and computer vision technology. This capability could help soldiers anticipate danger in a tunnel system, autonomous vehicles avoid collision, and first responders safely traverse a building. Non-line-of-sight (NLOS) imaging techniques may be divided into active and passive methods. Active methods use controlled illumination of the hidden scene and are thus generally less stealthy. Furthermore, most active methods use pulsed laser illumination and single-photon detection, thus requiring costly, high-power equipment. An excellent, comprehensive review of active NLOS imaging systems may be found in [1].

This paper focuses on passive NLOS imaging with an ordinary digital camera in the configuration shown in Figure 1. Like most passive methods, we emphasize the exploitation of occluding structures. A camera obscura or pinhole camera is a simple example of a useful occluding structure. Here, incident light from a given direction illuminates a unique point on the observation plane, creating a projection of the scene outside of the camera on the observation plane. Useful naturally occuring occluding structures might be the aperture formed by an open window [2] or the 'inverse pinhole' created when an object changes position between subsequent measurements [2], [3]. In [4], a 2D reconstruction of a hidden scene was formed using a single photograph from an ordinary digital camera and an occluder with known shape. Joint estimation of unknown occluder shape and the hidden scene has been demonstrated using motion of the hidden scene in [5] and deep matrix factorization in [6]. More recent works addressing similar acquisition configurations have improved inversion techniques [7], [8], [9], broadened the set of occluding objects [10], and even shown that reconstructions are possible without the benefit of an occluding object when a trained network is employed [9], [11].

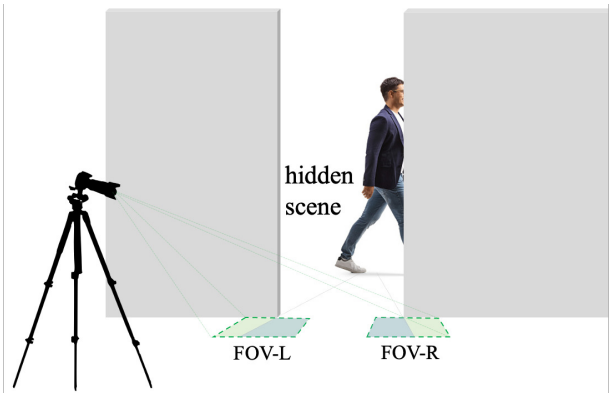


Fig. 1: An illustration of the doorway camera acquisition strategy. An ordinary digital camera takes a single photograph of the penumbrae in FOV-L and FOV-R. This photograph is used to form a 2D reconstruction of the hidden scene.

Our most direct inspiration is the *corner camera* of Bouman et al. [12]. The corner camera uses the vertical edge that occurs naturally at the corner of every wall and

• *W. Krska, S. W. Seidel, C. Saunders, and V. K. Goyal are with the Department of Electrical and Computer Engineering, Boston University, Boston, MA, 02215 USA. S. W. Seidel and C. Yu are with the Charles Stark Draper Laboratory, Cambridge, MA 02139 USA. R. Czajkowski and J. Murray-Bruce are with the Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620 USA. E-mail: {wkrska,sseidel,goyal}@bu.edu*

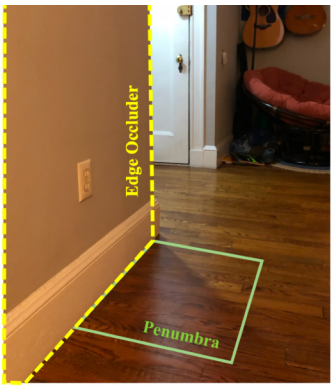


Fig. 2: An annotated photograph shows a single vertical edge occluder (yellow). Light from around the corner is cast onto the floor on the visible side forming a fan-like penumbra pattern (green).

doorway, so the question of knowledge of the occluder shape is obviated. Ambient light from around the occluding edge (on the hidden side of the wall) is cast onto the floor on the visible side forming a fan-line *penumbra* shadow pattern observed on the floor, as shown in Figure 2. Photographs of the penumbra region may be used to form reconstructions of the hidden scene that are highly resolved in azimuthal angle around the corner. This was first demonstrated in [12], where temporal differences between subsequent measurement frames were processed to produce 1D (in angle) reconstructions of motion in the hidden scene. Subsequently, 1D reconstruction of an entire hidden scene (i.e., both moving and non-moving) from a single photograph while also jointly estimating the unknown, non-uniform floor albedo was demonstrated [13]. Two-dimensional reconstruction (where the second dimension is range from the corner) using a single vertical edge was demonstrated in [14], although range estimation was found to be highly susceptible to model mismatch, even in a controlled environment. In active NLOS imaging, a vertical edge has been exploited to control hidden scene illumination with selectivity in azimuthal angle [15].

In this paper, we explore the *doorway camera* formed by two adjacent vertical edges, as shown in Figure 1. Unlike the single edge in [12], which results in poor radial resolution while requiring high signal-to-noise ratio (SNR) [14], the azimuthal resolution of two separate edges enables robust 2D reconstruction of the hidden scene. The idea of combining two corner cameras was initially demonstrated in [12]. There, an angular location was extracted from each 1D reconstruction using manual contour tracing and combined to triangulate the 2D position of a single hidden target in motion. Even if contour tracing can be performed perfectly, with the simple triangulation approach in [12], n target contours at each corner camera correspond to n^2 possible target positions in the hidden scene. As in [12], we perform temporal differencing and seek a reconstruction of change in the hidden scene. However, unlike [12], we jointly process the data from the two penumbrae. Our reconstruction approach produces two unique 2D views of change in the hidden scene, one from the perspective of each occluding edge. We demonstrate our algorithm at far lower SNR than is required for accurate 2D reconstruction in [14] Our reconstruction technique is also capable of resolving multiple moving objects in the foreground while simultaneously describing illumination changes in the background due to occlusion in the hidden scene. Our key contributions include:

- Cramér–Rao bound (CRB) analysis (Section 3) to demonstrate the utility of the doorway occluder for 2D hidden scene reconstruction. Our analysis illustrates the large reduction in location uncertainty provided by the second edge in all parts of the hidden scene and motivates our use of the biangular coordinate system.
- An inversion algorithm (Section 4) to jointly estimate *two* views (i.e., one from each occluding edge) of the hidden scene in a novel 2D reconstruction grid.
- Demonstration of our reconstruction algorithm (Section 5) on a variety of synthetic and experimental hidden scenes.

2 Forward Model

The doorway NLOS acquisition strategy in Figure 1 consists of two vertical edges, each with an adjacent swath of floor where the penumbra is measured, called FOV-L and FOV-R for left and right when facing the doorway. Considered alone, each of the vertical edges may be treated as a single vertical edge occluder, as in [12], [13], [14]. In this section we review the forward model for a single vertical edge occluder and discuss the two-edge doorway scenario.

2.1 The Single Edge Occluder

In Figure 3 we show the bird's eye view of a single edge occluder. We will assume a thin planar occluding wall, although additional modeling could account for non-negligible wall thickness. Point $\mathbf{p} = (r, \theta)$ is on the floor in the camera FOV at range r from the corner and azimuthal angle θ measured away from the occluding wall. The location of a point in the hidden scene (yellow) is parameterized in cylindrical coordinates, with range ρ measured from the corner along the floor, azimuthal angle α measured around the corner in the plane of the floor, and height z. Assuming the camera is pointed straight down at a Lambertian floor, and excluding foreshortening factors for simplicity, the brightness $L_{\mathrm{o}}(\mathbf{p})$ of point $\mathbf{p}$ may be written as a sum of all light incident on point $\mathbf{p}$ multiplied by the albedo $a(\mathbf{p})$ at point $\mathbf{p}$:

$$L_{\mathrm{o}}(\mathbf{p}) = a(\mathbf{p})\left(L_{\mathrm{v}}(\mathbf{p}) + L_{\mathrm{h}}(\mathbf{p})\right),$$

where $L_{\mathrm{v}}(\mathbf{p})$ is incident light originating on the visible side and $L_{\mathrm{h}}(\mathbf{p})$ is incident light originating on the hidden side. We simplify this to

$$L_{\mathrm{o}}(\mathbf{p}) = L_{\mathrm{v}} + L_{\mathrm{h}}(\mathbf{p}) \tag{1}$$

by assuming that $a(\mathbf{p})$ and $L_{\mathrm{v}}(\mathbf{p})$ do not vary with $\mathbf{p}$ within the camera FOV and normalizing to $a(\mathbf{p}) = 1$. These are justified by the floor albedo being constant or known; and any visible-side light sources being relatively far from the camera FOV.

We parameterize the hidden scene in cylindrical coordinates with range ρ, azimuthal angle α (measured into the

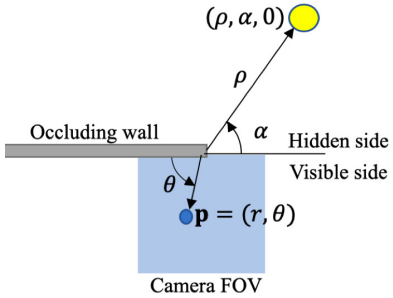


Fig. 3: Bird's eye view of the single corner camera.

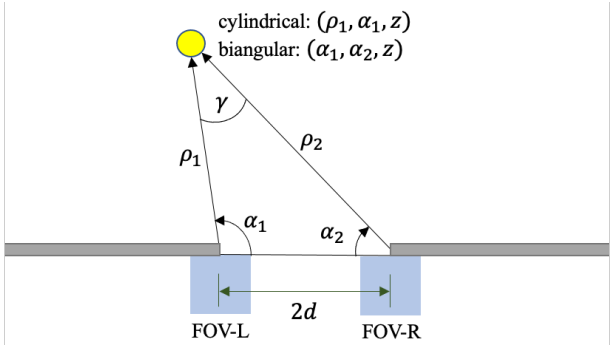


Fig. 4: The doorway camera. A point in the hidden scene is shown in both cylindrical(ρ_1, α_1, z) and biangular (α_1, α_2, z) coordinates.

hidden scene and around the corner), and height h. Incident light originating from the hidden side $L_{\mathrm{h}}(\mathbf{p})$ may be written as a sum over different directions of incident light L_{i}:

$$L_{\mathrm{h}}(r,\theta) \overset{(a)}{=} \int_0^{\theta}\int_0^{\infty}\int_0^{\infty} L_{\mathrm{i}}(\rho,\alpha,z)\rho\,\mathrm{d}z\,\mathrm{d}\rho\,\mathrm{d}\alpha$$
$$\overset{(b)}{=} \int_0^{\theta}\int_0^{\infty}\int_0^{\infty} \frac{S_{\mathrm{h}}(\rho,\alpha,z)}{d^2(r,\theta,\rho,\alpha)+z^2}\rho\,\mathrm{d}z\,\mathrm{d}\rho\,\mathrm{d}\alpha, \quad (2)$$

with integral bounds in (a) set up from $\alpha = 0$ to $\alpha = \theta$ to sum over all *un-occluded* hidden light sources. In (b), we rewrite L_{i} in terms of the radiosity of the hidden scene $S(\rho, \alpha, z)$ at point (ρ, α, z), attenuated by radial falloff where

$$d^2(r,\theta,\rho,\alpha) = r^2 + \rho^2 - 2r\rho\cos(\pi - \theta + \alpha) \quad (3)$$

is the squared distance between point $\mathbf{p}$ and hidden scene point (ρ, α, z) measured along the floor.

In this work, as in [12], we seek to reconstruct *change* in the hidden scene by subtracting a background measurement $L_{\mathrm{o}}^{\mathrm{b}}(\mathbf{p}) = L_{\mathrm{v}}^{\mathrm{b}} + L_{\mathrm{h}}^{\mathrm{b}}(\mathbf{p})$ from each new measurement frame $L_{\mathrm{o}}^{t}(\mathbf{p}) = L_{\mathrm{v}}^{t} + L_{\mathrm{h}}^{t}(\mathbf{p})$:

$$L_{\mathrm{o}}^{t}(\mathbf{p}) - L_{\mathrm{o}}^{\mathrm{b}}(\mathbf{p}) = (L_{\mathrm{v}}^{t} - L_{\mathrm{v}}^{\mathrm{b}}) + (L_{\mathrm{h}}^{t}(\mathbf{p}) - L_{\mathrm{h}}^{\mathrm{b}}(\mathbf{p}))$$
$$= b + L_{\mathrm{h}}^{\Delta}(\mathbf{p}), \quad (4)$$

where $L_{\mathrm{h}}^{\Delta}(\mathbf{p})$ is the change in illumination at point $\mathbf{p}$ due to change (i.e., motion) in the hidden scene and b quantifies the change in illumination due to the visible side. The algorithm proposed in Section 4 seeks to invert (4) from the two edge occluder measurements that can be acquired near a doorway. For each vertical edge in the doorway, we seek to estimate the change in the hidden scene $S_{\mathrm{h}}^{\Delta}(\rho, \alpha, z)$, as well as the illumination change b from the perspective of that corner.

2.2 The Doorway Occluder and Biangular Coordinates

In Figure 4, we show a bird's eye view of the doorway occluder, consisting of two vertical edge occluders separated by a known distance $2d$ along the x-axis. We acquire measurements (M pixels in a square array) at each of FOV-L and FOV-R, each of which may be described by the forward model in Section 2.1. As in Section 2.1, ranges ρ_1 and ρ_2 measure scene distance from the corner along the floor and azimuthal angles α_1 and α_2 are measured around each occluding edge into the hidden scene. As previously demonstrated in [12], [13], [14], each separate edge occluder primarily provides azimuthal information (i.e., the ability to resolve the hidden scene in α). Thus, in this work, we adopt the *biangular* coordinate system [16] and describe a point on the ground plane of the hidden scene by (α_1, α_2). Ranges ρ_1 and ρ_2 may be written in terms of α_1 and α_2 alone by applying the law of sines:

$$\rho_1 = \frac{2d\sin\alpha_2}{\sin\gamma} \quad \text{and} \quad \rho_2 = \frac{2d\sin\alpha_1}{\sin\gamma}, \quad (5)$$

where $\gamma = \pi - \alpha_1 - \alpha_2$, as marked in Figure 4.

3 Cramér–Rao Bound for Localization

The Cramér–Rao bound provides a lower bound on the variance of an unbiased estimate of unknown parameters. It allows us to explore how much information about our unknown parameters is contained in our noisy camera measurements under different acquisition strategies. The CRB for a hidden point target's unknown parameters in polar coordinates for the single edge-occluder scenario is given in [14]. In this section, we consider the problem of localizing a point target in biangular coordinates for the doorway acquisition strategy. Our goal is to provide a qualitative comparison among acquisition strategies, to show the merit of two edge occluders over one or none.

Take the hidden target to be a hypothetical point emitter with brightness c_{s}, located in biangular coordinates at $(\alpha_1, \alpha_2, z) = (\phi_1, \phi_2, 0)$. The equivalent in cylindrical coordinates is $(\rho_1, \alpha_1, z) = (\rho_{1s}, \phi_1, 0)$, where

$$\rho_{1s} = \frac{2d\sin\phi_2}{\sin\gamma_{\mathrm{s}}} \quad (6)$$

from (5), with $\gamma_{\mathrm{s}} = \pi - \phi_2 - \phi_2$. This point corresponds to hidden scene

$$S_{\mathrm{h}}(\rho_1,\alpha_1,z) = \frac{c_{\mathrm{s}}}{\rho_{1\mathrm{s}}}\delta(\rho_1 - \rho_{1s})\delta(\alpha_1 - \phi_1)\delta(z).$$

Evaluating (2), the outgoing radiosity at point $\mathbf{p}$ in FOV-L due to the hidden scene is

$$L_{\mathrm{h}}(\mathbf{p}) = \frac{c_s H(\theta - \phi_1)}{r^2 + \rho_{1s}^2 - 2r\rho_{1s}\cos(\phi_1 + \pi - \theta)}, \quad (7)$$

where $H(x)$ is the Heaviside step function. Substituting (6) into (7), we rewrite (7) in biangular coordinates:

$$L_{\mathrm{h}}(\mathbf{p}) = \frac{c_s H(\theta - \phi_1)}{r^2 + (\frac{2d\sin\phi_2}{\sin\gamma_{\mathrm{s}}})^2 - \frac{4dr\sin\phi_2}{\sin\gamma_{\mathrm{s}}}\cos(\phi_1 + \pi - \theta)}. \quad (8)$$

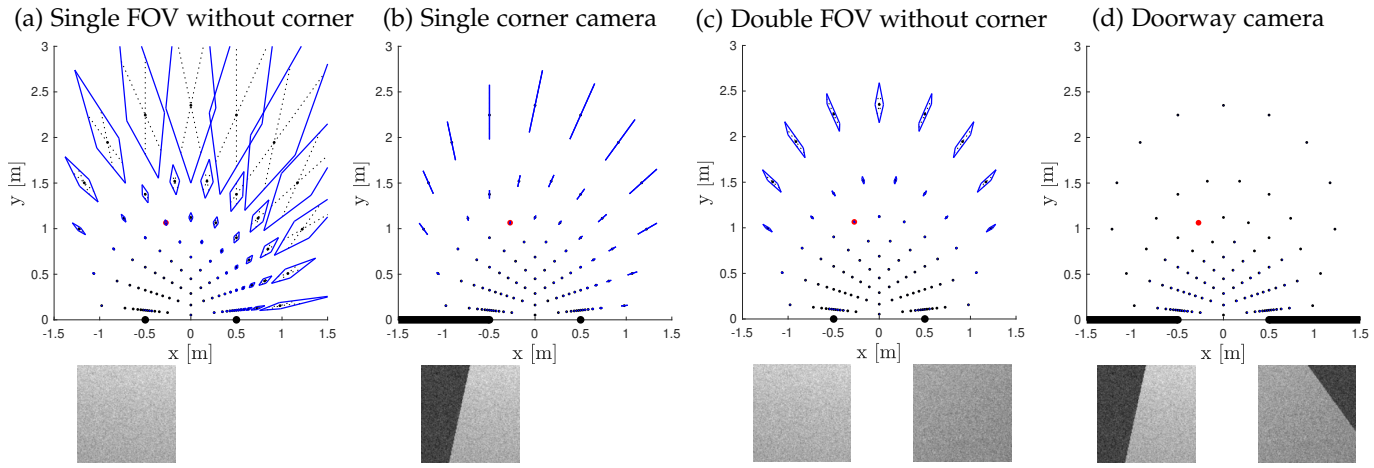


Fig. 5: Uncertainty regions (top row) for four different acquisition strategies. Uncertainty regions have a width of $2\sqrt{\mathrm{CRB}_{\mathrm{door}}(\phi_1)}$ in coordinate α_1 and a width of $2\sqrt{\mathrm{CRB}_{\mathrm{door}}(\phi_2)}$ in coordinate α_2. Two black dots on the x-axis mark the points around which α_1 and α_2 are measured; in this analysis, these points are spaced one meter apart. The bottom row of images show sample measurements when the target is located at the red dot, at the same noise level used to compute uncertainty regions. Each camera FOV is 0.2 m wide and the number of pixels per FOV is $M = 125^2$.

Assuming no visible-side light for simplicity, the measurement at pixel m is $y_m = i_m + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Summing L_{h} over the spatial extent $\mathcal{P}_m$ of the camera pixel gives

$$i_m = \int_{\mathbf{p}\in\mathcal{P}_m} \frac{c_s H(\theta - \phi_1)}{r^2 + (\frac{2d\sin\phi_2}{\sin\gamma_{\mathrm{s}}})^2 - \frac{4dr\sin\phi_2}{\sin\gamma_{\mathrm{s}}}\cos(\phi_1 + \pi - \theta)}\,\mathrm{d}\mathbf{p}. \tag{9}$$

Similarly, if pixel m is in FOV-R,

$$i_m = \int_{\mathbf{p}\in\mathcal{P}_m} \frac{c_s H(\theta - \phi_2)}{r^2 + (\frac{2d\sin\phi_1}{\sin\gamma_{\mathrm{s}}})^2 - \frac{4dr\sin\phi_1}{\sin\gamma_{\mathrm{s}}}\cos(\phi_2 + \pi - \theta)}\,\mathrm{d}\mathbf{p}. \tag{10}$$

Under our Gaussian model, the Fisher information about unknown parameters $(c_{\mathrm{s}}, \phi_1, \phi_2)$ in measurements $\{y_m\}_{m=1}^{2M}$ is given by $\mathbf{F} = (1/\sigma^2)\left(\nabla\mathbf{I}^{\mathsf{T}}\nabla\mathbf{I}\right)$, where

$$\nabla\mathbf{I} = \begin{bmatrix} \frac{\partial i_1}{\partial c_{\mathrm{s}}} & \frac{\partial i_1}{\partial \phi_1} & \frac{\partial i_1}{\partial \phi_2} \\ \frac{\partial i_2}{\partial c_{\mathrm{s}}} & \frac{\partial i_2}{\partial \phi_1} & \frac{\partial i_2}{\partial \phi_2} \\ \vdots & \vdots & \vdots \\ \frac{\partial i_{2M}}{\partial c_{\mathrm{s}}} & \frac{\partial i_{2M}}{\partial \phi_1} & \frac{\partial i_{2M}}{\partial \phi_2} \end{bmatrix}.$$

The CRB on parameters c_{s}, ϕ_1 and ϕ_2 may be computed as follows:

$$\mathrm{CRB}_{\mathrm{door}}(c_{\mathrm{s}}) = [F^{-1}]_{1,1}, \tag{11a}$$
$$\mathrm{CRB}_{\mathrm{door}}(\phi_1) = [F^{-1}]_{2,2}, \tag{11b}$$
$$\mathrm{CRB}_{\mathrm{door}}(\phi_2) = [F^{-1}]_{3,3}. \tag{11c}$$

where the subscript 'door' indicates that these results correspond to the two-edge doorway scenario.

In Figure 5 we use these CRB results to plot uncertainty regions at different points on a biangular grid. Uncertainty regions have a width of $2\sqrt{\mathrm{CRB}_{\mathrm{door}}(\phi_1)}$ in coordinate ϕ_1 and a width of $2\sqrt{\mathrm{CRB}_{\mathrm{door}}(\phi_2)}$ in coordinate ϕ_2. Two black dots on the x-axis mark the points around which ϕ_1 and ϕ_2 are measured; in this analysis, these points are spaced one meter apart. In addition to plotting the uncertainty regions for $\mathrm{CRB}_{\mathrm{door}}$ (d), we compute and plot similar bounds for the following cases for comparison:

- **Single FOV without corner**: A single photograph of FOV-L with no occluding edge.
- **Single corner camera**: A single photograph of FOV-L with an occluding edge. This models an occluding wall rather than a doorway.
- **Double FOV without corner**: Photographs of both FOV-L and FOV-R, each with no occluding edge. Here we explore the effects of having two spaced out photographs.

The bottom row of Figure 5 includes sample measurements for each case with the same noise level used to compute the uncertainty regions shown above. In the single corner camera and doorway camera cases, the penumbra due to the point target is clearly visible in the measurement. In the cases with no corner, the radial falloff patterns used to localize the hidden target are faintly visible, most noticeably when comparing FOV-L and FOV-R in the double FOV case.

We note that the largest uncertainty regions occur in the single FOV without corner case, where a single measurement is collected with no occluding edge (a). Uncertainty grows in both directions as test points move away from FOV-L. The lack of rotational symmetry is due to using the uncertainty in biangular coordinates despite the collected data being only from FOV-L. When a single corner camera is used (b), an edge is introduced and uncertainty regions effectively collapse to lines as ϕ_1 uncertainty becomes vanishingly small. Here we also note that ϕ_2 uncertainty is relatively unchanged with the introduction of the edge, consistent with the result in [14] that the vertical edge primarily enhances azimuthal resolving power. Uncertainty regions in the double FOV with no corner case (c) shrink compared to the single FOV no corner case. In addition to collecting measurements at twice as many pixels, the two FOVs are spaced apart along the x-direction giving them moderately improved localization power, especially for hidden points closer to the x-axis. The uncertainty regions in

the double FOV with no corner case (c) are too small to observe the spatial variation in uncertainty for this case. In Supplementary Note 1, we plot these regions with a width of $500\sqrt{\mathrm{CRB}_{\mathrm{door}}(\phi_1)}$ in coordinate ϕ_1 and a width of $500\sqrt{\mathrm{CRB}_{\mathrm{door}}(\phi_2)}$ in coordinate ϕ_2 to make these spatial variations more clear.

In the doorway camera acquisition strategy (d) that we focus on in this paper, the azimuthal resolution provided by both occluding edges is combined to effectively shrink uncertainty regions down to points. The reconstruction grid proposed in Section 4 is motivated by the observation that each measurement FOV provides a 'view' of the hidden scene that is primarily resolved in angle around the adjacent occluding edge.

4 Inverse Formulation and Algorithm

4.1 Linearity and Occlusion

The model in Section 2.1 is linear and hence does not include the effect of some part of the hidden scene occluding another. When imaging from a single vantage point, one may neglect such occlusion because there is no attempt to see through foreground objects. In the doorway camera, an object may be influencing FOV-L but not FOV-R, or vice versa. To maintain the simplicity of a linear forward model, we choose to combine measurements at both corners to simultaneously reconstruct a view from each FOV vantage point. This also allows for other angular variation of appearance, such as self-occlusion of objects and directional lighting variations. As described in Section 4.2, these reconstructions are coupled to improve the physical realism of the model and reduce the number of decision variables at the expense of making the model bilinear. The formation of a unified reconstruction from the two views is discussed in Section 4.4.

4.2 Discrete Forward Model

Our reconstruction approach assumes a hidden scene composed of vertical planar facets placed on a biangular grid. Each facet has unknown brightness in red, green, and blue. The geometry that gives rise to our biangular reconstruction grid is illustrated in Figure 6a. Here we have plotted biangular 'graph paper,' composed of 21 lines equally spaced in angle α_1 (blue) and 21 lines equally spaced in angle α_2 (red). The 20 intervals in each angle create $\sum_{n=1}^{20-1} n = 19 \cdot 20/2 = 190$ cells, known as the 19th triangular number, some of which are outside the plotted area; cells near the origin are extremely small. A subset of the points of intersection of the lines are marked with black dots, with those outside the plotted area and very close to the origin omitted. Facets in our reconstruction grid rest on the gray lines formed by connecting these black dots to bisect the cells in the direction transverse to the origin. As shown in Figure 6a, these facets form concentric rings about the doorway opening. Note that although facets are longer at ranges further from the doorway, all facets share the same angular widths in α_1 and α_2.

In practice, our reconstruction grid is discretized more finely than Figure 6a, with 4° angular resolution for each of α_1 and α_2. Using all $44 \cdot 45/2 = 990$ potential facet locations at this angular resolution would result in a large number of small facets near the x-axis. Instead, we limit our attention to α_1 and α_2 in $[\pi/8, 7\pi/8]$, and we add facets parallel to the y-axis to extend from the most extreme facets to the front wall. We also omit facets that fall within the ellipse with foci at the doorway edges and minor axis of length 1 m. The resulting reconstruction grid with $N = 230$ potential facet locations is shown in Figure 6b. Facets shown in red are assumed to represent moving scene content in the foreground and are constrained to be positive. Facets shown in purple represent the scene background and may be positive or negative to allow the reconstruction to also describe negative change in the measurement when a moving object blocks light from the scene behind it. The color triplets that describe the N vertical facets of 1 m height in Figure 6b are stored in the rows of $\mathbf{v} \in \mathbb{R}^{N\times 3}$.

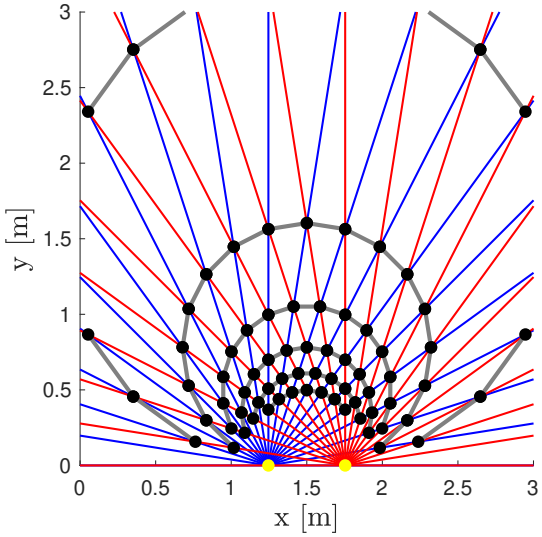
The concatenated measurement in the pth color channel $\mathbf{x}^p = [\mathbf{x}^p_{\mathrm{L}}; \mathbf{x}^p_{\mathrm{R}}] \in \mathbb{R}^{2M}$ includes M pixels from each FOV and may be modeled as follows:

$$\mathbf{x}^p = \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{bmatrix} \mathbf{b}^p + \begin{bmatrix} \mathbf{A}_{\mathrm{L}} & 0 \\ 0 & \mathbf{A}_{\mathrm{R}} \end{bmatrix} \begin{bmatrix} \mathrm{diag}(\mathbf{c}_{\mathrm{L}}) \\ \mathrm{diag}(\mathbf{c}_{\mathrm{R}}) \end{bmatrix} \mathbf{v}^p + \mathbf{q}. \quad (12)$$

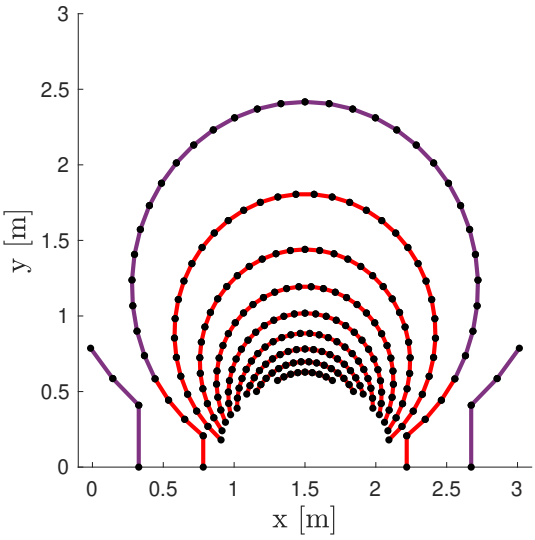
(In the following, superscripts are used for color channel indexes and subscripts are used for left and right indications and other indexing.) The first term describes the effects of far-field visible side light (i.e., b in (4)), where $\mathbf{b}^p = \begin{bmatrix} b^p_{\mathrm{L}} & b^p_{\mathrm{R}} \end{bmatrix}^{\mathsf{T}}$ with b_{L} and b_{R} representing the intensity at the left and right measurements and $\mathbf{1} \in \mathbb{R}^M$ denoting a vector of ones. The second term describes light from the hidden scene, where $\mathbf{A}_{\mathrm{L}} \in \mathbb{R}^{M\times N}$ and $\mathbf{A}_{\mathrm{R}} \in \mathbb{R}^{M\times N}$ are the left and right light transport models as described in (2). The vector $\mathbf{v}^p$ is the pth column of $\mathbf{v}$, corresponding to the pth color channel, and $\mathbf{q}$ describes noise and model mismatch. Multiplicative factors $\mathbf{c}_{\mathrm{L}}, \mathbf{c}_{\mathrm{R}} \in \mathbb{R}^N$ account for differences in observed brightness between left and right corners. Note that rather than estimate six unknown values per hidden pixel (i.e., red, green, and blue brightness values for each FOV), we choose to estimate five: a single RGB triplet (constrained to have unit norm) and an intensity scaling factor for each FOV. A choice to reconstruct six variables (i.e., RGB triples for each view) without additional regularization to promote similarity between views, is effectively the same as using two isolated single corner cameras to form 2D reconstructions. As shown in [14], and our own results in Figure 5b, a single edge occluder leaves considerable uncertainty in range. This formulation creates a strong coupling between the two different reconstruction 'views.' The left corner's view of the hidden scene is $\mathbf{v}_{\mathrm{L}} = \mathrm{diag}(\mathbf{c}_{\mathrm{L}})\mathbf{v}$; the right corner's view is $\mathbf{v}_{\mathrm{R}} = \mathrm{diag}(\mathbf{c}_{\mathrm{R}})\mathbf{v}$. When both corners 'see' the same facet, their views (of the hidden scene specifically, not the values of their measurement pixels) are constrained to be the same color but may have different brightness. When the hidden scene is self-luminous and conforms perfectly to our reconstruction grid, we expect $\mathbf{c}_{\mathrm{L}} = \mathbf{c}_{\mathrm{R}}$. In practice, the differences in entries of $\mathbf{c}_{\mathrm{L}}$ and $\mathbf{c}_{\mathrm{R}}$ describe the effects of off-grid facet tilts, illumination foreshortening, and occlusions.

4.3 Inversion Algorithm

We seek to reconstruct change in the hidden scene between subsequent frames given the difference measurement: $\mathbf{y}_t =$



(a) A coarse illustration of the reconstruction grid concept with 9° angular resolution.



(b) The full reconstruction grid with 4° angular resolution and $N = 230$ planar facets. Foreground facets are red; background facets are purple.

Fig. 6: The reconstruction grid is composed of planar facets resting on a biangular grid.

$\mathbf{x}_t - \mathbf{x}_0$. Here $\mathbf{x}_t$ is the measurement acquired after motion has occurred at time t, and $\mathbf{x}_0$ is an earlier measurement used to subtract light contributions from stationary scenery. We propose to estimate $\mathbf{v} \in \mathbb{R}^{N\times 3}$, $\mathbf{c}_\mathrm{L} \in \mathbb{R}^N$, $\mathbf{c}_\mathrm{R} \in \mathbb{R}^N$, and $\mathbf{b} = [\mathbf{b}^1\ \mathbf{b}^2\ \mathbf{b}^3] \in \mathbb{R}^{2\times 3}$ by solving the following problem:

$$\min_{\mathbf{v},\mathbf{c}_\mathrm{L},\mathbf{c}_\mathrm{R},\mathbf{b}} \left\{ \lambda_\mathrm{group} \sum_{n=1}^{N} \left\| \begin{bmatrix} (\mathbf{c}_\mathrm{L})_n \\ (\mathbf{c}_\mathrm{R})_n \end{bmatrix} \right\|_2 + \frac{1}{2} \sum_{p=1}^{3} \left\| \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{bmatrix} \mathbf{b}^p + \begin{bmatrix} \mathbf{A}_\mathrm{L} & 0 \\ 0 & \mathbf{A}_\mathrm{R} \end{bmatrix} \begin{bmatrix} \mathrm{diag}(\mathbf{c}_\mathrm{L}) \\ \mathrm{diag}(\mathbf{c}_\mathrm{R}) \end{bmatrix} \mathbf{v}^p - \mathbf{y}^p \right\|_2^2 \right\} \tag{13a}$$

$$\text{subject to} \quad \|\mathbf{v}_{i*}\|_2 = 1, \quad i \in \{1, 2, \ldots, N\}, \tag{13b}$$

$$\begin{bmatrix} \mathbf{S}_\mathrm{FG}\mathbf{c}_\mathrm{L} \\ \mathbf{S}_\mathrm{FG}\mathbf{c}_\mathrm{R} \end{bmatrix} \geq 0, \tag{13c}$$

where $\mathbf{v}_{i*}$ is the ith row of $\mathbf{v}$. The second term in the cost function (13a) is an ℓ_2-squared data fidelity, summed over the three color channels. Tuning parameter λ_group controls the amount of regularization introduced by the group sparsity term, with groups defined to be the pairs of entries in $\mathbf{c}_\mathrm{L}$ and $\mathbf{c}_\mathrm{R}$ corresponding to the same facet location in the two views. This term is designed to promote co-located scene content between left and right scene estimates.

Matrix $\mathbf{S}_\mathrm{FG} \in \mathbb{R}^{Q\times N}$ returns only the entries of $\mathbf{c}_\mathrm{L}$ or $\mathbf{c}_\mathrm{R}$ that correspond to facets in the foreground (i.e., red facets in Figure 6b). When differencing is performed, moving objects introduce a positive change in light in the foreground while reducing the light that returns from the more distant scenery behind the objects. The elementwise nonnegativity constraint in (13c) forces our estimate of foreground facets to be nonnegative while allowing the distant outer ring of facets (i.e., purple facets in Figure 6b) to describe negative changes in our data due to occlusion in the hidden scene.

With $\mathbf{c}_\mathrm{L}$, $\mathbf{c}_\mathrm{R}$, and $\mathbf{b}$ fixed, (13a) is a least-squares problem whose solution can be projected to satisfy the normalization constaints (13b); with $\mathbf{v}$ and $\mathbf{b}$ fixed, (13a) and (13c) give a constrained group lasso problem [17]. We perform the optimization using projected gradient methods [18].

4.4 Forming a Unified Reconstruction

Left and right reconstruction views $\mathbf{v}_\mathrm{L}$ and $\mathbf{v}_\mathrm{R}$ may be combined into a single unified reconstruction $\mathbf{v}_\mathrm{u}$ of the hidden scene

$$\mathbf{v}_\mathrm{u} = \mathbf{M} \odot \max(\mathbf{v}_\mathrm{L}, \mathbf{v}_\mathrm{R}), \tag{14}$$

where $\max(\cdot,\cdot)$ returns an element-wise maximum, $\odot$ is the Hadamard product, and $\mathbf{M}$ is a mask with elements given by

$$\mathbf{M}_{i,j} = \begin{cases} 0, & \text{if } [(\mathbf{c}_\mathrm{L})_i < t] \vee [(\mathbf{c}_\mathrm{R})_i < t]; \\ 1, & \text{otherwise,} \end{cases} \tag{15}$$

where t is some small threshold. Effectively, this sets to zero any pixels that are not nonzero in *both* views. When a pixel is nonzero in both views, we return the maximum of the two. In our experimental results, we observed that $\mathbf{v}_\mathrm{u}$ rejects artifacts due to model mismatch that arise in only one view.

5 Experimental Results

Our algorithm has been tested in a variety of conditions, including different hidden scenes and lighting conditions.[1]

In Section 5.1, we share results achieved using synthetic data, allowing us to compare to a perfect ground truth reference. In Section 5.2, we demonstrate that our reconstruction technique works well on real data acquired in the laboratory. Complete views of all reconstruction results included in

1. All code and data required to reproduce the results in this paper may be found at https://github.com/wkrska/DoorwayCameraCode.

this work are shared in Supplementary Note 2. Additional synthetic results for different noise levels, a scenario with three targets of the same color, a non-Lambertian hidden scene, mismatched model height, different camera FOV sizes, and different numbers of camera pixels are included in Supplementary Note 3.

5.1 Synthetic Experiments

Synthetic data for preliminary tests was generated with the Blender rendering engine [19]. Blender was chosen for its ability to create repeatable, physically accurate measurement photos without running into the 'inverse crime' of using the same model to generate synthetic data and test an inversion technique. It offers great control over scene objects with perfect reference ground truth. A virtual $2\,\mathrm{m} \times 2\,\mathrm{m}$ room and various objects were created to match the scale of our laboratory experiments. All scenes were rendered with up to five bounces per light path and up to 8000 light paths per pixel, and all surface materials were modeled as Lambertian. In Section 5.1.1, use a simple example to introduce the concept of our two-view reconstruction. In Section 5.1.2, we show a more challenging two-object scenario with challenging lighting conditions. In Section 5.1.3, we successfully reconstruct a hidden scene with four hidden targets and show the rendered data used to form the reconstruction.

5.1.1 Simple demonstration

Figure 7 shows results for a simple scenario to illustrate the different products produced by our algorithm. Unlike results included later in this section, the walls and hidden object are self-luminous to keep this initial demonstration simple. The two figure columns correspond to the views at the left and right occluding edges. As shown in row (a), the stationary hidden scene consists of red and blue side walls with a green back wall. Before our new measurement frame, a red cylinder has entered the hidden scene, as shown in row (b). Rows (c) and (d) show the positive and negative ground truth change between the stationary background (a) and the new frame (b). Note that the positive change is due to the introduction of the red cylinder; the negative change is due to the occlusion of the green back wall. We show the positive part of our reconstructions $(\mathbf{v}_\mathrm{L})^+$ and $(\mathbf{v}_\mathrm{R})^+$ in row (e) and negative parts $|(\mathbf{v}_\mathrm{L})^-|$ and $|(\mathbf{v}_\mathrm{R})^-|$ in row (f). Observe that the positive reconstruction shows the foreground cylinder at the the same nonzero pixels in both views. The negative reconstruction shows the part of the green back wall that is occluded by the foreground object for each view. Because of the distance between the two edges (marked by yellow dots), the occluded regions share some nonzero pixels but do not completely overlap.

5.1.2 Two object example

In Figures 8 and 9, we demonstrate our algorithm on a more life-like scenario. Here, the room and objects are not self-luminous, but rather illuminated by two large light sources located on the wall containing the doorway. This type of lighting causes additional occlusion within the hidden scene as moving objects cast shadows on the back walls. As in Figure 7, Figure 8 shows ground truth LOS renderings of the hidden scene, positive and negative LOS change in

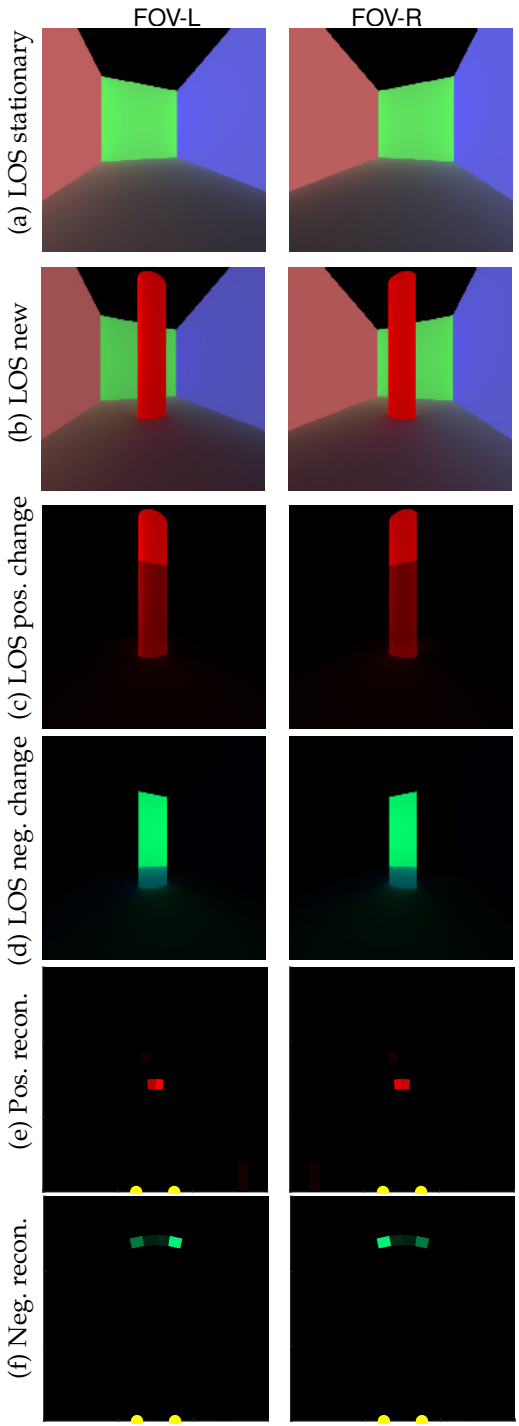


Fig. 7: Algorithm products and LOS renderings at each FOV for a simple self-luminous hidden scene. Rows show (a) LOS views of the stationary hidden scene taken from the point of view of the corner; (b) LOS views of the scene when a red object has entered; (c) the positive change between (b) and (a); (d) the negative change between (b) and (a); (e) the reconstructed positive change; and (f) the reconstructed negative change.

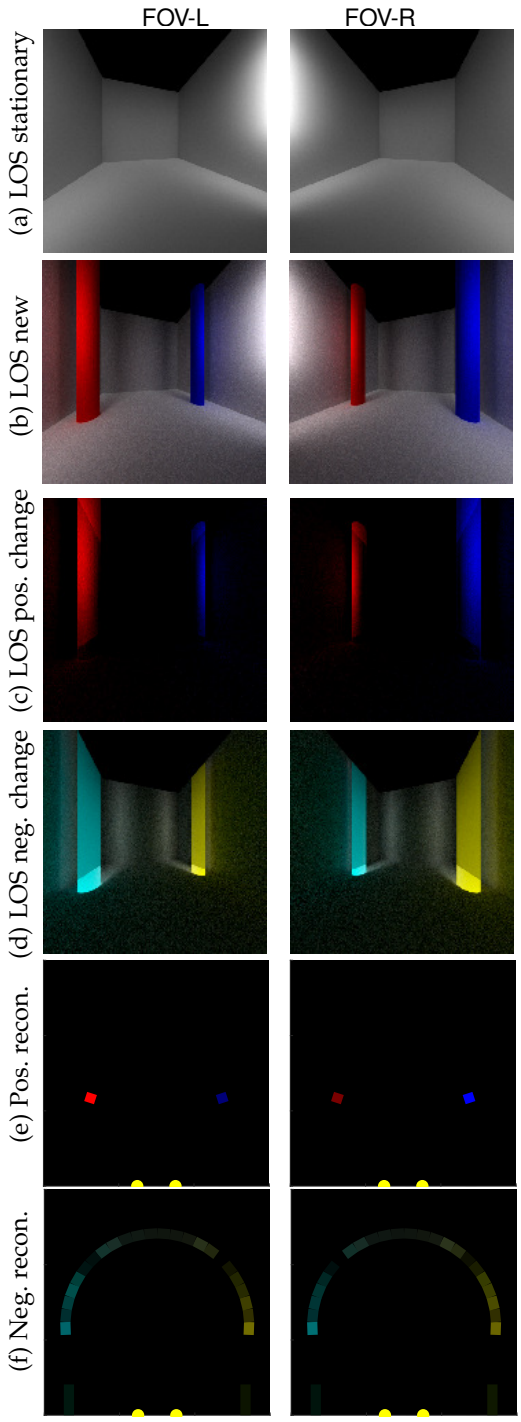
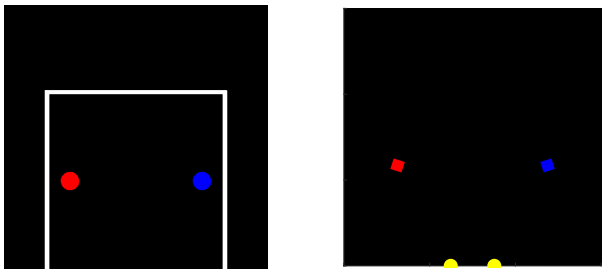


Fig. 8: LOS renderings of the hidden scene and algorithm products for the result in Figure 9. For row descriptions, see Figure 7.



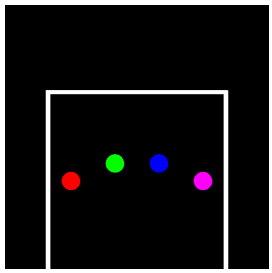
(a) Ground truth (b) Reconstruction $\mathbf{v}_u$

Fig. 9: A bird's eye view of the ground truth and unified reconstruction $\mathbf{v}_u$ for the scenario in Figure 8. Ground truth and reconstruction dimension are both $3\,\text{m} \times 3\,\text{m}$. Nonzero facets are outlined in white for improved visibility against the background. The true location of stationary walls are traced in white for reference in (a). The result in (b) was produced using $t = 0.2$.

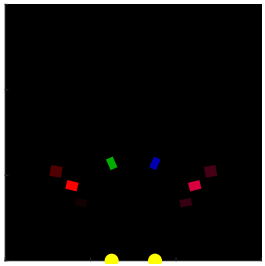
the hidden scene, and positive and negative reconstruction results. The bird's eye ground truth is shown in Figure 9, along with unified reconstruction $\mathbf{v}_u$. The stationary scene, shown in row (a) of Figure 8, consists of white walls. Two cylindrical objects, one red and one blue, enter the scene before the new measurement frame, as shown in row (b). Positive left and right reconstruction results, shown in row (e), both clearly resolve both red and blue objects, with meaningful differences between them. Note that the left result reconstructs the red object more brightly, and the blue object more faintly, than the the right reconstruction. This is consistent with the positive ground truth LOS change, where the left view 'sees' a larger part of the red object and a smaller part of the blue. The negative reconstruction results, row (f), closely match the negative ground truth LOS change in the hidden scene, row (d). The unified reconstruction shown in Figure 9b closely matches the ground truth shown in Figure 9a.

5.1.3 Four object example

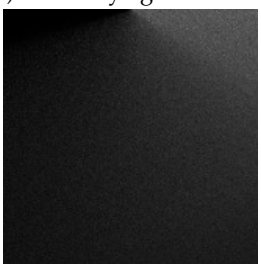
Figure 10 shows results for a four-object scenario. The bird's eye view of the hidden scene is shown in Figure 10a, with the location of stationary walls outlined in white. Stationary scene measurements at FOV-L and FOV-R are shown in Figures 10c and 10d. Measurements acquired after the four objects have entered the scene are in Figures 10e and 10f. In Figures 10g and 10h, the difference between the stationary measurement and new frame measurement is scaled between 0 and 1 and plotted as an RGB image. Although a faint colored hue is visible in the new measurement frame, it is mostly indistinguishable from the stationary scene measurement with the naked eye. The color-magnified difference measurement clearly contains penumbra content, although the hidden scene is not obvious simply by inspection. In fact, the darkest part of both difference measurements is the region closest to the hidden scene, suggesting that the majority of the change between frames is actually due to occlusion within the scene, rather than light introduced by the hidden objects. The unified reconstruction produced using these measurements is shown in Figure 10b



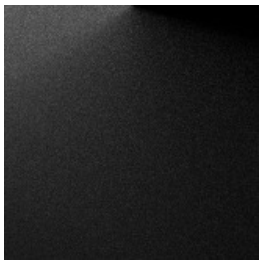
(a) Bird's eye ground truth



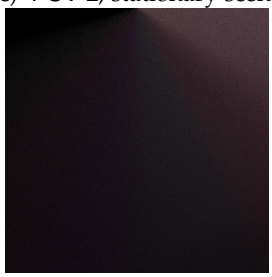
(b) $\mathbf{v}_u$



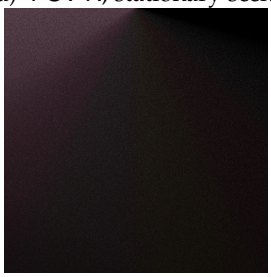
(c) FOV-L, stationary scene



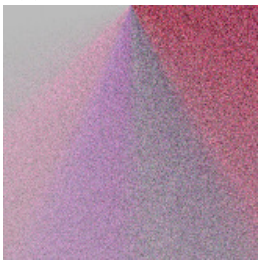
(d) FOV-R, stationary scene



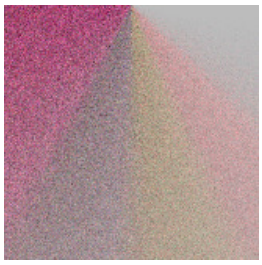
(e) FOV-L, new frame



(f) FOV-R, new frame



(g) FOV-L, color-magnified difference



(h) FOV-R, color-magnified difference

Fig. 10: The unified reconstruction result $\mathbf{v}_u$, produced with $t = 0.2$, for a four target scenario (a) are shown in (b). Measurements at FOV-L and FOV-R are shown in (c) and (d) for the stationary scene and in (e) and (f) for the new measurement frame. The color-magnified difference measurement is shown in (g) and (h).

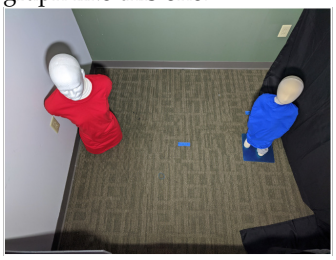
and closely matches the ground truth with correct object location and color.

5.2 Laboratory Experiments

Our laboratory setup is shown in Figure 11. We constructed a 2 m × 2 m hidden volume consisting of two existing walls (white and green), a curtain, and a 0.5 m wide foam-board doorway. As shown in Figure 11a, doorway walls were not extended vertically to the ceiling so a mock ceiling was added to prevent light from passing over the doorway and striking the measurement. The hidden scene was illuminated by a single floodlight placed on top of the mock ceiling. Mannequins dressed in different colored clothing were introduced to the hidden scene, as shown in Figure 11c. Two white squares were centered on the floor adjacent to each edge in FOV-L and FOV-R, and the penumbrae were captured in a single photograph with a DLSR camera pointed directly down at FOV-L and FOV-R, as shown in Figure 11b.[2] For each new frame, we captured a single 1.6 second exposure with the camera set to an f/5.6 aperture and ISO100. Camera hardware with better low-light performance would allow for a significantly shorter acquisition time. FOV-L and FOV-R were compressed as JPEG files, manually cropped from the original photograph, rectified, and downsampled to contain $M = 64^2$ pixels each.

(a) A photograph of the experimental setup taken from the visible side.

(b) Both FOV-L and FOV-R are extracted from a single photograph like this one.



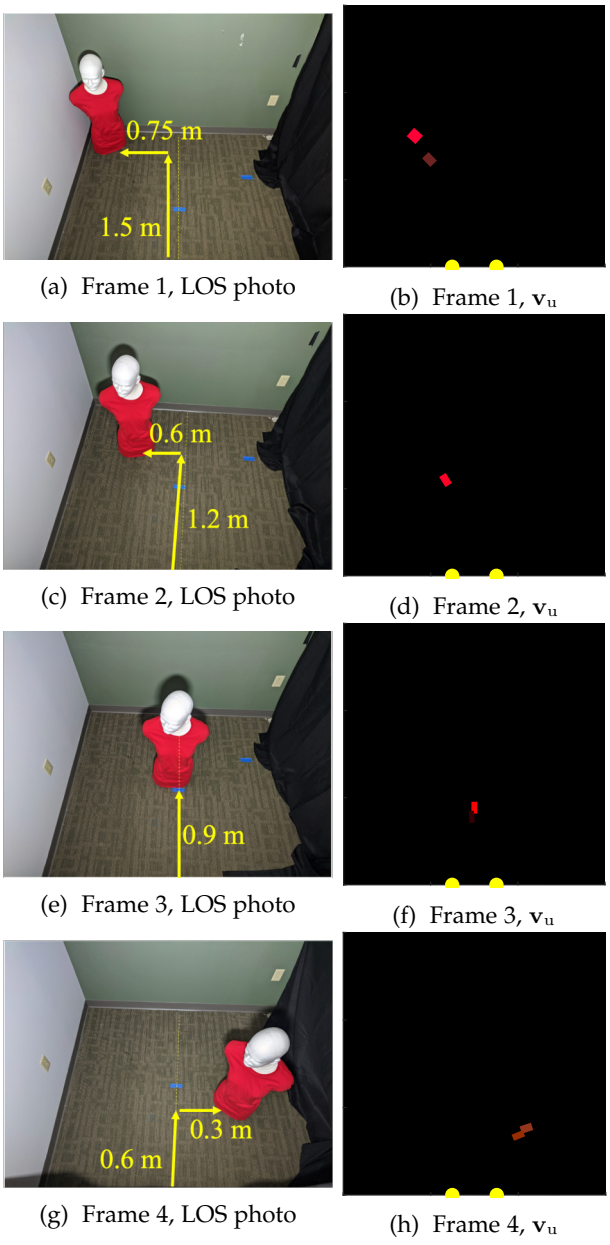
(c) A photograph of the hidden side of the doorway.

Fig. 11: Photographs of the laboratory setup show the test doorway (a), a sample measurement of FOV-L and FOV-R (b), and a LOS view of the hidden scene with two mannequins (c).

5.2.1 *Single object example*

In Figure 12, we show a sequence of results obtained when a mannequin, dressed in a red shirt, moves diagonally through the hidden scene. The first figure column shows an LOS photograph of the hidden scene, with the mannequin's position with respect to the center of the doorway annotated in yellow. The right column shows the corresponding unified reconstruction results $\mathbf{v}_u$ for each position. In the first three positions, the unified reconstruction accurately resolves the hidden target in location and color. In the fourth row, the target is more coarsely resolved. We suspect that this may be due to the unmodeled thickness of our foam core walls. Reconstruction error due to unmodeled wall thickness would be greatest further around one of the

2. This paper assumes that the camera is centered over FOV-L and FOV-R, looking directly down at the floor, so that the floor area seen by each camera pixel is approximately the same. The setup in Figure 1, where the camera views the floor from an angle, is achievable when differences in projected pixel area are incorporated into the model.



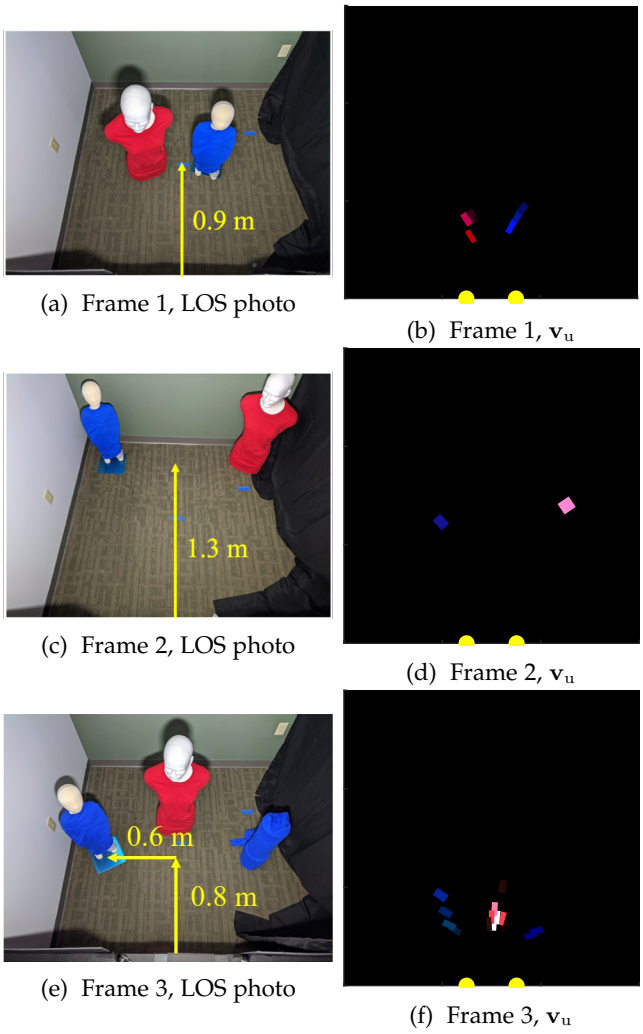
(a) Frame 1, LOS photo (b) Frame 1, $\mathbf{v}_u$

(c) Frame 2, LOS photo (d) Frame 2, $\mathbf{v}_u$

(e) Frame 3, LOS photo (f) Frame 3, $\mathbf{v}_u$

(g) Frame 4, LOS photo (h) Frame 4, $\mathbf{v}_u$

Fig. 12: Experimental results for four frames of a single target scenario. Results were produced using $t = 0$. Reconstruction results show a 3 m $\times$ 3 m region of the hidden scene. Some pixels in the reconstructions are saturated for improved visual clarity.



(a) Frame 1, LOS photo (b) Frame 1, $\mathbf{v}_u$

(c) Frame 2, LOS photo (d) Frame 2, $\mathbf{v}_u$

(e) Frame 3, LOS photo (f) Frame 3, $\mathbf{v}_u$

Fig. 13: Three frames of a multi-target scenario tested in the lab. Ground truth photographs are shown on the left and unified reconstructions $\mathbf{v}_u$ shown on the right. Results were produced using $t = 0.2$. The measurements corresponding to Frame 2 may be found in Figure 14. Some pixels in the reconstructions are saturated for improved visual clarity.

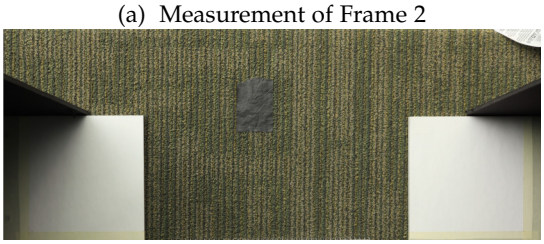
corners (i.e., one of the angles α_1 or α_2 is small and the other is large).

5.2.2 *Multi-object example*

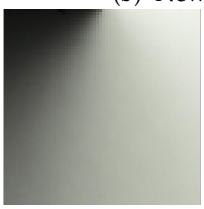
Figure 13 shows results for a scenario with multiple hidden targets. Figure rows correspond to different measurement frames. Frames 1 and 2 are acquired while the hidden scene contains two moving objects, one blue and one red. In the third frame, a third object (blue) has entered the hidden scene, and we test our ability to reconstruct multiple objects of the same color. Figure columns are (1) ground truth LOS photographs of the hidden scene; and (2) unified reconstruction results $\mathbf{v}_u$. Results in all measurement frames exhibit a close match to the ground truth of the new object. Note that in the third frame, the red and blue mannequins are reconstructed to be brighter than the blue cylinder. This accurately reflects their larger size and bright white heads.

The measured data for Frame 2 is shown in Figure 14. A cropped portion of the single stationary scene photograph, taken of the floor on the visible side before objects entered the hidden scene, is shown in Figure 14a. The photograph taken after the objects have entered the hidden scene is shown in Figure 14b. Cropped, rectified, and downsampled FOV-L and FOV-R are shown in Figures 14c and 14d for the stationary scene, and in Figure 14e for the new frame. In both the uncropped and cropped photographs, the stationary and new frame measurement are indistinguishable with the naked eye. In Figures 14g and 14h, we show the difference between the stationary and new measurement, scaled between 0 and 1 and plotted as an RGB image. Here, penumbra patterns are faintly visible with more noise at FOV-R because it has a more direct view of the bright white

(a) Measurement of Frame 2

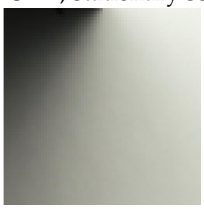


(b) New measurement frame



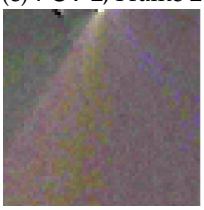
(c) FOV-L, stationary scene

(d) FOV-R, stationary scene

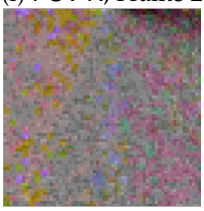


(e) FOV-L, Frame 2

(f) FOV-R, Frame 2



(g) FOV-L, color-magnified difference



(h) FOV-R, color-magnified difference

Fig. 14: Measurements corresponding to Frame 2 in Figure 13. The stationary measurement (a) and new measurement frame (b) are indistinguishable to the naked eye, although penumbra patterns may be seen in the their differences (g) and (h).

left wall in the hidden scene. These noisy difference measurements are what we use to produce the reconstruction result shown in Figure 13d.

6 Conclusion

In this work we introduce the doorway camera as a way to exploit the two vertical edges in a doorway, for robust passive 2D NLOS imaging. Unlike previous work that uses two separate corner cameras in stereo, our method jointly processes the data from two penumbra measurements and produces two unique 2D views of change in the hidden scene. Using the Cramér–Rao bound for a single hidden target, we demonstrate the utility of the two-edge doorway occluder over a single vertical edge occluder. Inspired by the azimuthal resolution provided by each edge, our novel inversion algorithm reconstructs the hidden scene on a biangular grid of planar facets. Synthetic and experimental results demonstrate the promise of our algorithm in a variety of scenarios, including challenging lighting conditions with as many as four moving objects.

Acknowledgment

This work was supported in part by the US National Science Foundation under grant number 1955219 and by the Draper Scholars program. Computing resources provided by Boston University Research Computing Service are greatly appreciated.

References

[1] D. Faccio, A. Velten, and G. Wetzstein, "Non-line-of-sight imaging," *Nature Reviews Physics*, vol. 2, no. 6, pp. 318–327, Jun 2020.

[2] A. Torralba and W. T. Freeman, "Accidental pinhole and pinspeck cameras: Revealing the scene outside the picture," *Int. J. Computer Vision*, vol. 110, no. 2, pp. 92–112, 2014.

[3] A. L. Cohen, "Anti-pinhole imaging," *Optica Acta*, vol. 29, no. 1, pp. 63–67, 1982.

[4] C. Saunders, J. Murray-Bruce, and V. K. Goyal, "Computational periscopy with an ordinary digital camera," *Nature*, vol. 565, no. 7740, pp. 472–475, 2019.

[5] A. B. Yedidia, M. Baradad, C. Thrampoulidis, W. T. Freeman, and G. W. Wornell, "Using unknown occluders to recover hidden scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2019.

[6] M. Aittala, P. Sharma, L. Murmann, A. Yedidia, G. W. Wornell, W. T. Freeman, and F. Durand, "Computational mirrors: Blind inverse light transport by deep matrix factorization," in *Advances in Neural Information Processing (NeurIPS)*, 2019, pp. 14 311–14 321.

[7] F. Wang, A. Eljarrat, J. Müller, T. R. Henninen, R. Erni, and C. T. Koch, "Multi-resolution convolutional neural networks for inverse problems," *Sci. Rep.*, vol. 10, no. 5730, 2020.

[8] C. Saunders and V. K. Goyal, "Fast computational periscopy in challenging ambient light conditions through optimized preconditioning," in *Proc. IEEE Int. Conf. Computational Photography*, 2021.

[9] R. Geng, Y. Hu, Z. Lu, C. Yu, H. Li, H. Zhang, and Y. Chen, "Passive non-line-of-sight imaging using optimal transport," *IEEE Trans. Image Process.*, vol. 31, pp. 110–124, Nov. 2021.

[10] T. Swedish, C. Henley, and R. Raskar, "Objects as cameras: Estimating high-frequency illumination from shadows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 2593–2602.

[11] P. Sharma, M. Aittala, Y. Y. Schechner, A. Torralba, G. W. Wornell, W. T. Freeman, and F. Durand, "What you can learn by staring at a blank wall," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 2330–2339.

[12] K. L. Bouman, V. Ye, A. B. Yedidia, F. Durand, G. W. Wornell, A. Torralba, and W. T. Freeman, "Turning corners into cameras: Principles and methods," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2017, pp. 2270–2278.

[13] S. W. Seidel, Y. Ma, J. Murray-Bruce, C. Saunders, W. T. Freeman, C. C. Yu, and V. K. Goyal, "Corner occluder computational periscopy: Estimating a hidden scene from a single photograph," in *Proc. IEEE Int. Conf. Computational Photography*, May 2019.

[14] S. W. Seidel, J. Murray-Bruce, Y. Ma, C. Yu, W. T. Freeman, and V. K. Goyal, "Two-dimensional non-line-of-sight scene estimation from a single edge occluder," *IEEE Trans. Comput. Imaging*, vol. 7, pp. 58–72, 2021.

[15] J. Rapp, C. Saunders, J. Tachella, J. Murray-Bruce, Y. Altmann, J.-Y. Tourneret, S. McLaughlin, R. M. A. Dawson, F. N. C. Wong, and V. K. Goyal, "Seeing around corners with edge-resolved transient imaging," *Nature Communications*, vol. 11, no. 5929, Nov. 2020.

[16] G. B. M. Zerr, "Biangular coordinates," *The American Mathematical Monthly*, vol. 17, no. 2, pp. 34–38, 1910.

[17] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.

[18] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
[19] Blender Online Community, *Blender - a 3D modelling and rendering package*, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. [Online]. Available: http://www.blender.org

William Krska Is currently an undergraduate student in computer engineering at Boston University, Boston, MA, USA. His research interests include computational imaging and high-performance computing.

Sheila W. Seidel received the B.S. degree in electrical engineering from Worcester Polytechnic Institute, Worcester, MA, USA, in 2013, and the M.S. degree in electrical engineering from the University of Massachusetts Amherst, Amherst, MA, USA, in 2016. She is a PhD candidate in electrical engineering at Boston University, Boston, MA, USA, and a Fellow at the Charles Stark Draper Laboratory, Cambridge, MA, USA. Her research interests include computational imaging, inverse problems, and statistical inference. She is a member of Tau Beta Pi and Eta Kappa Nu.

Charles Saunders received the M.Eng. degree in acoustical engineering from the University of Southampton, UK, in 2016, and the PhD degree in electrical engineering from Boston University, Boston, MA, USA, in 2021. He is currently a Postdoctoral Research Scientist at Meta. His research interests include computational imaging, 3D sensing, and optimization.

Robinson Czajkowski received the B.S. degree in mathematics, with a minor in astrophysics, from the University of Toledo, OH, USA. He is working towards the PhD degree at the University of South Florida, Tampa, FL, USA.

Christopher Yu received the B.S. degree at the Massachusetts Institute of Technology, Cambridge, MA, USA, and the M.S. and Ph.D. degrees at Princeton University, Princeton, NJ, USA, all in electrical engineering. He is the Software and Algorithms Engineering Director at Draper Laboratory, Cambridge, MA, USA, since 2004. At Draper, he has taken on various roles including Signal Processing and Communications Group Leader; Signals, Sensors, and Navigation Division Leader; Secure and Assured Software Systems Division Leader; Distinguished Member of Technical Staff; Education Office Director; and IRAD Director. Prior to Draper, Dr. Yu was technical staff at Bell Laboratories in Murray Hill, NJ, USA, for 7 years working in the areas of video compression and networks.

John Murray-Bruce (S'14—M'17) received the M.Eng. and Ph.D. degrees in electrical and electronic engineering (EEE) from the Imperial College London, UK, in 2012 and 2016, respectively. He is currently an assistant professor in the Department of Computer Science and Engineering of the University of South Florida, Tampa, FL, USA. From 2017 to 2019, he was a Postdoctoral Research Associate with the Electrical and Computer Engineering Department, Boston University, Boston, MA, USA. From 2012 to 2017, he was a Research Assistant with the Communications and Signal Processing Group, Imperial College London. His primary research interests are in computational imaging and inverse problems, as well as signal interpolation, approximation, and sampling theory. He is a member of the Computational Imaging Technical Committee of the IEEE Signal Processing Society. In 2012, he was the recipient of the Institute of Engineering and Technology Prize for the "best all round performance" in the M.Eng. EEE degree.

Vivek K Goyal (S'92–M'98–SM'03–F'14) received the B.S. degree in mathematics and the B.S.E. degree in electrical engineering from the University of Iowa, where he received the John Briggs Memorial Award for the top undergraduate across all colleges. He received the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, where he received the Eliahu Jury Award for outstanding achievement in systems, communications, control, or signal processing.

He was a Member of Technical Staff in the Mathematics of Communications Research Department of Bell Laboratories, Lucent Technologies, 1998–2001; and a Senior Research Engineer for Digital Fountain, Inc., 2001–2003. He was with the Massachusetts Institute of Technology 2004–2013, where he was the Esther and Harold E. Edgerton Associate Professor of Electrical Engineering. He was an adviser to 3dim Tech, Inc. (winner of the 2013 MIT $100K Entrepreneurship Competition Launch Contest Grand Prize), and was subsequently with Nest, an Alphabet company, 2014–2017.

Dr. Goyal is a member of Phi Beta Kappa and Tau Beta Pi. He was awarded the 2002 IEEE Signal Processing Society (SPS) Magazine Award, the 2014 IEEE Int. Conf. Image Processing Best Paper Award, and the IEEE SPS Best Paper Award in 2017 and 2019. Work he supervised won student best paper awards at the IEEE Data Compression Conf. in 2006 and 2011, the 2012 IEEE Sensor Array and Multichannel Signal Processing Workshop, and the 2018 IEEE Int. Conf. Image Processing, as well as the 2020 IEEE SPS Young Author Best Paper Award and the 2021 IEEE SPS Best PhD Dissertation Award. He was a Co-chair of the SPIE Wavelets and Sparsity conference series 2006–2016. He current serves on the Editorial Boards of IEEE TRANS. COMPUTATIONAL IMAGING, SIAM J. IMAGING SCIENCES, and *Foundations and Trends in Signal Processing*. He is a Fellow of Optica and a co-author of *Foundations of Signal Processing* (Cambridge University Press, 2014).