

Two-Dimensional Non-Line-of-Sight Scene Estimation from a Single Edge Occluder

Sheila W. Seidel, *Student Member, IEEE*, John Murray-Bruce, *Member, IEEE*, Yanting Ma, *Member, IEEE*, Christopher Yu, *Member, IEEE*, William T. Freeman, *Fellow, IEEE*, and Vivek K Goyal, *Fellow, IEEE*

Abstract—Passive non-line-of-sight imaging methods are often faster and stealthier than their active counterparts, requiring less complex and costly equipment. However, many of these methods exploit motion of an occluder or the hidden scene, or require knowledge or calibration of complicated occluders. The edge of a wall is a known and ubiquitous occluding structure that may be used as an aperture to image the region hidden behind it. Light from around the corner is cast onto the floor forming a fan-like penumbra rather than a sharp shadow. Subtle variations in the penumbra contain a remarkable amount of information about the hidden scene. Previous work has leveraged the vertical nature of the edge to demonstrate 1D (in angle measured around the corner) reconstructions of moving and stationary hidden scenery from as little as a single photograph of the penumbra. In this work, we introduce a second reconstruction dimension: range measured from the edge. We derive a new forward model, accounting for radial falloff, and propose two inversion algorithms to form 2D reconstructions from a single photograph of the penumbra. Performances of both algorithms are demonstrated on experimental data corresponding to several different hidden scene configurations. A Cramér–Rao bound analysis further demonstrates the feasibility (and utility) of the 2D corner camera.

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

I. INTRODUCTION

THE ability to form non-line-of-sight (NLOS) images would be useful in a variety of situations. It could help soldiers anticipate danger as they navigate a tunnel system, autonomous vehicles avoid collision, and first responders as they enter buildings. Current NLOS imaging methods may be *active*, based predominantly on the transient imaging framework first proposed in [1], [2] and requiring control of hidden scene illumination, or *passive*, where only light sources already present are used. The earliest active NLOS imaging systems

combined a femtosecond laser with a 2 picosecond resolution streak camera [2], [3]; newer systems using single-photon avalanche diode (SPAD) detectors and time-correlated single photon counting (TCSPC) modules provide a less expensive alternative. These systems have been used extensively for both line of sight imaging [4]–[6] and NLOS applications [7]–[19]. Recently, SPAD-based NLOS imaging systems have demonstrated faster processing using confocal scanning [16], reconstruction algorithms based on wave properties [17], [18], and color reconstructions using multiple wavelengths of illumination [19]. The system demonstrated in [20] scans a pulsed laser along an arc at the base of a vertical edge occluder, like the edge we use in this paper, to achieve reconstructions with high azimuthal resolution. An active non-SPAD-based system for tracking a hidden object was demonstrated in [21] using intensity-only measurements.

Compared to active methods, passive NLOS imaging techniques may be less expensive and stealthier, with lower power requirements and faster data acquisition. These passive methods leverage occluding structures and light sources already present in the environment [22]. Useful structures may be the aperture formed by a partially open window or door, or the ‘accidental pinhole’ formed when a once-present object is moved [23], [24]. Using an ordinary digital camera, Saunders *et al.* formed NLOS color reconstructions when the form of the occluder was known [25]. Other methods use the motion of the hidden scene to discern the shape of an unknown occluder [26], or deep matrix factorization to simultaneously reconstruct an unknown hidden scene and occluder [27]. Unlike other occluders used in NLOS imaging systems [25]–[29], a wall edge has a known shape and is ubiquitous. In this case, light is cast onto the visible floor around the occluding edge forming a *penumbra*. Photographs of the penumbra may be used to produce angularly resolved reconstructions of the hidden scene. This concept was introduced in [30], where smoothed differences between consecutive video frames were used to form one-dimensional reconstructions of changes in the hidden scene, even when the video frames were visually indistinguishable. Our previous work demonstrated 1D reconstruction of both moving and stationary hidden-scene components from a single photograph, while simultaneously estimating unknown nonuniform floor albedo [31].

In this paper, we go beyond 1D reconstructions by combining the resolving power of a vertical edge with information from the relationship between intensity and radial distance. To see why this is possible—though potentially difficult—we first examine the effects of certain idealized occluders. Consider a

This work was supported in part by a Draper Fellowship, by the Defense Advanced Research Projects Agency REVEAL Program under contract number HR0011-16-C-0030, and by the US National Science Foundation under Grants 1815896 and 1955219.

S. W. Seidel and V. K. Goyal are with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (sseidel@bu.edu; v.goyal@ieee.org).

J. Murray-Bruce is with the Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620 USA (murray-bruce@usf.edu).

Y. Ma is with Mitsubishi Electric Research Laboratories, Cambridge, MA 02139 USA (yma@merl.com).

C. Yu is with Charles Stark Draper Laboratory, Cambridge, MA 02139 USA (cyu@draper.com).

W. T. Freeman is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (billf@mit.edu).

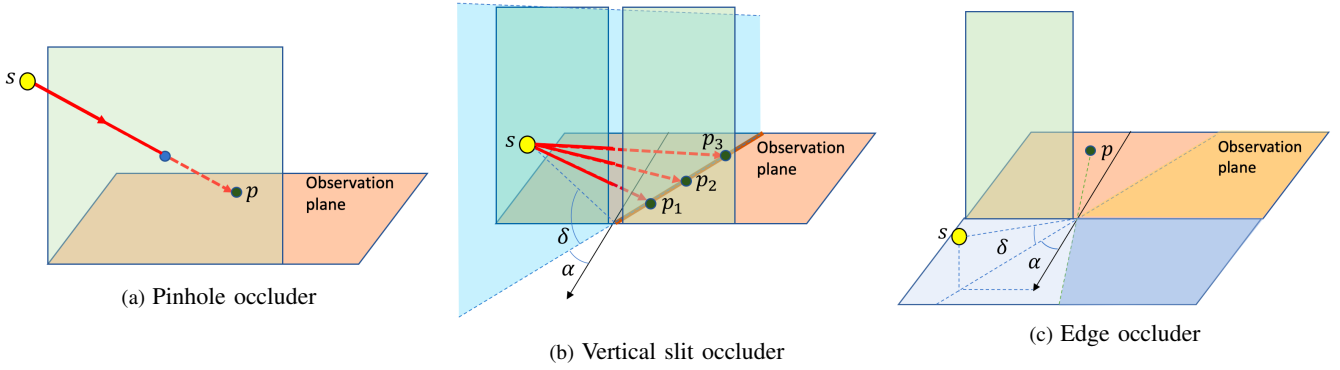


Fig. 1: The edge occluder (c) may be better understood by considering other well-known occluders, such as the pinhole (a) and vertical slit (b).

pinhole in a vertical plane, as shown in Fig. 1a. The direction (i.e., azimuth and elevation angles) of a point source s of unknown brightness is easily recoverable, but its range is not. If the scene is extended rather than a single point source, a sharp, classical pinhole projection of the scene is obtained. When the occluder has a vertical slit, as shown in Fig. 1b, a slice of the 3D world, a portion of which is shown in light blue, is mapped to a line (brown) on the observation plane. Here, the azimuthal angle α of incident light is easily recoverable. If there is a single omnidirectional point source in the planar slice at azimuthal angle α , its elevation angle δ and range can be recovered from the variation of light intensity, due to radial falloff and path length differences, at p_1 , p_2 , and p_3 . For an extended scene, the blurring in elevation angle δ is extremely ill-conditioned for inversion.

The edge occluder, shown in Fig. 1c, may be thought of as ‘half’ of a slit occluder. For a point source, the sharp shadow on the observation plane makes the azimuthal angle α again easily recoverable. The elevation angle δ and range are easier to recover than in the slit case because they influence the gradations of brightness over the entire unobscured portion of the observation plane (i.e., the yellow region in Fig. 1c). Imaging of an extended scene is more difficult than in the slit case because of range-dependent blurring over both α and δ .

Localization of a point source is both easier to achieve and easier to analyze. We study this in Section III, with the main objective of precisely characterizing how much easier the estimation of azimuthal angle α is than the estimation of range. Our greater interest is in imaging extended hidden scenes, and in principle, we could attempt a 3D scene reconstruction. However, Cramér–Rao bound analyses presented in Sections III and VI-A suggest that this 3D reconstruction problem would be ill-conditioned. As such, earlier works [30], [31] attempt only 1D reconstructions. We seek 2D reconstructions depending on angle and range (omitting height) in cylindrical coordinates centered at the vertical edge. This is motivated primarily by real-world scenes being dominated by objects that are approximately vertically oriented. The specifics of the 3D-to-2D mapping are given in Section II-A. Alternative inverse problem formulations for the same imaging configuration are discussed in Section VI-A.

Our key contributions include:

- A new forward model that describes a single photograph as a combination of light originating from a range of angles and depths in the hidden scene (Section II).
- Cramér–Rao bound (CRB) analysis (Section III) to demonstrate the limits of exploiting measurement of visible penumbras for 2D hidden scene reconstruction. Our analysis shows that while range estimation is possible, it is inherently difficult relative to angle estimation.
- Two different inversion algorithms, proposed in Section IV).
- Experimental demonstration of our 2D reconstruction algorithm on a variety of colored hidden scenes (Section V).

II. FORWARD MODEL

A. Light Transport

Consider the NLOS imaging scenario in Fig. 2, where the hidden scene consists of a green cylinder. Here we show a tripod, but in practice, the camera could be mounted anywhere that has a line-of-sight view of the floor adjacent to the corner. For example, it could be mounted on an autonomous vehicle or on a soldier’s helmet. As shown in Fig. 2, we parameterize the hidden scene in cylindrical coordinates with range ρ , angle α , and height z . A point $\mathbf{p} = (r, \theta)$ on the floor in the camera field of view is described by its range r and angle measured from the wall θ . Assuming the camera looks straight down at a Lambertian floor,¹ and the effects of all foreshortening factors are negligible, the brightness $L_o(\mathbf{p})$ of point $\mathbf{p}$ on the floor is the albedo $f(\mathbf{p})$ at point $\mathbf{p}$ multiplied by the sum of all incident light:

$$L_o(\mathbf{p}) = f(\mathbf{p}) (L_v(\mathbf{p}) + L_h(\mathbf{p})), \quad (1)$$

where $L_v(\mathbf{p})$ is the incident ambient light originating from the visible side, and $L_h(\mathbf{p})$ is the incident light originating from the hidden side.

¹For the sake of demonstration, this paper assumes the camera is centered above the observation region, looking directly down at the floor. At this viewing angle, the floor area seen by each camera pixel is approximately the same. The setup in Fig. 2, where the camera views the floor from an angle, is realizable when differences in projected pixel area are included in the model.

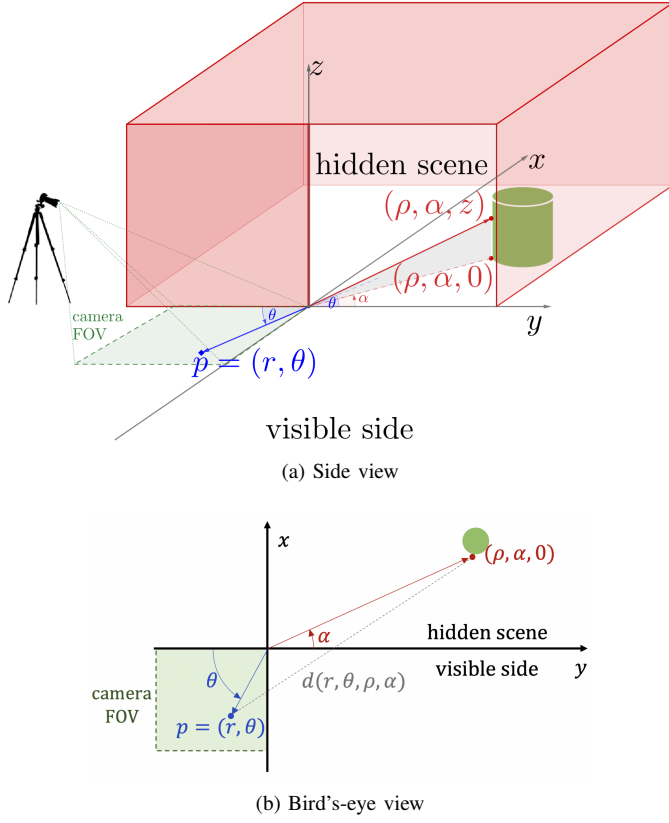


Fig. 2: Acquisition setup and depiction of problem geometry. A point p in the camera's field of view is represented in polar coordinates, while the 3D hidden scene is represented in cylindrical coordinates.

The measured photograph is an array of size $M_x \times M_y$, with $M = M_x M_y$ total pixels. The measurement y_m of camera pixel m is equal to the total radiosity of camera pixel $\mathcal{P}_m$, which consists of all points $\mathbf{p}$ on the measurement plane that are focused on camera pixel m . Thus,

$$y_m = \int_{\mathbf{p} \in \mathcal{P}_m} L_o(\mathbf{p}) d\mathbf{p},$$

where we have ignored the constant scaling factor associated with the camera because we are not attempting to estimate a physically meaningful overall scaling factor for the hidden scene. Because all camera pixels have equal projected area $\kappa_{\text{cam}} \stackrel{\text{def}}{=} \text{area}(\mathcal{P}_m)$ on the measurement plane, and because $\kappa_{\text{cam}} = 1$ without loss of generality, we can use (1) to write

$$\begin{aligned} y_m &\approx \kappa_{\text{cam}} L_o(r_m, \theta_m) \\ &= f(r_m, \theta_m) (L_v(r_m, \theta_m) + L_h(r_m, \theta_m)), \end{aligned} \quad (2)$$

where (r_m, θ_m) is the center of camera pixel $\mathcal{P}_m$.

By adopting a cylindrical coordinate parameterization of the hidden scene, the hidden scene contribution $L_h(r, \theta)$ is the sum of incident light L_i coming from different directions:

$$\begin{aligned} L_h(r, \theta) &= \int_0^\theta \int_0^\infty \int_0^\infty L_i(\rho, \alpha, z) \rho dz d\rho d\alpha \\ &= \int_0^\theta \int_0^\infty \int_0^\infty \frac{S_h(\rho, \alpha, z)}{d^2(r, \theta, \rho, \alpha) + z^2} \rho dz d\rho d\alpha, \end{aligned} \quad (3a)$$

where $S_h(\rho, \alpha, z)$ is the radiosity of a hidden scene location (ρ, α, z) , assumed to have no directional dependence over the extent of the camera FOV, and

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

is the squared distance between point $\mathbf{p}$ on the visible floor and a hidden scene (floor) point $(\rho, \alpha, 0)$. The inversion of (3) to estimate the 3D scene $S_h(\rho, \alpha, z)$ from $L_h(r, \theta)$ may be possible with suitable discretization and regularization. While we do not study the conditioning of this problem formally, it is presumably quite poorly conditioned. Instead, we would like to write an approximation of (3) that has a 2D hidden scene representation and then estimate that representation.

B. 2D Scene Model

Inspired by [30], [31], the presence of the vertical edge occluding our view of the hidden scene makes the α dependence of $S_h(\rho, \alpha, z)$ the easiest dimension to estimate. We choose to represent the dependence on ρ rather than on z to recover a 2D plan view representation of the hidden scene. To this end, we rewrite (3a) as

$$L_h(r, \theta) = \int_0^\theta \int_0^\infty \frac{\rho}{d^2(r, \theta, \rho, \alpha)} \bar{S}_h(r, \theta, \rho, \alpha) d\rho d\alpha, \quad (4a)$$

where

$$\bar{S}_h(r, \theta, \rho, \alpha) = \int_0^\infty \frac{S_h(\rho, \alpha, z)}{1 + (z/d(r, \theta, \rho, \alpha))^2} dz. \quad (4b)$$

For (4) to be a forward operator to invert for recovery of a 2D representation of the scene, we would like to replace $\bar{S}_h$ with an approximation with no dependence on r and θ . If scene content is mostly short relative to the horizontal dimensions, $z \ll d(r, \theta, \rho, \alpha)$ where $S_h(\rho, \alpha, z)$ is appreciable. The 2D scene representation

$$\bar{S}_h(\rho, \alpha) = \int_0^\infty S_h(\rho, \alpha, z) dz \quad (5)$$

thus satisfies

$$\bar{S}_h(\rho, \alpha) \approx \bar{S}_h(r, \theta, \rho, \alpha)$$

for all (r, θ) in the camera FOV. Using this approximation, the inversion we study in Section IV is of

$$L_h(r, \theta) = \int_0^\theta \int_0^\infty \frac{\rho}{d^2(r, \theta, \rho, \alpha)} \bar{S}_h(\rho, \alpha) d\rho d\alpha. \quad (6)$$

When a target has appreciable height, we expect there to be some distortion. In simulation, we found that taller targets were estimated to be farther from the occluding edge than shorter targets.

To model occlusion in the hidden scene, we assume that all the contributions to our measurement from a given angle α come from a single range ρ . This roughly corresponds to a hidden scene composed of opaque vertical objects resting on the ground. Under this assumption, we write $\bar{S}_h(\rho, \alpha) = \delta(\rho - \rho_h(\alpha)) s_h(\alpha)$ as a separable function of range $\rho_h(\alpha) \geq 0$ and angle $\alpha \in (0, \pi]$, where $s_h(\alpha)$ denotes the dependence of

scene radiosity on α , and $\delta(\cdot)$ is the Dirac delta function. Then (6) becomes

$$\begin{aligned} L_h(r, \theta) &= \int_0^\theta \int_0^\infty \frac{\rho}{d^2(r, \theta, \rho, \alpha)} \delta(\rho - \rho_h(\alpha)) s_h(\alpha) d\rho d\alpha \\ &= \int_0^\theta \frac{\rho_h(\alpha)}{d^2(r, \theta, \rho_h(\alpha), \alpha)} s_h(\alpha) d\alpha. \end{aligned}$$

Thus, substituting

$$L_h(r_m, \theta_m) = \int_0^{\theta_m} \frac{\rho_h(\alpha)}{d^2(r_m, \theta_m, \rho_h(\alpha), \alpha)} s_h(\alpha) d\alpha \quad (7)$$

into (2) we obtain the model

$$\begin{aligned} y_m \approx f(r_m, \theta_m) &\left(L_v(r_m, \theta_m) \right. \\ &\left. + \int_0^{\theta_m} \frac{\rho_h(\alpha)}{d^2(r_m, \theta_m, \rho_h(\alpha), \alpha)} s_h(\alpha) d\alpha \right) \end{aligned} \quad (8)$$

for the hidden scene and visible scene contributions to camera measurement m .

C. Discrete Forward Model

For the sake of concisely demonstrating our core ideas, this paper focuses on estimating half of the hidden scene, i.e., the region $\alpha \in (0, \pi/2]$.² We discretize the hidden region into N equiangular wedges identified by their center angles

$$\alpha_n = \frac{(n - \frac{1}{2})\pi}{2N}, \quad n = 1, 2, \dots, N,$$

and associate a single unknown range value $\rho_h(\alpha_n)$ with each wedge. Then the pair $(\rho_h(\alpha_n), \alpha_n)$ defines a (unique) position in the hidden space for each $n \in \{1, 2, \dots, N\}$. Now gathering these variables into the hidden-scene radiosity vector $\mathbf{s}_h = [s_h(\alpha_1), s_h(\alpha_2), \dots, s_h(\alpha_N)]^\top$ and range vector $\boldsymbol{\rho}_h = [\rho_h(\alpha_1), \rho_h(\alpha_2), \dots, \rho_h(\alpha_N)]^\top$ gives the discrete, non-linear forward model

$$\mathbf{y} = \mathbf{a} \odot \mathbf{f} + \mathbf{f} \odot (\mathbf{V} \odot \mathbf{D}(\boldsymbol{\rho}_h)) \mathbf{s}_h + \boldsymbol{\epsilon}, \quad (9)$$

where $\mathbf{y} = [y_1, y_2, \dots, y_m] \in \mathbb{R}^M$ denotes the vectorized camera photograph, $\mathbf{a} \in \mathbb{R}^M$ is the discretization of ambient light contribution L_v , $\mathbf{f} \in \mathbb{R}^M$ is the floor albedo, $\mathbf{V} \in \mathbb{R}^{M \times N}$ is a binary-valued visibility matrix (with the entry $[\mathbf{V}]_{m,n}$ equalling 0 if the path joining p_m and $(\rho(\alpha_n), \alpha_n)$ is occluded by the wall, otherwise it is equal to 1), and $\odot$ is Hadamard (or element-wise) product. Note that when no target is present at angle α_n , $s_h(\alpha_n) = 0$. The matrix $\mathbf{D}(\boldsymbol{\rho}_h) \in \mathbb{R}^{M \times N}$ has elements

$$[\mathbf{D}(\boldsymbol{\rho}_h)]_{m,n} = \frac{\rho_h(\alpha_n)}{d^2(r_m, \theta_m, \rho_h(\alpha_n), \alpha_n)},$$

and $\boldsymbol{\epsilon}$ models noise and other possible model mismatch.

Inverse Problem: Our goal is to recover a 2D (plan view) reconstruction $(\mathbf{s}_h, \boldsymbol{\rho}_h)$ of a hidden scene S_h from a single

²Our proposed methods could be extended to reconstruct the *full* hidden scene, i.e., $\phi_s \in (0, \pi]$, by: (i) extending the camera's FOV to also include measurements from the positive y -axis of the visible side floor surface (see Fig. 2); and (ii) carefully accounting for the thickness of the occluding wall if necessary.

photograph $\mathbf{y}$ of the penumbra created on a visible floor surface using (9).

Before presenting our approaches for solving (9), we study the feasibility (and certain limits) of realizing the 2D corner camera. Specifically, by evaluating the CRBs for hypothetical cases where the hidden scene comprises only a few hidden point targets, we demonstrate the merits of the occluding wall (or corner occluder) for hidden scene recovery.

III. CRAMÉR–RAO BOUND FOR TARGET LOCALIZATION

In the subsections that follow, we present CRB analysis to demonstrate the merit and challenge of an edge occluder for 2D plan-view reconstruction of a hidden scene. To truly understand the effect of the edge, we perform our analysis both for the edge occluder scenario, and the scenario where no edge is in place. We start with the former.

In (7), measurement y_m , with the edge in place, is approximated by the intensity at the center of the pixel. Now, we leave the more precise integral across points $\mathbf{p} = (r, \theta)$, in floor patch $\mathcal{P}_m$, in place and assume no ambient light contributions, i.e. $L_v = 0$. Under an additive white Gaussian noise (AWGN) model, the noisy camera measurement is given by

$$y_m = \int_{\mathbf{p} \in \mathcal{P}_m} \int_0^{\theta_m} \frac{\rho_h(\alpha)}{d^2(r, \theta, \rho_h(\alpha), \alpha)} s_h(\alpha) d\alpha d\mathbf{p} + \epsilon, \quad (10)$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$.

A. Single Hidden Target

Assume the hidden target is a hypothetical point emitter, located at the point $(\rho_s, \phi_s, 0)$ on the ground, i.e. $S_h(\rho, \alpha, z) = c_s \delta(\rho - \rho_s) \delta(\alpha - \phi_s) \delta(z)$, where $\phi_s \in (0, \pi/2]$. Evaluating (3a), the outgoing radiosity from a point $\mathbf{p} = (r, \theta)$ is

$$L_o(\mathbf{p}) = f(\mathbf{p}) \frac{c_s \rho_s H(\theta - \phi_s)}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \pi - \theta)}, \quad (11)$$

where $H(x)$ is the Heaviside step function.

Assuming a uniform albedo $f(\mathbf{p}) = 1$, the measurement at pixel m is $y_m = i_m + \epsilon$, with

$$i_m = \int_{\mathbf{p} \in \mathcal{P}_m} \frac{c_s \rho_s H(\theta - \phi_s)}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \pi - \theta)} d\mathbf{p}. \quad (12)$$

We will now analyze how small changes in target location (ρ_s, ϕ_s) and brightness c_s propagate to small changes in the camera measurements. This will allow us to analyze how much information about these unknown parameters is contained in our noisy measurements y_m . The derivatives of i_m with respect to the unknown parameters are as follows:

$$\begin{aligned} \frac{\partial i_m}{\partial c_s} &= \int_{\mathbf{p} \in \mathcal{P}_m} \frac{\rho_s H(\theta - \phi_s)}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \pi - \theta)} d\mathbf{p}, \\ \frac{\partial i_m}{\partial \rho_s} &= c_s \int_{\mathbf{p} \in \mathcal{P}_m} \frac{(r^2 - \rho_s^2) H(\theta - \phi_s)}{(r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \pi - \theta))^2} d\mathbf{p}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial i_m}{\partial \phi_s} &= -c_s \rho_s \int_{\mathbf{p} \in \mathcal{P}_m} \frac{\delta(\theta - \phi_s)}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \pi - \theta)} \\ &\quad + \frac{2r\rho_s \sin(\phi_s + \pi - \theta) H(\theta - \phi_s)}{(r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \pi - \theta))^2} d\mathbf{p}. \end{aligned}$$

Interchanging the integral and derivative is justified since the definite integral i_m is finite. We define the following matrix:

$$\nabla \mathbf{I} = \begin{bmatrix} \frac{\partial i_1}{\partial c_s} & \frac{\partial i_1}{\partial \rho_s} & \frac{\partial i_1}{\partial \phi_s} \\ \frac{\partial i_2}{\partial c_s} & \frac{\partial i_2}{\partial \rho_s} & \frac{\partial i_2}{\partial \phi_s} \\ \vdots & \vdots & \vdots \\ \frac{\partial i_M}{\partial c_s} & \frac{\partial i_M}{\partial \rho_s} & \frac{\partial i_M}{\partial \phi_s} \end{bmatrix} \quad (13)$$

and note that under our Gaussian model, the Fisher information matrix, which describes the amount of information our data contains about unknown parameters, for estimating (c_s, ρ_s, ϕ_s) from the noisy measurements $\{y_m\}_{m=1}^M$ is given by

$$\begin{aligned} \mathbf{F} &= \frac{1}{\sigma^2} (\nabla \mathbf{I}^T \nabla \mathbf{I}), \\ &= \frac{1}{\sigma^2} \begin{bmatrix} \sum_m \left(\frac{\partial i_m}{\partial c_s} \right)^2 & \sum_m \frac{\partial i_m}{\partial c_s} \frac{\partial i_m}{\partial \rho_s} & \sum_m \frac{\partial i_m}{\partial c_s} \frac{\partial i_m}{\partial \phi_s} \\ \sum_m \frac{\partial i_m}{\partial \rho_s} \frac{\partial i_m}{\partial c_s} & \sum_m \left(\frac{\partial i_m}{\partial \rho_s} \right)^2 & \sum_m \frac{\partial i_m}{\partial \rho_s} \frac{\partial i_m}{\partial \phi_s} \\ \sum_m \frac{\partial i_m}{\partial \phi_s} \frac{\partial i_m}{\partial c_s} & \sum_m \frac{\partial i_m}{\partial \phi_s} \frac{\partial i_m}{\partial \rho_s} & \sum_m \left(\frac{\partial i_m}{\partial \phi_s} \right)^2 \end{bmatrix}. \end{aligned} \quad (14)$$

The CRB provides a lower bound for the achievable uncertainty of an unbiased estimator and is related to the reciprocal of the Fisher information:

$$\text{CRB}_{\text{EO}}(c_s) = \sigma^2 [\mathbf{F}^{-1}]_{1,1}, \quad (15a)$$

$$\text{CRB}_{\text{EO}}(\rho_s) = \sigma^2 [\mathbf{F}^{-1}]_{2,2}, \quad (15b)$$

$$\text{CRB}_{\text{EO}}(\phi_s) = \sigma^2 [\mathbf{F}^{-1}]_{3,3}, \quad (15c)$$

where the EO subscript indicates that these CRB results are for the edge occluder scenario. In other words, greater information about the unknown parameters corresponds to lower estimate uncertainty.

Without the occluding edge, corresponding CRBs ($\text{CRB}_{\text{no-EO}}(c_s)$, $\text{CRB}_{\text{no-EO}}(\rho_s)$, and $\text{CRB}_{\text{no-EO}}(\phi_s)$) for estimating the same “out-of-view” target parameters follow similarly. Without the occlusion described in (10), the measurement by the m th camera pixel is

$$\begin{aligned} y_m^{\text{no-EO}} &\stackrel{\text{def}}{=} \int_{\mathbf{p} \in \mathcal{P}_m} \frac{c_s \rho_s}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \theta)} d\mathbf{p} + \epsilon \\ &= i_m^{\text{no-EO}} + \epsilon. \end{aligned}$$

Using the derivatives of $i_m^{\text{no-EO}}$ with respect to the hidden target’s parameters:

$$\begin{aligned} \frac{\partial i_m^{\text{no-EO}}}{\partial c_s} &= \int_{\mathbf{p} \in \mathcal{P}_m} \frac{\rho_s}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \theta)} d\mathbf{p}, \\ \frac{\partial i_m^{\text{no-EO}}}{\partial \rho_s} &= c_s \int_{\mathbf{p} \in \mathcal{P}_m} \frac{(r^2 - \rho_s^2)}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \theta)} d\mathbf{p}, \end{aligned}$$

and

$$\frac{\partial i_m^{\text{no-EO}}}{\partial \phi_s} = c_s \rho_s \int_{\mathbf{p} \in \mathcal{P}_m} \frac{2r\rho_s \sin(\theta + \phi_s)}{r^2 + \rho_s^2 - 2r\rho_s \cos(\phi_s + \theta)} d\mathbf{p},$$

the Fisher information matrix $\mathbf{F}_{\text{no-EO}}$, along with CRBs ($\text{CRB}_{\text{no-EO}}(c_s)$, $\text{CRB}_{\text{no-EO}}(\rho_s)$, and $\text{CRB}_{\text{no-EO}}(\phi_s)$) may be computed using the approach outlined in (13, 14, 15).

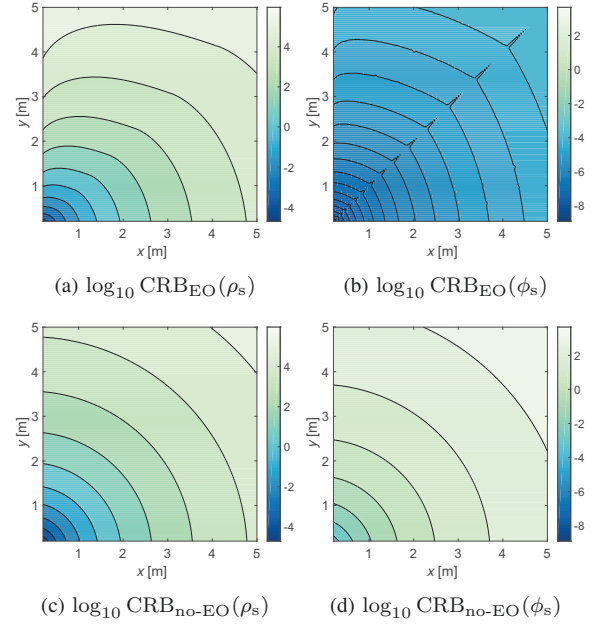


Fig. 3: Variation of the CRBs for estimating a single hidden target for different target locations. The number of measurement pixels is $M = 155^2$ and the measurement FOV is $0.2 \text{ m} \times 0.2 \text{ m}$, with fixed noise variance $\sigma^2 = 10$.

Contour plots of computed CRBs for various ground truth target positions with respect to the origin (corner) are shown in Fig. 3, for the corner (CRB_{EO}) and no corner cases ($\text{CRB}_{\text{no-EO}}$). Comparing Figs. 3a and 3c, achievable target range estimates MSE has marginal dependence on the presence of a corner, when estimating a single point target. On the other hand, Figs. 3b and 3d suggest that CRBs for angle estimates with the corner are around five to seven orders of magnitude smaller when compared to the no-corner case.

Fixing the target’s range at $\rho_s = 1 \text{ m}$, Fig. 4 summarizes the dependence of the computed CRBs on $\phi_s \in [\pi/64, 63\pi/64]$ rads. First, Fig. 4a shows that $\text{CRB}_{\text{EO}}(\rho_s)$ and $\text{CRB}_{\text{no-EO}}(\rho_s)$ are nearly equal at very shallow target angles, because the shadowed region in the occluded case is very small (the measurements for the corner and no corner cases are almost the same). However, with measurement noise variance fixed and ϕ_s increasing, $\text{CRB}_{\text{EO}}(\rho_s)$ diverges because the in-shadow region—which cannot possibly be informative about the occluded target’s distance—grows, while $\text{CRB}_{\text{no-EO}}(\rho_s)$ changes only marginally (reaching a maximum at $\pi/4$ rads before decreasing again). At the deepest angle, $\text{CRB}_{\text{EO}}(\rho_s)$ is roughly 28 times $\text{CRB}_{\text{no-EO}}(\rho_s)$. Figs. 4b and 4c indicate that $\text{CRB}_{\text{no-EO}}(\phi_s)$ has relatively mild dependence on the true target angle ϕ_s , with symmetry around $\pi/4$. The observed partial symmetry, in Fig. 4c, about $\pi/4$, with $\phi_s \in [\pi/8, 3\pi/8]$ is because, in contrast to range estimation, the in-shadow region is also informative (subject to prevalent noise levels) about the target’s angular position. The asymmetry (for $\phi_s \notin [\pi/8, 3\pi/8]$) is explained by a fixed noise variance (i.e., measurement SNR reduces with increasing target angle). Overall, the variation in $\text{CRB}_{\text{EO}}(\phi_s)$ is small relative to the roughly five orders of magnitude improvement due to the occluding wall.

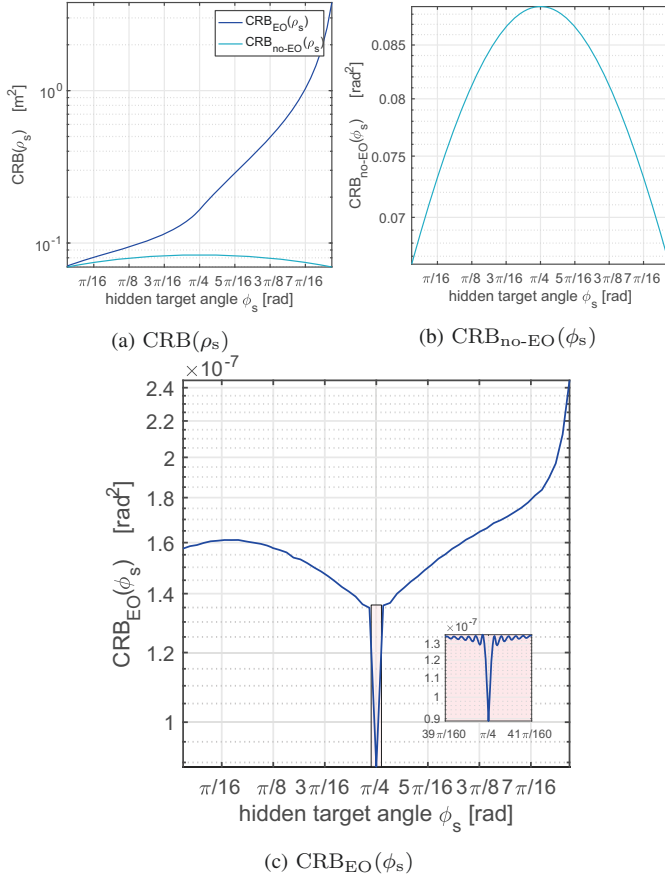


Fig. 4: Variation of CRB of the hidden target estimate in response to varying the target's angular position. (a) CRBs for range estimates with and without a corner. (b) CRB for angle estimates without a corner. (c) CRB for angle estimates with a corner camera. Camera FOV = $0.2 \text{ m} \times 0.2 \text{ m}$, $\sigma^2 = 10$, and $\rho_s = 1 \text{ m}$.

Second, with the target's angle held constant ($\phi_s = \pi/3$ rads) while its distance from the corner increases from zero, Fig. 5a shows that $\text{CRB}_{\text{EO}}(\rho_s)$ and $\text{CRB}_{\text{no-EO}}(\rho_s)$ are both small for a close target, but increase dramatically with target's distance. The uninformative nature of the in-shadow measurements for range estimation causes $\text{CRB}_{\text{EO}}(\rho_s)$ to be higher than $\text{CRB}_{\text{no-EO}}(\rho_s)$, whereas the presence of the corner makes $\text{CRB}_{\text{EO}}(\phi_s)$ at least five orders of magnitude lower than $\text{CRB}_{\text{no-EO}}(\phi_s)$ (see Fig. 5b). Under our measurement scenario, a target 3 m from the corner (with $\phi_s = \pi/3$) for instance has $\sqrt{\text{CRB}_{\text{EO}}(\phi_s)} \approx 10^{-5/2} = 0.003$ rads, while $\sqrt{\text{CRB}_{\text{no-EO}}(\phi_s)} \approx 10^{-3/4} = 0.178$ rads.

Our study for a single point target demonstrates overwhelming improvement in the estimation of ϕ_s due to the occluding wall, with marginal negative impact on the expected estimation quality of ρ_s . This is because the occluding wall effectively separates light paths arising from different angles in the hidden region. Phrased differently, the exact proportion of shadowed-to-nonshadowed regions within the camera's FOV is informative about the angular location of the hidden target.

Using the CRBs, one can compute theoretical spatial uncertainty regions for a hidden target. These are regions within which the majority of a target's estimates are expected to fall. Specific examples for a camera FOV of $0.15 \text{ m} \times 0.15 \text{ m}$

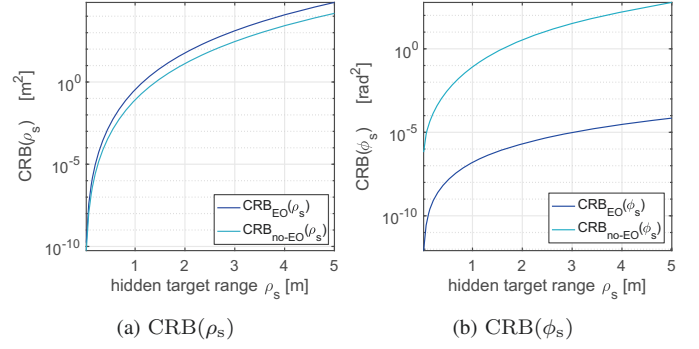


Fig. 5: Variation of CRB of the hidden target estimate in response to varying the target's distance from the corner ρ_s . FOV = $[0.2 \text{ m} \times 0.2 \text{ m}]$, $\sigma^2 = 10$, $\phi_s = \pi/3$ rads.

and $M = 155^2$ pixels, and at SNR levels resembling real experimental measurements are shown in Fig. 6. Assuming unbiased estimators that achieve the CRB, these bubbles depict regions within which two standard deviations of a target's estimate are expected to fall. We observe that the uncertainty regions are very different with and without the occluding wall: almost circular for the latter, while the angular uncertainties are virtually imperceptible for the former. The presence of the corner collapses these bubbles into lines, with the length of each line representing the uncertainty in range, while angular uncertainties are almost completely removed.

B. Multiple Hidden Targets

Although our single point-target CRB analysis showed the incredible benefits of the occluding wall in ϕ_s estimation, estimation of ρ_s was actually shown to be slightly more challenging, especially for hidden targets at greater angular depths. The benefit of the occluding wall in range estimation is realized when the hidden scene is more complicated. We extend our single point target CRB analysis to include a second hidden point target to demonstrate this effect.

In Fig. 7, Target 1 is fixed at $(\rho_1, \phi_1) = (1 \text{ m}, \pi/4 \text{ rads})$ while Target 2 is held at $\rho_2 = 2 \text{ m}$ and moved in angle ϕ_2 . The CRB for both parameters and targets are compared for scenarios with and without the corner in place. Fig. 7a shows that $\text{CRB}(\rho_1)$ and $\text{CRB}(\rho_2)$ are, generally, over an order of magnitude smaller when the corner is in place, the only exception being when both targets are at or very near the same angle. In this case, it becomes difficult to isolate the two targets in range. Just like the single-target scenario, $\text{CRB}_{\text{EO}}(\phi_1)$ and $\text{CRB}_{\text{EO}}(\phi_2)$ are seen to be many orders of magnitude smaller (than $\text{CRB}_{\text{no-EO}}(\phi_1)$ and $\text{CRB}_{\text{no-EO}}(\phi_2)$) in Fig. 7b. This significantly improved angular resolution depends on the ability to separate angular derivatives due to each target, which becomes more challenging when they are very close to each other in angle, causing the peak at $\phi_2 = \pi/4$ in Fig. 7b. When the two targets are at or near the same angle, the no-corner case, which relies exclusively on radial falloff, shows improvement due to contributions from each target adding constructively in the measurement. Though that improvement is marginal relative to improvement from having a corner.

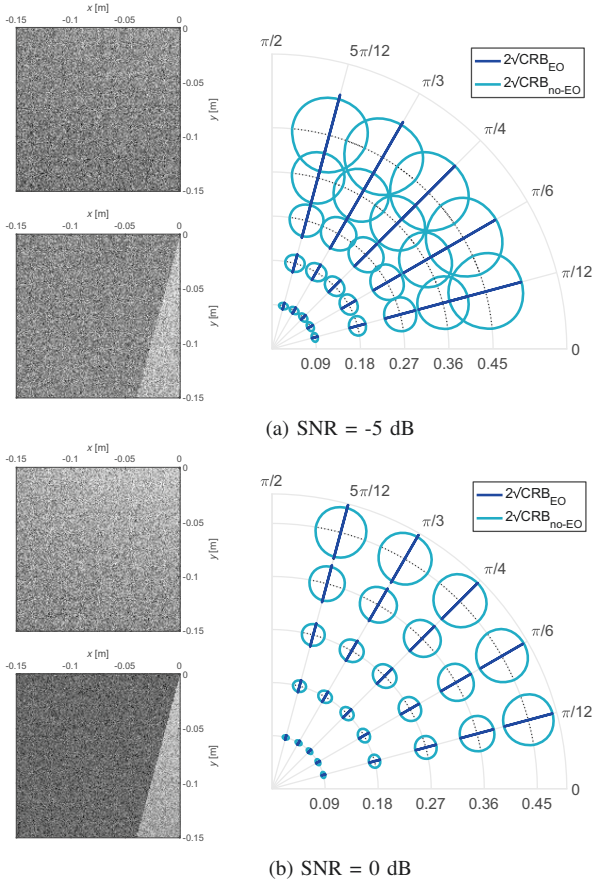


Fig. 6: $2 \times \sqrt{\text{CRB}}$ uncertainty regions (right) for various measurement SNR levels. Each uncertainty region is an ellipse (in polar coordinates) with minor and major axis length set to $4 \times \sqrt{\text{CRB}}$ for the corresponding dimension. Camera FOV = $0.15\text{m} \times 0.15\text{m}$, one typical realization of the camera measurement made at the corresponding SNR assuming no occluding wall (insets: top) and with an occluding wall (insets: bottom). The number of camera pixels $M = 155^2$.

Even when the angular location of both targets is given, $\text{CRB}(\rho; \phi_1, \phi_2)$ is still substantially lower for the corner camera case, as shown in Fig. 7c. This may be explained by the fact that light from the shallowest (in angle) target in the hidden scene affects a larger angular wedge in the measurement than the less shallow target. The difference between these two wedges is a swath of pixels affected only by the shallowest target, making range estimation for that target easier. In contrast, light from both targets without the corner strikes all pixels in the measurement, making separating each radial falloff pattern more challenging.

In addition, note that knowing the angular location of both targets has a negligible effect on $\text{CRB}(\rho; \phi_1, \phi_2)$ when the wall is in place, likely due to the fact that angular uncertainty is already so low in that situation. Inspired by these observations, we introduce an algorithm that alternates between estimating ϕ and ρ in Section IV-C.

IV. INVERSE PROBLEMS & ALGORITHMS

In this section, we present two approaches to form a plan-view reconstruction from a photograph of the penumbra. The

first method, described in Section IV-B, discretizes the hidden scene into a polar grid of pixels; thus transforming our inverse problem into a linear problem of estimating the intensity of each *polar pixel*. While this method is straightforward, we demonstrate improved reconstructions using a second approach that solves the inverse problem introduced in II-C by alternating between estimating angular and range information. First, by exploiting the high angular resolution provided by the corner an initial estimate of the scene is formed, as a function of angle. From this initial profile of the scene, the number of hidden targets is estimated. Finally, we alternate between estimating a single range for each target (i.e., *learning* the true forward model), and updating the angular profile.

A. Floor Albedo and Ambient Light

Jointly estimating $\mathbf{f}$ along with a 1D projection of the hidden scene has been studied in [31], with the assumption that ambient—or visible side—light contribution to the measurements $\mathbf{a} \approx c_1 \mathbf{1}$ is approximately constant over the camera FOV. This work assumes uniform floor albedo $f(r, \theta)$ (i.e., $\mathbf{f} = \mathbf{1}$), though we remark that both inversion methods can be similarly extended to handle the case of spatially varying floor albedo $\mathbf{f}$. This is by no means trivial and we leave it for a future work. As previously mentioned, it is also assumed that the floor surface is Lambertian. Given the intended tripod viewing angle shown in Fig. 2a, specular mirror-like reflections would provide little useful information about the hidden scene. When specular reflections are present, these regions of the floor may be isolated and those pixels removed from the data. This was not necessary in the experiments that follow. In addition, because ambient light contributions in the camera measurements are slowly varying across the camera field of view, they can be approximately decomposed into a sum of light contributions from sources near the measurement surface, $\mathbf{a}_{\text{NF}}$, and those in the far-field, $\mathbf{a}_{\text{FF}}$. The far-field contribution is roughly constant over the camera FOV, $\mathbf{a} \approx c_1 \mathbf{1} + c_2 \mathbf{a}_{\text{NF}}$, where c_1 and c_2 are constants that lead to dimensionless pixel values. The term $\mathbf{a}_{\text{NF}}$ can be measured, or computed from our knowledge of the position of the visible side, so that the only unknown needed to describe $\mathbf{a}$ is $\mathbf{c} = [c_1, c_2]$. In the presence of ambient light, the inverse problem becomes estimating $(\mathbf{s}_h, \rho_h, \mathbf{c})$ from measurements $\mathbf{y}$, under the model

$$\mathbf{y} = \mathbf{A}\mathbf{c} + (\mathbf{V} \odot \mathbf{D}(\rho_h))\mathbf{s}_h + \epsilon, \quad (16)$$

where $\mathbf{A} = [\mathbf{1}, \mathbf{a}_{\text{NF}}]$.

B. A Linear Model and Inverse Algorithm

Equation (16) is linear in $\mathbf{s}_h$ and nonlinear in ρ_h . However, by discretizing the possible values of each element of $\rho_h(\alpha_n)$, we can formulate a new system that is linear in all unknown parameters. Specifically, let $\{\rho_1, \rho_2, \dots, \rho_L\}$ be the set of allowed ranges. Then the Cartesian product $\{\rho_1, \rho_2, \dots, \rho_L\} \times \{\alpha_1, \alpha_2, \dots, \alpha_N\}$ gives a 2D polar partitioning of the hidden region, with each element (ρ_ℓ, α_n) defining a hidden-scene *polar pixel*. Shown in Fig. 8 is a (coarse) 5×6 polar grid

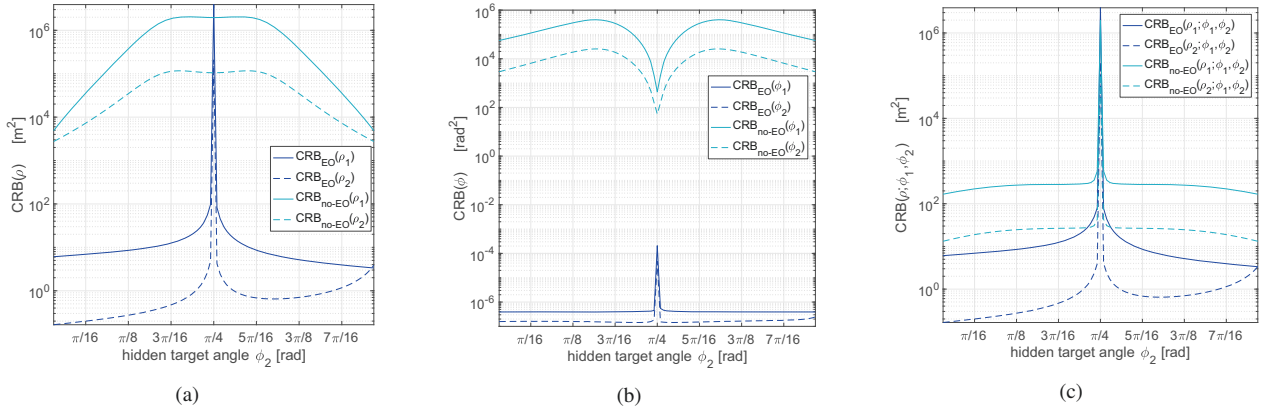


Fig. 7: Variation of the CRBs for estimating two hidden targets for different target positions. The number of measurement pixels, $M = 155^2$ and the measurement FOV is $0.2\text{m} \times 0.2\text{m}$, with fixed noise variance $\sigma^2 = 10$. Target 1 is fixed at $(\rho_1, \phi_1) = (1\text{m}, \pi/4)$; target 2 is at range $\rho_2 = 2\text{m}$ and moved in angle ϕ_2 .

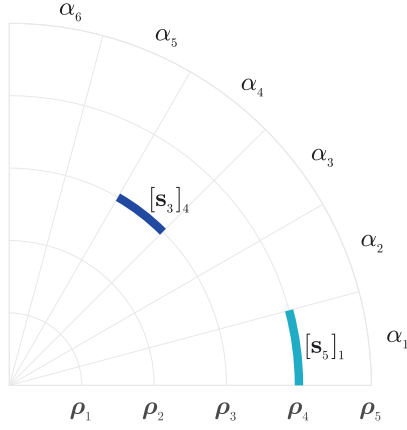


Fig. 8: Polar partitioning of the hidden space to obtain polar pixels at six discrete angles and five discrete range values. The n -th column of each submatrix $\bar{\mathbf{D}}_l$, of $\bar{\mathbf{D}}$, describes propagation of light from hidden scene polar pixel (ρ_l, α_n) , to the measurement plane.

discretization of the hidden space. Under this partitioning, the forward model (9) becomes

$$\mathbf{y} = \mathbf{A}\mathbf{c} + \bar{\mathbf{D}}\bar{\mathbf{s}}_{\mathbf{h}} + \boldsymbol{\epsilon}, \quad (17)$$

where $\bar{\mathbf{D}} = [\mathbf{V} \odot \mathbf{D}(\rho_1 \mathbf{1}), \mathbf{V} \odot \mathbf{D}(\rho_2 \mathbf{1}), \dots, \mathbf{V} \odot \mathbf{D}(\rho_L \mathbf{1})]$, and $\bar{\mathbf{s}}_{\mathbf{h}} = \text{vec}([\mathbf{s}_{\mathbf{h}1}, \mathbf{s}_{\mathbf{h}2}, \dots, \mathbf{s}_{\mathbf{h}L}]) \in \mathbb{R}_+^{NL}$ with $[\mathbf{s}_{\mathbf{h}\ell}]_n$ representing the radiosity of pixel (n, ℓ) at range ρ_ℓ and angular bin n .

Although (17) is linear in all unknown parameters and (9) is not, there is an important difference. Built into (9) is the constraint that only a single hidden object per angle contributes to the measurement. This constraint is based on the assumption that the scene is composed of opaque vertical facets, so light from objects that are behind other objects is blocked from reaching the corner. In contrast, this constraint is *not* built into (17). In this case, to model the fact that the vast majority of pixels in the hidden scene either do not contain a target or are occluded from the camera FOV by another visible to the camera FOV, we promote sparsity in our estimate of $\bar{\mathbf{s}}_{\mathbf{h}}$,

resulting in the ℓ_1 -regularized problem

$$[\hat{\mathbf{s}}_{\mathbf{h}}, \hat{\mathbf{c}}] = \arg \min_{\bar{\mathbf{s}}_{\mathbf{h}}, \mathbf{c}} \left[\frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{c} - \bar{\mathbf{D}}\bar{\mathbf{s}}_{\mathbf{h}}\|_2^2 + \lambda \|\bar{\mathbf{s}}_{\mathbf{h}}\|_1 \right], \quad (18)$$

where λ is the regularization parameter. The optimization problem (18) is efficiently solved using the FISTA algorithm [32].

We evaluate the linear model approach for the hidden scene and measurement in Fig. 9a.³ Reconstructions at range resolutions of $L = 10$ and $L = 40$ are shown in Fig. 9b and Fig. 9c for angular resolution $N = 90$. Both reconstructions exhibit two clusters of pixels with intensities larger than zero, corresponding to the two hidden objects in the scene. While the relative order of the objects is correct, the yellow-blue stripe is estimated to be closer than its true location in both reconstructions. Both targets are reconstructed with mostly correct color content, though several angular bins have different range estimates across the three different color channels causing some misalignment in the reconstructions. Although both targets are at an approximately constant range across their angular extent, this is not the case in both reconstructions, particularly in the more coarse reconstruction of Fig. 9b. Our nonlinear, more physically-inspired, model addresses some of these challenges.

Remark 1: Under the assumption of at most one target in any angular bin, each sub-vector $\mathbf{s}_{\mathbf{h}\ell}$ in $\bar{\mathbf{s}}_{\mathbf{h}}$ is either 1-sparse or zero. Combining this with the existence of only a few targets means that there is a small number of 1-sparse groups in $\bar{\mathbf{s}}_{\mathbf{h}}$, i.e., sparsity both within and across groups. This could be incorporated by solving a Sparse-Group Lasso problem [33]:

$$\arg \min_{\bar{\mathbf{s}}_{\mathbf{h}}, \mathbf{c}} \left[\frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{c} - \bar{\mathbf{D}}\bar{\mathbf{s}}_{\mathbf{h}}\|_2^2 + \lambda_1 \sum_{\ell=1}^L \|\mathbf{s}_{\mathbf{h}\ell}\|_2 + \lambda_2 \|\bar{\mathbf{s}}_{\mathbf{h}}\|_1 \right]. \quad (19)$$

Empirically, we found no compelling evidence that solving (19) is superior to solving (18). Consequently, all results for the linear inverse problem (17) are based on solving (18), separately, for each color channel.

³The corresponding experimental setup will be described in detail in Section V, after we present our second, preferred, algorithm.

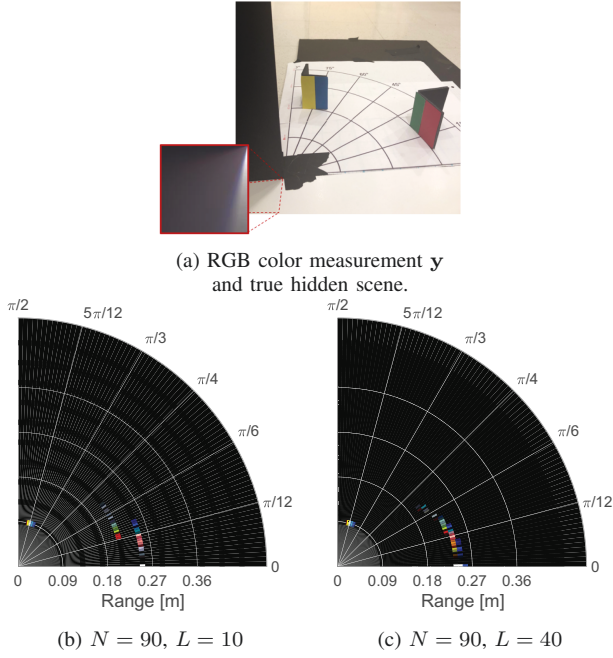


Fig. 9: Demonstration of linear inversion algorithm for hidden scene and measurement shown in (a). The hidden region is discretized into N angles and L ranges.

C. Nonlinear Modeling and Inversion

In many practical scenarios, the hidden scene is composed of only a few hidden targets of interest, with each target having some angular extent and being roughly at a constant distance from the corner. Solving (18) with fine range and angular discretization is computationally expensive; similarly, finely discretizing the angular dimension and estimating a unique range value $\rho(\alpha_n)$ for each hidden-scene angle α_n , using (16), is unnecessarily ambitious. Alternatively, we can assume that there is an unknown number $N_t \ll N$ of disjoint targets to be estimated, each with unknown range and radiosity. Mathematically,

$$\bar{S}_h(\rho, \alpha) = \sum_{j=1}^{N_t} s(\alpha) \delta(\rho - \bar{\rho}_j) u\left(\frac{\alpha - \bar{\alpha}_j}{\Delta_j}\right), \quad (20)$$

with the j th hidden target having angular position $\bar{\alpha}_j$, angular extent Δ_j , and range $\bar{\rho}_j$; $u(\cdot)$ is the zero-centered unit rectangular function. The radiosity $s(\alpha)$ is piecewise constant, taking value $[s]_n$ for $\alpha \in [(n-1)\pi/(2N), n\pi/(2N))$, such that $\mathbf{s} \in \mathbb{R}^N$ is the vector representing the discretization of $s(\alpha)$.

Angular bins containing no detected targets are attributed to background, and the minimal light coming from those regions will be assumed to be coming from a large distance $\rho_{\max}$. Under the model (20), instead of having N different range values, contiguous angular bins will have the same range $\bar{\rho}_j$ if they contain the same target. Letting $\bar{\rho} = [\bar{\rho}_1, \bar{\rho}_2, \dots, \bar{\rho}_{N_t}]^T$:

$$\mathbf{y} = \mathbf{A}\mathbf{c} + (\mathbf{V} \odot \mathbf{D}(\bar{\rho}))\mathbf{s} + \boldsymbol{\epsilon}, \quad (21)$$

where for any $m = 1, 2, \dots, M$ and $n = 1, 2, \dots, N$,

$$[\mathbf{D}(\bar{\rho})]_{m,n} = \begin{cases} \frac{\bar{\rho}_j}{d^2(r_m, \theta_m, \bar{\rho}_j, \alpha_n)}, & \text{when target } j \text{ is present at angle } \alpha_n; \\ \frac{\rho_{\max}}{d^2(r_m, \theta_m, \rho_{\max}, \alpha_n)}, & \text{when no target is present at angle } \alpha_n. \end{cases} \quad (22)$$

Note that target j is present at angle α_n when $\alpha_n \in [\bar{\alpha}_j - \Delta_j/2, \bar{\alpha}_j + \Delta_j/2]$. We propose to estimate $\mathbf{s}$, $\bar{\rho}$ and $\mathbf{c}$ by solving

$$\min_{\mathbf{s}, \bar{\rho}, \mathbf{c}} \left(\underbrace{\frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{c} - (\mathbf{V} \odot \mathbf{D}(\bar{\rho}))\mathbf{s}\|_2^2}_{\text{data fidelity}} + \underbrace{\lambda_1 \|\mathbf{W}\mathbf{s}\|_1 + \lambda_2 \|\mathbf{B}\mathbf{s}\|_2^2 + \iota_{[0, \infty)^N}(\mathbf{s})}_{\text{regularizers for } \mathbf{s}} + \underbrace{\iota_{[c, \infty)^{N_t}}(\bar{\rho})}_{\text{regularizer for } \bar{\rho}} \right), \quad (23)$$

where $\mathbf{W}$ is a wavelet transform matrix (we use the Daubechies wavelet of order 4), $\mathbf{B}$ returns the difference between subsequent entries in $\mathbf{s}$ that are attributed to hidden-scene background terms, λ_1 and λ_2 are tuning parameters, and

$$\iota_{\mathcal{C}}(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \in \mathcal{C}; \\ \infty, & \text{otherwise} \end{cases}$$

is the indicator function for a set $\mathcal{C}$. In (23), the regularizers for $\mathbf{s}$ promote sparsity in the wavelet basis, smoothness in hidden-scene background contributions, and positivity in $\mathbf{s}$, respectively. The regularizer for range $\bar{\rho}$ enforces range estimates to be at least $c > 0$ (a small constant). This optimization problem is solved using an alternating approach described below.

1) Initialize $\mathbf{s}$ and $\mathbf{c}$ by solving

$$[\mathbf{s}^0, \mathbf{c}^0] = \arg \min_{\mathbf{s}, \mathbf{c}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{c} - (\mathbf{V} \odot \mathbf{D}(\bar{\rho}_0))\mathbf{s}\|_2^2 + \lambda \|\mathbf{W}\mathbf{s}\|_1 + \iota_{[0, \infty)^N}(\mathbf{s}), \quad (24)$$

with $\bar{\rho}^0 = \rho_{\text{FF}} \mathbf{1}$, initialized to represent a single target ($N_t = 1$) in the far field ($\rho_{\text{FF}} \gg 0$). Our motivation to first estimate $\mathbf{s}$ is because, given $\bar{\rho}^0$, the resulting problem is well-conditioned (Section III).

2) Determine number of targets N_t by comparing $[\mathbf{s}^0]_n$ to the threshold $\kappa_n = \alpha/(2\ell + 1) \sum_{i=n-\ell}^{n+\ell} [\mathbf{s}^0]_i$, where $\alpha \in \mathbb{R}_+$ and (odd) filter length $(2\ell + 1) \in \mathbb{Z}_+$ are tuneable parameters. It is assumed that $[\mathbf{s}^{t-1}]_n = 0$ for $n \notin \{1, 2, \dots, N\}$. Consecutive threshold crossings in $\mathbf{s}^0$ represent the edges of a single target.

3) Update $\bar{\rho}^t$ by

$$[\bar{\rho}^t, \mathbf{z}^t] = \arg \min_{\bar{\rho}, \mathbf{z}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{c}^{t-1} - (\mathbf{V} \odot \mathbf{D}(\bar{\rho}))\mathbf{s}^{t-1}\|_2^2 + \iota_{[c, \infty)^{N_t}}(\bar{\rho}) + \iota_{[0, \infty)^{N_t}}(\mathbf{z}), \quad (25)$$

where $\mathbf{z} = [z_1, z_2, \dots, z_{N_t}]^T \in \mathbb{R}_+^{N_t}$ is such that for any $n = 1, 2, \dots, N$,

$$[\mathbf{D}(\bar{\rho})]_{m,n} = \frac{z_j \bar{\rho}_j}{d^2(r_m, \theta_m, \bar{\rho}_j, \alpha_n)}$$

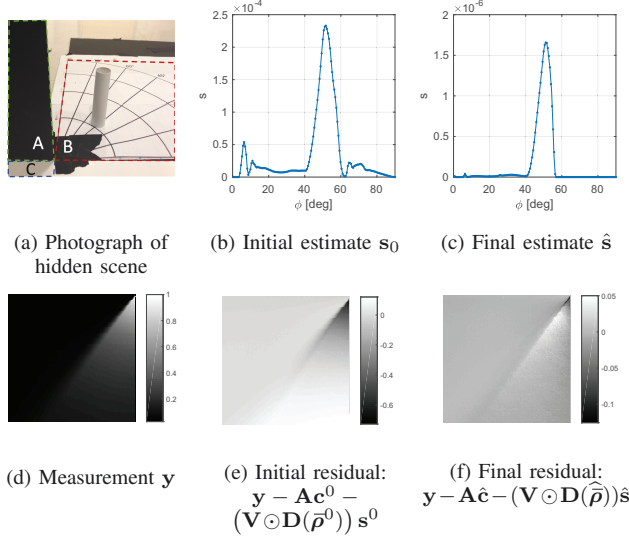


Fig. 10: Demonstration of model mismatch with scene initialization. The hidden scene was a narrow white cylinder with a diameter of 2.5cm, 18cm away from the corner, at $\phi = 45^\circ$, as shown in (a). When all light is assumed to originate in the far field, the initial estimate of $\mathbf{s}_0$ (b) does not describe the radial falloff that is present in the measurement $\mathbf{y}$ (d), as shown in the initial residual (e). After the algorithm converges, the range estimate is updated allowing for a more accurate estimate of $\mathbf{s}$ as shown in (c) with a much smaller residual (f).

when $\alpha_n \in [\bar{\alpha}_j - \Delta_j/2, \bar{\alpha}_j + \Delta_j/2]$. The introduction of $\mathbf{z}$ couples the minimization problems (25) and (26), permitting radiosities $\mathbf{s}^{t-1}$ to be scaled, appropriately, as $\bar{\rho}$ is updated.

4) **Update $\mathbf{s}$ and $\mathbf{c}$** by solving

$$[\mathbf{s}^t, \mathbf{c}^t] = \arg \min_{\mathbf{s}, \mathbf{c}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{c} - (\mathbf{V} \odot \mathbf{D}(\bar{\rho}^t))\mathbf{s}\|_2^2 + \lambda_1 \|\mathbf{W}\mathbf{s}\|_1 + \lambda_2 \|\mathbf{B}\mathbf{s}\|_2^2 + \iota_{[0, \infty)^N}(\mathbf{s}). \quad (26)$$

5) **Increment iteration** counter t by one.

6) **Repeat** steps 3, 4 and 5 until convergence.

7) **Return** $\hat{\rho} \leftarrow \bar{\rho}^t$, and $\hat{\mathbf{s}} \leftarrow \mathbf{s}^t$.

Steps 1, 3, and 4 are solved using projected gradient methods [32]. Fig. 10 illustrates several algorithm steps for a scene containing a single hidden cylinder, shown in Fig. 10a, resulting in measurement $\mathbf{y}$, shown in Fig. 10d. The final estimate of $\mathbf{s}$ (Fig. 10c) does not contain the artifacts seen in the initial estimate $\mathbf{s}_0$ (Fig. 10b) because, instead of assuming the hidden scene is in the far field, the model has been updated to include the effects of radial falloff due to a target at estimated distance $\hat{\rho}$. In fact, the residual due to the initial far field assumption (Fig. 10e) clearly contains unmodeled radial falloff, whereas the final residual (Fig. 10f) exhibits a much better overall fit.

D. Nonlinear RGB Model Inversion

The algorithm described in Section IV-C, which operates on a single measurement channel, may be adapted to operate on color (RGB) data. In this case, the camera measures $\mathbf{y}_R$,

$\mathbf{y}_G$, and $\mathbf{y}_B$ corresponding to each color channel. Although our goal is still to estimate range values $\bar{\rho} \in \mathbb{R}_+^{N_t}$, we now seek radiosity estimates $\hat{\mathbf{s}}_R$, $\hat{\mathbf{s}}_G$, and $\hat{\mathbf{s}}_B$, as well as estimates of ambient light $\hat{\mathbf{c}}_R$, $\hat{\mathbf{c}}_G$, and $\hat{\mathbf{c}}_B$. These estimates are obtained by solving (23) with substitutions

$$\begin{aligned} \mathbf{y} &\rightarrow \tilde{\mathbf{y}} = \text{vec}([\mathbf{y}_R, \mathbf{y}_G, \mathbf{y}_B]) \in \mathbb{R}^{3M}, \\ \mathbf{A} &\rightarrow \tilde{\mathbf{A}} = \text{diag}([\mathbf{1}, \mathbf{a}_R], [\mathbf{1}, \mathbf{a}_G], [\mathbf{1}, \mathbf{a}_B]) \in \mathbb{R}^{3M \times 6}, \\ \mathbf{V} \odot (\mathbf{D}(\bar{\rho})) &\rightarrow \tilde{\mathbf{D}}(\bar{\rho}) = (\mathbf{V} \odot \mathbf{D}(\bar{\rho})) \otimes \mathbf{I} \in \mathbb{R}^{3M \times 3N}, \\ \mathbf{B} &\rightarrow \tilde{\mathbf{B}} = \mathbf{B} \otimes \mathbf{I} \in \mathbb{R}^{3(N-1) \times 3N}, \end{aligned}$$

where $\mathbf{I}$ is the 3×3 identity matrix. The optimization becomes

$$\begin{aligned} \min_{\bar{\rho}, \tilde{\mathbf{s}}, \tilde{\mathbf{c}}} & \left(\frac{1}{2} \|\tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{c}} - \tilde{\mathbf{D}}(\bar{\rho})\tilde{\mathbf{s}}\|_2^2 + \lambda_2 \|\tilde{\mathbf{B}}\tilde{\mathbf{s}}\|_2^2 \right. \\ & \left. + \lambda \|\mathbf{W}\tilde{\mathbf{s}}\|_1 + \iota_{[0, \infty)^{3N}}(\tilde{\mathbf{s}}) + \iota_{[c, \infty)^{N_t}}(\bar{\rho}) \right), \end{aligned} \quad (27)$$

which, like before, is solved using an alternating approach, performing initial thresholding, or target counting, on $\tilde{\mathbf{s}} = \frac{1}{3}(\mathbf{s}_R^0 + \mathbf{s}_G^0 + \mathbf{s}_B^0)$, with (25) modified to update $\tilde{\mathbf{z}} = [\mathbf{z}_R; \mathbf{z}_G; \mathbf{z}_B] \in \mathbb{R}^{3N_t}$ instead of $\mathbf{z}$:

$$\begin{aligned} [\bar{\rho}^t, \tilde{\mathbf{z}}^t] &= \arg \min_{\bar{\rho}, \tilde{\mathbf{z}}} \frac{1}{2} \|\tilde{\mathbf{y}} - \mathbf{A}\tilde{\mathbf{c}}^{t-1} - (\tilde{\mathbf{D}}(\bar{\rho}))\tilde{\mathbf{s}}^{t-1}\|_2^2 \\ &+ \iota_{[c, \infty)^{N_t}}(\bar{\rho}) + \iota_{[0, \infty)^{3N_t}}(\tilde{\mathbf{z}}). \end{aligned} \quad (28)$$

The scene $\tilde{\mathbf{s}}$ is updated by solving (26) in parallel for each color channel. This concatenation of the color channel measurements enforces consensus among channels in the range estimate and angular extent of a given hidden target, thus avoiding the spurious range estimates observed in the ℓ_1 -regularized solutions of the linear inverse problem formulation (Fig. 9).

V. EXPERIMENTAL EVALUATION USING REAL DATA

Performances of the algorithms presented in Sections IV-C and IV-D were evaluated in a variety of conditions, using the scaled-down laboratory setup shown in annotated photograph Fig. 10a. A tripod-mounted *FLIR Grasshopper3* camera model GS3-U3-41S4C-C equipped with a Tamron M118FM16 lens was used to photograph the floor (C) on the visible side of occluding wall (A). The camera was positioned approximately 1 meter off the ground, directly above the observation region. A tuneable light source consisting of an array of LEDs, positioned behind the occluding wall, was used to illuminate the hidden scene region (B). The camera's exposure time was adjusted for each experimental run to avoid saturation while still using the full dynamic range. Camera automatic gain correction and gamma correction were turned off. The floor in region (B) was covered with a large paper polar grid for reference. Black tape secures the corner of this paper to keep it flush with the floor. In this work, we reconstruct a region that extends $\pi/2$ radians into the hidden scene. In principle, the full π radians of hidden scene may be reconstructed by extending the photograph region (C) to the right.

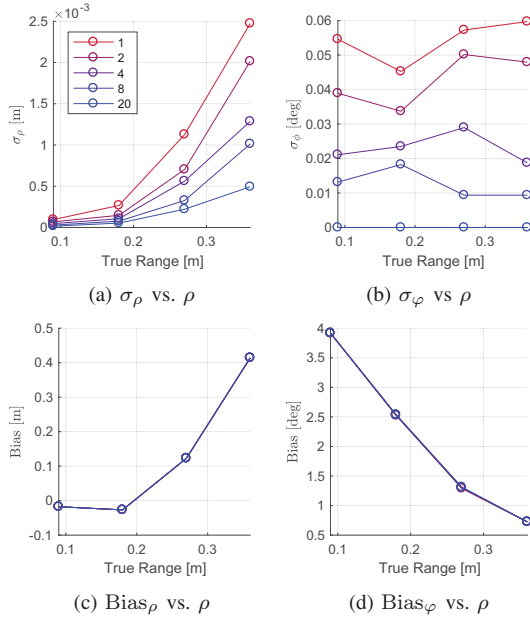


Fig. 11: Evaluation of algorithm performance for a single target at four different ranges, in five different noise conditions, placed at $\phi = 45^\circ$. The standard deviation of the range estimate (a), σ_ρ , increases with increasing range ρ , and the standard deviation of the angular estimate (b), σ_ϕ is small at all ranges. Bias for the range estimate (c) increases with increasing range, while the bias for the angular estimate remains small at all ranges.

A. Empirical Performance Evaluation: Single Target

To evaluate performance, a single white cylindrical target, shown in Fig. 10a, was placed at different positions (ρ_1, ϕ_1) in range and angle. For each position, with a camera FOV of $0.16 \text{ m} \times 0.16 \text{ m}$, 150 snapshots of the visible floor were taken. By combining 1, 2, 4, 8, and 20 randomly selected snapshots (without replacement), we emulated decreasing measurement noise levels. Estimates' bias and variance were computed using the recovery results from 60 repetitions of each configuration. In each trial, scalar range parameter estimates $\hat{\varphi}_1$ and angular profiles $\hat{s}$ are recovered for the target, as shown in Fig. 10c. We use the peak value of $\hat{s}$ (after up-sampling) as a proxy for $\hat{\varphi}_1$, to compute its bias and variance. We take the measured center, in angle, of the target as the true φ_1 . While it is expected that they are close, this measured center of the cylinder may not exactly match the brightest illuminated region of the cylinder.

1) *Varying Range*: Fig. 11 shows estimate bias and standard deviation computed for $\varphi_1 = 45^\circ$ and ranges $\rho_1 = 0.09 \text{ mm}$, 0.18 mm , 0.27 mm , and 0.36 mm . As shown in Fig. 11a, range estimate standard deviation increases in noisier conditions (i.e., fewer combined frames) and at greater ranges. Fig. 11b shows that, as predicted by the CRB analyses in Section III, the standard deviation of estimate $\hat{\varphi}_1$ remains small at every position in range.

Fig. 11c and Fig. 11d show the bias for range and angle estimates respectively, at the four ranges in the same five noise conditions. For both range and angle estimates, bias is constant at a given range, regardless of the noise level. In both cases, the bias is orders of magnitude larger than the corresponding standard deviation. For the range estimate, we attribute this

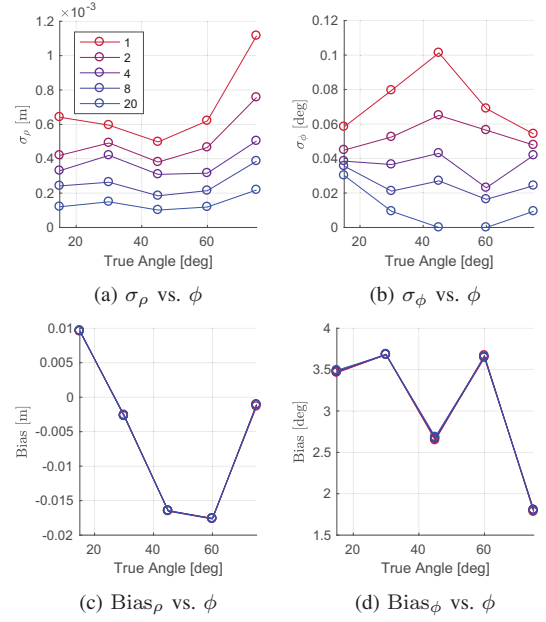


Fig. 12: Evaluation of algorithm performance for a single target at five different angles, in five different noise conditions, placed at $\rho = 0.18 \text{ m}$. The standard deviation of the range estimate (a), σ_ρ , is greatest when $\phi = 75^\circ$, when the fewest pixels on the floor are exposed to penumbra. The standard deviation of the angular estimate (b), σ_ϕ is small at all angles.

bias to model mismatch due to unmodelled reflections, nonzero target height, and edge imperfections. As shown in Fig. 11d, angular bias is much smaller, and may correctly reflect the fact that the brightest part of the cylinder changes in angle, as the cylinder moves with respect to the fixed hidden scene illumination.

2) *Varying Angle*: Fig. 12 shows estimate standard deviation and bias for fixed range $\rho_1 = 0.18 \text{ m}$ a set of angles $\phi_1 = 15^\circ, 30^\circ, 45^\circ, 60^\circ$, and 75° . We similarly observe lower estimate standard deviation in less noisy conditions, greater standard deviation in range than angle, and substantially higher bias than standard deviation for both range and angle.

3) *Varying Ambient Light*: Measurements were also taken of this same target in a fixed position, with different levels of ambient light. A constant light source on the visible side introduced ambient light while a light source on the hidden side was tuned to vary penumbra brightness. Fig. 13a, Fig. 13b, Fig. 13c, and Fig. 13d show measurements as the penumbra becomes faint to the point of not being visible to the naked eye; Fig. 13e, Fig. 13f, Fig. 13g, and Fig. 13h show the corresponding reconstructions with the true target location marked by a red dot. All four reconstructions correctly pick out the target in angle demonstrating robustness to a surprising amount of ambient light, although the higher SNR case is both sharper in angle and more accurate in range estimation.

B. Color Reconstructions

The RGB nonlinear inversion algorithm was tested on scenes with colored objects in several different configurations. Testing was also performed on a multi-object, colored scene

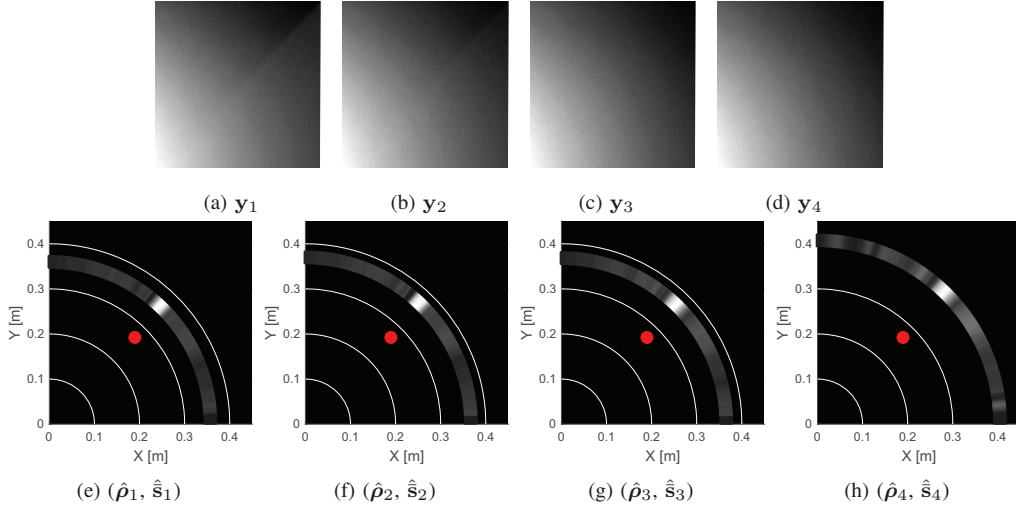


Fig. 13: Demonstration of performance degradation as the penumbra becomes fainter. The true location of the hidden object, a white cylinder, is shown in red. When the penumbra is brighter, (a) and (e), the estimated range $\hat{\rho}$ is closer to the truth, and $\hat{s}$ contains a sharp peak at the true angular location of the hidden object. When the penumbra is more faint, (d) and (h), the estimated range $\hat{\rho}$ is further from the truth and the peak in $\hat{s}$ less sharp.

in the presence of increasingly bright ambient light to demonstrate algorithm robustness to low SNR conditions.

1) *Multiple Targets*: Measurement and hidden scene pairs are shown in Fig. 14a and Fig. 14c, where the same two colored objects have been placed in reverse positions. Reconstructions Fig. 14b and Fig. 14d show that in both cases, the two targets are accurately found in angle, and placed in range correctly with respect to each other. High angular resolution is demonstrated in both reconstructions, with the red-green object correctly portrayed to have a slightly wider red section, just like the yellow-blue object has a slightly wider yellow section. The scenario and measurement shown in Fig. 14e tests our algorithm on a scene that includes three targets instead of two. The reconstruction shown in Fig. 14f accurately picks out all three targets in angle and places them at ranges that are correct with respect to each other.

2) *Varying Ambient Light*: In Fig. 15, we demonstrate algorithm robustness to increasing amounts of ambient light. Here, the hidden scene, arranged as shown in Fig. 15e, consists of the yellow-blue target and a more distant, in both range and angle, white cylinder. With an angular location close to $\pi/2$, very few pixels in the measurement are exposed to light from the white cylinder making range estimation more challenging. Still, all but the lowest SNR reconstructions correctly place the white cylinder at a greater range than the yellow-blue target. All four reconstructions demonstrate high angular resolution, even resolving the sharp boundary between the yellow and blue portions of the yellow-blue target.

VI. DISCUSSION

A. Technical Conclusions

We proposed and tested two inversion algorithms: one based on a conventional linear model and the other on an alternating approach that more directly inverts the nonlinear forward model (9). Both make use of regularization to solve an ill-conditioned problem and demonstrate high angular resolution

and significantly coarser range resolution in reconstruction results, owing to the conditioning of range estimation. While the linear model (17) enjoys simplicity, it omits the opacity assumption that is naturally embedded in the nonlinear model (9), thus allowing multiple nonzero pixels in a single angle. In addition, without enforcing a single range per target, this method also raises questions about how to effectively promote coincident pixels across the color channels. The second (non-linear) method benefits from the natural separation of range and angle estimation problems, enabling a highly effective alternating recovery algorithm.

The linear approach discretizes the hidden scene into a polar grid. Even with fine angular discretization and sparsity-enforcing priors, estimating a range per angle when a hidden scene likely contains only a few targets is unnecessarily challenging. In contrast, the initialization step of the alternating algorithm that we propose for solving the nonlinear problem (9) allows us to exploit excellent angular resolution to count the number of targets and estimate only one range per detected hidden target. In this light, the range update step can be interpreted as *learning* the forward model to ultimately allow for better angular reconstructions, as demonstrated in Fig. 10. With the few unknown ranges as *parameters*, the alternating approach enjoys less model mismatch than the linear inversion algorithm, because the range parameters are not discretized.

The fact that the alternating algorithm treats unknown ranges as parameters also lends itself to a natural three-channel RGB extension. Estimating a single range per fixed angular extent enforces consensus across color channels. In contrast, the linear inversion algorithm operates separately on three color channels and may place RGB values for the same object at different ranges or angles, as shown in Fig. 9. One way this may ultimately be improved is by forming a reconstruction in the YUV color space. Enforcing sparsity on component Y (i.e., the ‘luma’, or ‘intensity’ component) would penalize intensities at multiple ranges in the same angular bin.

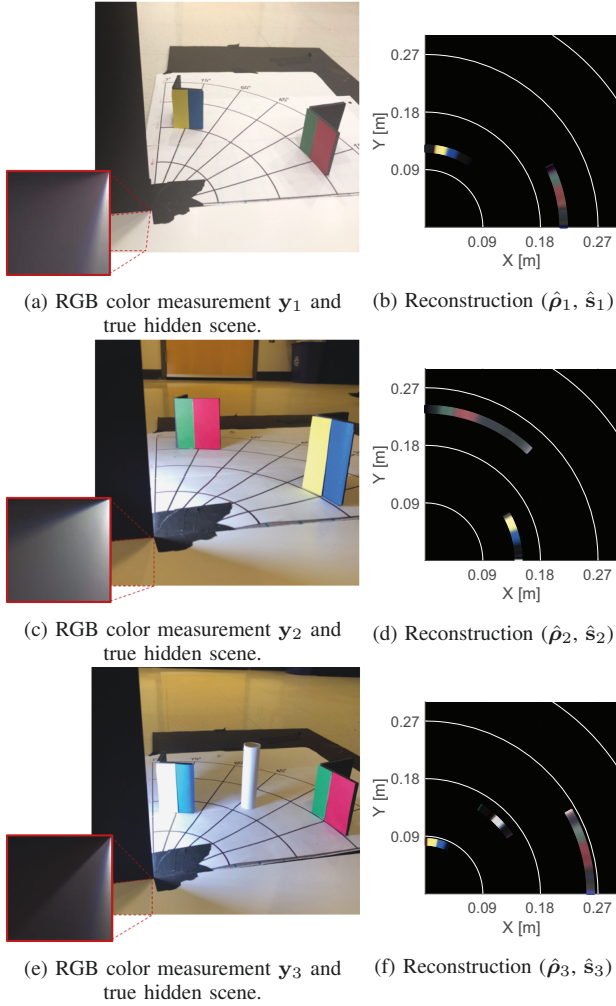


Fig. 14: Demonstration of color reconstruction with three different scenes containing multiple hidden objects. The widths of yellow, blue, white, green, and red objects are 2.9 cm, 2.7 cm, 2.5 cm, 2.9 cm, and 4.2 cm respectively. The black arcs on the floor in (a), (c), and (e) correspond to the ranges marked in white in (b), (d), and (f).

Both algorithms seek to reconstruct a 2D plan view of the 3D hidden scene. Angular information is provided by the edge occluder, while target range information is contained in the subtle radial falloff curvature measured by the unoccluded camera pixels. Shining a spotlight on a hidden target only changes the intensity scaling of this curve. However, moving a target in range changes the shape of this curvature, allowing us to distinguish between targets at different positions. Although both the height and range of a single hidden point source may be recoverable in a noiseless setting, in practice, we contend with measurement noise, ambient light on the visible side, and more complicated hidden scenes. As demonstrated by the CRB analysis in Section III, even when the target is assumed to be in the ground plane, range information in our measured data is subtle. In fact, when a point source's vertical position z_s is treated as another unknown parameter, the CRB for estimating both source range ρ_s and vertical position z_s is even larger, as shown for a point source at $z_s = 0.2$ m in Fig. 16a and Fig. 16b. When it is assumed that the source's

vertical coordinate z_s or range ρ_s is known, as is the case in Fig. 16c and Fig. 16d, uncertainty is reduced. Furthermore, comparing columns in Fig. 16 illustrates that, at least when a point source is not very high off the ground, range is the easier of the two parameters to estimate. These observations support our choice to reconstruct range (along the floor) rather than height under the assumption that targets rest on the *ground* with heights that are small compared to their range.

When a target has significant vertical extent, our assumption is violated and our algorithms may be expected to estimate targets to be further than they truly are. We demonstrate this using synthetic data in Fig. 17. Here the simulated target spans 35° - 55° within the hidden scene and is fixed at range $\rho = 0.15$ m. Target height extent η is varied from 0 m to 0.3 m and target range ρ is estimated using the alternating algorithm for each target height. Fortunately, even in an extreme case where the target's height extent η is approximately equal to its range ($\rho \approx \eta \approx 0.15$ m), the relative error in the estimated range is still less than 20%. In fact, in many of our experimental results, targets had some vertical extent but were placed closer to the corner in reconstruction results, likely due to this effect being overwhelmed by other sources of model mismatch.

B. Future Outlook

Experimental results presented in Section V were obtained using an experimental setup with target distances on the order of half a meter, but we believe both inversion algorithms could work with a larger experimental setup, given comparable SNR and larger camera FOV. CRB analysis may be extended to determine the effect of camera FOV on estimate variance for targets at a given range. We conjecture that for a given target range, there may be an optimal camera FOV for range recovery, although generally speaking a larger camera FOV makes angular estimation more challenging. In the alternating algorithm, this trade-off may be managed by taking one photo with a smaller FOV to use in scene's initialization and update steps, (24) and (26), and another photo with larger FOV for range estimation step, (25). Though at the scale of the experiments in this paper, this was not necessary.

The alternating algorithm may be further adapted to handle the common scenario of a few hidden objects with heights known *a priori*. Imagine a scene composed of people or cars that we observed entering the hidden scene from the visible side. If the heights of hidden objects are known and we assume a constant radiosity across height for a fixed angle, we may model the radiosity of the hidden scene as

$$S_h(\rho, \alpha, z) = \begin{cases} S_h^c(\rho, \alpha), & \text{for } z \in [0, \eta(\rho, \alpha)]; \\ 0, & \text{otherwise,} \end{cases}$$

where at hidden scene floor location (ρ, α) , an object of height $\eta(\rho, \alpha)$ has constant radiosity $S_h^c(\rho, \alpha)$ over its full height. Now, instead of recovering $\hat{S}_h(\rho, \alpha)$ of the hidden scene, as defined in (5), we seek to recover $S_h^c(\rho, \alpha)$. With this explicit modeling of the z dependence of $S_h(\rho, \alpha, z)$, our expression

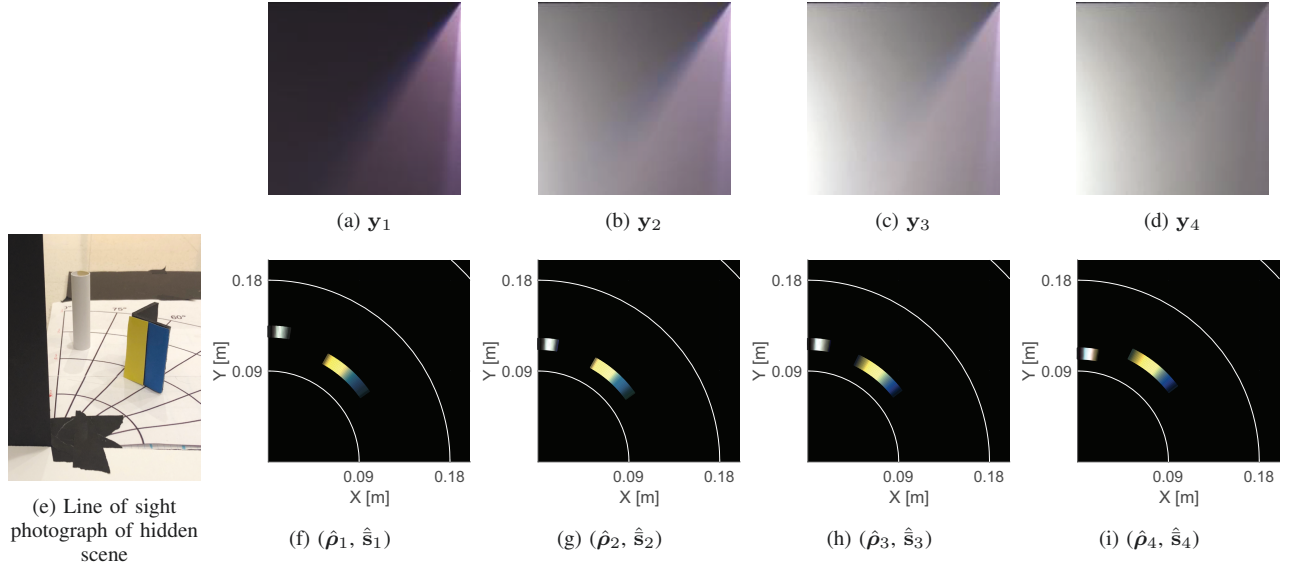


Fig. 15: Demonstration of RGB alternating inversion algorithm as SNR is reduced. The white cylinder (width 0.025 m) and yellow-blue stripes (widths 0.029 m and 0.027 m, respectively) were arranged as shown in (e), at 0.21 m and 0.16 m respectively. Note the increased difficulty due to the white cylinder's deep placement, in angle, into the scene. In this location, only a small fraction of the measurement pixels are exposed to light from the cylinder. Still, only the lowest SNR reconstruction fails to place the white cylinder at a greater range than the yellow-blue object. In all cases, the two objects are resolved in great angular detail. Black arcs on the floor correspond to the ranges marked in white on the reconstructions.

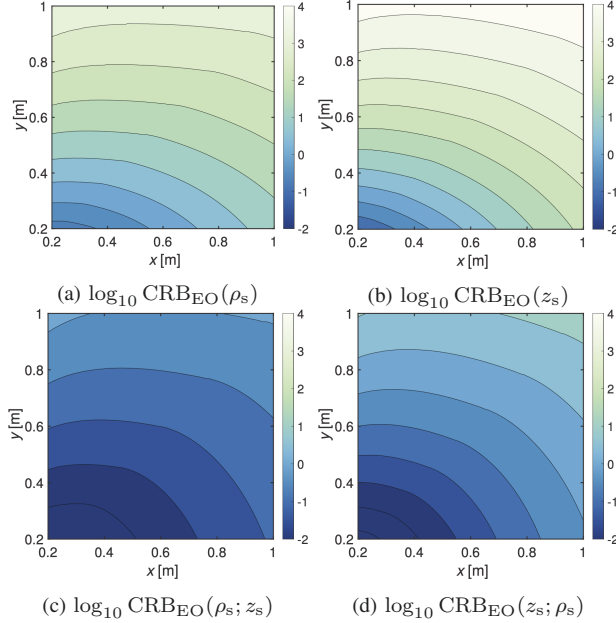


Fig. 16: Variation in CRB_{EO} for parameters ρ and z when both are unknown (first row) and when the other parameter is assumed known (second row). The true position of the hidden point source is varied over $[0.2, 1]^2$ in the (x, y) -plane, with a fixed vertical position $z_s = 0.2$ m. The number of measurement pixels is $M = 107^2$ and the measurement FOV is $0.2 \text{ m} \times 0.2 \text{ m}$ and noise variance $\sigma^2 = 10$.

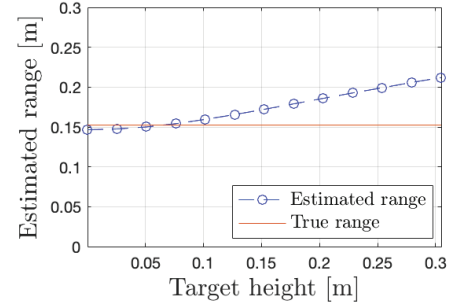


Fig. 17: Demonstration of range estimation error as a function of target height extent η . Targets with greater height extent are estimated to be further away. The simulated target spans 35° – 55° within the hidden scene and is fixed at range $\rho = 0.15$ m. Target height extent η is varied from 0 m to 0.3 m, and target range ρ is estimated using the alternating algorithm. Simulated measurements are noiseless with $M = 155^2$ measurement pixels and the measurement FOV is $0.15 \text{ m} \times 0.15 \text{ m}$.

scene), (3a), may be rewritten:

$$\begin{aligned} L_h(r, \theta) &= \int_0^\theta \int_0^\infty \int_0^{\eta(\rho, \alpha)} \frac{S_h^c(\rho, \alpha)}{d^2(r, \theta, \rho, \alpha) + z^2} \rho \, dz \, d\rho \, d\alpha \\ &= \int_0^\theta \int_0^\infty \frac{S_h^c(\rho, \alpha)}{d(r, \theta, \rho, \alpha)} \arctan\left(\frac{\eta(\rho, \alpha)}{d(r, \theta, \rho, \alpha)}\right) \rho \, d\rho \, d\alpha. \end{aligned} \quad (29)$$

In the alternating method for inverting our nonlinear model, using a vector of known target heights $\bar{\eta} = [\bar{\eta}_1, \bar{\eta}_2, \dots, \bar{\eta}_{N_t}] \in \mathbb{R}^{N_t}$ to recover $S_h^c(\rho, \alpha)$ instead of $\bar{S}_h(\rho, \alpha)$ can be appropriately incorporated by replacing (22) with

$$[\mathbf{D}(\bar{\rho})]_{m,n} = \frac{\bar{\rho}_j \arctan(\bar{\eta}_j / d(r_m, \theta_m, \bar{\rho}_j, \alpha_n))}{d(r_m, \theta_m, \bar{\rho}_j, \alpha_n)} \quad (30)$$

for incident light on the floor (originating from the hidden

when $\alpha_n \in [\bar{\alpha}_j - \Delta_j/2, \bar{\alpha}_j + \Delta_j/2]$.

Outside the laboratory, ambient light on the visible side might be more difficult to describe, and hidden scenes may be more complicated than a collection of a few targets. As our CRB results indicate, range information is subtle, making it easily overwhelmed by model mismatch. For these reasons, our 2D reconstruction algorithms, in their current form, might lack robustness outside of the controlled laboratory setting. However, there are ways the ideas in this paper could be realizable outside of the laboratory. For example, our algorithms might be adapted to produce reconstructions of only the moving part of the hidden scene. Like [30], we could greatly reduce model mismatch through background subtraction. Unlike [30], the ideas presented in this paper would allow us to produce 2D, instead of 1D, reconstructions of these moving targets.

The breadth of recent work on edge-resolved NLOS imaging [30], [31], including a recent active method [20], allows us to envision some of the ideas in this paper as part of a larger hybrid edge-resolved NLOS imaging system. Such a system might include both active and passive components, allowing for greater robustness under real world conditions. In this paper, range is an unknown part of the forward model that we *learn*. In a hybrid system, an active module might provide coarse (in angle) range information that could be plugged directly into our forward model so it would not need to be learned. The passive module could then quickly provide high resolution (in angle) RGB information about the hidden scene.

VII. CONCLUSION

In this work we explore 2D reconstruction of the region hidden behind a wall using a single photograph of the floor on the visible side. Unlike previous work, which has assumed all light sources to be in the far field, we propose a more complete forward model to describe radial falloff, enabling 2D reconstructions of the hidden scene. Using the Cramér–Rao bound for a single target, we demonstrate the utility and difficulty of using penumbra measurements for 2D reconstruction. We propose an alternating nonlinear inversion algorithm for 2D reconstruction and provide a comparison to a more conventional linear inversion algorithm. Experimental results demonstrate the promise and robustness of both methods.

ACKNOWLEDGMENT

The authors thank Charles Saunders for his help with data acquisition. Computing resources provided by Boston University Research Computing Service and support provided by the Draper Fellowship program are greatly appreciated.

REFERENCES

- [1] A. Kirmani, H. Jeelani, V. Montazerhodjat, and V. K. Goyal, “Diffuse imaging: Creating optical images with unfocused time-resolved illumination and sensing,” *IEEE Signal Process. Lett.*, vol. 19, no. 1, pp. 31–34, 2012.
- [2] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, “Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging,” *Nat. Commun.*, vol. 3, no. 745, 2012.
- [3] O. Gupta, T. Willwacher, A. Velten, A. Veeraraghavan, and R. Raskar, “Reconstruction of hidden 3D shapes using diffuse reflections,” *Opt. Express*, vol. 20, no. 17, pp. 19 096–19 108, 2012.
- [4] A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, “First-photon imaging,” *Science*, vol. 343, no. 6166, pp. 58–62, Jan. 2014.
- [5] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, “Photon-efficient computational 3d and reflectivity imaging with single-photon detectors,” *IEEE Trans. Comput. Imaging*, vol. 1, no. 2, pp. 112–125, Jun. 2015.
- [6] J. Rapp and V. K. Goyal, “A few photons among many: Unmixing signal and noise for photon-efficient active imaging,” *IEEE Trans. Comput. Imaging*, vol. 3, no. 3, pp. 445–459, 2017.
- [7] K. Xu, W. Jin, S. Zhao, J. Liu, H. Guo, S. Qiu, and D. Wu, “Image contrast model of non-line-of-sight imaging based on laser range-gated imaging,” *Opt. Eng.*, vol. 53, no. 6, p. 061610, 2013.
- [8] M. Laurenzis, F. Christnacher, J. Klein, M. B. Hullin, and A. Velten, “Study of single photon counting for non-line-of-sight vision,” in *Proc. SPIE Advanced Photon Counting Techniques IX*, vol. 9492, 2015, p. 94920K.
- [9] M. Laurenzis, F. Christnacher, and A. Velten, “Study of a dual mode SWIR active imaging system for direct imaging and non-line-of-sight vision,” in *Proc. SPIE Laser Radar Technology and Applications XX and Atmospheric Propagation XII*, vol. 9465, 2015, p. 946509.
- [10] M. Buttafava, J. Zeman, A. Tosi, K. Eliceiri, and A. Velten, “Non-line-of-sight imaging using a time-gated single photon avalanche diode,” *Opt. Express*, vol. 23, no. 16, pp. 20 997–21 011, 2015.
- [11] G. Gariepy, F. Tonolini, R. Henderson, J. Leach, and D. Faccio, “Detection and tracking of moving objects hidden from view,” *Nat. Photon.*, vol. 10, no. 1, pp. 23–26, 2016.
- [12] J. Klein, M. Laurenzis, and M. Hullin, “Transient imaging for real-time tracking around a corner,” in *SPIE Proc. Electro-Optical Remote Sensing X*, 2016, p. 998802.
- [13] S. Chan, R. E. Warburton, G. Gariepy, J. Leach, and D. Faccio, “Non-line-of-sight tracking of people at long range,” *Opt. Express*, vol. 25, no. 9, pp. 10 109–10 117, 2017.
- [14] C.-y. Tsai, K. N. Kutulakos, S. G. Narasimhan, and A. C. Sankaranarayanan, “The geometry of first-returning photons for non-line-of-sight imaging,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 7216–7224.
- [15] F. Heide, M. O’Toole, K. Zang, D. B. Lindell, S. Diamond, and G. Wetzstein, “Non-line-of-sight imaging with partial occluders and surface normals,” *ACM Trans. Graphics*, vol. 38, no. 3, 2019, art. 22.
- [16] M. O’Toole, D. B. Lindell, and G. Wetzstein, “Confocal non-line-of-sight imaging based on the light-cone transform,” *Nature*, vol. 555, pp. 338–341, Mar. 2018.
- [17] X. Liu, I. Guillén, M. La Manna, J. H. Nam, S. A. Reza, T. Huu Le, A. Jarabo, D. Gutierrez, and A. Velten, “Non-line-of-sight imaging using phasor-field virtual wave optics,” *Nature*, vol. 572, no. 7771, pp. 620–623, 2019.
- [18] D. B. Lindell, G. Wetzstein, and M. O’Toole, “Wave-based non-line-of-sight imaging using fast f - k migration,” *ACM Trans. Graph.*, vol. 38, no. 4, 2019, art. 116.
- [19] G. Musarra, A. Lyons, E. Conca, Y. Altmann, F. Villa, F. Zappa, M. Padgett, and D. Faccio, “Non-line-of-sight three-dimensional imaging with a single-pixel camera,” *Physical Review Applied*, vol. 12, no. 1, Jul 2019.
- [20] J. Rapp, C. Saunders, J. Tachella, J. Murray-Bruce, Y. Altmann, J.-Y. Tournet, S. McLaughlin, R. M. A. Dawson, F. N. C. Wong, and V. K. Goyal, “Seeing around corners with edge-resolved transient imaging,” 2020, arXiv:2002.07118 [eess.IV].
- [21] J. Klein, C. Peters, J. Martín, M. Laurenzis, and M. B. Hullin, “Tracking objects outside the line of sight using 2D intensity images,” *Scientific Reports*, vol. 6, no. 1, p. 32491, Oct. 2016.
- [22] C. Thrampoulidis, G. Shulkind, F. Xu, W. T. Freeman, J. H. Shapiro, A. Torralba, F. N. C. Wong, and G. W. Wornell, “Exploiting occlusion in non-line-of-sight active imaging,” *IEEE Trans. Comput. Imaging*, vol. 4, no. 3, pp. 419–431, Sept 2018.
- [23] A. L. Cohen, “Anti-pinhole imaging,” *Optica Acta*, vol. 29, no. 1, pp. 63–67, 1982.
- [24] 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.
- [25] 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.
- [26] 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. Computer Vision and Pattern Recognition*, June 2019.

- [27] 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.
- [28] J. Murray-Bruce, C. Saunders, and V. K. Goyal, "Occlusion-based computational periscopy with consumer cameras," in *SPIE Wavelets and Sparsity XVIII*, vol. 11138, 2019, pp. 286–297.
- [29] F. Xu, G. Shulkind, C. Thrampoulidis, J. H. Shapiro, A. Torralba, F. N. C. Wong, and G. W. Wornell, "Revealing hidden scenes by photon-efficient occlusion-based opportunistic active imaging," *Opt. Express*, vol. 26, no. 8, pp. 9945–9962, Apr. 2018.
- [30] 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. 23rd IEEE Int. Conf. Computer Vision*, 2017, pp. 2270–2278.
- [31] 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, pp. 1–9.
- [32] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm," *SIAM J. Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.
- [33] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "A sparse-group lasso," *J. Computational and Graphical Statistics*, vol. 22, no. 2, pp. 231–245, 2013.