Turning Corners into Cameras: Principles and Methods
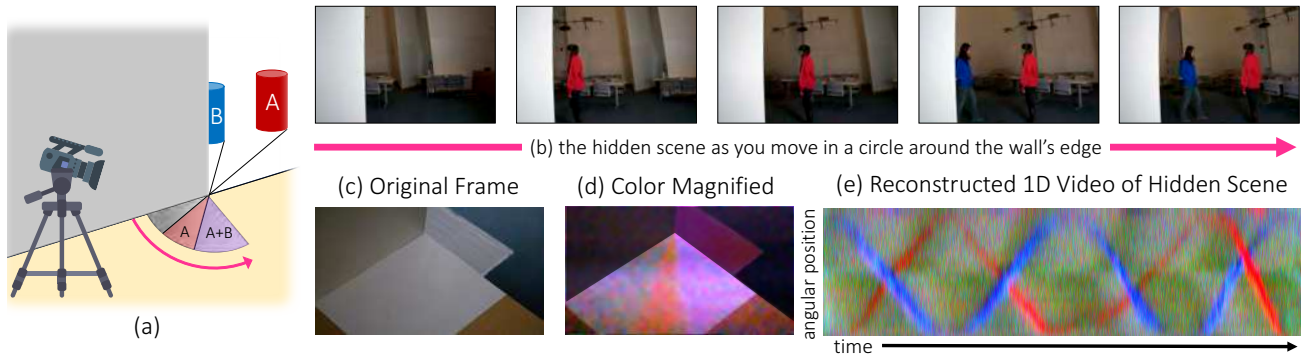
Katherine L. Bouman[1]   Vickie Ye[1]   Adam B. Yedidia[1]   Frédo Durand[1]

Gregory W. Wornell[1]   Antonio Torralba[1]   William T. Freeman[1,2]

[1]Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology   [2]Google Research

Figure 1: We construct a 1-D video of an obscured scene using RGB video taken with a consumer camera. The stylized diagram in (a) shows a typical scenario: two people—one wearing red and the other blue—are hidden from the camera's view by a wall. Only the region shaded in yellow is visible to the camera. To an observer walking around the occluding edge (along the magenta arrow), light from different parts of the hidden scene becomes visible at different angles (see sequence (b)). Ultimately, this scene information is captured in the intensity and color of light reflected from the corresponding patch of ground near the corner. Although these subtle irradiance variations are invisible to the naked eye (c), they can be extracted and interpreted from a camera position from which the entire obscured scene is hidden from view. Image (d) visualizes these subtle variations in the highlighted corner region. We use temporal frames of these radiance variations on the ground to construct a 1-D video of motion evolution in the hidden scene. Specifically, (e) shows the trajectories over time that specify the angular position of hidden red and blue subjects illuminated by a diffuse light.

Abstract

We show that walls, and other obstructions with edges, can be exploited as naturally-occurring "cameras" that reveal the hidden scenes beyond them. In particular, we demonstrate methods for using the subtle spatio-temporal radiance variations that arise on the ground at the base of a wall's edge to construct a one-dimensional video of the hidden scene behind the wall. The resulting technique can be used for a variety of applications in diverse physical settings. From standard RGB video recordings, we use edge cameras to recover 1-D videos that reveal the number and trajectories of people moving in an occluded scene. We further show that adjacent wall edges, such as those that arise in the case of an open doorway, yield a stereo camera from which the 2-D location of hidden, moving objects can be recovered. We demonstrate our technique in a number of indoor and outdoor environments involving varied floor surfaces and illumination conditions.

1. Introduction

The ability to see around obstructions would prove valuable in a wide range of applications. As just two examples, remotely sensing occupants in a room would be valuable in search and rescue operations, and the ability to detect hidden, oncoming vehicles and/or pedestrians would be valuable in collision avoidance systems [2]. Although often not visible to the naked eye, in many environments, light from obscured portions of a scene is scattered over many of the observable surfaces. This reflected light can be used to recover information about the hidden scene (see Fig. 1). In this work, we exploit the vertical edge at the corner of a wall to construct a "camera" that sees beyond the wall. Since vertical wall edges are ubiquitous, such cameras can be found in many environments.

The radiance emanating from the ground in front of a corner, e.g., at the base of a building, is influenced by many factors: the albedo, shape, and BRDF of its surface, as well as the light coming from the full hemisphere above it. Assuming the ground has a significant diffuse component, a majority of the reflected light comes from the surroundings

that are easily seen from the observer's position next to the occluding wall (the visible region is shaded in yellow in Fig. 1(a)). However, emitted and reflected light from behind the corner, hidden from the observer, also has a small effect on the ground's radiance in the form of a subtle gradient of light encircling the corner; this is not a shadow, but is instead what is referred to as a *penumbra*.

The faint penumbra on the ground is caused by the reflection of an increasing amount of light from the hidden scene. To illustrate this, imagine standing with your shoulder up against the building's wall (refer to the leftmost picture of Fig. 1(b)). At this position you are unable to see any of the scene behind the corner. However, as you slowly move away from the wall, walking along the magenta circle shown in Fig. 1(a), you see an increasing amount of the scene. Eventually, the hidden scene comes fully into view. Similarly, different points on the ground reflect light integrated from differently-sized fractions of the hidden scene.

Now imagine someone has entered the hidden portion of the scene. This person would introduce a small change to the light coming from an angular slice of the room. From behind the corner this change would often not be perceptible to the naked eye. However, it would result in a subtle change to the penumbra; see Fig. 1(c) and (d). We use these subtle changes, recorded from standard video cameras, to construct a 1-D version of how the hidden scene beyond the corner evolves with time; see Fig. 1(e).

Section 2 summarizes related work that puts the present contribution in context. Section 3 shows how, using our proposed methods, it is possible to identify the number and location of people in a hidden scene. Section 4 shows how parallax created by a pair of adjacent edges, such as in a doorway, can be used to triangulate the 2D position of moving people over time. Experimental results (in the paper and supplemental material) are shown for a number of indoor and outdoor environments with varied flooring, including carpet, tile, hardwood, concrete, and brick.

2. Related Work

In this section we describe previous non-line-of-sight (NLoS) methods. Previous methods used to see past or through occluders have ranged from using WiFi signals [1] to exploiting random specular surfaces [21, 4]. In this summary, we emphasize a few active and passive approaches that have previously been used to see past occluders and image hidden scenes.

Recovery under Active Illumination: Past approaches to see around corners have largely involved using time-of-flight (ToF) cameras [14, 20, 10, 6]. These methods involve using a laser to illuminate a point that is visible to both the observable and hidden scene, and measuring how long it takes for the light to return [20, 15]. By measuring the

light's time of flight, one can infer the distance to objects in the hidden scene, and by measuring the light's intensity, one can often learn about the reflectance and curvature of the objects [13]. Past work has used ToF methods to infer the location [7], size and motion [12, 5], and shape [17] of objects in the hidden scene. These methods have also been used to count hidden people [19].

ToF cameras work well in estimating the depths of hidden objects, however, they have some limitations. First, they require specialized and comparatively expensive detectors with fine temporal resolution. Second, they are limited in how much light they can introduce in the scene to support imaging. Third, they are vulnerable to interference from ambient outdoor illumination. By contrast, our proposed real-time passive technique operates in unpredictable indoor and outdoor environments with inexpensive consumer cameras, without additional illumination.

In [9] a laser is used to indirectly illuminate an object behind an occluder. Using a standard camera the authors are then able to identify the position of the hidden object. Similar to our proposed work, [9] uses a standard camera; however, their proposed system has a number of limitations. Namely, they require controlled conditions where the geometry of the unknown moving object is rigid, and its shape and material are either known or can be closely modeled by a single oriented surface element. In contrast, our method requires minimal prior information, is completely passive, and has been shown to work in many natural settings.

Passive Recovery: Other work has previously considered the possibility of using structures naturally present in the real world as cameras. Naturally occurring pinholes (such as windows) or pinspecks have been previously used for non-line-of-sight imaging [16, 3]. In addition, specular reflections off of human eyes have been used to image hidden scenes [11]. Although these accidental cameras can be used to reconstruct 2-D images, they require a more specialized accidental camera scenario than the simple edges we propose to use in this work.

The technique presented in [18] also detects and visualizes small, often imperceptible, color changes in video. However, in this work, rather than just visualize these tiny color changes, we interpret them in order to reconstruct a video of a hidden scene.

3. Edge Cameras

An edge camera system consists of four components: the visible and hidden scenes, the occluding edge, and the ground, which reflects light from both scenes. We refer to the (ground) plane perpendicular to the occluding edge as the *observation plane*. By analyzing subtle variations in the penumbra at the base of an edge, we are able to deduce a hidden subject's pattern of motion.

The reflected light from a surface at point p, with normal $\hat{n}$, is a function of the incoming light L_i' as well as the surface's albedo a and BRDF β. Specifically,

$$L_o'(p, \hat{v}_o) = a(p) \int L_i'(p, \hat{v}_i)\, \beta(\hat{v}_i, \hat{v}_o, \hat{n})\, \gamma(\hat{v}_i, \hat{n})\, \mathrm{d}\hat{v}_i, \tag{1}$$

where $\hat{v}_i$ and $\hat{v}_o$ denote the incoming and outgoing unit vectors of light at position $p = (r, \theta)$, respectively, and $\gamma(\hat{v}_i, \hat{n}) = \hat{v}_i \cdot \hat{n}$. We parameterize p in polar coordinates, with the origin centered at the occluding edge and $\theta = 0$ corresponding to the angle parallel to the wall coming from the corner (refer to Fig. 2). For simplicity, we assume the observation plane is Lambertian, and that the visible and hidden scene are modeled as light emitted from a large celestial sphere, parameterized by right ascension α and declination δ. Under these assumptions, we simplify (1):

$$L_o'(r, \theta) = a(r, \theta) \int_{\alpha=0}^{2\pi} \int_{\delta=0}^{\pi/2} L_i(\alpha, \delta)\, \mathrm{d}\alpha\, \mathrm{d}\delta \tag{2}$$

where $L_i = L_i'\gamma$. Furthermore, since the occluding edge blocks light from $[\pi + \theta, 2\pi]$ at the radial line θ,

$$L_o'(r, \theta) = a(r, \theta) \left[L_v + \int_{\phi=0}^{\theta} L_h(\phi)\, \mathrm{d}\phi \right] \tag{3}$$

for $L_v = \int_{\alpha=0}^{\pi} \int_{\delta=0}^{\pi/2} L_i(\alpha, \delta)\, \mathrm{d}\alpha\, \mathrm{d}\delta$ and $L_h(\phi) = \int_{\delta=0}^{\pi/2} L_i(\pi + \phi, \delta)\, \mathrm{d}\delta$. By inspecting (3) we can see that the intensity of light on the penumbra is explained by a constant term, L_v, which is the contribution due to light visible to the observer (shaded in yellow in Fig. 1(a)), and a varying angle dependent term which integrates the light in the hidden scene, L_h. For instance, a radial line at $\theta = 0$ only integrates the light from the scene visible to the observer, while the radial line $\theta = \pi/2$ reflects the integral of light over the entire visible and hidden scenes.

Then, if we assume that $\frac{d}{d\theta} a(r, \theta) \approx 0$[1], the derivative of the observed penumbra recovers the 1-D angular projection of the hidden scene:

$$\frac{\mathrm{d}}{\mathrm{d}\theta} L_o'(r, \theta) \approx a(r, \theta)\, L_h(\theta). \tag{4}$$

But what happens if someone walks into the hidden scene at time t, changing $L_h^0(\theta)$ to $L_h^t(\theta)$? In this case, the spatial derivative of the temporal difference encodes the angular change in lighting:

$$\frac{d}{d\theta} \left[L_o'^t(r, \theta) - L_o'^0(r, \theta) \right] = a(r, \theta) \left[L_h^t(\theta) - L_h^0(\theta) \right] \tag{5}$$

---

[1] In practice, we subtract a background frame to substantially remove per-pixel albedo variations. Refer to Section 3.1.1
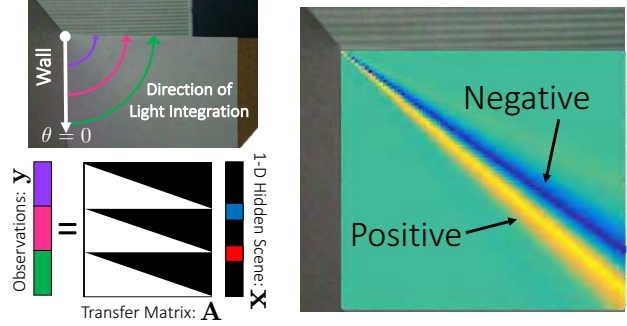
(a) Constructing Transfer Matrix   (b) Sample Estimation Gain Image

Figure 2: In (a), the transfer matrix, $\mathbf{A}$, is shown for a toy situation in which observations lie along circles around the edge. In this case, $\mathbf{A}$ would simply be a repeated lower triangular matrix. (b) contains an example estimation gain image, which describes the matrix operation performed on observations $\mathbf{y}^{(t)}$ to estimate $\mathbf{x}^{(t)}$. As predicted, the image indicates that we are essentially performing an angular derivative in recovering a frame of the 1-D video.

In other words, the angular derivative of the penumbra's difference from the reference frame is a signal that indicates the angular change in the hidden scene over time. In practice, we obtain good results assuming $a(r, \theta) = 1$ and using the cameras' native encoded intensity values while subtracting the temporal mean as a background frame (see Section 3.1.1).

3.1. Method

Using a video recording of the observation plane, we generate a 1-D video indicating the changes in a hidden scene over time. These 1-D angular projections of the hidden scene, viewed over many time-steps, reveal the trajectory of a moving object behind the occluding edge.

Likelihood: At each time t, we relate the observed M-pixels on the projection plane, $\mathbf{y}^{(t)}$, to the 1-D angular projection of the hidden scene, $L_h^{(t)}(\theta)$. We formulate a discrete approximation to our edge camera system by describing the continuous image $L_h^{(t)}(\theta)$ using N terms, $\mathbf{x}^{(t)}$. The observations $\mathbf{y}^{(t)}$ then relate to the unknown parameters $\mathbf{x}^{(t)}$ and $L_v^{(t)}$ by a linear matrix operation:

$$\mathbf{y}^{(t)} = L_v^{(t)} + \mathbf{A}\mathbf{x}^{(t)} + \mathbf{w}^{(t)}, \qquad \mathbf{w}^{(t)} \sim \mathcal{N}(0, \lambda^2 \mathbf{I}),$$

where the $M \times N$ matrix $\mathbf{A}$ is defined by the geometry of the system. More explicitly, each row m of $\mathbf{A}$ integrates the portion of the hidden scene visible from observation m, $\mathbf{y}_m^{(t)}$. In the simplified case of observations that lie on a circle around the occluding edge, $\mathbf{A}$ would simply be a constant lower-triangular matrix; see Fig. 2(a).

Let $\widetilde{\mathbf{A}}$ be the column augmented matrix $[\mathbf{1}\ \mathbf{A}]$. We can then express the likelihood of an observation given $\mathbf{x}^{(t)}$ and $L_v^{(t)}$ as:

$$p(\mathbf{y}^{(t)} | \mathbf{x}^{(t)}, L_v^{(t)}) = \mathcal{N}\left(\widetilde{\mathbf{A}} \left[L_v^{(t)}\ \mathbf{x}^{(t)T} \right]^T, \lambda^2 \mathbb{1} \right). \tag{6}$$

Prior: The signal we are trying to extract is very small relative to the total light intensity on the observation plane. Therefore, to improve the quality of results, we enforce spatial smoothness of $\mathbf{x}^{(t)}$. We use a simple L2 smoothness regularization over adjacent parameters in $\mathbf{x}^{(t)}$. This corresponds, for a gradient matrix $\mathbf{G}$, to using the prior

$$p(\mathbf{x}^{(t)}) \propto \prod_{n=1}^{N-1} \exp\left[-\frac{1}{2\sigma_1^2}\|\mathbf{x}^{(t)}[n] - \mathbf{x}^{(t)}[n-1]\|_2^2\right]$$

$$\prod_{n=1}^{N} \exp\left[-\frac{1}{2\sigma_2^2}\|\mathbf{x}^{(t)}[n]\|_2^2\right] \quad (7)$$

$$= \mathcal{N}(\mathbf{0}, \sigma_1^2(\mathbf{G}^T\mathbf{G})^{-1} + \sigma_2^2 \mathbb{1}). \quad (8)$$

Inference: We seek a maximum a posteriori (MAP) estimate of the hidden image coefficients, $\mathbf{x}^{(t)}$, given M observations, $\mathbf{y}^{(t)}$, measured by the camera. By combining the defined Gaussian likelihood and prior distributions, we obtain a Gaussian posterior distribution of $\mathbf{x}^{(t)}$ and $L_v^{(t)}$,

$$p(\mathbf{x}^{(t)}, L_v^{(t)}|\mathbf{y}^{(t)}) = \mathcal{N}\left(\left[\hat{L}_v^{(t)}\ \hat{\mathbf{x}}^{(t)T}\right]^T, \Sigma^{(t)}\right)$$

$$\Sigma^{(t)} = \left[\lambda^{-2}\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}} + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\mathbf{G}^T\mathbf{G}}{\sigma_1^2} + \frac{\mathbb{1}}{\sigma_2^2} \end{pmatrix}\right]^{-1}$$

$$\left[\hat{L}_v^{(t)}\ \hat{\mathbf{x}}^{(t)T}\right]^T = \Sigma^{(t)}\lambda^{-2}\widetilde{\mathbf{A}}^T\mathbf{y}^{(t)} \quad (9)$$

where the maximum a posteriori estimate is given by $\hat{\mathbf{x}}^{(t)}$.

To better understand the operation that is being performed to obtain the 1-D reconstruction, we visualize each row of the matrix $\Sigma^{(t)}\lambda^{-2}\widetilde{\mathbf{A}}^T$. We refer to each reshaped row of this matrix as the *estimation gain image*. An example estimation gain image is shown in Fig. 2b. As expected, the matrix operation is computing an angular derivative over the observation plane. Note that although earlier we assumed $\frac{d}{d\theta}a(r,\theta) \approx 0$, in reality the albedo simply needs to be orthogonal to the zero-mean pie-wedges in each estimation gain image. We expect violations from this assumption to be small.

3.1.1 Implementation Details

Rectification: All of our analysis thus far has assumed we are observing the floor parallel to the occluding edge. However, in most situations, the camera will be observing the projection plane at an angle. In order to make the construction of the matrix $\mathbf{A}$ easier, we begin by rectifying our images using a homography. In these results, we assume the ground is perpendicular to the occluding edge, and estimate the homography using either a calibration grid or regular patterns, such as tiles, that naturally appear on the ground. Alternatively, a known camera calibration could be used.

Background Subtraction: Since we are interested in identifying temporal differences in a hidden scene due to a moving subject, we must remove the effect of the scene's background illumination. Although this could be accomplished by first subtracting a background frame, L_o^0, taken without the subject, we avoid requiring the availability of such a frame. Instead, we assume the subject's motion is roughly uniform over the video, and use the video's mean image in lieu of a true background frame. We found that in sequences containing people moving naturally, background subtraction using the average video frame worked well.

Temporal Smoothness: In addition to spatial smoothness we could also impose temporal smoothness on our MAP estimate. $\hat{\mathbf{x}}^{(t)}$. This helps to further regularize our result, at the cost of some temporal blurring. However, to emphasize the coherence among results, we do not impose this additional constraint. Each 1-D image, $\mathbf{x}^{(t)}$, that we show is independently computed. Results obtained with temporal smoothness constraints are shown in the supplemental material.

Parameter Selection: The noise parameter λ^2 is set for each video as the median variance of estimated sensor noise. The regularization parameters σ_1 and σ_2 are empirically set to 0.1 for all results.

3.2. Experiments and Results

Our algorithm reconstructs a 1-D video of a hidden scene from behind an occluding edge, allowing users to track the motions of obscured, moving objects. In all results shown, the subject was not visible to an observer at the camera.

We present results as space-time images. These images contain curves that indicate the angular trajectories of moving people. All results, unless specified otherwise, were generated from standard, compressed video taken with a SLR camera. Please refer to the supplemental video for full sequences and additional results.

3.2.1 Environments

We show several applications of our algorithm in various indoor and outdoor environments. For each environment, we show the reconstructions obtained when one or two people were moving in the hidden scene.

Indoor: In Fig. 1(e) we show a result obtained from a video recorded in a mostly dark room. A large diffuse light illuminated two hidden subjects wearing red and blue clothing. As the subjects walked around the room, their clothing reflected light, allowing us to reconstruct a 1-D video of colored trajectories. As correctly reflected in our reconstructed video, the subject in blue occludes the subject in red three times before the subject in red becomes the occluder.

Fig. 3 shows additional examples of 1-D videos recovered from indoor edge cameras. In these sequences, the environment was well-lit. The subjects occluded the bright ambient
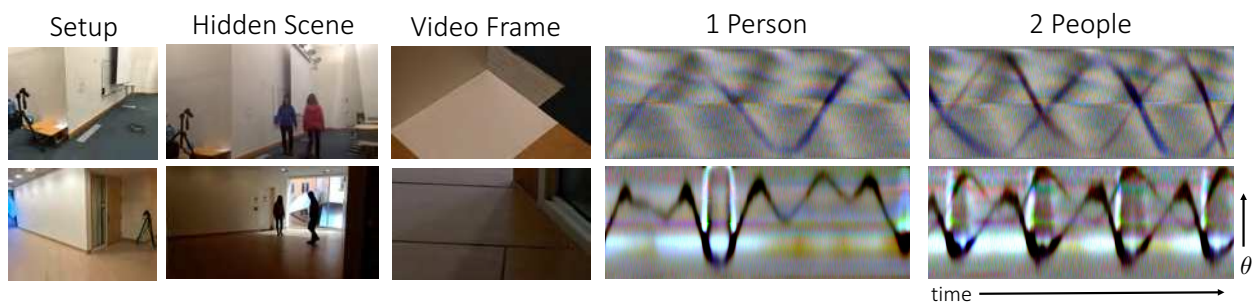
Figure 3: One-dimensional reconstructed videos of indoor, hidden scenes. Results are shown as space-time images for sequences where one or two people were walking behind the corner. In these reconstructions, the angular position of a person, as well as the number of people, can be clearly identified. Bright vertical line artifacts are caused by additional shadows appearing on the penumbra. We believe horizontal line artifacts result from sampling on a square grid.
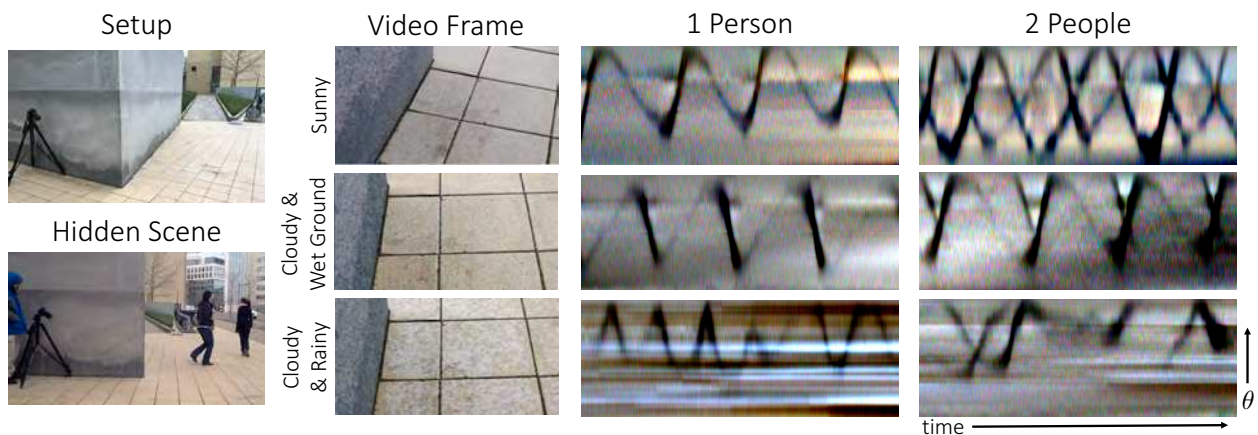

Figure 4: 1-D reconstructed videos of a common outdoor, hidden scene under various weather conditions. Results are shown as space-time images. The last row shows results from sequences taken while it was beginning to rain. Although artifacts appear due to the appearing raindrops, motion trajectories can be identified in all reconstructions.

light, resulting in the reconstruction's dark trajectory. Note that in all the reconstructions, it is possible to count the number of people in the hidden scene, and to recover important information such as their angular position and speed, and the characteristics of their motion.

Outdoor: In Fig. 4 we show the results of a number of videos taken at a common outdoor location, but in different weather conditions. The top sequences were recorded during a sunny day, while the bottom two sequences were recorded while it was cloudy. Additionally, in the bottom sequence, raindrops appeared on the ground *during* recording, while in the middle sequence the ground was fully saturated with water. Although the raindrops cause artifacts in the reconstructed space-time images, you can still discern the trajectory of people hidden behind the wall.

3.2.2 Video Quality:

In all experiments shown thus far we have used standard, compressed video captured using a consumer camera. However, video compression can create large, correlated noise that may affect our signal. We have explored the effect video quality has on results. To do this, we filmed a common scene using 3 different cameras: an iPhone 5s, a Sony Alpha 7s, and a uncompressed RGB Point Grey. Fig. 5 shows the results of this experiment assuming different levels of i.i.d. noise. Each resulting 1-D image was reconstructed from a

single frame. The cell phone camera's compressed videos resulted in the noisiest reconstructions, but even those results still capture key features of the subject's path.

3.2.3 Velocity Estimation

The derivative of a person's trajectory over time, $\theta^{(t)}$, indicates their angular velocity. Fig. 6 shows an example of the estimated angular velocity obtained from a single edge camera when the hidden subject was walking roughly in a circle. Note that the person's angular size and speed are both larger when the person is closer to the corner. Such cues can help approximate the subject's 2-D position over time.

3.3. Estimated Signal Strength

In all of our presented reconstructions we show images with an intensity range of 0.1. As these results were obtained from 8-bit videos, our target signal is less than 0.1% of the video's original pixel intensities.

To better understand the signal measurement requirements, we have developed a simple model of the edge camera system that both explains experimental performance and enables the study of asymptotic limits.

We consider three sources of emitted or reflected light: a cylinder (proxying for a person), a hemisphere of ambient light (the surrounding scene), and an infinitely tall half-plane (the occluding wall). If all surfaces are Lambertian, the
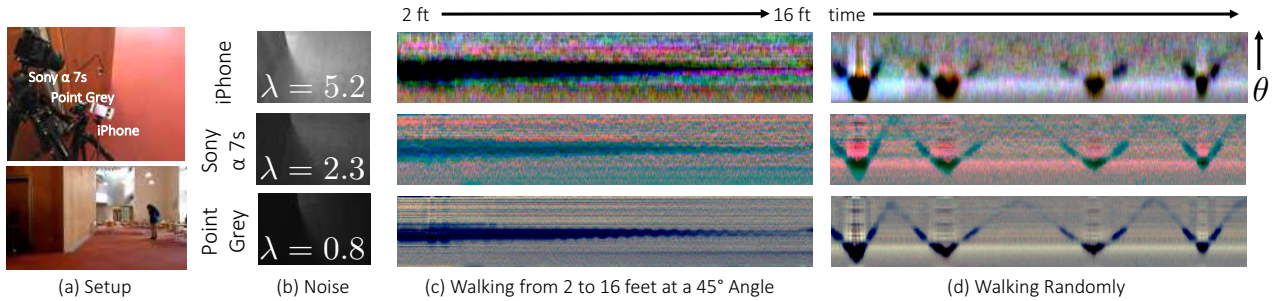
Figure 5: The result of using different cameras on the reconstruction of the same sequence in an indoor setting. Three different 8-bit cameras (an iPhone 5s, a Sony Alpha 7s, and an uncompressed RGB Point Grey) simultaneously recorded the carpeted floor. Each camera introduced a different level of sensor noise. The estimated standard deviation of per-pixel sensor noise, λ, is shown in (b). We compare the quality of two sequences in (c) and (d). In (c), we have reconstructed a video from a sequence of a single person walking directly away from the corner from 2 to 16 feet at a 45 degree angle from the occluded wall. This experiment helps to illustrate how signal strength varies with distance from the corner. In (d), we have done a reconstruction of a single person walking in a random pattern. In (c) the hidden person does not change in angular position. Thus, for these results, we subtract an average background frame computed from a different portion of the video sequence.
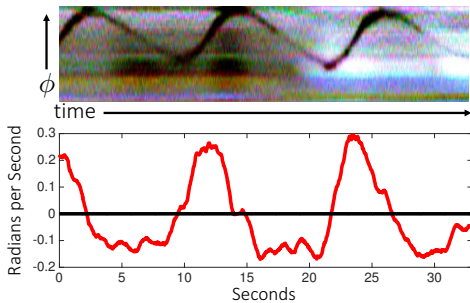

Figure 6: A subject's reconstructed angular velocity relative to the corner as a function of time. In this sequence, a person was walking in circles far from the corner.

brightness change of the observation plane due the presence of the cylinder around the corner can be computed analytically for this simple system. See the supplementary document.

For reasonable assumed brightnesses of the cylinder, hemisphere, and half-plane (150, 300, and 100, respectively, in arbitrary linear units), the brightness change on the observation plane due to the cylinder will be an extremum of -1.7 out of a background of 1070 units. This is commensurate with our experimental observations of $\sim 0.1\%$ change of brightness over the penumbra region. Our model shows novel asymptotic behavior of the edge camera. Namely, at large distances from the corner, brightness changes in the penumbra decrease faster than would otherwise be expected from a 1-D camera. This is because the arrival angle of the rays from a distant cylinder are close to grazing with the ground, lessening their influence on the penumbra. However, within 10 meters of the corner, such effects are small.

4. Stereo Edge Cameras

Although the width of a track recovered in the method of the previous section can give some indication of a hidden person's relative range, more accurate methods are possible by exploiting adjacent walls. For example, when a hidden scene is behind a doorway, the pair of vertical doorway
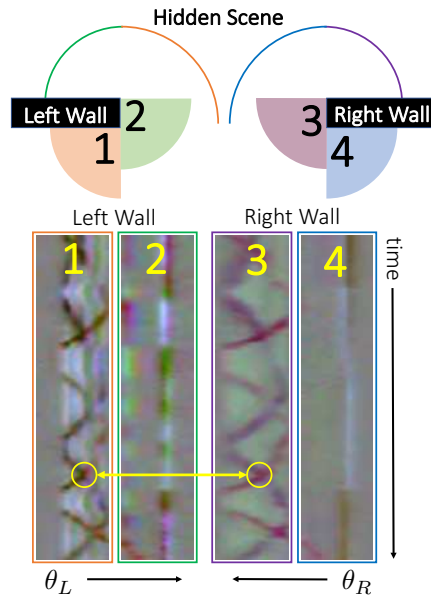

Figure 7: The four edges of a doorway contain penumbras that can be used to reconstruct a $180°$ view of a hidden scene. The top diagram indicates the penumbras and the corresponding region they describe. Parallax occurs in the reconstructions from the left and right wall. This can be seen in the bottom reconstruction of two people hidden behind a doorway. Numbers/colors indicate the penumbras used for each $90°$ space-time image.

wall edges yield a pair of corner cameras. By treating the observation plane at the base of each edge as a camera, we can obtain stereo 1-D images that we can then use to triangulate the absolute position of a subject over time.

4.1. Method

A single edge camera allows us to reconstruct a $90°$ angular image of an occluded scene. We now consider a system composed of four edge cameras, such as an open doorway, as illustrated in Fig. 7. Each side of the doorway contains two adjacent edge cameras, whose reconstructions together create a $180°$ view of the hidden scene.

The two sides of the doorway provide two views of the same hidden scene, but from different positions. This causes
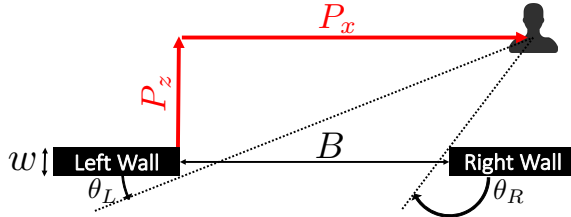
Figure 8: A hidden person will introduce an intensity change on the left and right wall penumbras at angles of $\theta_L^{(t)}$ and $\theta_R^{(t)}$, respectively. Once these angles have been identified, we can recover the hidden person's two-dimensional location using Eq. 11.

an offset in the projected angular position of the same person (see Fig. 8). Our aim is to use this angular parallax to triangulate the location of a hidden person over time. Assume we are observing the base of a doorway, with walls of width w separated by a distance B. A hidden person will introduce an intensity change on the left and right wall penumbras at angles of $\theta_L^{(t)}$ and $\theta_R^{(t)}$, respectively. From this correspondence, we can triangulate their 2-D location.

$$P_z^{(t)} = \frac{B - \eta^{(t)}}{\cot \theta_L^{(t)} + \cot \theta_R^{(t)}} \tag{10}$$

$$P_x^{(t)} = P_z^{(t)} \cot \theta_L^{(t)} \tag{11}$$

$$\eta^{(t)} = \begin{cases} w \cot(\theta_R) & P_x \leq 0 \\ 0 & 0 \leq P_x \leq B \\ w \cot(\theta_L) & P_x \geq B \end{cases} \tag{12}$$

where (P_x, P_z) are the x- and z-coordinate of the person. We define the top corner of the left doorway, corner 1 in Fig. 7, as $(P_x, P_z) = (0, 0)$.

Assuming the wall is sufficiently thin compared to the depth of moving objects in the hidden scene, the $\eta^{(t)}$ term can be ignored. In this case, the relative position of the person can be reconstructed without any knowledge of the absolute geometry of the doorway (e.g. B or w). In all results shown in this paper, we have made this assumption.

Identifying Trajectories: While automatic contour tracing methods exist [8], for simplicity, in our stereo results, we identify the trajectories of objects in the hidden scene manually by tracing a path on the reconstructed space-time images.

4.2. Experiments and Results

We demonstrate the ability of our method to localize the two-dimensional position of a hidden object using four edge cameras, such as in a doorway. We present a series of experiments in both controlled and uncontrolled settings. Full sequences, indicating the ground truth motions, and additional results can be found in the supplemental material.

Controlled Environment: To demonstrate the ability to infer depth from stereo edge cameras we constructed a controlled experiment. A monitor displaying a slowly moving
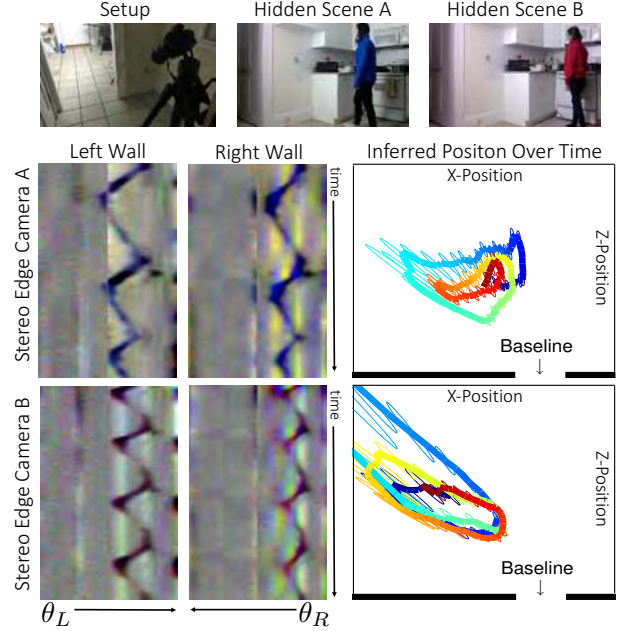

Figure 9: The results of our stereo experiments in a natural setting. Each sequence consists of a single person walking in a roughly circular pattern behind a doorway. The 2-D inferred locations over time are shown as a line from blue to red. Error bars indicating one standard deviation of error have been drawn around a subset of the points. Our inferred depths capture the hidden subject's cyclic motion, but are currently subject to large error. A subset of B's inferred 2-D locations have been cut out of this figure, but can be seen in full in the supplemental material.

green line was placed behind two walls, separated by a baseline of 20 cm, at a distance of roughly 23, 40, 60, and 84 cm. Fig. 10(b) shows sample space-time reconstructions of each 180° edge camera. The depth of the green line was then estimated from manually identified trajectories obtained from these space-time images. Empirically estimated error ellipses are shown in red for a subset of the depth estimates.

Natural Environment: Fig. 9 shows the results of estimating 2-D positions from doorways in natural environments. The hidden scene consists of a single person walking in a circular pattern behind the doorway. Although our reconstructions capture the cyclic nature of the subject's movements, they are sensitive to error in the estimated trajectories. Refer to Section 4.3. Ellipses indicating empirically estimated error have been drawn around a subset of the points.

4.3. Error Analysis

There are multiple sources of error that can introduce biases into location estimates. Namely, inaccuracy in localizing the projected trajectories, and mis-calibration of the scene cause error in the estimates. We discuss the effects of some of these errors below. Further derivations and analysis can be seen in our supplemental material.

Trajectory Localization: Because P_z scales inversely with $cot(\theta_L) + cot(\theta_R)$, small errors in the estimated projected angles of the person in the left and right may cause
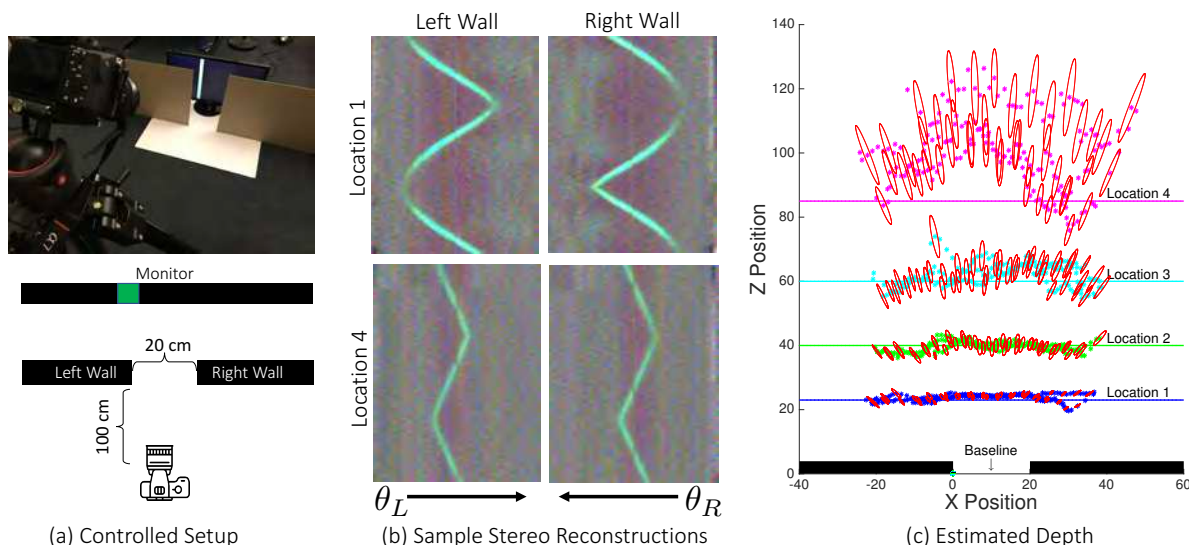
Figure 10: Controlled experiments were performed to demonstrate the ability to infer depth from stereo edge cameras. A monitor displaying a moving green line was placed behind an artificial doorway (a) at four locations corresponding to 23, 40, 60, and 84 cm, respectively. (b) shows sample reconstructions done of the edge cameras for the left and right wall when the monitor was placed at 23 and 84 cm. Using tracks obtained from these reconstructions, the 2-D position of the green line in each sequence was estimated over time (c). The inferred position is plotted with empirically computed error ellipses (indicating one standard deviation of noise).

large errors in the estimated position of the hidden person, particularly at larger depths. Assuming Gaussian uncertainty in the left and right angular trajectories, σ_{θ_L} and σ_{θ_R}, the uncertainty in the estimated position of the hidden person will not be Gaussian. However, the standard deviation of empirical distributions through sampling, as seen in Figs. 9 and 10, can be informative. Additionally, by using standard error propagation of independent variables, we can compute a first order approximation of the uncertainty. For instance, the uncertainty in the z position, σ_{P_z}, is

$$\sigma_{P_z} = B\sqrt{\frac{\sigma_{\theta_L}^2 \csc^4 \theta_L + \sigma_{\theta_R}^2 \csc^4 \theta_R}{(\cot \theta_L + \cot \theta_R)^4}} \quad (13)$$
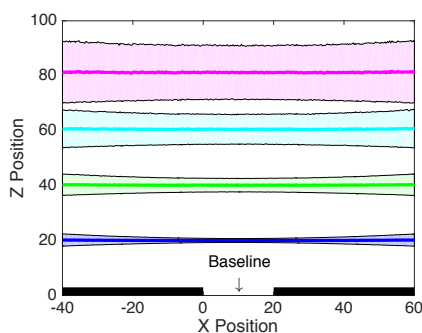

Figure 11: The empirical means plus or minus one standard deviation of the estimated P_z as a function of its x-coordinate, assuming true P_z of 20, 40, 60, and 80. Here, the two corner location errors at each of the boundaries of the doorway are independent and subject to $\sigma_{\Delta x}^2 = \sigma_{\Delta z}^2 = 0.04$.

Corner Identification: Misidentifying the corner of each occluding edge will cause systematic error to the estimated 2-D position. To determine how erroneously identifying a corner affects our results, we consider the following situation: a doorway of baseline $B = 20$ obscuring a bright object at angular position θ in an otherwise dark scene.

Assuming the offset from the true corner location is drawn from an independent Gaussian distribution, we can calculate the error between the estimated and true angular position, and then subsequently use these offsets to calculate the error in depth. Fig. 11 shows the error as a function of depth for a stereo camera setup in which the corner offset has been drawn from a Gaussian distribution with variance 0.04.

5. Conclusion

We show how to turn corners into cameras, exploiting a common, but overlooked, visual signal. The vertical edge of a corner's wall selectively blocks light to let the ground nearby display an angular integral of light from around the corner. The resulting penumbras from people and objects are invisible to the eye – typical contrasts are 0.1% above background – but are easy to measure using consumer-grade cameras. We produce 1-D videos of activity around the corner, measured indoors, outdoors, in both sunlight and shade, from brick, tile, wood, and asphalt floors. The resulting 1-D videos reveal the number of people moving around the corner, their angular sizes and speeds, and a temporal summary of activity. Open doorways, with two vertical edges, offer stereo views inside a room, viewable even away from the doorway. Since nearly every corner now offers a 1-D view around the corner, this opens potential applications for automotive pedestrian safety, search and rescue, and public safety. This ever-present, but previously unnoticed, 0.1% signal may invite other novel camera measurement methods.

References

[1] F. Adib and D. Katabi. See through walls with wifi! *ACM*, 43(4):75–86, 2013. 2

[2] P. Borges, A. Tews, and D. Haddon. Pedestrian detection in industrial environments: Seeing around corners. 2012. 1

[3] A. L. Cohen. Anti-pinhole imaging. *Optica Acta: International Journal of Optics*, 29(1):63–67, 1982. 2

[4] R. Fergus, A. Torralba, and W. Freeman. Random lens imaging. 2006. 2

[5] G. Gariepy, F. Tonolini, R. Henderson, J. Leach, and D. Faccio. Detection and tracking of moving objects hidden from view. *Nature Photonics*, 2015. 2

[6] F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3222–3229, June 2014. 2

[7] A. Kadambi, H. Zhao, B. Shi, and R. Raskar. Occluded imaging with time-of-flight sensors, 2016. ACM Transactions on Graphics. 2

[8] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988. 7

[9] 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*, 6, 2016. 2

[10] M. Laurenzis, A. Velten, and J. Klein. Dual-mode optical sensing: three-dimensional imaging and seeing around a corner. *Optical Engineering*, 2017. 2

[11] K. Nishino and S. Nayar. Corneal imaging system: Environment from eyes. *International Journal of Computer Vision*, 70(1):23–40, 2006. 2

[12] R. Pandharkar, A. Velten, A. Bardagjy, B. M. Lawson, E., and R. Raskar. Estimating motion and size of moving non-line-of-sight objects in cluttered environments, 2011. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on (pp. 265-272). 2

[13] D. Shin, A. Kirmani, V. Goyal, and J. Shapiro. Computational 3d and reflectivity imaging with high photon efficiency. *Image Processing (ICIP), 2014 IEEE International Conference*, 2014. 2

[14] D. Shin, A. Kirmani, V. Goyal, and J. Shapiro. Photon-efficient computational 3-d and reflectivity imaging with single-photon detectors. *IEEE Transactions on Computational Imaging*, 2015. 2

[15] S. Shrestha, F. Heide, W. Heidrich, and G. Wetzstein. Computational imaging with multi-camera time-of-flight systems. *ACM Transactions on Graphics (TOG)*, 2016. 2

[16] A. Torralba and W. T. Freeman. Accidental pinhole and pinspeck cameras: Revealing the scene outside the picture. *Computer Vision and Pattern Recognition (CVPR). IEEE.*, pages 374–381, 2012. 2

[17] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. Bawendi, and R. Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications*, 3(3):745, 2012. ACM Transactions on Graphics. 2

[18] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. *IEEE Signal Processing Letters*, 2012. 2

[19] L. Xia, C. Chen, and J. Aggarwal. Human detection using depth information by kinect. *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2011. 2

[20] F. Xu, D. Shin, D. Venkatraman, R. Lussana, F. Villa, F. Zappa, V. Goyal, F. Wong, and J. Shapiro. Photon-efficient computational imaging with a single-photon camera. *Computational Optical Sensing and Imaging*, 2016. 2

[21] Z. Zhang, P. Isola, and E. Adelson. Sparklevision: Seeing the world through random specular microfacets. 2014. 2