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, MIT [2]Google Research

Please see our supplementary videos for a brief overview of our methods as well as additional results.

Contents

1 Simple Perturbation Model

This section analyzes the signal on the observation plane.

We model three sources of light in the image, shown in Fig. 1: a hemisphere of external illumination of uniform brightness A; a wall of brightness B; and an imaging subject of brightness C. The wall is assumed to be infinitely tall (relative to the scale of the imaged region of the ground), and the height, width, and location of the cylindrical subject are as shown in the figure. We assume that all surfaces are Lambertian.

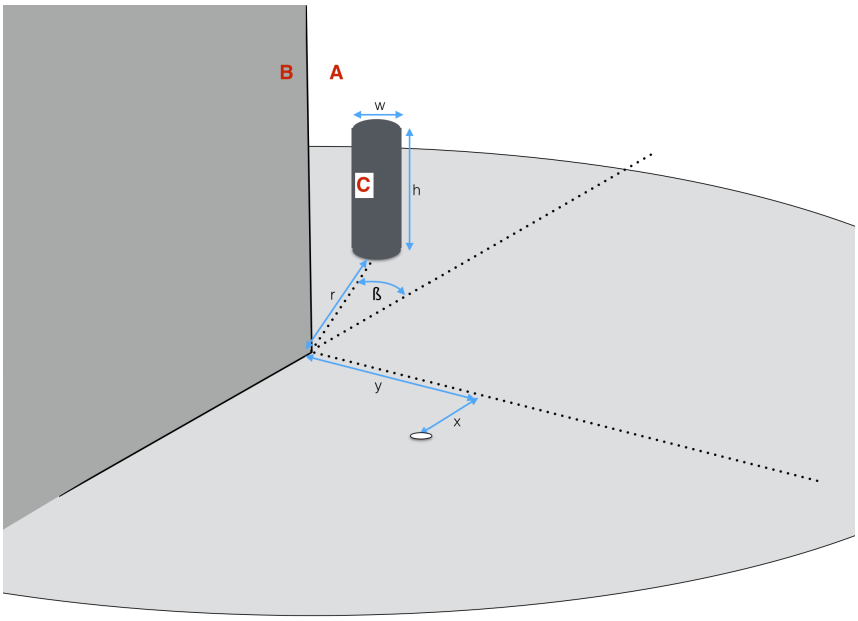

Figure 1: Simple model of a corner camera. We seek formulas for the image intensities at a point (x, y) in the observation plane. We have a wall of brightness B, a subject of brightness C, and a uniform background of brightness A.

While this model is quite simple, it allows us to examine the signal strength–the change of ground brightness due to the presence of the subject–for many cases of interest. We can calculate

- the signal as a function of a person's (the subject's) distance away from the corner, and position relative to the corner;
- the signal as a function of position on the observation plane;
- the approximate effects of a sunlit subject, shaded subject, and cloudy or sunny days.
- the effect of the wall's brightness on the corner camera.

1.1 Photometric signal of the corner camera

We assume a moving subject. We want to calculate the ground plane intensities as a function of position, both with and without the subject present. The difference in the brightnesses observed on the ground plane will then tell us the requirements for our photometric measurements to enable imaging from a corner camera.

1.1.1 Intensity contributions of the wall of brightness B

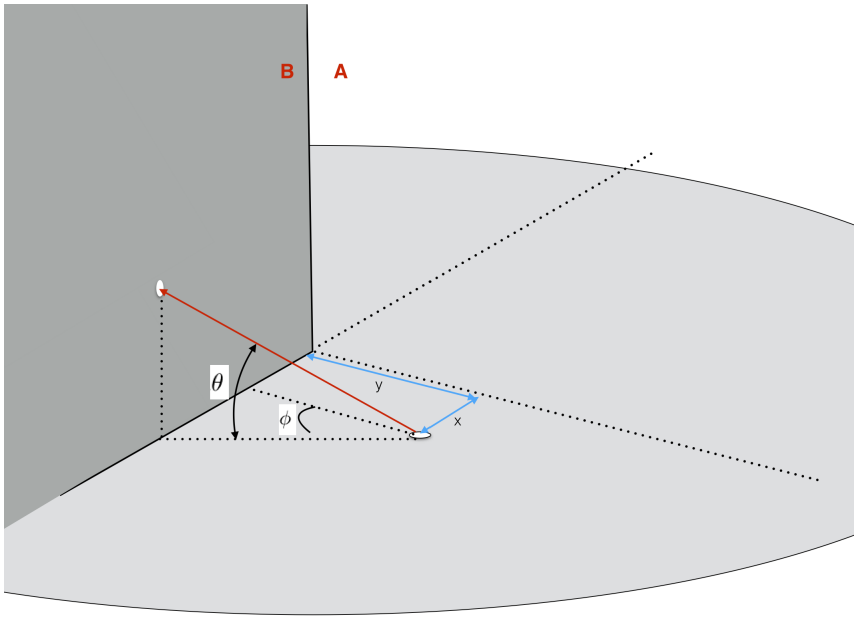

Figure 2: Brightness contribution of one patch of the wall to the brightness of the ground plane depends on the angle between the red arrow and the ground's surface normal.

Consider a small patch on the wall at position x', z'. To calculate its brightness contribution to the brightness on the ground at position (x, y), we need to calculate the normal vector from the ground position to the wall patch, $\hat{N}_{gw}$. The dot product of that normal vector with the ground's normal controls the brightness contribution from the point (x', z').

But if the wall is Lambertian, and the angle at which we view it doesn't change its brightness, B, then we can just as easily, and more simply, perform the integration from the observation plane point (x, y). There we just have to integrate the brightness B over all directions from (x, y) where we see the wall. Let θ be the angle of a ray to the wall with respect to the ground plane normal, where the z-axis has $\theta = \frac{\pi}{2}$. Let ϕ be the azimuthal angle to the point on the wall, where the y-axis direction has $\phi = 0$.

Then the brightness contribution from the entire wall, I_B, is the following double-integral

$$
\begin{aligned}
I_B &= B \int_{-\frac{\pi}{2}}^{\arctan(\frac{x}{y})} \mathrm{d}\phi \int_0^{\frac{\pi}{2}} \sin(\theta)\mathrm{d}\theta && (1)\\
&= B \int_{-\frac{\pi}{2}}^{\arctan(\frac{x}{y})} \mathrm{d}\phi && (2)\\
&= B \int_{-\frac{\pi}{2}}^{\arctan(\frac{x}{y})} \mathrm{d}\phi && (3)\\
&= B \left(\arctan\left(\frac{x}{y}\right) + \frac{\pi}{2} \right) && (4)
\end{aligned}
$$

1.1.2 Intensity contributions of the non-wall environment of brightness A

Similarly, we want to integrate the background brightness A over all angles where we see A and not B from the observation plane patch at (x, y). Calling this brightness I_A, this is the same integral as above, except we are integrating the constant A, not B, and over the complement of the azimuthal angles as before. So the result is

$$I_A = A(2\pi - (\arctan(\frac{x}{y}) + \frac{\pi}{2})) \quad (5)$$

$$= A(\frac{3}{2}\pi - \arctan(\frac{x}{y})) \quad (6)$$

$$(7)$$

1.2 Intensity contributions of the subject of brightness C

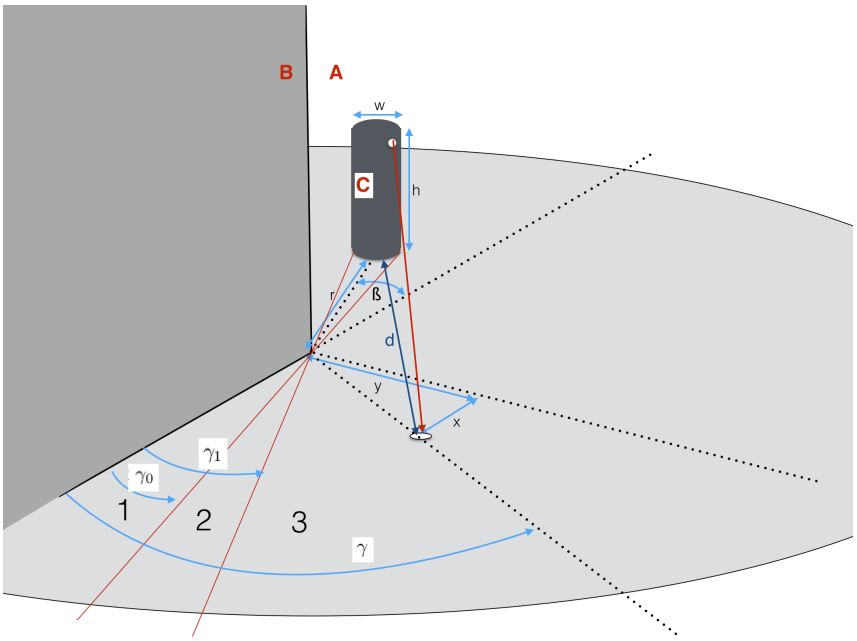

Figure 3: Calculation of the contribution to the brightness of the subject.

The geometry of the edge and the subject divides the observation plane into three regions, with different intensity functions in each region. Let $\gamma = \arctan(\frac{y}{x})$. Let $\gamma_0 = \beta - \frac{w}{r}$ and $\gamma_1 = \beta + \frac{w}{r}$. Points of the observation plane for $\gamma < \gamma_0$ (using the small angle approximation for the angular extent of the cylinder) don't see the cylinder and form region 1 of the ground plane. Points where $\gamma > \gamma_1$ see a full view of the subject cylinder and form region 3 of the ground plane. Points (x, y) for $\gamma_0 < \gamma(x, y) < \gamma_1$ form region 2 of the ground plane. We derive the formulas for the intensities within each region.

1.2.1 I_{C1}, for ground plane region 1

Points on the ground plane in region 1 never see the subject and so their intensities are governed by the function $I_{C1}(x, y) = I(x, y) = I_A(x, y) + I_B(x, y)$.

1.2.2 I_{C3}, for ground plane region 3

Points in region 3 have a full view of the subject, not clipped by the wall edge. The subject has brightness C over its cross-section, and occludes the background brightness A over the same cross-section. So its brightness contribution, I_C, will be the brightness difference $C - A$, integrated over the cross-section of the subject. Then the total brightness at the point (x, y), with the subject in the scene, $I_T(x, y)$, will be

$$I_T(x, y) = I_A + I_B + I_{C3} \tag{8}$$

Then the total brightness at the point (x, y), without the subject present, $I(x, y)$, will be

$$I(x, y) = I_A + I_B \tag{9}$$

Since it is also assumed to be Lambertian, it will be easier to perform this integration from the point of view of the ground plane patch at (x, y). We first want to calculate the distance, d, from the ground plane patch at (x, y) to the subject.

The x-coordinate difference from the subject to the patch, d_x is

$$d_x = r\cos(\beta) + x \tag{10}$$

The y-coordinate difference from the subject to the patch, d_y is

$$d_y = r\sin(\beta) + y \tag{11}$$

Then we have, for the distance d from the subject to the patch,

$$\begin{aligned} d &= \sqrt{d_x^2 + d_y^2} && (12)\\ &= \sqrt{(r\cos(\beta) + x)^2 + (r\sin(\beta) + y)^2} && (13) \end{aligned}$$

Now, we integrate the brightness difference, $C - A$, over the cylinder of width and height w and h at a distance d from the patch at (x, y). We simplify the cylinder to be a w by h rectangle, as seen from the patch at (x, y). If we assume the ground plane is Lambertian, then intensities arriving at the patch are modulated by a factor of $\sin(\theta)$:

$$\begin{aligned} I_{C3} &= (C - A)\int_0^{\arctan(\frac{h}{d})} \sin(\theta)\mathrm{d}\theta \int_k^{k+\frac{w}{d}} \mathrm{d}\phi && (14)\\ &= (C - A)\frac{w}{d}(1 - \cos(\arctan(\frac{h}{d}))) && (15)\\ &= (C - A)\frac{w}{d}(1 - \frac{d}{\sqrt{d^2 + h^2}}) && (16) \end{aligned}$$

where, for simplicity in the integration limit for ϕ, we assume that $w \ll d$, and k an offset in ϕ that doesn't matter in the result. And we used $\cos(\arctan(x)) = \frac{1}{\sqrt{1+x^2}}$.

1.2.3 I_{C2}, for ground plane region 2

Points in region 2 only see a partial image of the subject. The integral over ϕ in Eq. (16) has a different bottom limit depending on the point's angular position within region 2.

The partial view of the subject will lead to this simple behavior of the intensities in region 2: at a constant radius away from the edge (and therefore a constant

distance d from the subject), the intensities in region 2 will lineaerly interpolate between the intensity values at the boundaries with regions 1 and 3.

For a given (x,y) in region 2:

1. Find $\gamma(x, y)$.
2. Find $\rho(x, y) = \sqrt{x^2 + y^2}$
3. Use $I_{C2}(\gamma, \rho) = \alpha I_{C2}(\gamma_0, \rho) + (1 - \alpha) I_{C2}(\gamma_1, \rho)$, where
4. $\alpha = \frac{\gamma - \gamma_0}{\gamma_1 - \gamma_0}$

The linear behavior of the limits of integration over ϕ in Eq. (16) allow this linear interpolation to hold.

1.3 Final Model

Brightness at (x,y) with subject present in scene, $I_T(x, y)$

$$\begin{aligned} I_T &= I_A + I_B + I_C && (17)\\ &= A(\frac{3}{2}\pi - \arctan(\frac{x}{y})) + && (18)\\ &\quad B(\arctan(\frac{x}{y}) + \frac{\pi}{2}) + && (19)\\ &\quad I_C(x, y) && (20) \end{aligned}$$

Brightness at (x,y) with no subject in scene, $I(x, y)$

$$\begin{aligned} I &= I_A + I_B && (21)\\ &= A(\frac{3}{2}\pi - \arctan(\frac{x}{y})) + && (22)\\ &\quad B(\arctan(\frac{x}{y}) + \frac{\pi}{2}) && (23) \end{aligned}$$

Fractional difference in brightness at (x,y) due to subject, $\frac{I_T(x,y)}{I_T(x,y)}$

$$\begin{aligned} \frac{I_T(x, y)}{I(x, y)} &= \frac{I_A + I_B + I_C}{I_A + I_B} && (24)\\ & && (25) \end{aligned}$$

1.4 Examples

1.4.1 Typical brightness change induced by a person 10 feet from corner

Let's examine values we consider to be typical for an outdoor, cloudy day scene. Let's say the surroundings have brightness $A = 300$, the wall has brightness $B = 100$, and the person has brightness $C = 150$ (arbitrary units, but roughly mapping to pixel intensities from a photograph). The person is $h = 5.5$ feet tall, and $w = 1.5$ feet wide, positioned at angle $\beta = \frac{\pi}{4}$, and at distance from the corner $r = 10$ feet.

Let's gaze upon the observation plane a distance $\rho = \sqrt{x^2 + y^2}$ from the corner. At the boundary between ground plane regions 1 and 2, we have, for the ground

plane image intensity,

$$\begin{aligned} I_{C1}(\rho,\gamma_0) &= I_A(x,y)+I_B(x,y) && (26)\\ &= A(\frac{3}{2}\pi-\arctan(\frac{x}{y}))+B(\arctan(\frac{x}{y})+\frac{\pi}{2}) && (27)\\ &= 300(\frac{3}{2}\pi-(\pi-\gamma_0))+100((\pi-\gamma_0)+\frac{\pi}{2}) && (28)\\ &= 300(\frac{\pi}{2}+\gamma_0)+100(\frac{3\pi}{2}-\gamma_0) && (29)\\ & && (30) \end{aligned}$$

Note this intensity doesn't depend on the distance away from the wall, at a constant angle $\gamma=\gamma_0$. The wall subtends a constant solid angle for all ground plane points along this straightline from the corner. Note that $\arctan(\frac{x}{y})=\pi-\gamma_0$.

Now let's examine the brightness at the region 2/region 3 boundary, at $\gamma=\gamma_1$. There we have

$$\begin{aligned} I_T &= I_A(x,y)+I_B(x,y)+I_{C3}(x,y) && (31)\\ &= I_{C1}(\rho,\gamma_0)+I_{C3}(x,y) && (32)\\ &= 300(\frac{\pi}{2}+\gamma_0)+100(\frac{3\pi}{2}-\gamma_0)+ && (33)\\ &= (C-A)\frac{w}{d}(1-\frac{d}{\sqrt{d^2+h^2}}) && (34) \end{aligned}$$

Remembering that $\gamma_0=\frac{\pi}{4}-\frac{w}{r}$ and $w=1.5$, $h=5.5$, $r=10$, $C=150$, $B=100$, $A=300$, we have

$$\begin{aligned} I_T &= 300(\frac{\pi}{2}+\frac{\pi}{4}-\frac{1.5}{10})+ && (35)\\ & \quad 100(\frac{3\pi}{2}-(\frac{\pi}{4}-\frac{1.5}{10}))+ && (36)\\ & \quad (-150)\frac{1.5}{d}(1-\frac{d}{\sqrt{d^2+5.5^2}}) && (37) \end{aligned}$$

Remember d is the distance from (x,y) to the cylinder, or $10+\sqrt{x^2+y^2}$. Let's look at $d=12$. For this distance, we have

$$\begin{aligned} I_T &= 300(2.356-0.15)+ && (38)\\ & \quad 100(4.7124-0.6354)+ && (39)\\ & \quad (-150)\frac{1.5}{12}(1-\frac{12}{\sqrt{12^2+5.5^2}}) && (40)\\ &= 662+ && (41)\\ & \quad 408+ && (42)\\ & \quad -1.7 && (43) \end{aligned}$$

That's in the right ballpark: The signal without the subject has brightness 1070. The signal with the subject has brightness 1068.3 requiring a measurement dynamic range of $\frac{1.7}{1070}$. Thus, we seek a signal change that is 0.16% of the DC signal. This is consistent (within a factor of two) with what we find to be typical for real images, $\frac{1}{1000}$,

Looking at the subject's signal as a function of ground plane distance away from the edge, ρ, we have $I_{C3}(\rho)=\frac{412.5}{(10+\rho)\sqrt{(10+\rho)^2+5.5^2}}$. For $\rho=0$, we have $I_{C3}(\rho)=3.6$ (a higher subject signal). For $\rho=4$, we have $I_{C3}(\rho)=1.96$ (a lower subject signal).

1.5 Rate of decay of the signal as the person moves away from the corner: $\frac{1}{d^4}$

As the person walks away from the corner, we lose signal strength for two independent, reinforcing reasons. One, from perspective projection, they subtend a smaller solid angle, with linear dimensions going as $\frac{1}{d}$ and areal dimensions contributing to brightness going down as $\frac{1}{d^2}$, where d is the distance from the corner. Second, as the person moves away from the corner, the light from the person comes in at a shallower and shallower angle, further reducing the signal intensity. The grazing angle to the horizontal will scale as $\frac{1}{d}$, but the brightness contribution will be $1 - \cos(\theta)$, which scales as $\frac{1}{d^2}$. Thus, the delta in brightness, from a person walking away from the corner, scales as $\frac{1}{d^4}$.

This is also apparent from applying a 2nd order Taylor series expansion to the square roots with d in them in Eq. (37).

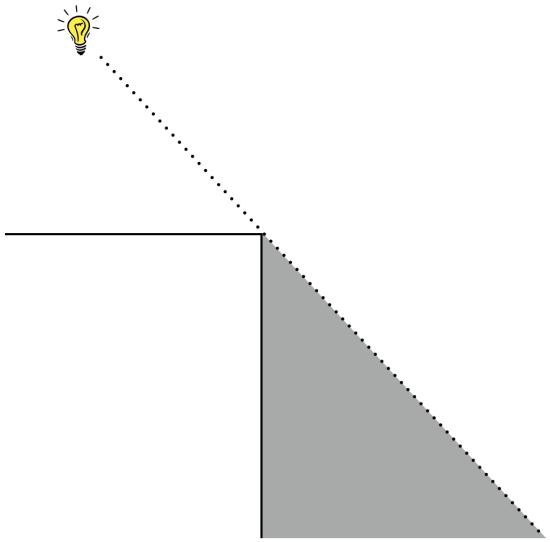

Figure 4: This figure shows the setup for the toy problem of interest. The scene consists of a single bright object, whose angular position θ we want to learn.

2 Corner Location Errors

One important source of error in the edge camera idea is the *corner location error*. When studying a movie of the observation plane, it's important to know where the corner of the wall is in order to make an accurate reconstruction. Corner location errors occur when the corner of the wall is erroneously chosen to be the wrong place. They introduce systematic error into the scene's reconstruction.

2.1 Edge Camera

Exactly how bad are corner location errors? To answer this question, we consider the situation shown in Fig. 4. Imagine a dark scene with a single bright object. We want to find the angular position of the bright object in the scene. We can do this by measuring θ: the angle of the shadow it casts against the wall. When we find the angle at which the observation plane goes from light to dark, we will know what θ is.

This story is simple in the case when there is no corner location error. But what about the case where there is such an error?

Fig. 5 shows this scenario. We can "sweep" the angle ϕ across the observation plane, and at the point where ϕ is midway between dark and light, we can presume that that will be the most likely inferred angular position of the object from the reconstructed space-time image. When there is no corner location error, we will naturally get $\theta = \phi$, but when there is a corner location error, ϕ will not equal θ, but will depend on θ and other parameters.

Fig. 6 plots intensity against the sweeping angle ϕ, both with and without a corner location error. Note that in the case where there is a corner location error, the maximum intensity value no longer takes on a maximum value of 1, but a value below 1. In the analysis that follows, we will call that value $l_{\max}$, and we will use $_{\max}/2$ as the "transition point" between light and dark. In other words, we will choose the ϕ that gives an intensity of $_{\max}/2$ as our estimate for θ.

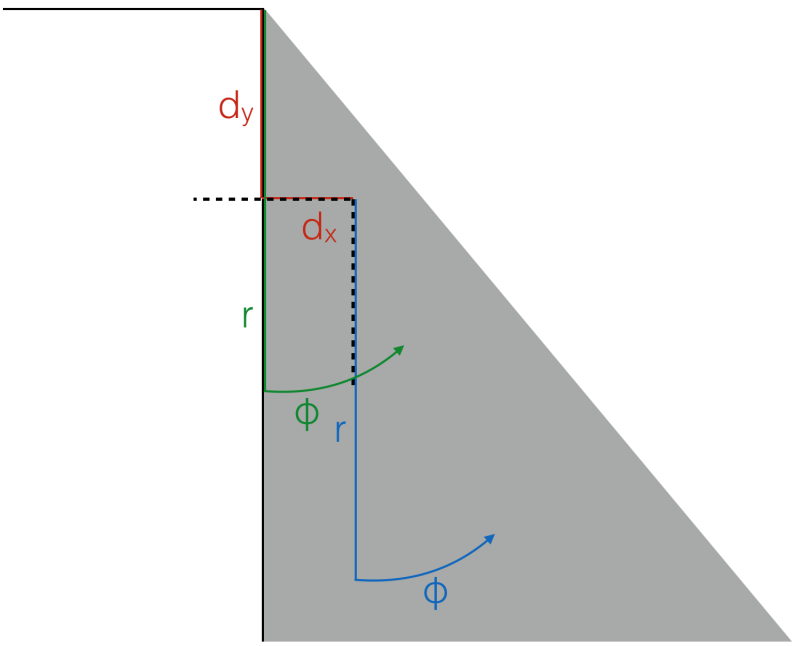

Figure 5: This figure shows the impact of a corner location error. In the error-free case, we would sweep ϕ across the observation plane (shown in green) hinging around the corner (the solid black line). But if made a corner location error, we would instead try to sweep ϕ across the observation plane erroneously (shown in blue) hinging around the false corner (shown with a dotted line).

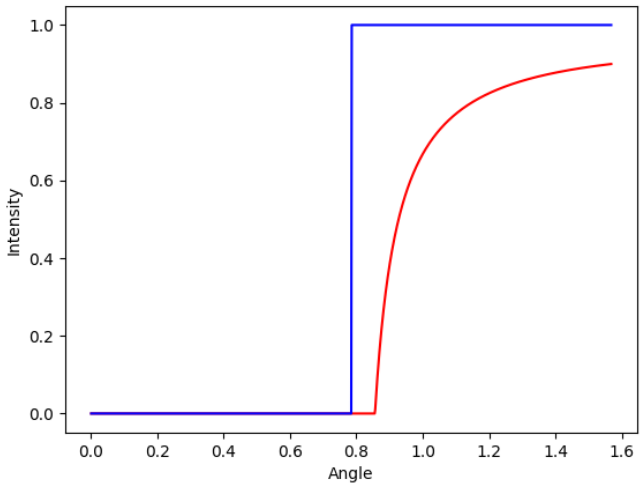

Figure 6: This plot shows how the observed intensity values vary with ϕ, in the case of correct corner location (in blue) and a corner location error ($(d_x, d_y) = (0.1, 0.2)$, in red). Note that in the case of a corner location error, the maximum value of the intensity does not reach 1. Note also that d_x and d_y are as a fraction of the radius of the observation plane, r, which here is taken to be 1.

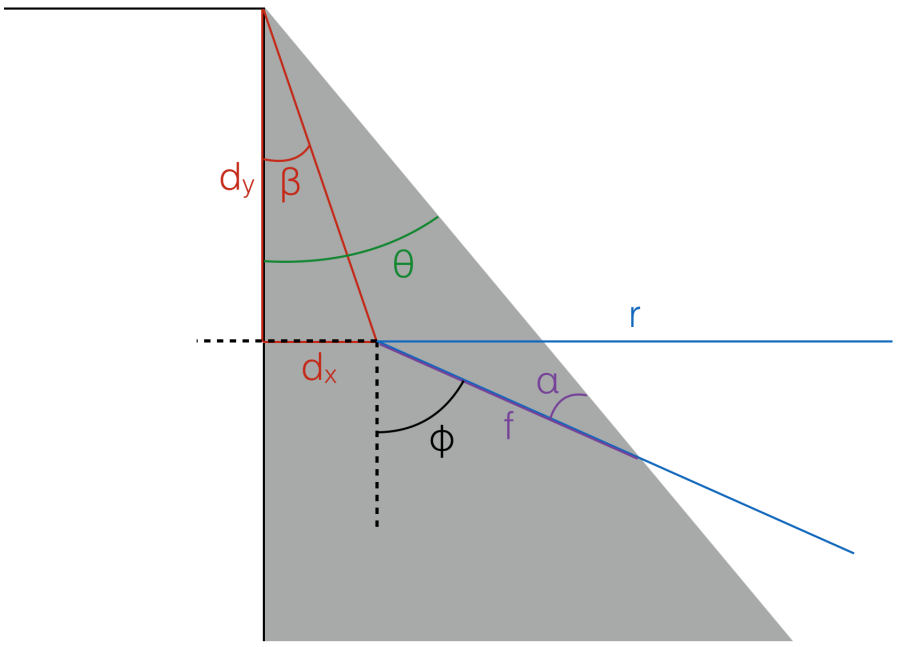

Figure 7: This plot is intended as a reference for the meanings of each of the variables used in the calculations of ϕ as a function of θ.

Fig. 7 is a detailed illustration of the situation, showing the names for the variables that we'll use in our analysis. As the figure shows, we are presuming a corner location error of (d_x, d_y) and a observation plane radius of r. We want to find what our estimate θ, ϕ, will be as a function of θ and in terms of d_x, d_y, and r.

Using Fig. 7 as a reference, we can make the following observations:

$$\beta = \tan^{-1}\left(\frac{d_x}{d_y}\right)$$

$$l_{\max} = r - d_y \tan(\theta) + d_x$$

$$f = r - \frac{l_{\max}}{2}$$

$$\alpha = \sin^{-1}\left(\frac{\sqrt{d_x^2 + d_y^2}\sin(\theta - \beta)}{f}\right)$$

$$\gamma = \pi - \alpha - \theta + \beta$$

$$\phi = \pi - \gamma + \beta$$

This is how ϕ is expressed in terms of the parameters of the problem (θ, d_x, d_y, r).

What sort of error does this introduce? In order to study this question, we assumed that d_x and d_y were normally distributed with means of 0 and small (relative to r^2) variances σ_x^2 and σ_y^2. We generated many sample (θ, ϕ) pairs for each θ between 0 and $\pi/2$. We then measured the empirical means and variances of these pairs. Fig 8 shows a few of our results.

Here were a few of our empirical findings:

1. The mean error was always 0 for all values of θ, σ_x and σ_y.

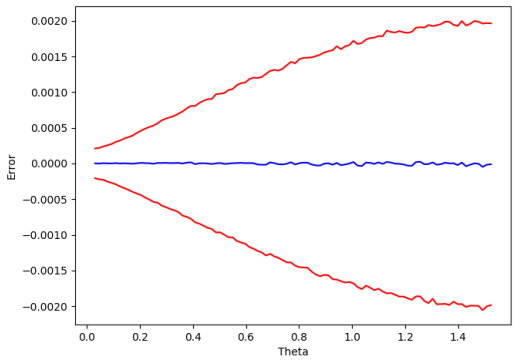

Figure 8: This plot shows the empirical mean (in blue) plus or minus one standard deviation (in red) of the error as a function of θ. Here, $\sigma_x = 10^{-4}$ and $\sigma_y = 10^{-3}$.

2. When $\sigma_x = \sigma_y$, the standard deviation of the error σ_ϵ was $2\sigma_x$ for all values of θ.

3. When $\sigma_x \neq \sigma_y$, the standard deviation of the error σ_ϵ varied between $2\sigma_x$ (for $\theta = 0$) and $2\sigma_y$ (for $\theta = \pi/2$).

2.2 Stereo Camera

Another situation in which it makes sense to study corner location errors is the case where there is a doorway just before the hidden scene, in which case we can use stereo vision to locate a moving object in two dimensions. We would like to know: what effect do corner location errors have on depth estimates, which are generally quite sensitive to noise? To be more precise, suppose that we call the axis along which the doorway lies the "x-axis," and suppose we call the perpendicular axis (of depth into the room) the "z-axis." Then, how much noise in the z dimension will a corner location error cause?

To give an approximate sense of how much error results in the recovered z position, we show the mean +/- one standard deviation in the z dimension as a function of the true x-position of the object in Fig. 9

Note that the empirical means are centered at the true depths of the objects. This does *not* mean that any single corner location error won't cause the depth of the reconstruction to be off systematically; it only means that on average, corner location errors that are normally distributed around the corners in question will push the reconstructed depths away as much as they pull them closer.

To see this systematic bias on its own, we can also study how a single corner error introduces systematic error in our reconstructions—after all, for a single experiment, we are likely to make a single corner error, and the resulting error in the depth calculations will extend across many x-coordinates as the subject of the experiment walks back and forth in the hidden scene. Figs. 10 and 11 show the systematic bias for two distinct *specific* corner location errors.

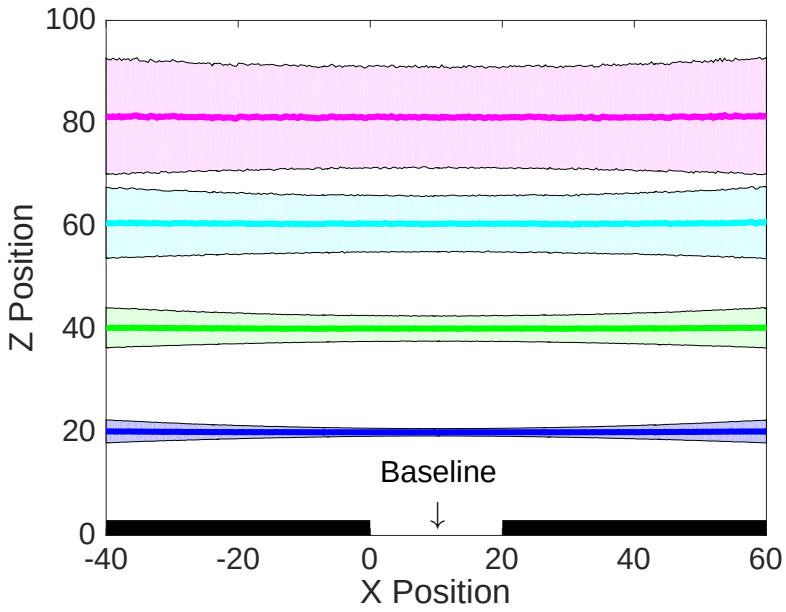

Figure 9: The empirical means plus or minus one standard deviation of the estimated P_z as a function of its x-coordinate, assuming true P_z of 20, 40, 60, and 80. Here, the two corner location errors at each of the boundaries of the doorway are independent and subject to $\sigma^2_{\Delta x} = \sigma^2_{\Delta z} = 0.04$. We sample from a set of 1000 corner errors to approximate the mean and standard deviations empirically.

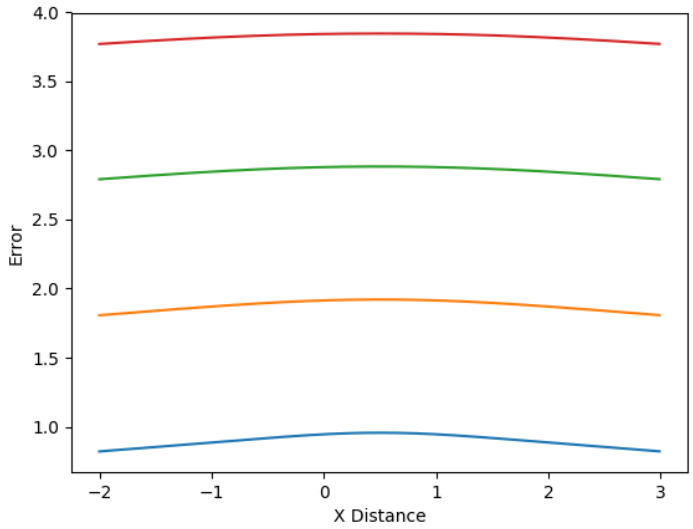

Figure 10: The reconstructed depths of objects at depths 1, 2, 3, and 4, given a corner error of $\Delta y_1 = \Delta y_2 = 0.02$.

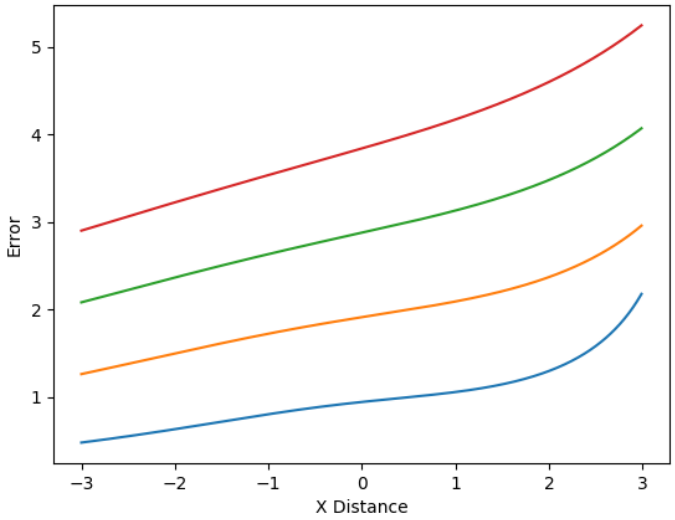

Figure 11: The reconstructed depths of objects at depths 1, 2, 3, and 4, given a corner error of $\Delta y_1 = -0.02$, $\Delta y_2 = 0.02$. Note that because of the different corner errors for each corner, there is the possibility of asymmetric behavior on either side of the doorway.

3 Temporal Smoothing

In addition to spatial smoothness we can also reduce noise temporally by imposing smoothness on our MAP estimate. $\hat{\mathbf{x}}^{(t)}$, or averaging adjacent frames in time. This helps to reduce noise, at the cost of some temporal blurring. However, to emphasis the coherence among results, we did not previously impose this additional constraint. Each 1-D image, $\mathbf{x}$, that we showed in the main paper was independently computed. In Figure 12 we show some results obtained with temporal smoothing. We either computed the 1-D video on averaged video frames, or applied Kalman Filtering/Smoothing to temporally smooth the result. Note that the Kalman Filtering/Smoothing causes the noise to blur out into vertical lines. Frame averaging results appear noiser only because there are less 1D reconstructed frames being shown.

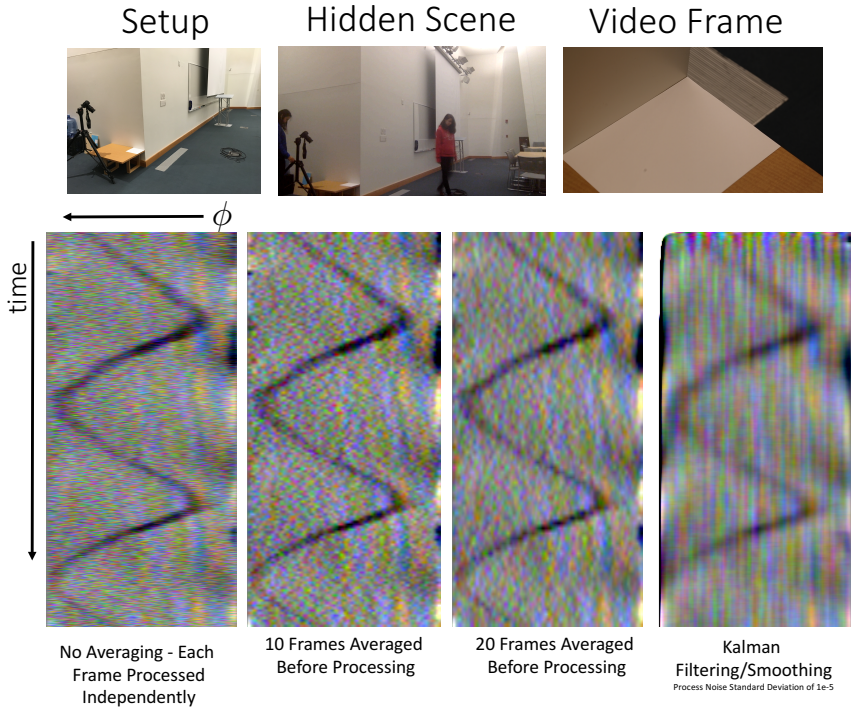

Figure 12: The result of imposing temporal smoothness or averaging adjacent frames in time to help in reducing noise.

3.0.1 Kalman Filtering/Smoothing

We model our scene as a linear dynamic system.

$$x_t = Fx_{t-1} + w_k \tag{44}$$
$$y_t = Ax_t + n_t \tag{45}$$
$$w_k \sim \mathcal{N}(0, R) \tag{46}$$
$$n_t \sim \mathcal{N}(0, \lambda \mathbf{I}) \tag{47}$$

The matrix F can be set as the identity matrix so changes over time are penalized. The marginals $p(x_t|y_1, ..., y_n)$ of this HMM can then be solved using the forward backward algorithm - or equivalently Kalman Filtering and Smoothing.

4 Inferred Position from Stereo Edge Cameras

In the submitted paper, the tracks in the lower left of Figure 9 were cut off due to space constraints. Here we show a larger version of the diagram, with more complete tracks.

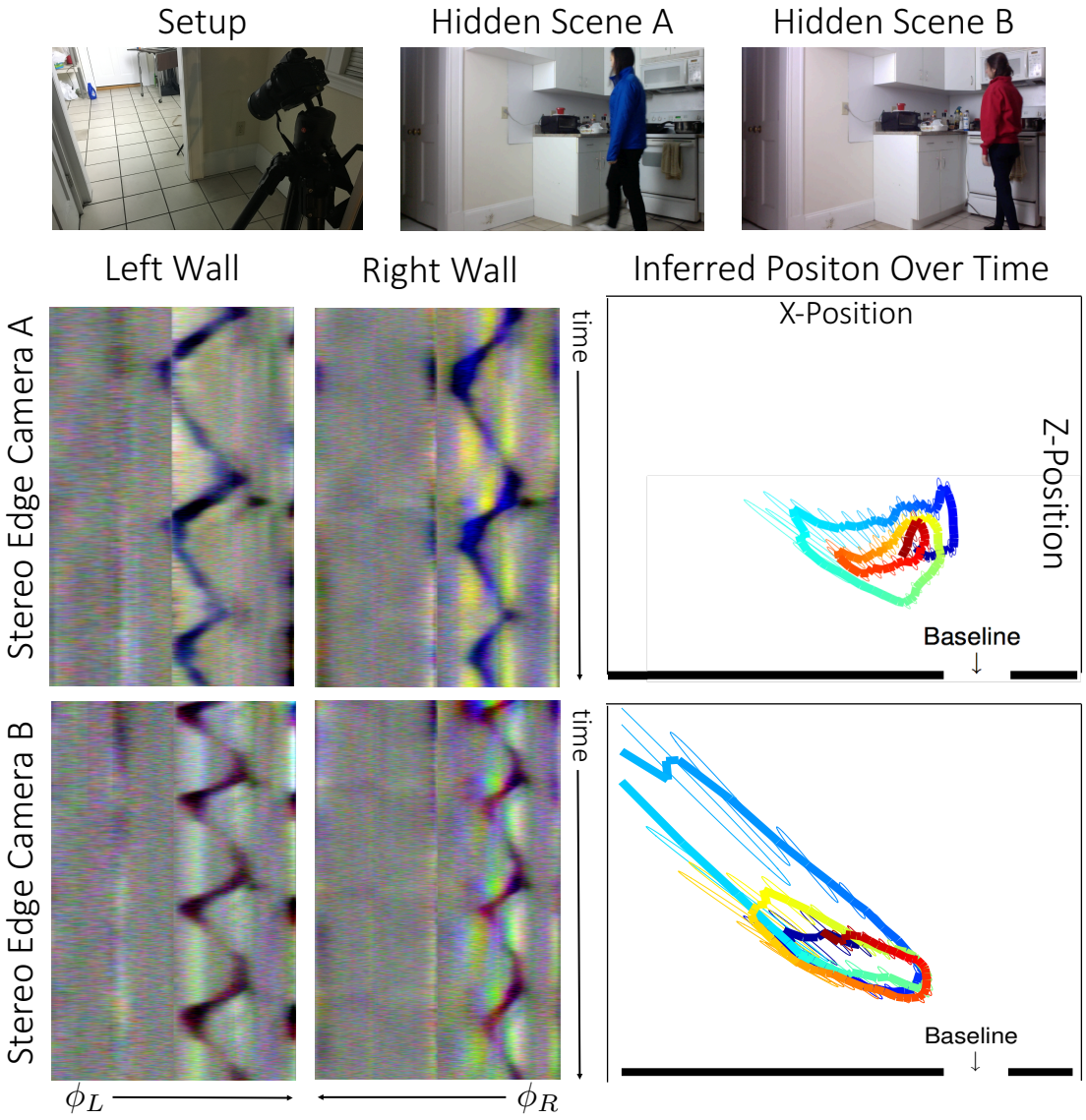

Figure 13: The results of our stereo experiments in a natural setting. Each sequence consists of a single person walking in a roughly circular pattern behind a doorway. The 2-D inferred locations over time are shown as a line from blue to red. Error bars indicating one standard deviation of error have 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.