Lecture 2

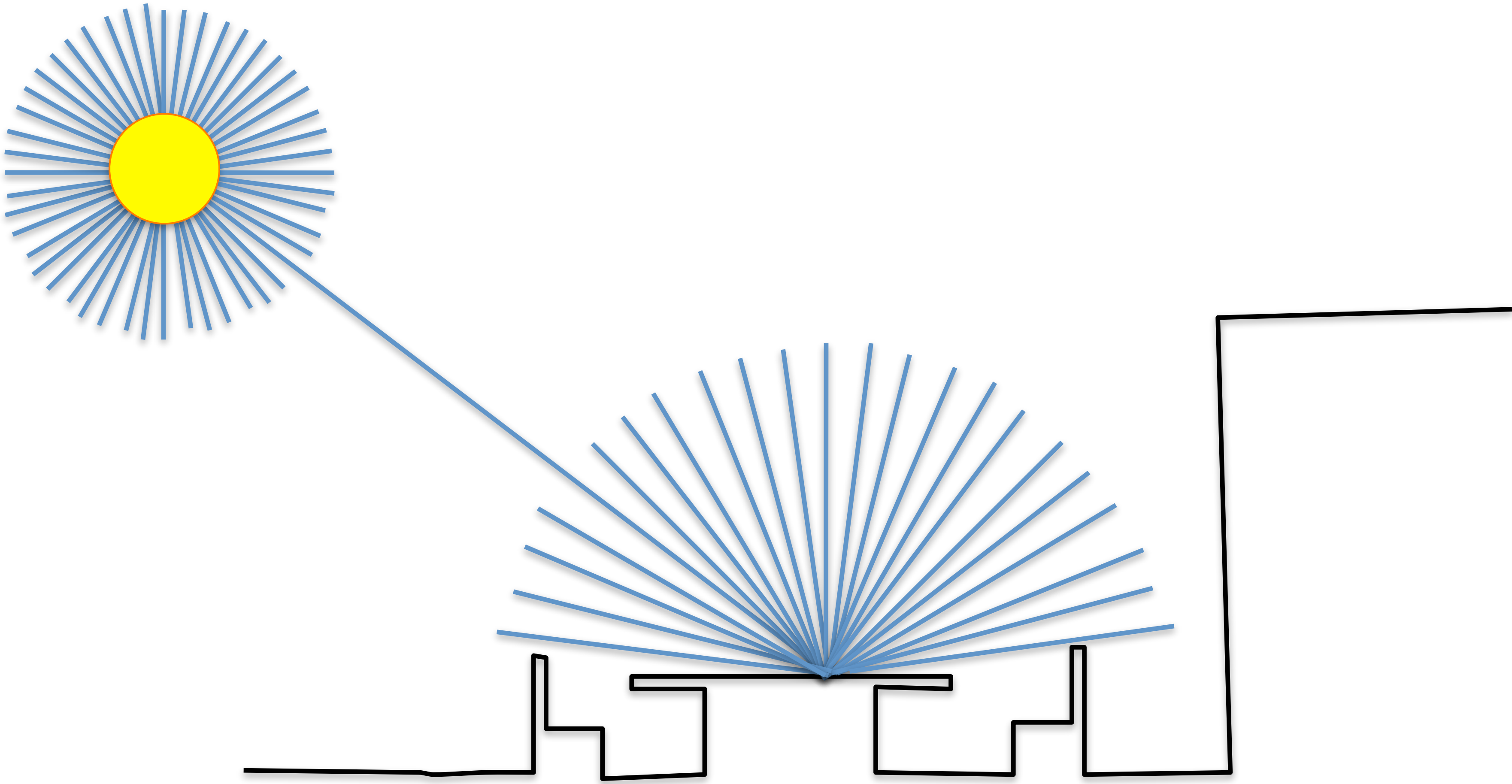
Image formation

6.869/6.819 Advances in Computer Vision

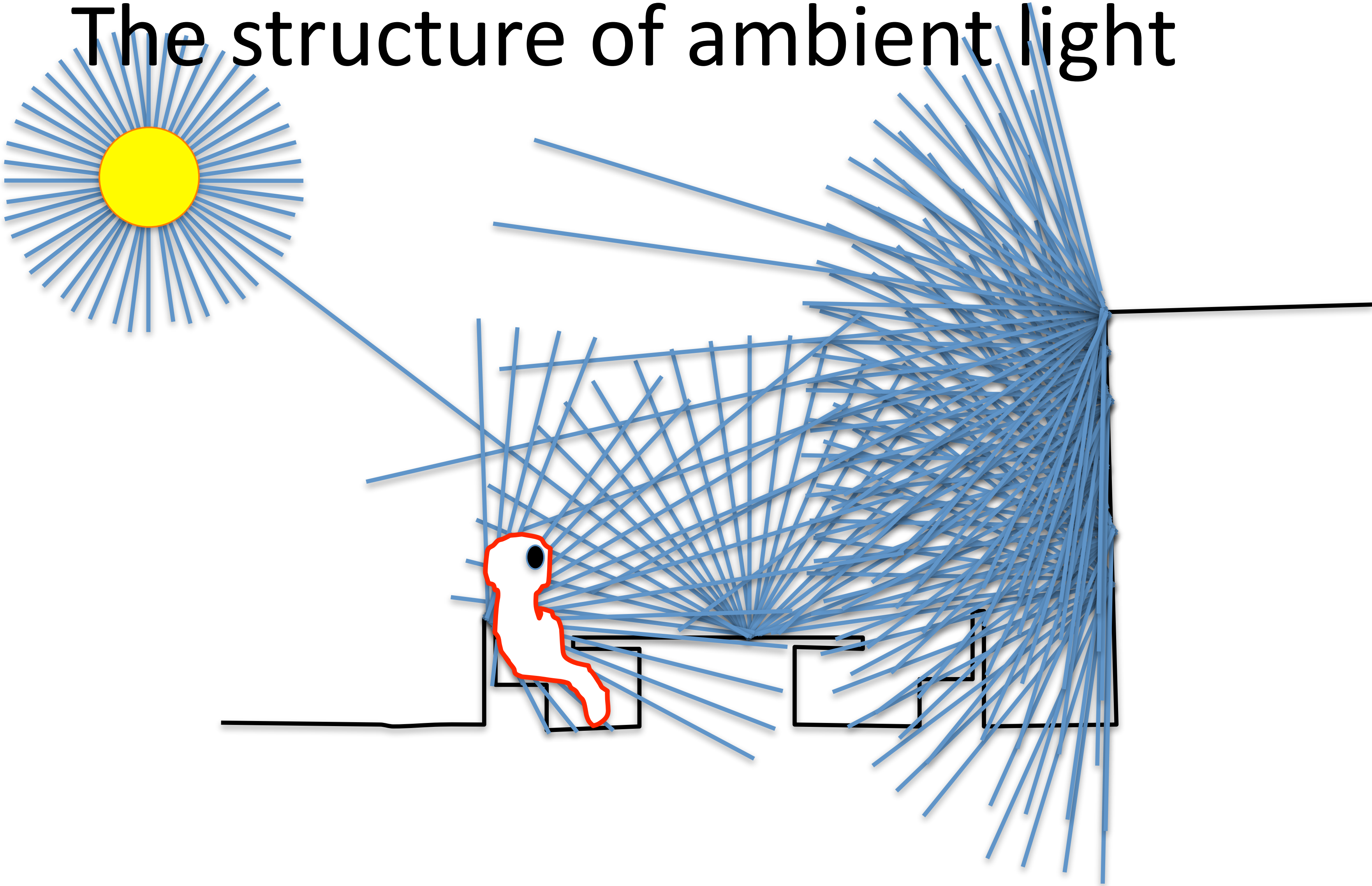
Spring 2022
Bill Freeman and Phillip Isola

Imaging lecture

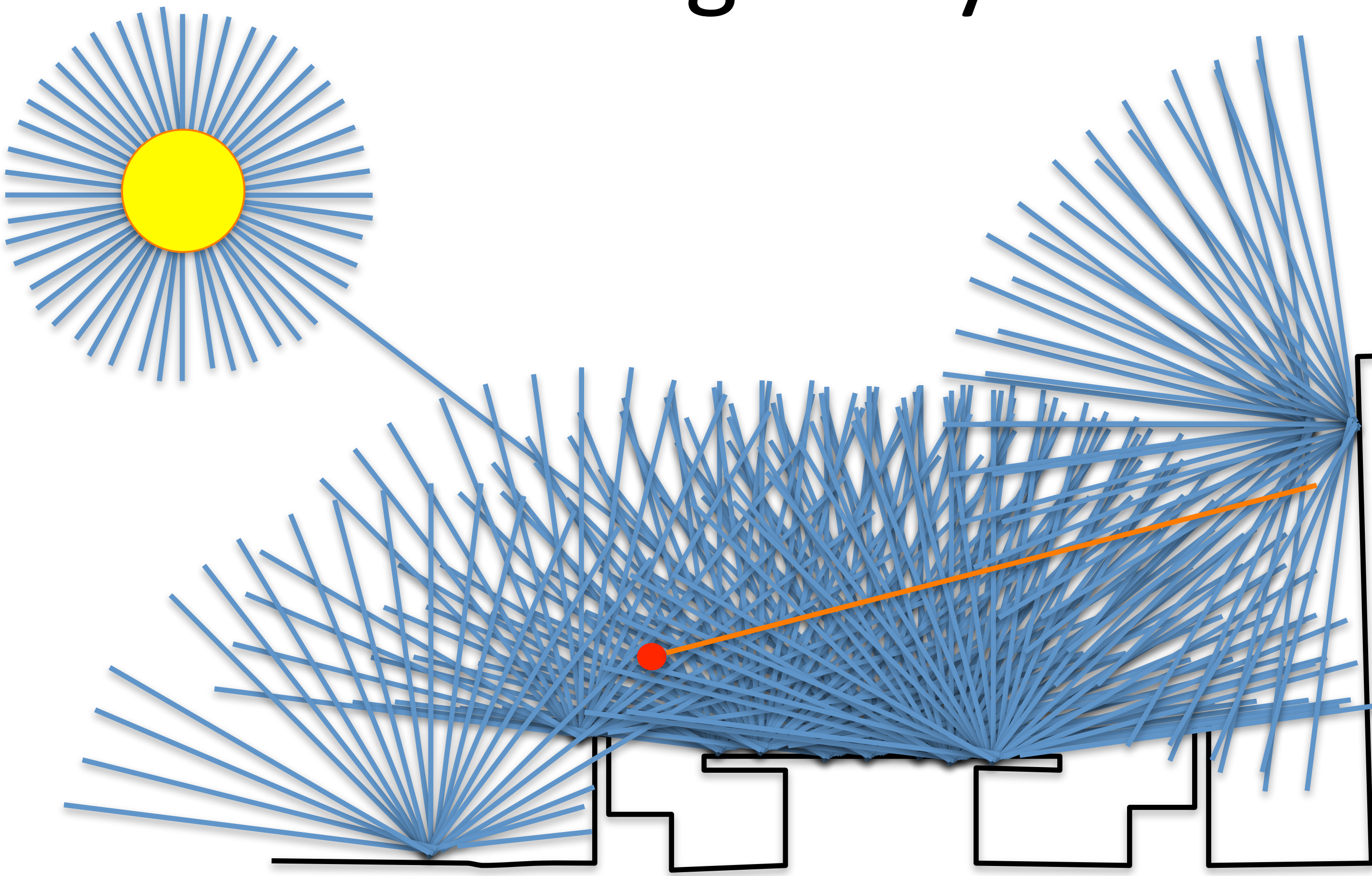
The structure of ambient light



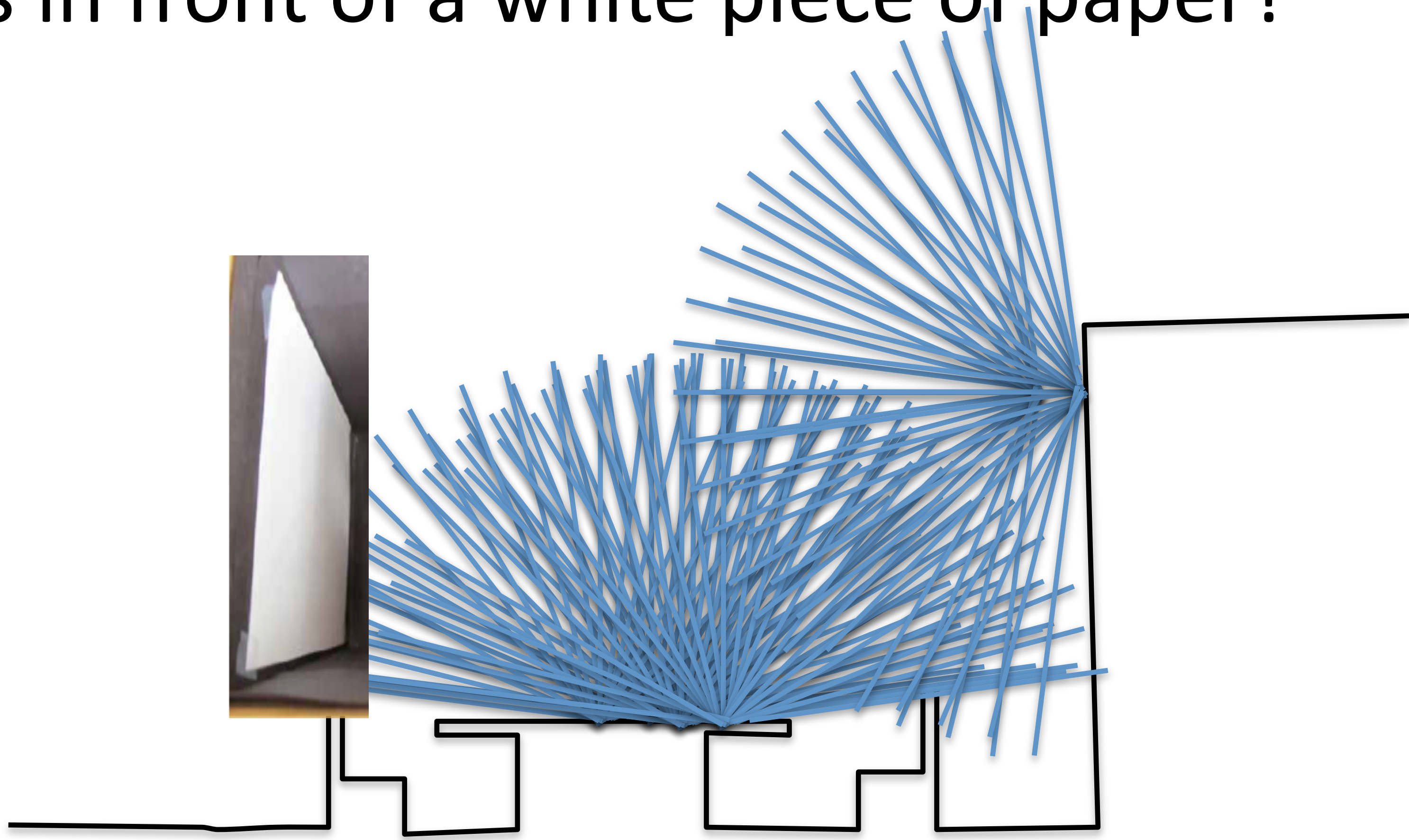
The structure of ambient light



All light rays

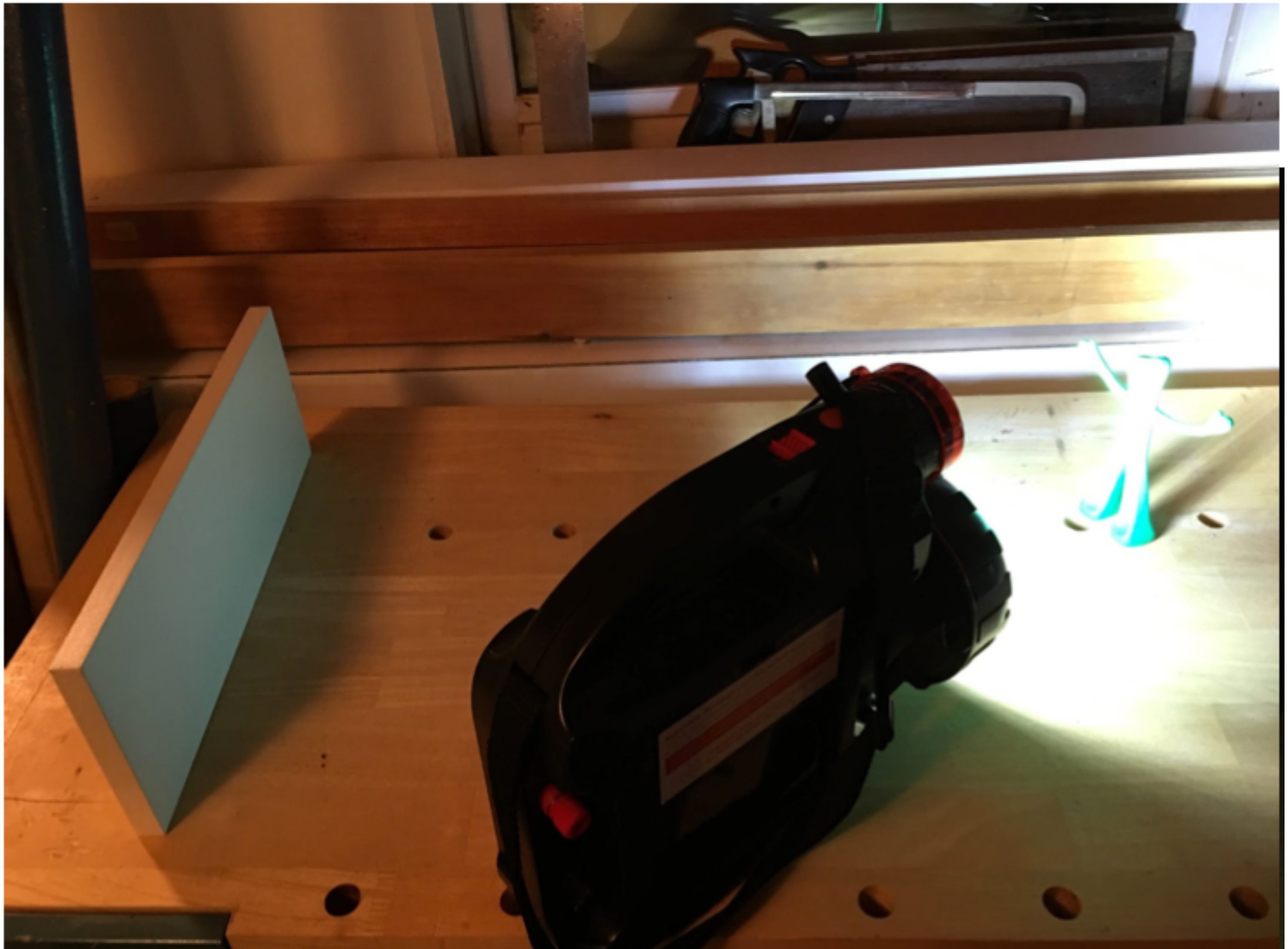


Why don't we generate an image when an object is in front of a white piece of paper?

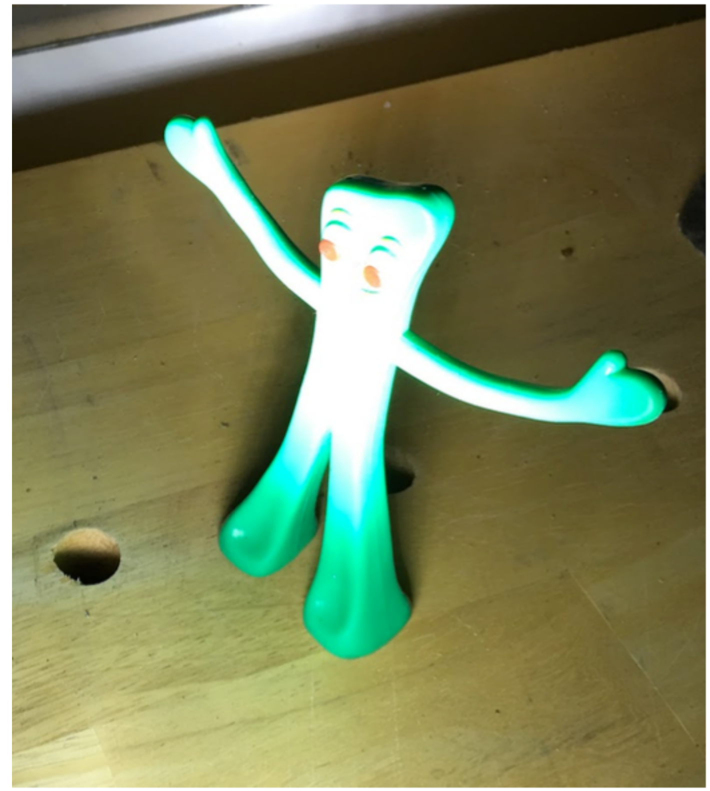


Why is there no picture appearing on the paper?

Let's check, do we get an image?



Let’s check, do we get an image? No



To make an image, we need to have only a subset of all the rays strike the sensor or surface

The camera obscura
The pinhole camera

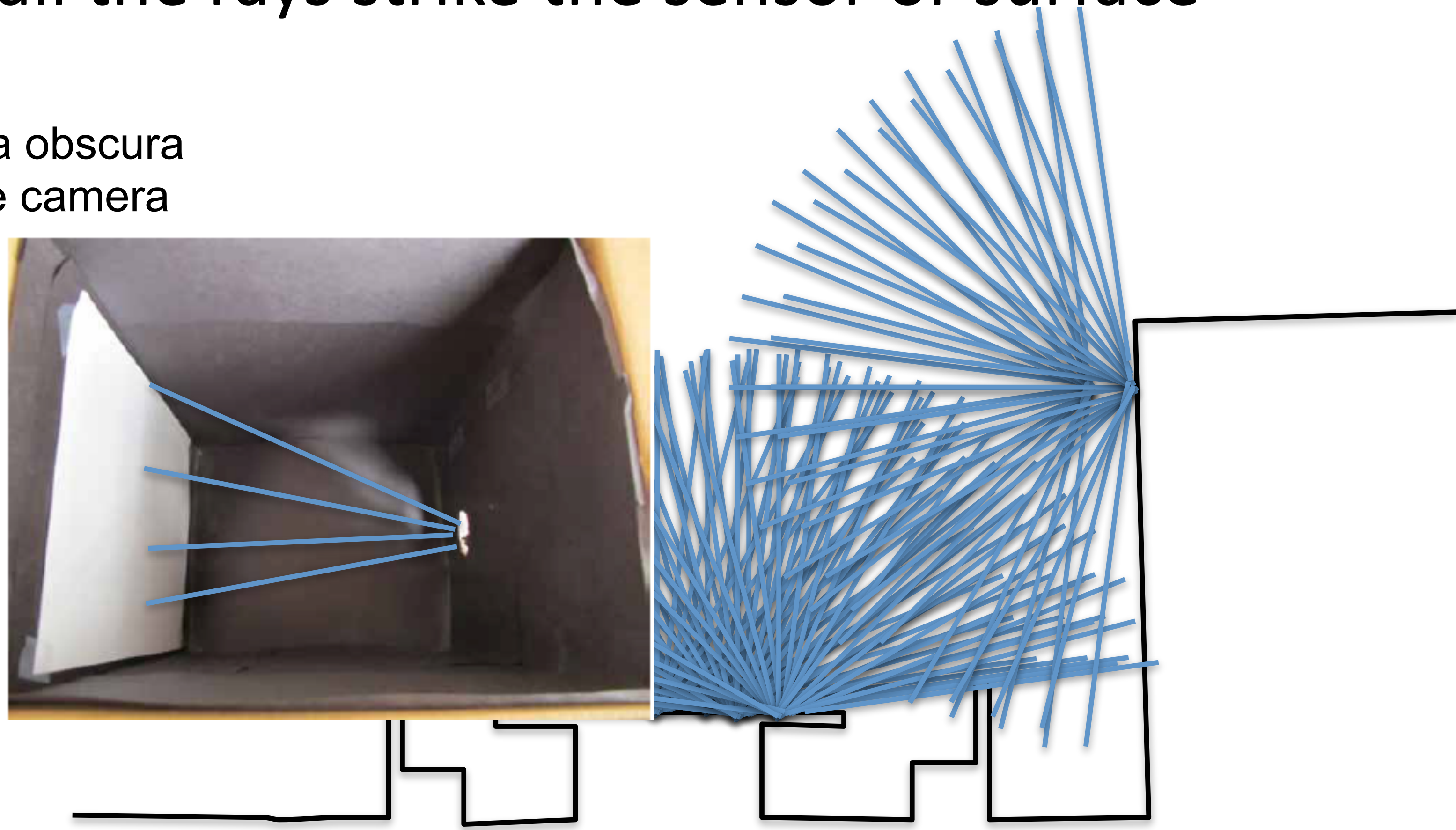
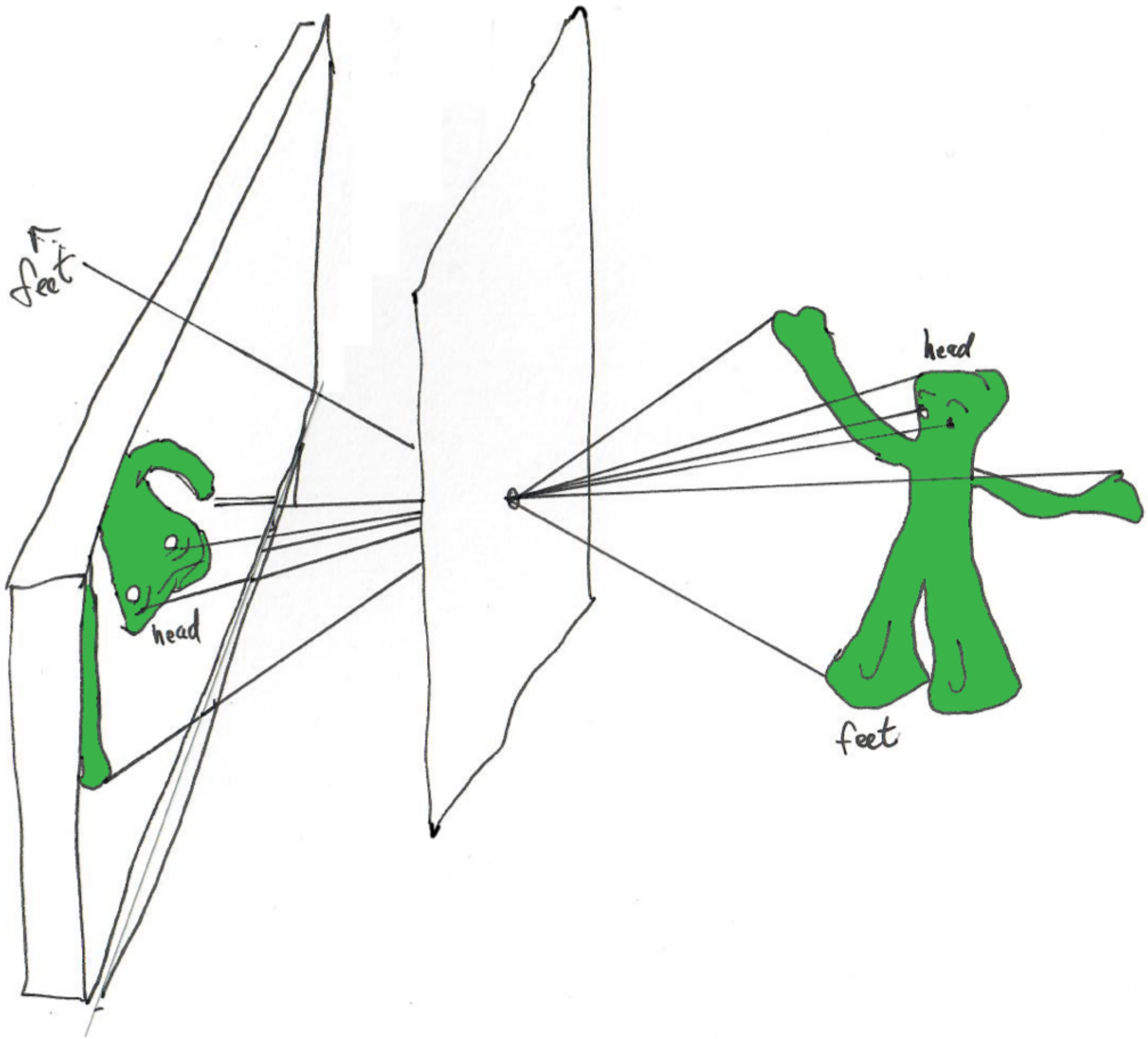
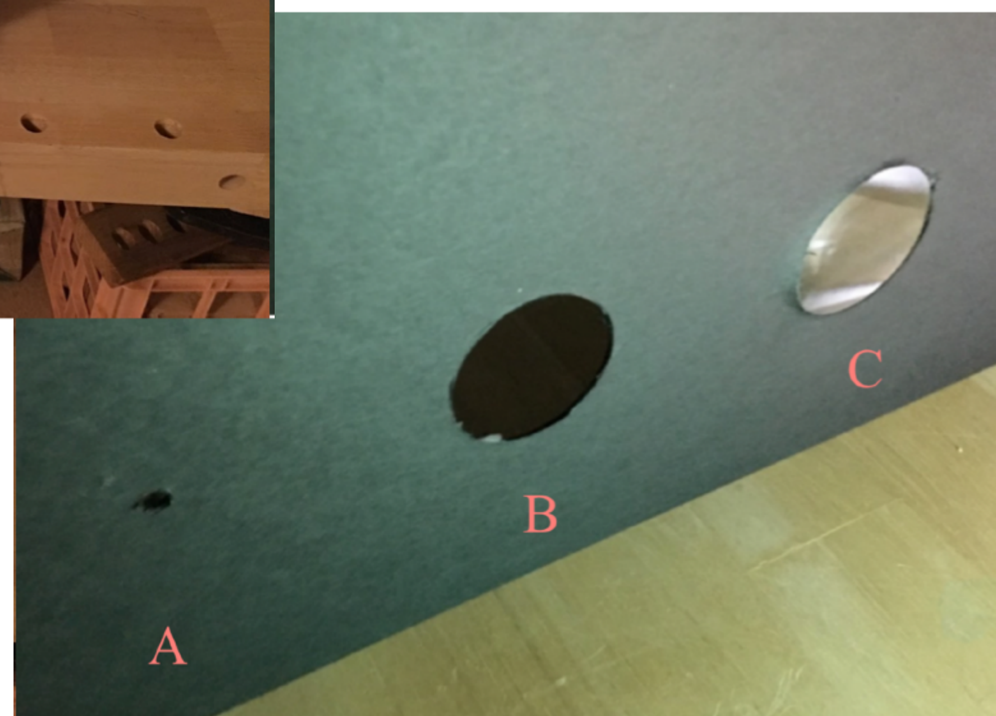


image is inverted



Let's try putting different occluders in between the object and the sensing plane



light on wall past pinhole

grocery bag pinhole camera

grocery bag pinhole camera

grocery bag pinhole camera

view from outside the bag

http://www.youtube.com/watch?v=FZyCFxsyx8o

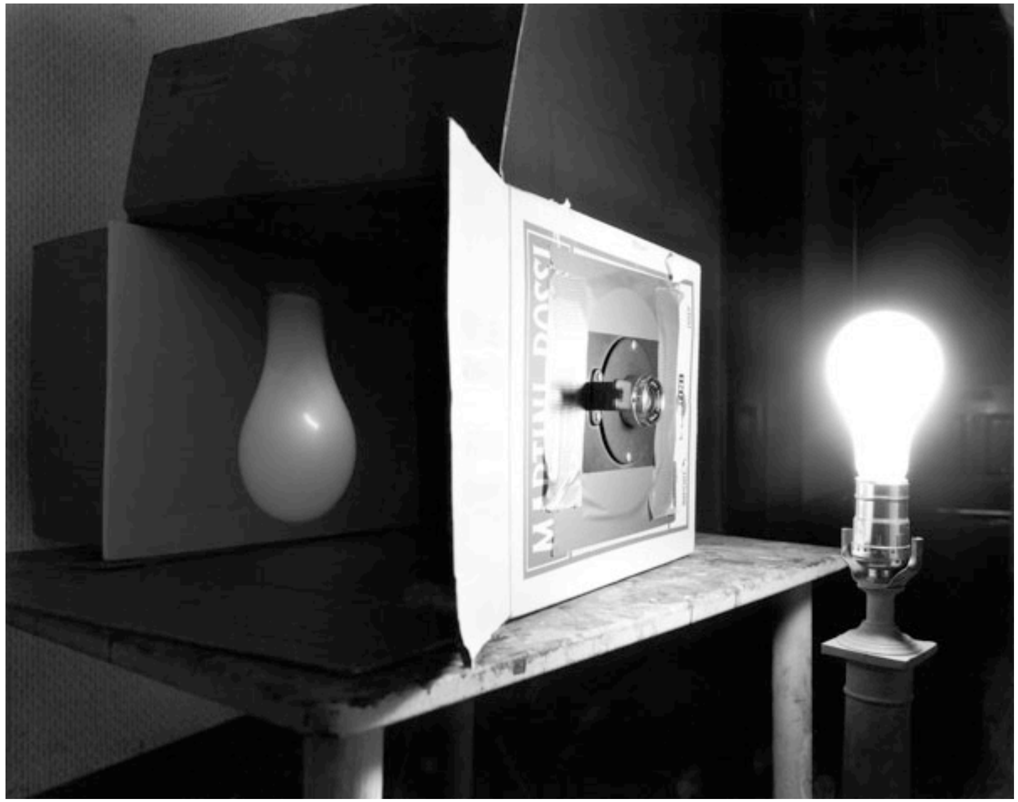
view from inside the bag

http://youtu.be/-rhZaAM3F44

me, with GoPro

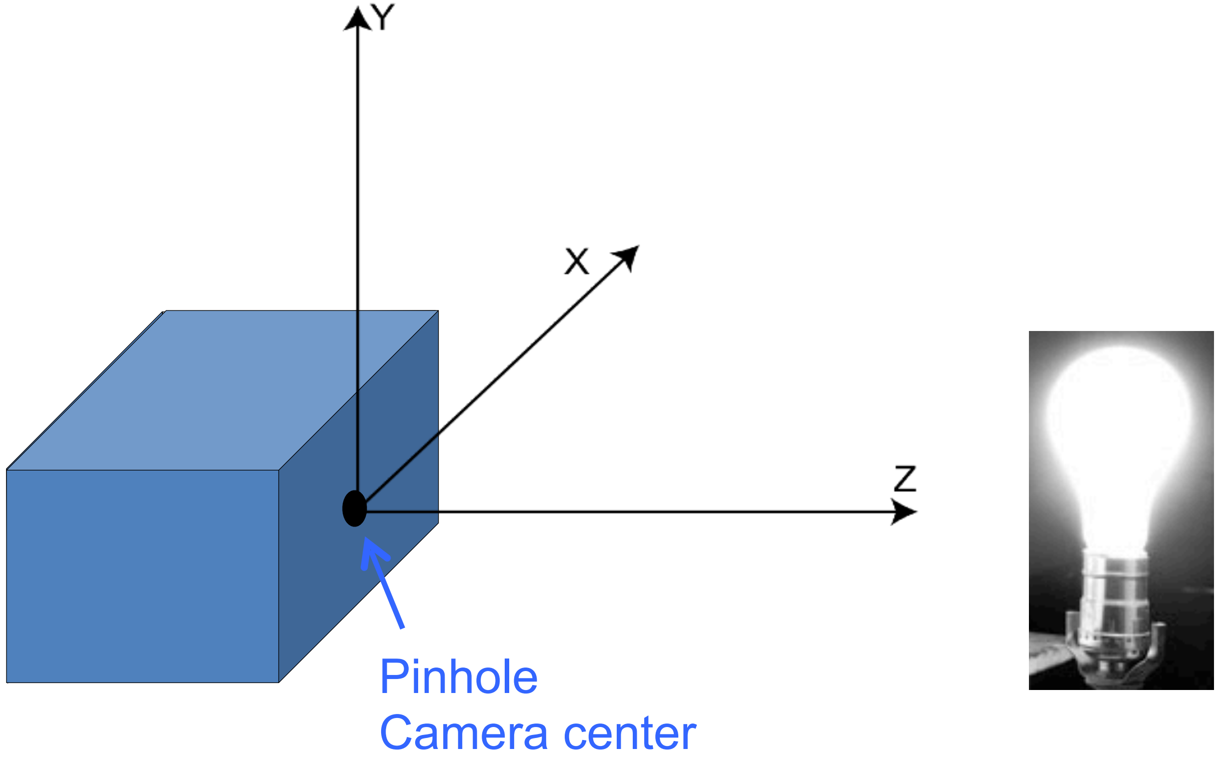
Recording from GoPro

Pinhole camera

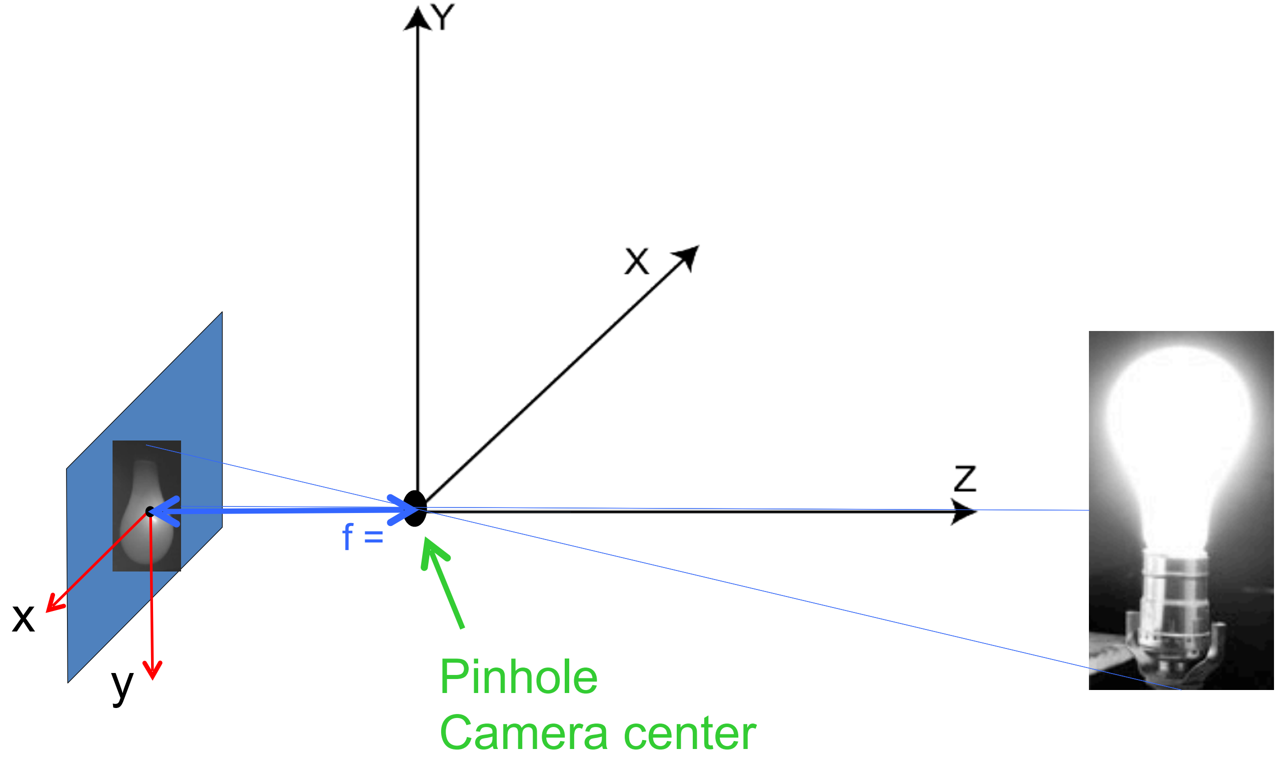


Photograph by Abelardo Morell, 1991

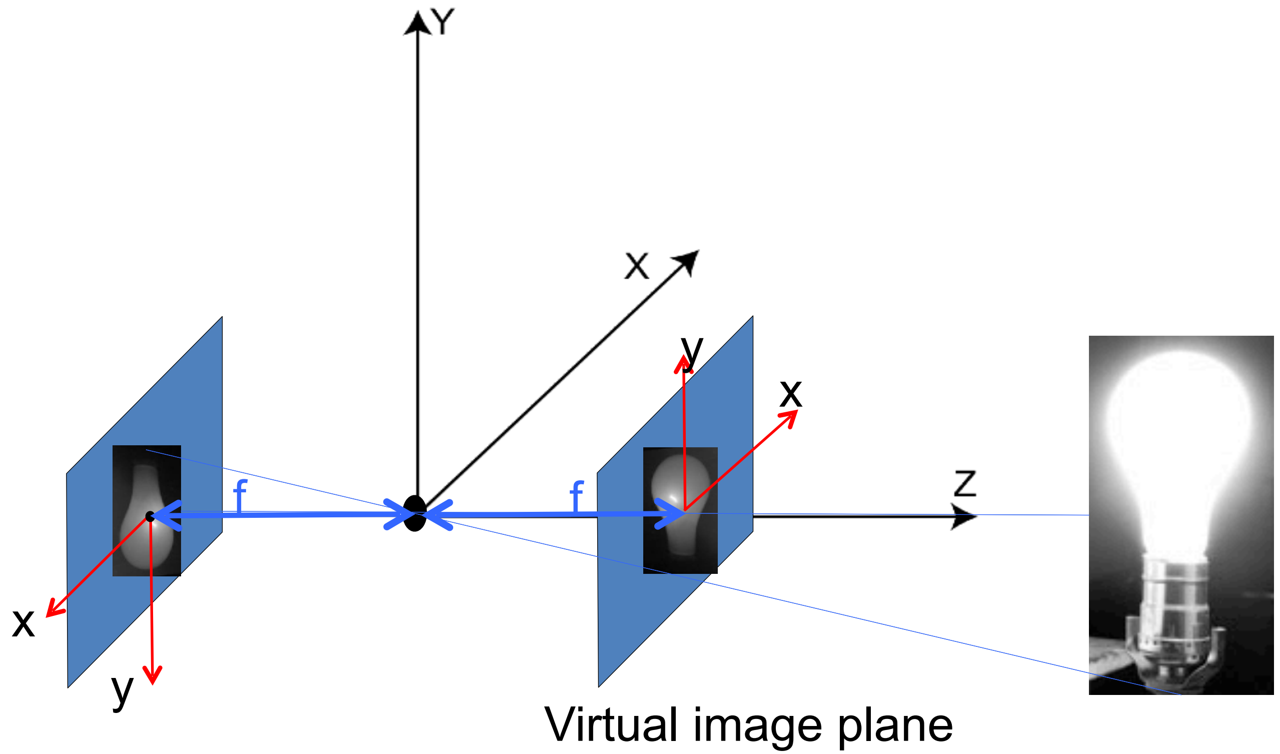
Perspective projection



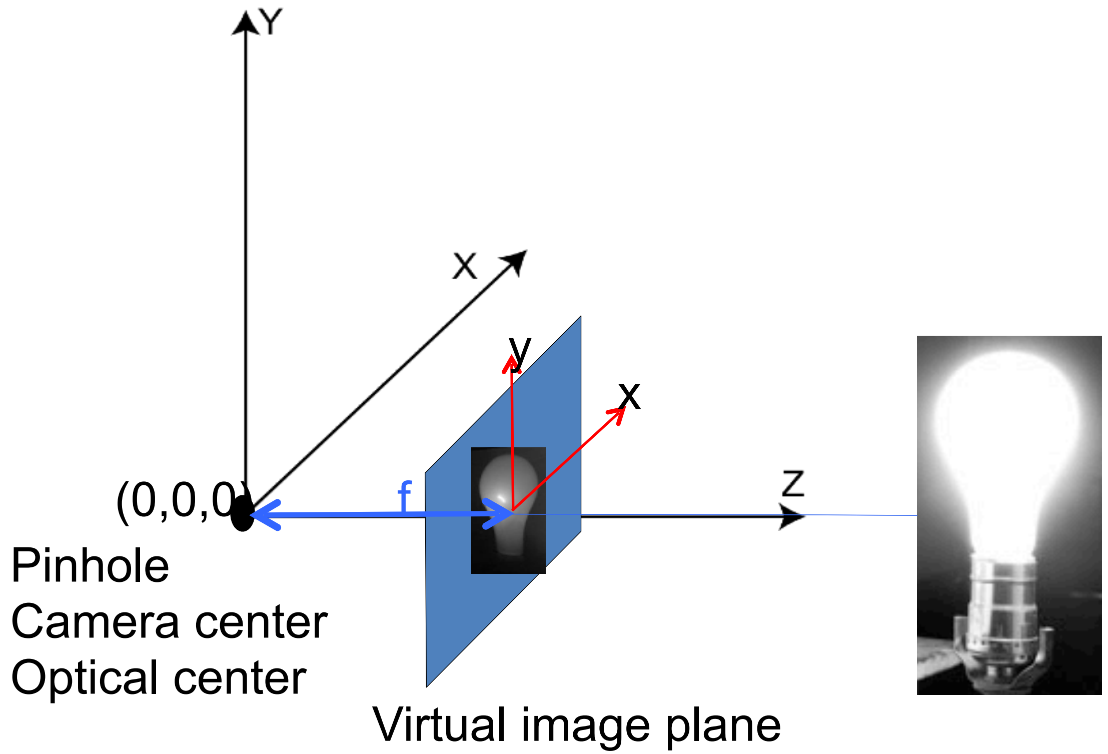
Perspective projection



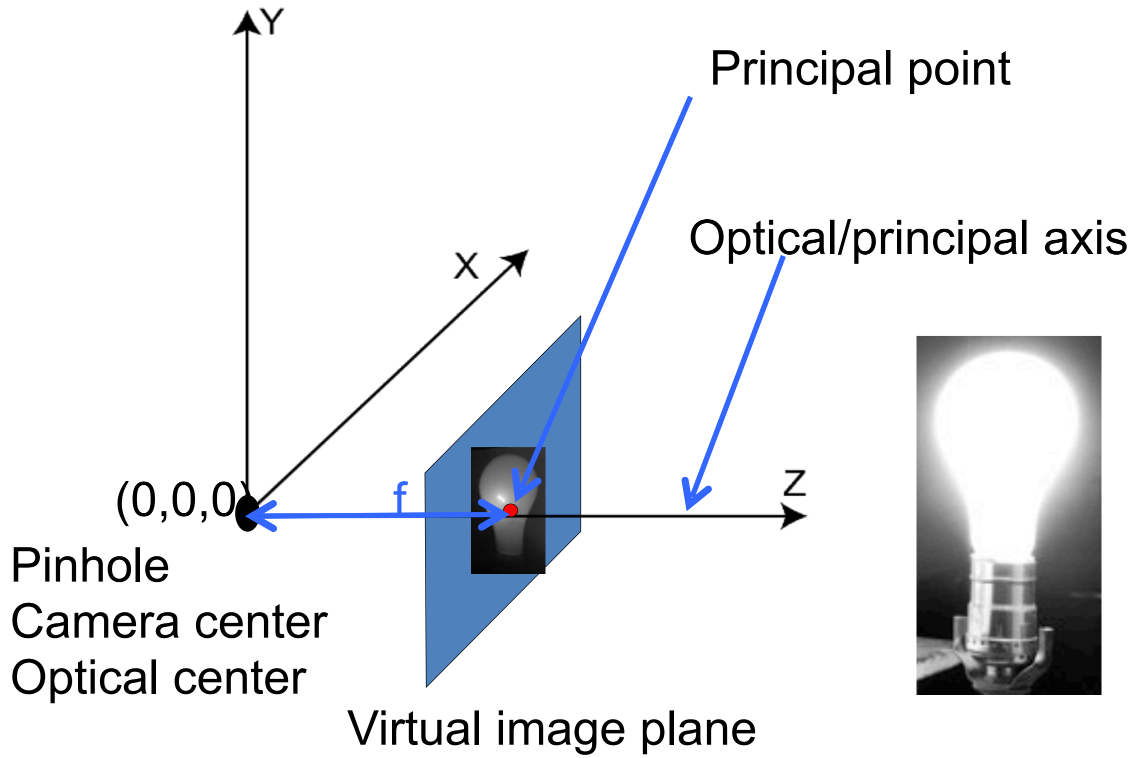
Perspective projection



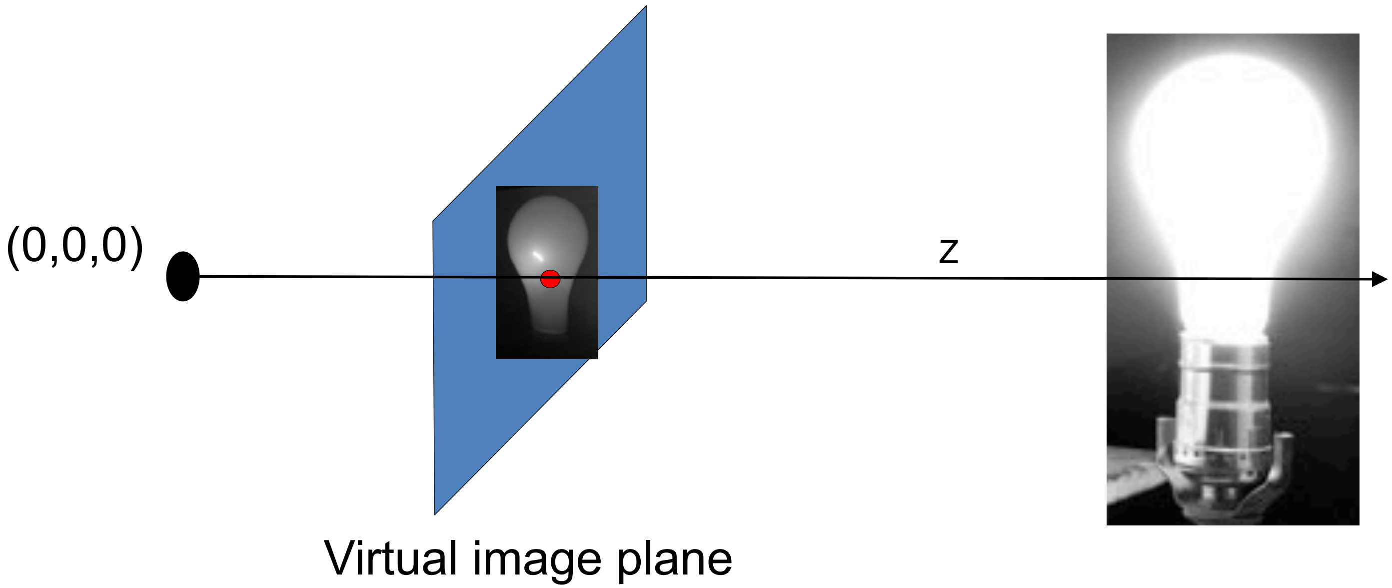
Perspective projection



Perspective projection



Perspective projection



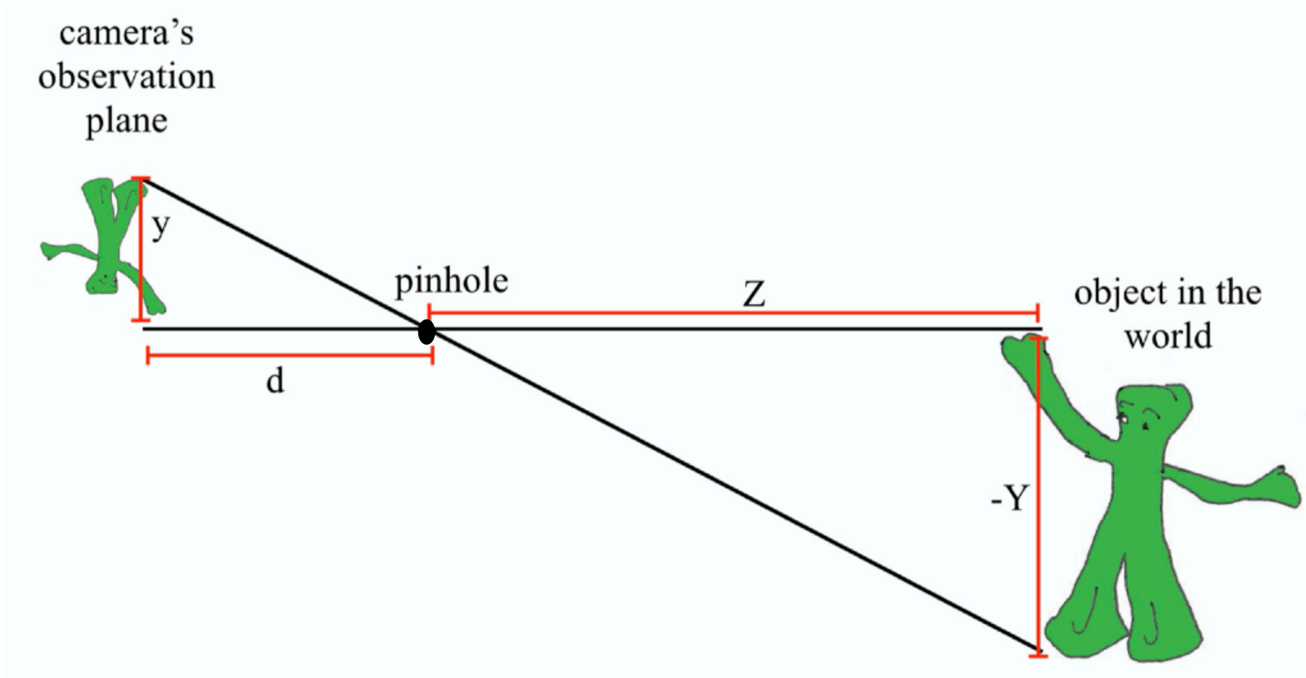
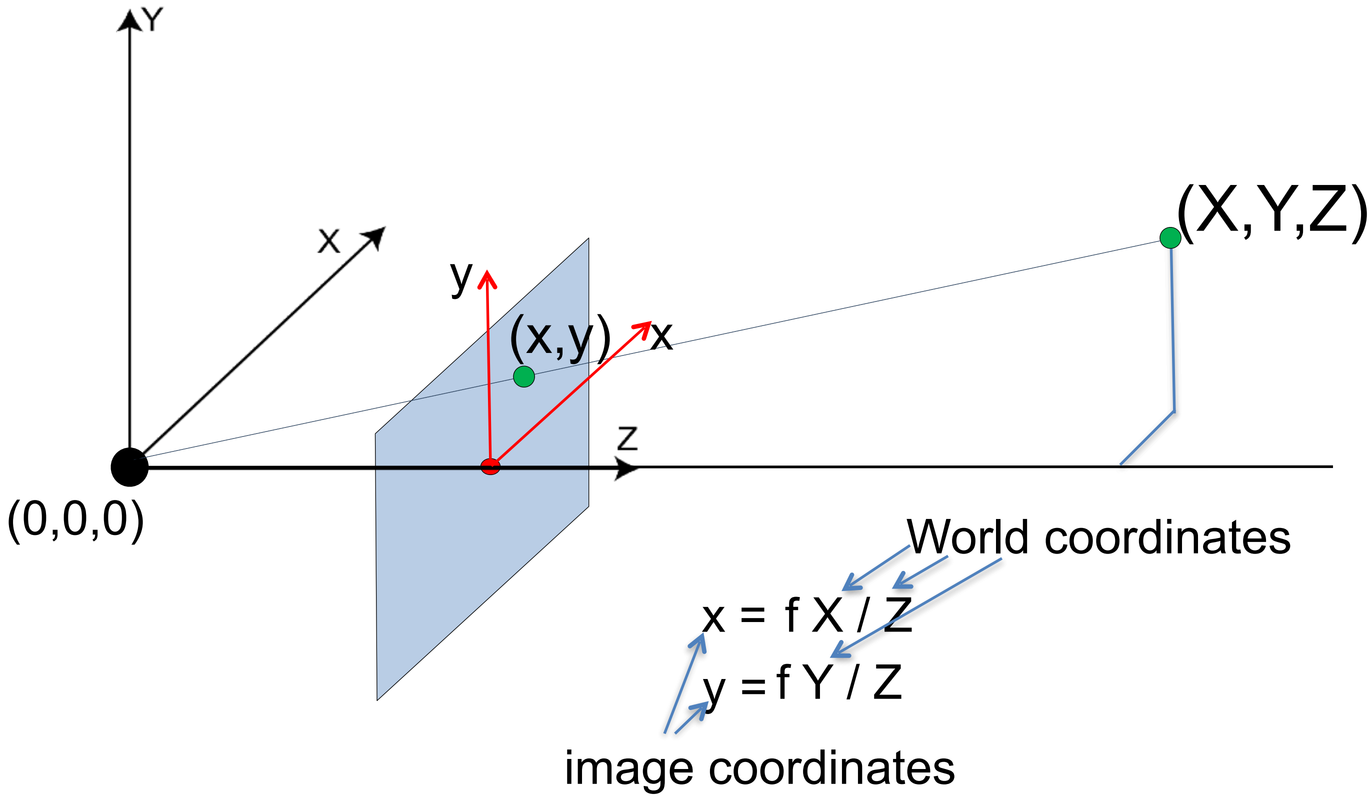
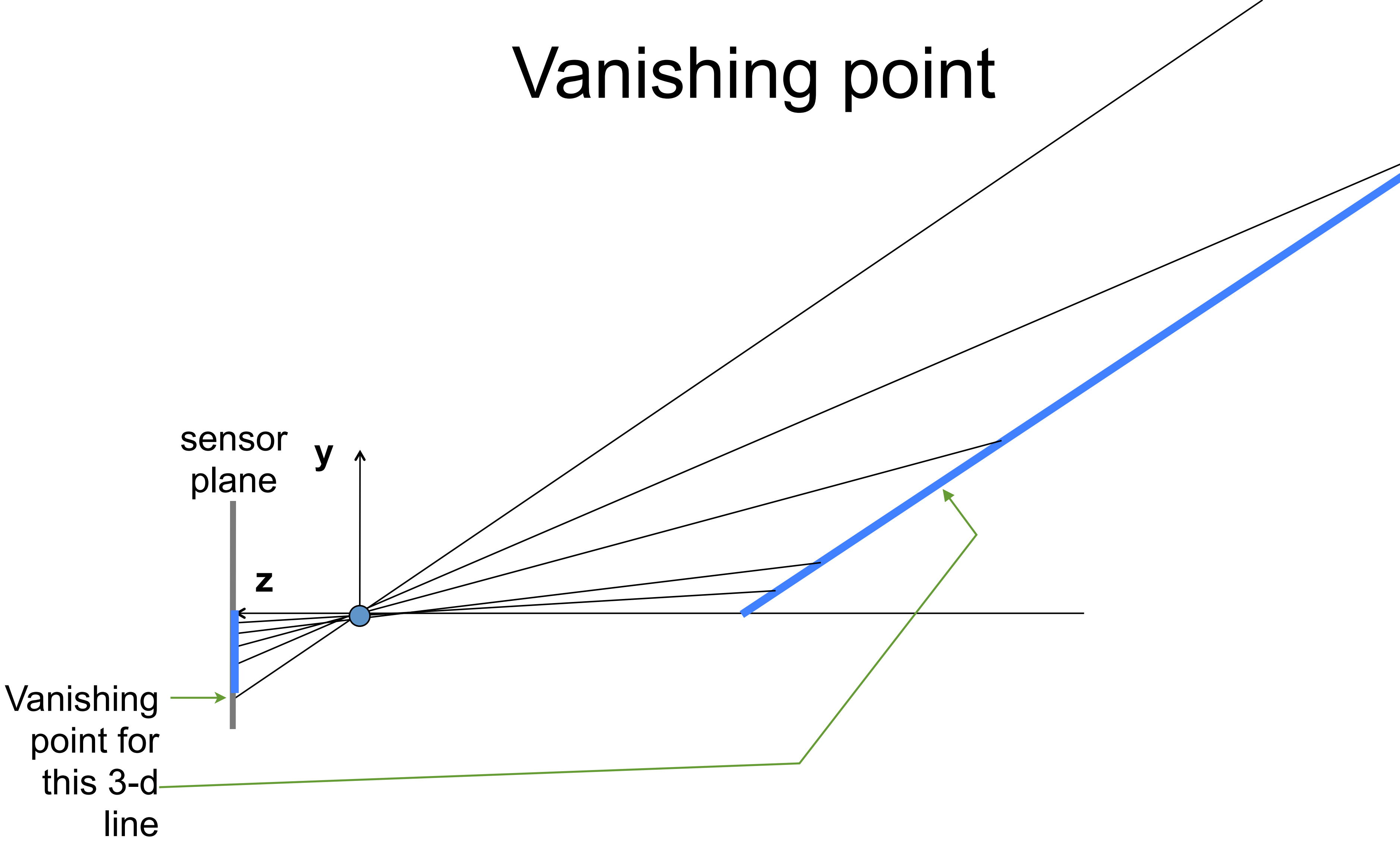


Figure 4.3: Perspective projection equations derived geometrically. From similar triangles, we have $y = -\frac{d}{Z}Y$.

Perspective projection



Vanishing point



Line in 3-space

$$X(t) = X_0 + at$$

$$Y(t) = Y_0 + bt$$

$$Z(t) = Z_0 + ct$$

Perspective projection of that line

$$x(t) = \frac{fX}{Z} = \frac{fX_0 + fat}{Z_0 + ct}$$

$$y(t) = \frac{fY}{Z} = \frac{fY_0 + fbt}{Z_0 + ct}$$

In the limit as $t \longrightarrow \pm\infty$ **we have (for** $c \neq 0$ **):**

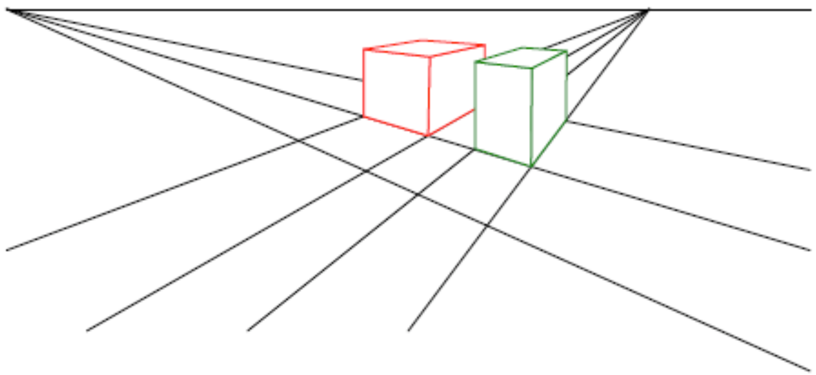
$$x(t \rightarrow \infty) \rightarrow \frac{fa}{c}$$

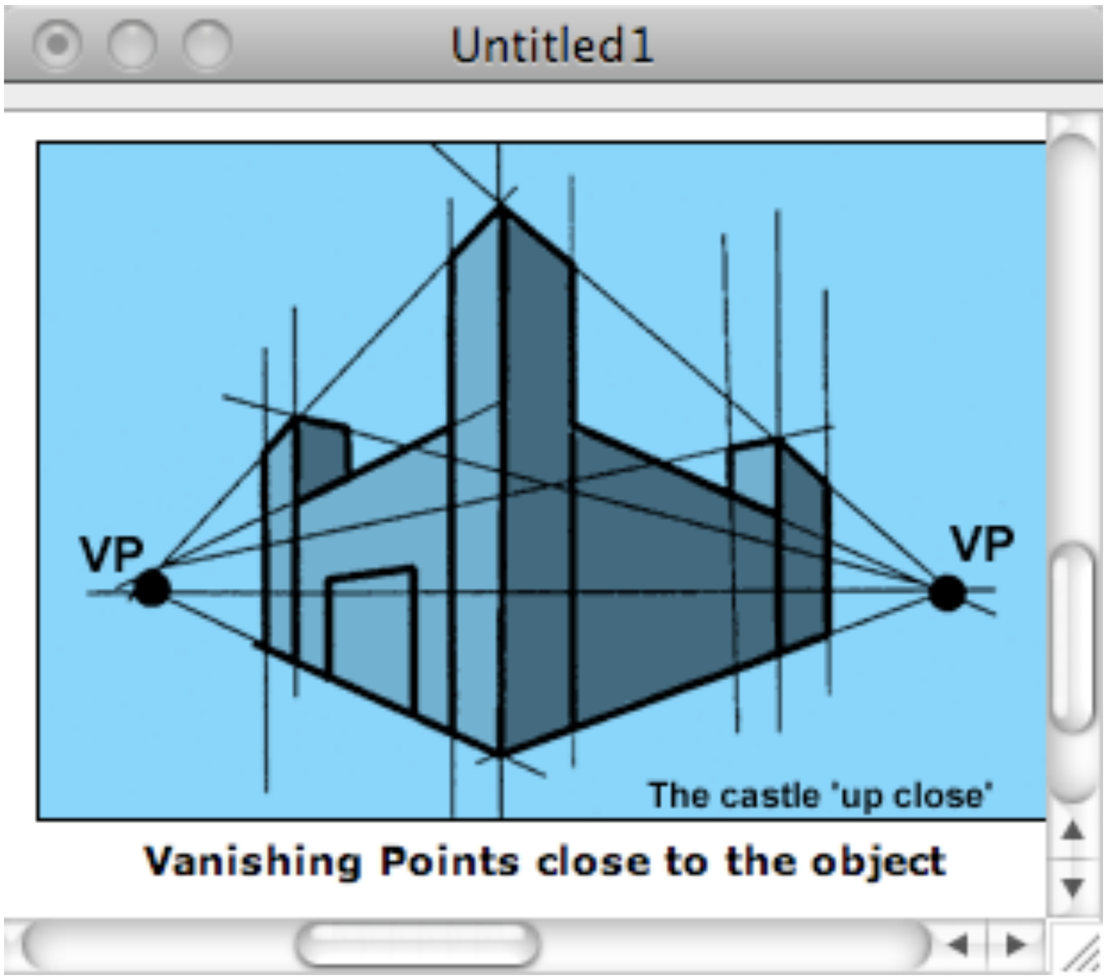
$$y(t \rightarrow \infty) \rightarrow \frac{fb}{c}$$

This tells us that any set of parallel lines (same a, b, c parameters) project to the same point (called the vanishing point).

Vanishing points

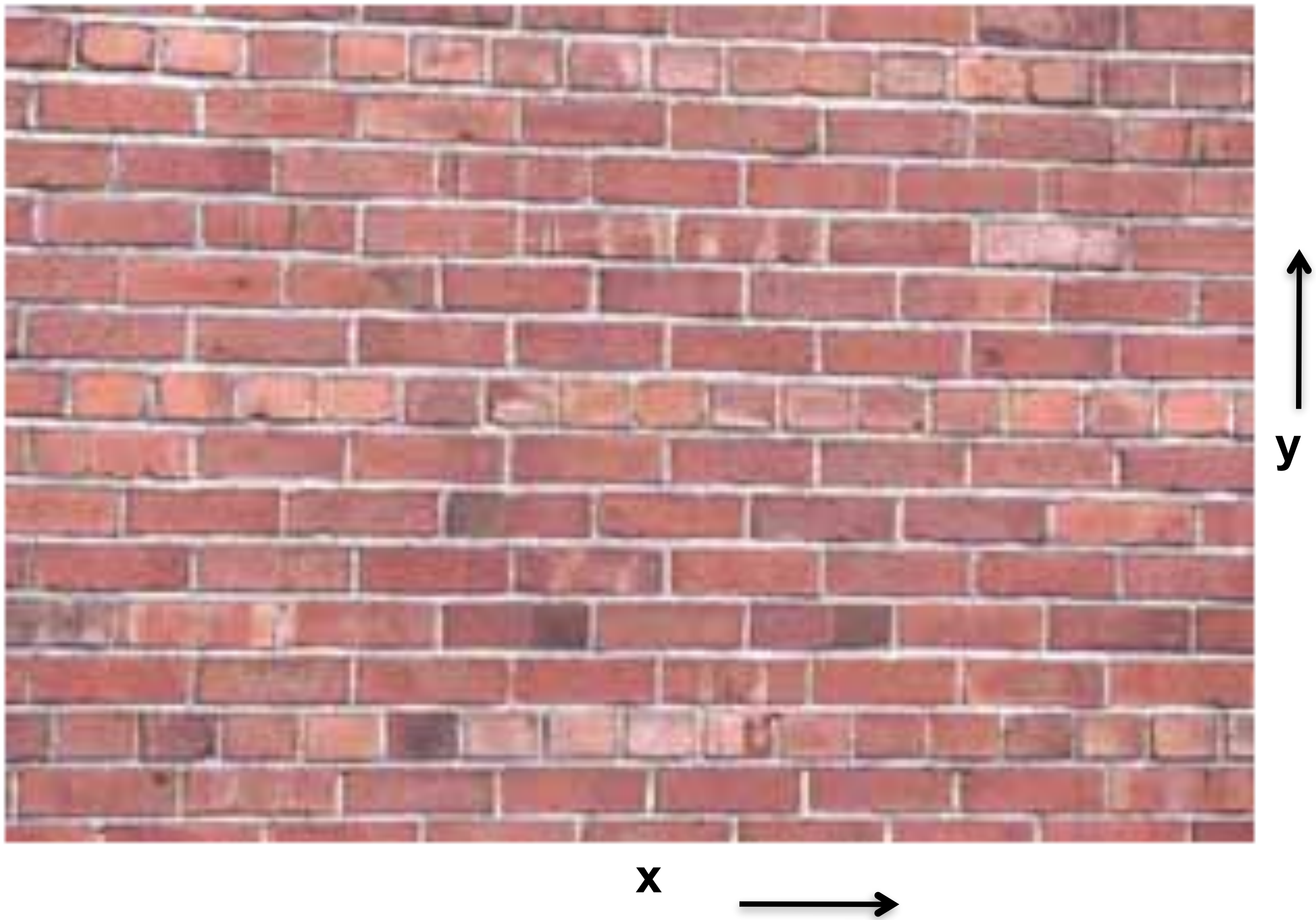
- Each set of parallel lines (=direction) meets at a different point
  - The *vanishing point* for this direction
- Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  - The line is called the *horizon* for that plane





http://www.ider.herts.ac.uk/school/courseware/graphics/two_point_perspective.html

What if you photograph a brick wall head-on?



Brick wall line in 3-space

$$X(t) = X_0 + at$$

$$Y(t) = Y_0$$

$$Z(t) = Z_0$$

Perspective projection of that line

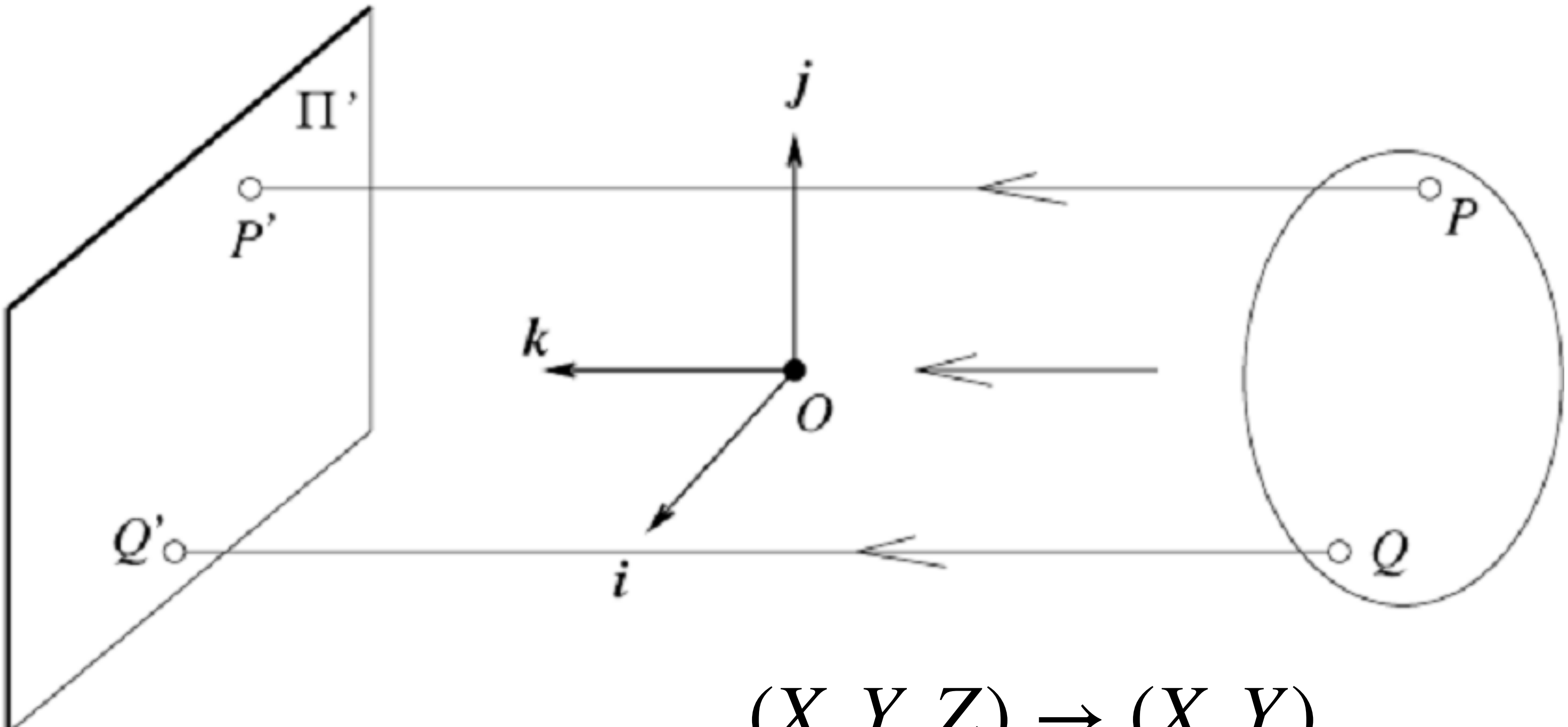
$$x(t) = \frac{fX}{Z} = \frac{fX_0 + fat}{Z_0}$$

$$y(t) = \frac{fY}{Z} = \frac{fY_0}{Z_0}$$

All bricks have same z_0. Those in same row have same y_0

Thus, a brick wall, photographed head-on, gets rendered as set of parallel lines in the image plane.

Other projection models: Orthographic projection



$$(X, Y, Z) \to (X, Y)$$

Approximation to this: telephoto lens with a very long focal length

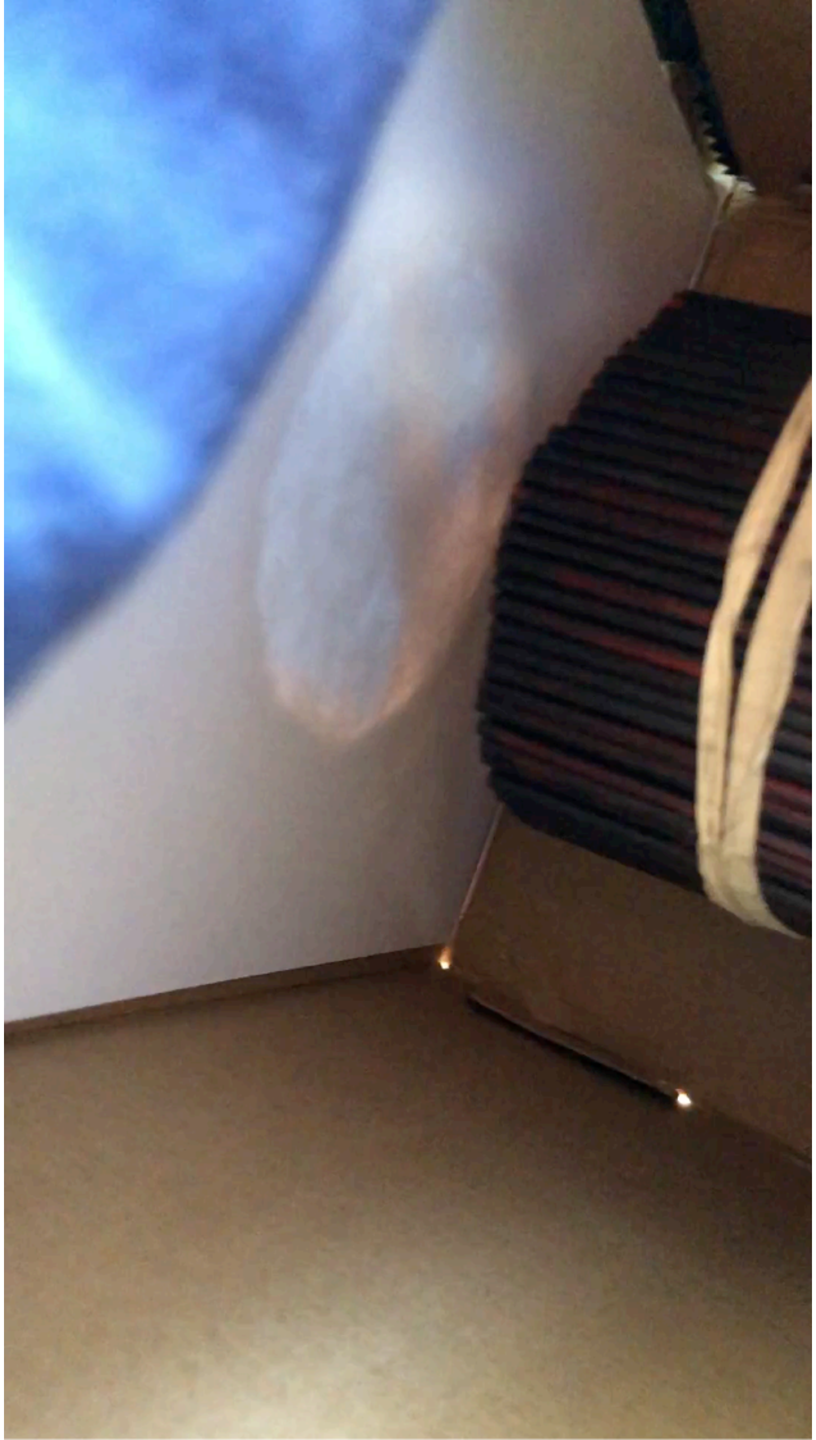
How else might you make a camera with this projection?

Straw camera

(a)

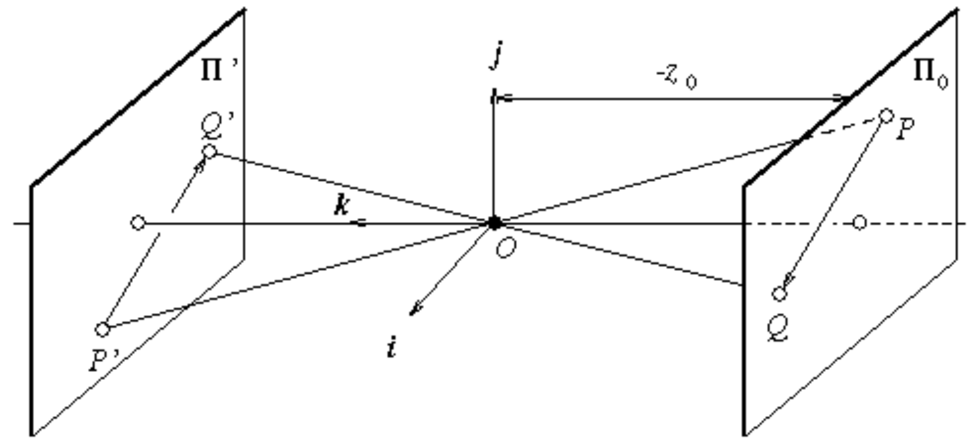
(b)

Straw camera



Other projection models: Weak perspective

- Issue
  - perspective effects, but not over the scale of individual objects
  - collect points into a group at about the same depth, then divide each point by the depth of its group
  - Adv: easy
  - Disadv: only approximate



$$(X, Y, Z) \rightarrow \left(\frac{fX}{Z_0}, \frac{fY}{Z_0} \right)$$

Three camera projections

3-d point → **2-d image position**

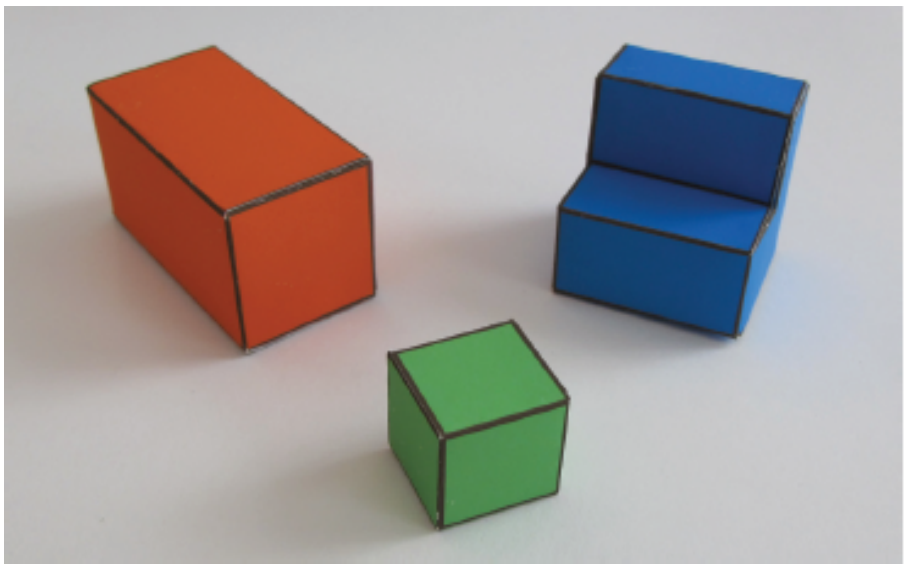
(1) Perspective: $(X, Y, Z) \rightarrow \left(\frac{fX}{Z}, \frac{fY}{Z}\right)$

(2) Weak perspective: $(X, Y, Z) \rightarrow \left(\frac{fX}{Z_0}, \frac{fY}{Z_0}\right)$

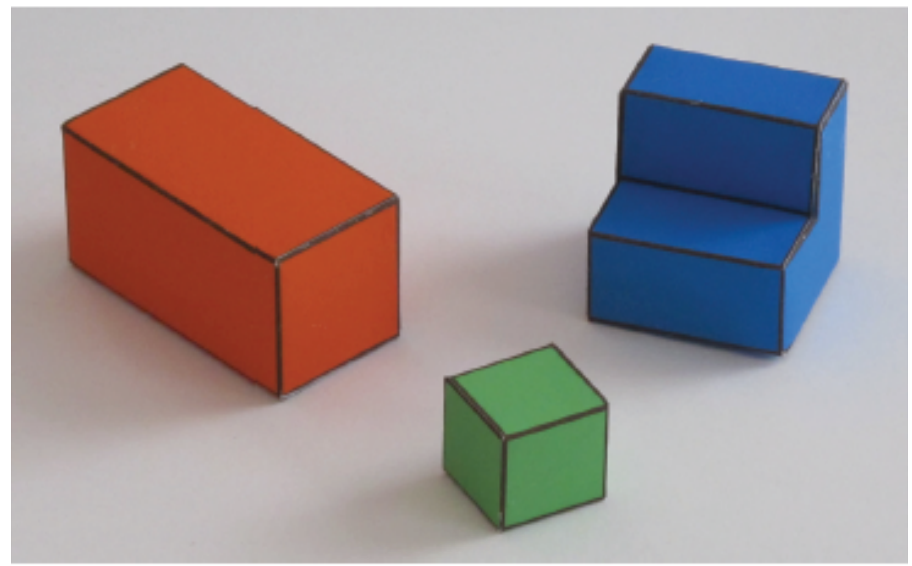
(3) Orthographic: $(X, Y, Z) \rightarrow (X, Y)$

which is perspective, which orthographic?

Perspective projection

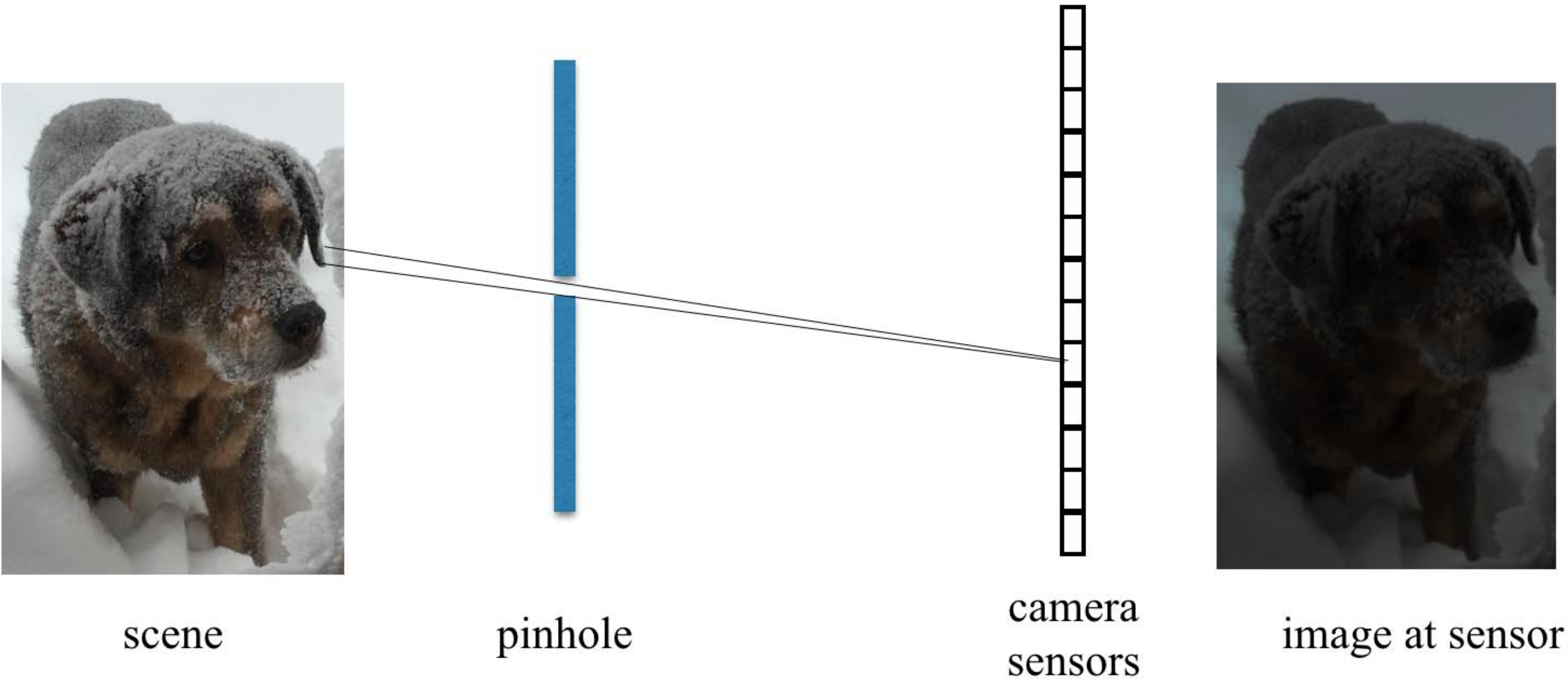


Parallel (orthographic) projection

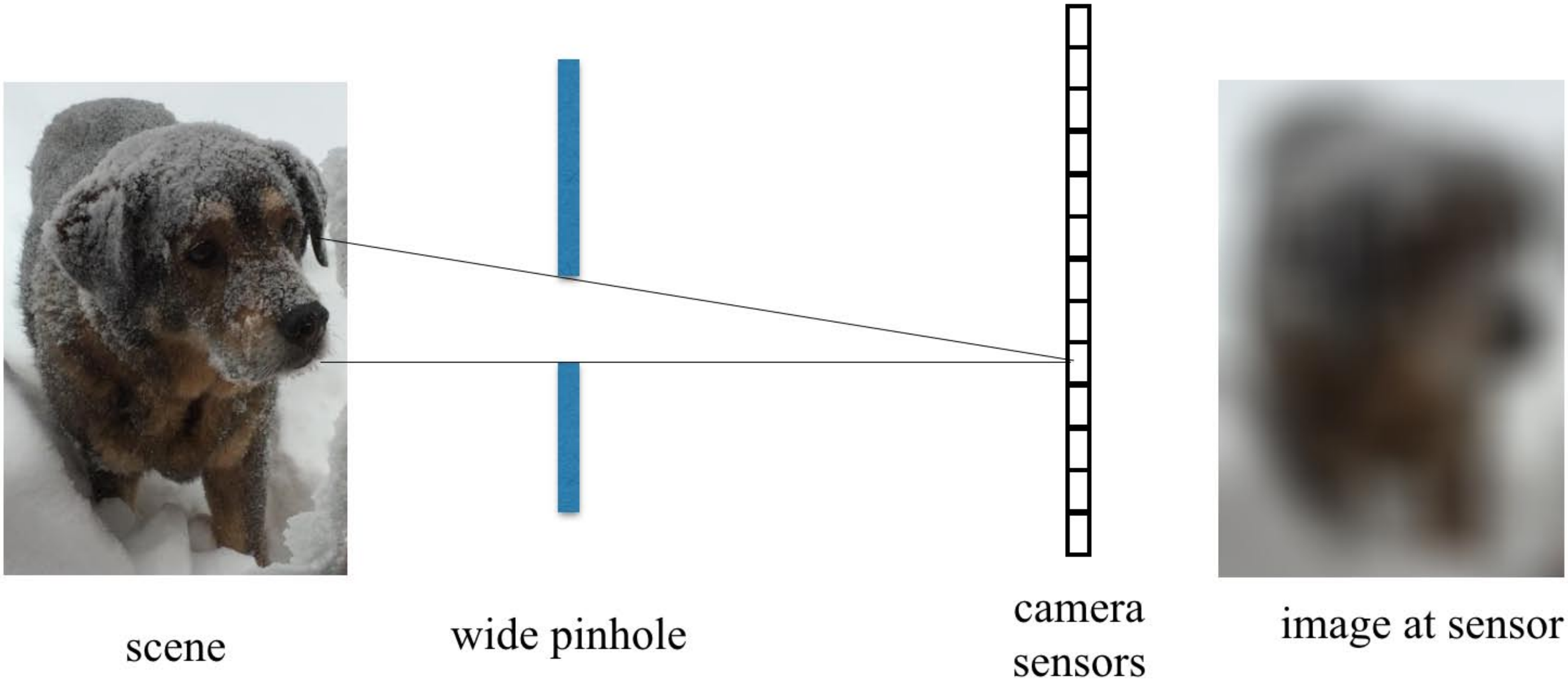


What are the drawbacks of pinhole cameras?

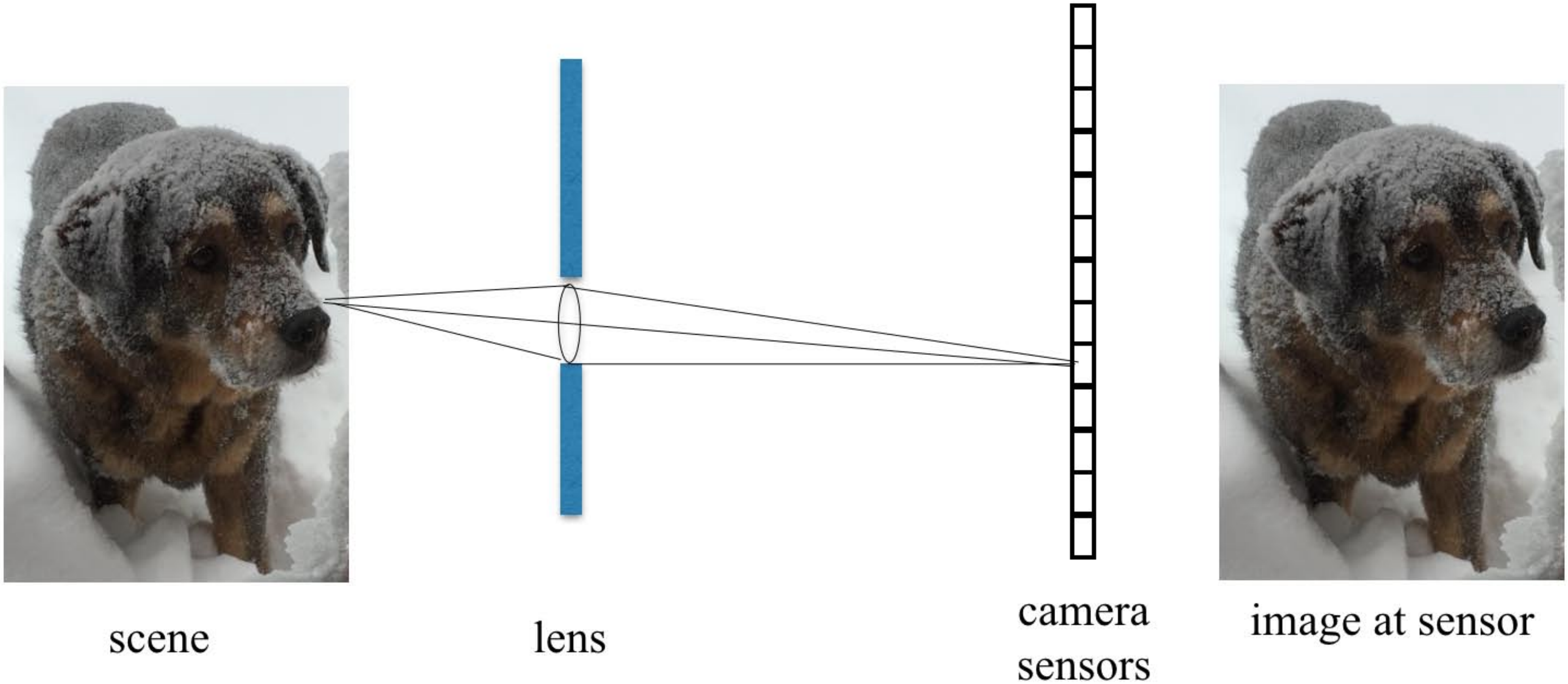
A problem: pinhole camera images are dark, or require long exposures



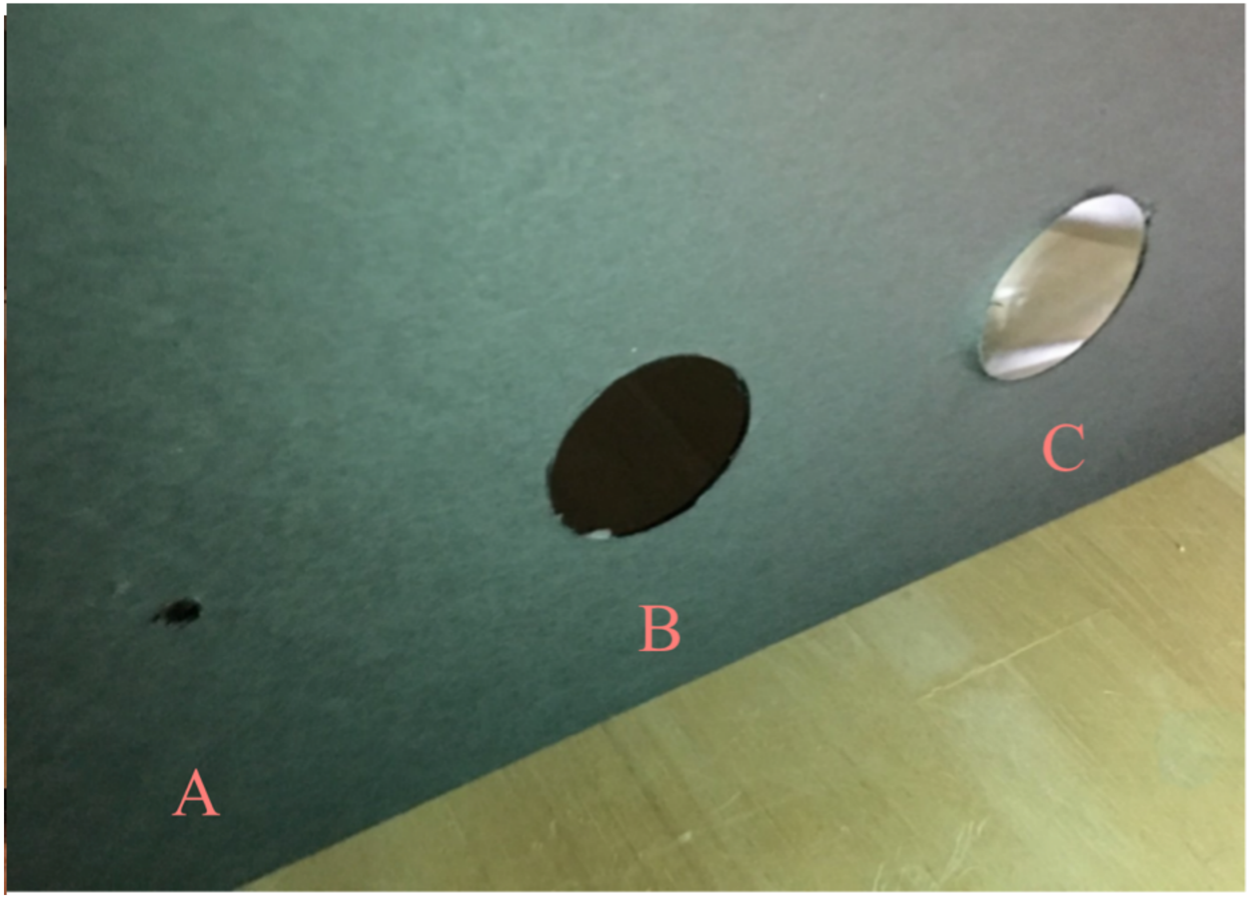
Large aperture gives a brighter image, but at the price of sharpness



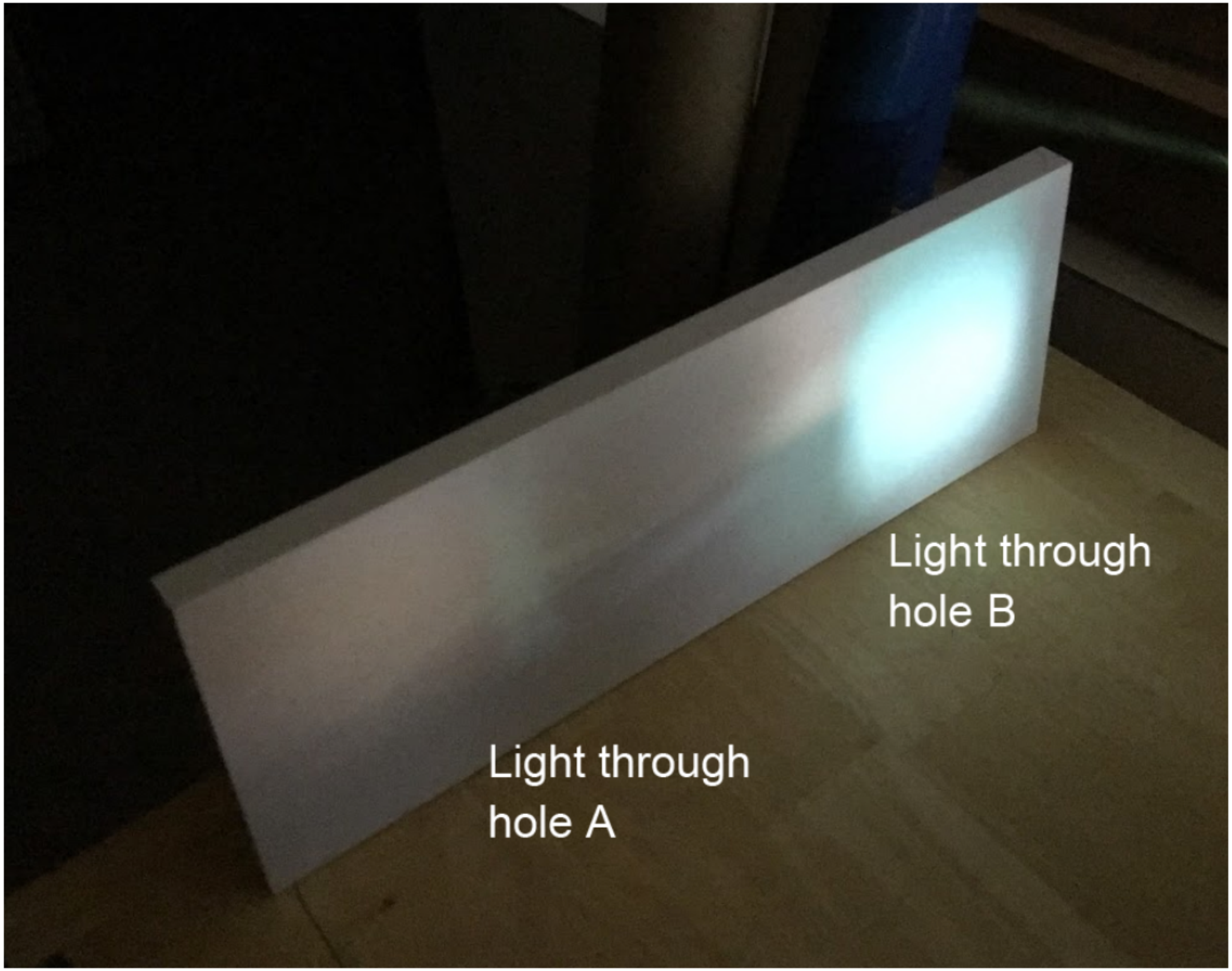
A lens allows a large aperture and a sharp image



Let's try putting different occluders in between the scene and the sensor plane

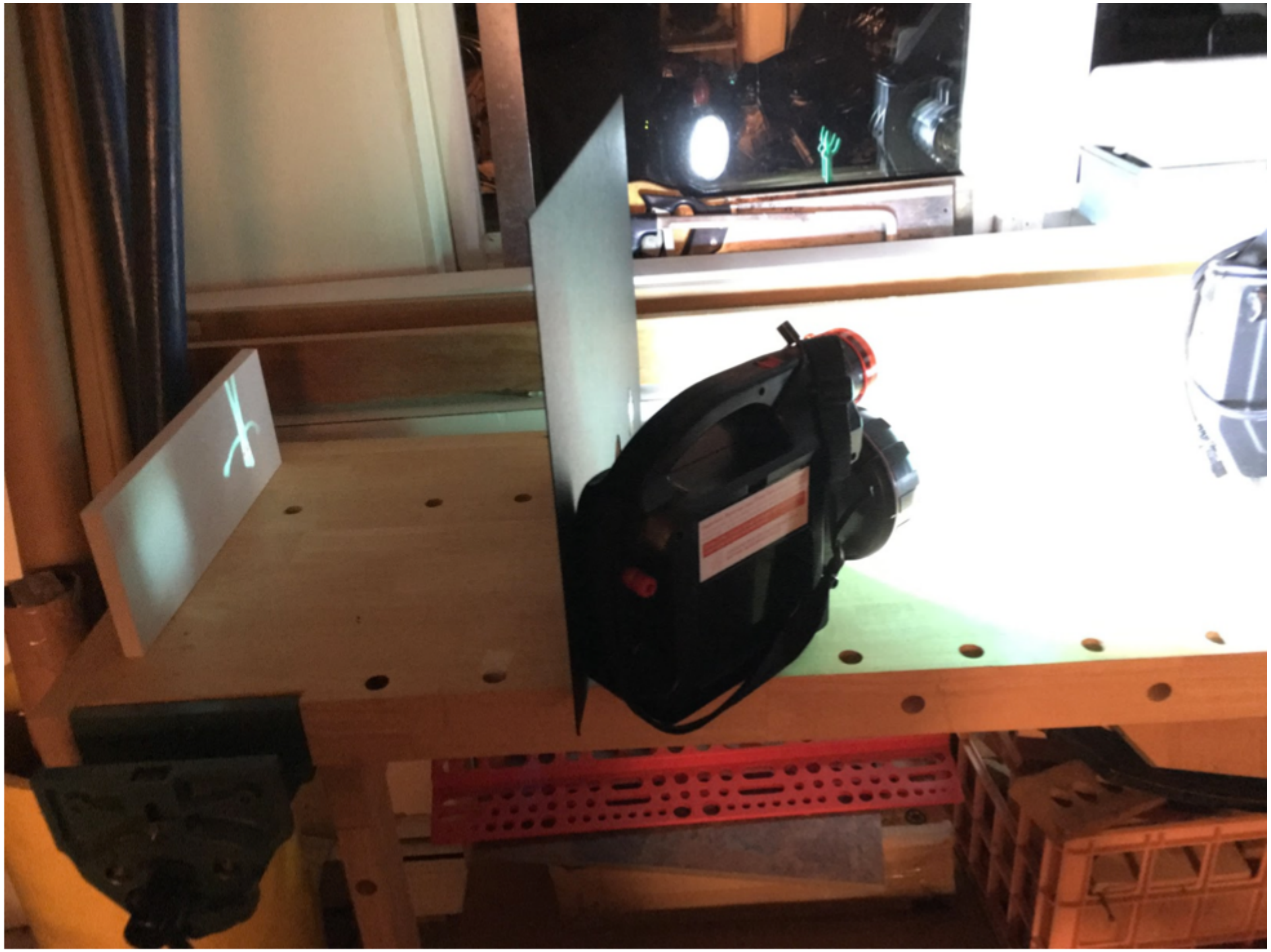


Influence of aperture size: with a small aperture, the image is sharp, but dim. A large aperture gives a bright, but blurry image.

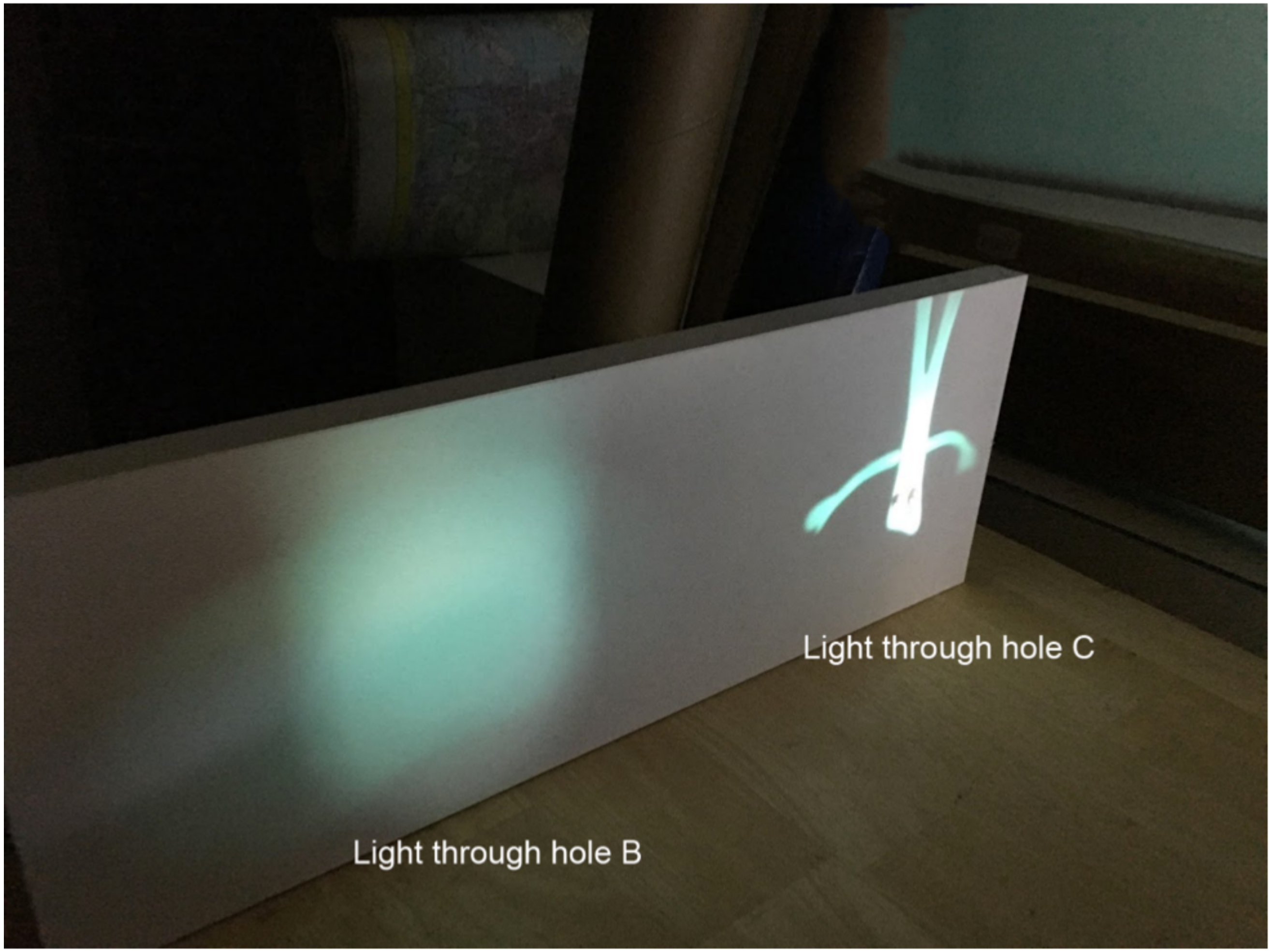


A lens can focus light from one point in the world to one point on the sensor plane.

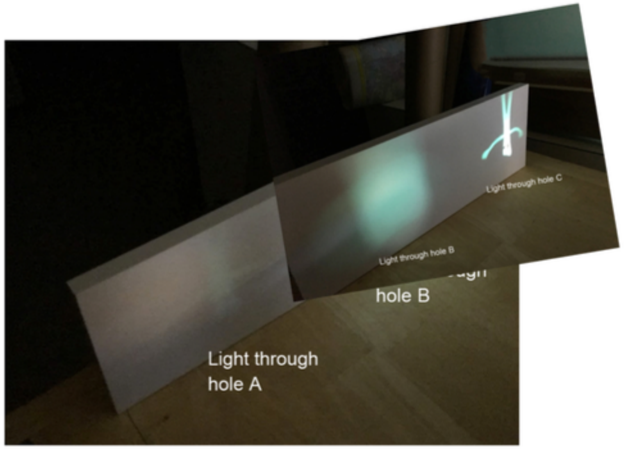
Images through large aperture, with and without lens present



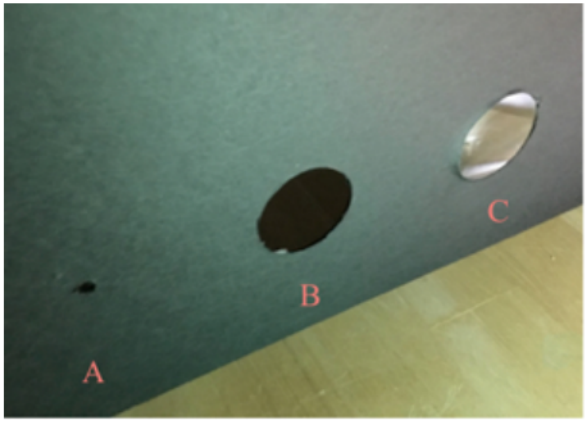
Images through large aperture, with and without lens present



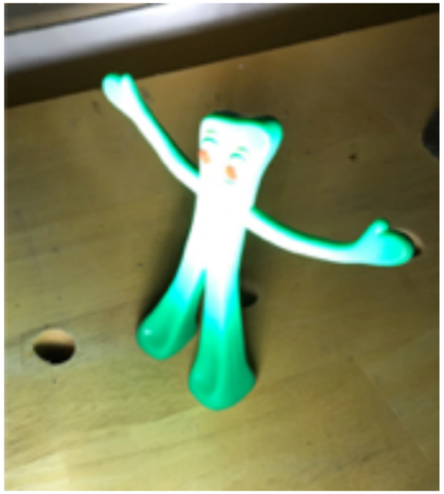
(a)


Light through hole C
Light through hole B
hole B
Light through
hole A

(b)

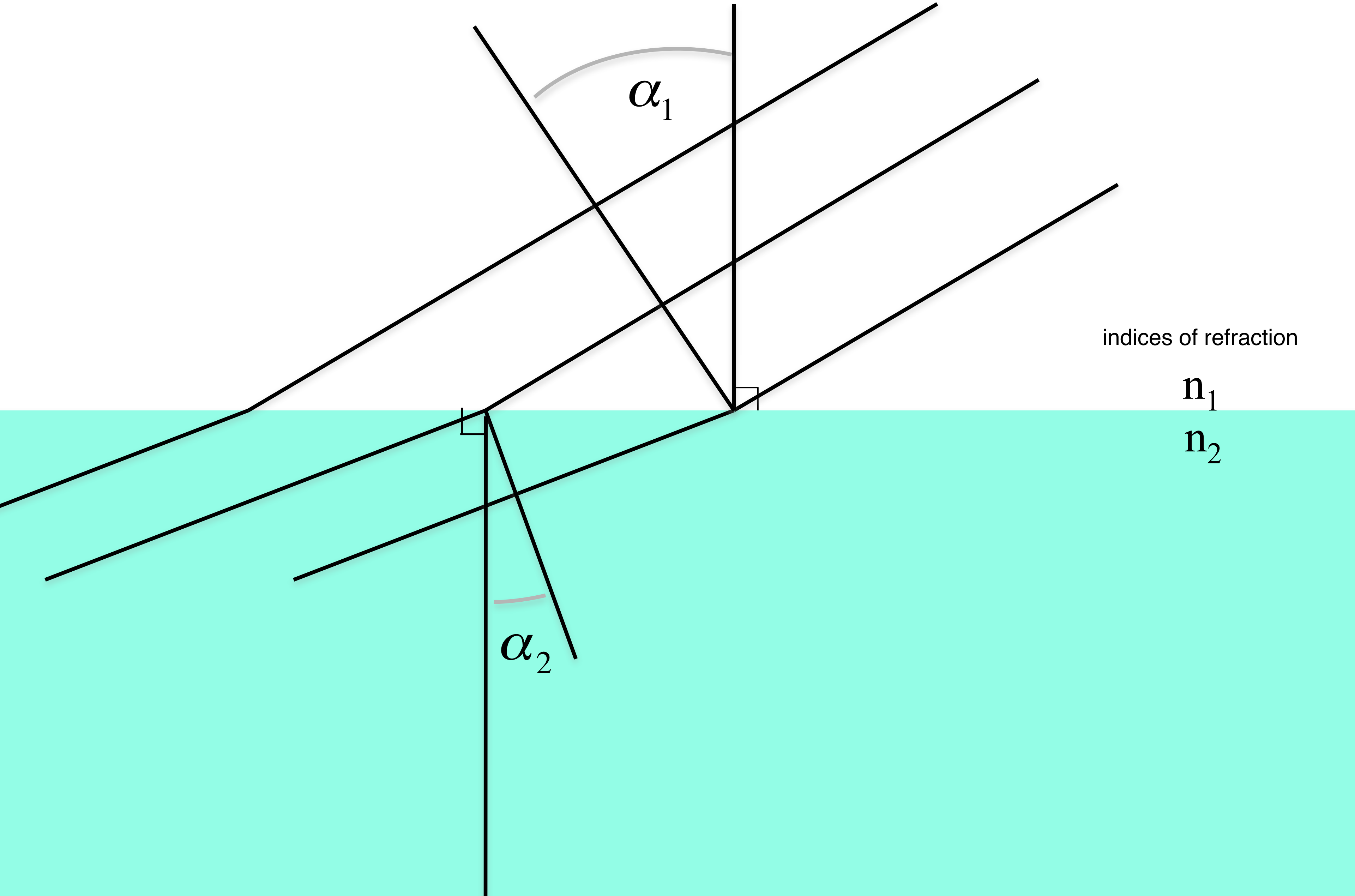

C
B
A

(c)



(d)

Light at a material interface



Light at a material interface

Speed, and thus wavelength of light, scales inversely with n. This requires that plane waves bend, according to

Snell's law of refraction

$$n_1 \sin(\alpha_1) = n_2 \sin(\alpha_2)$$

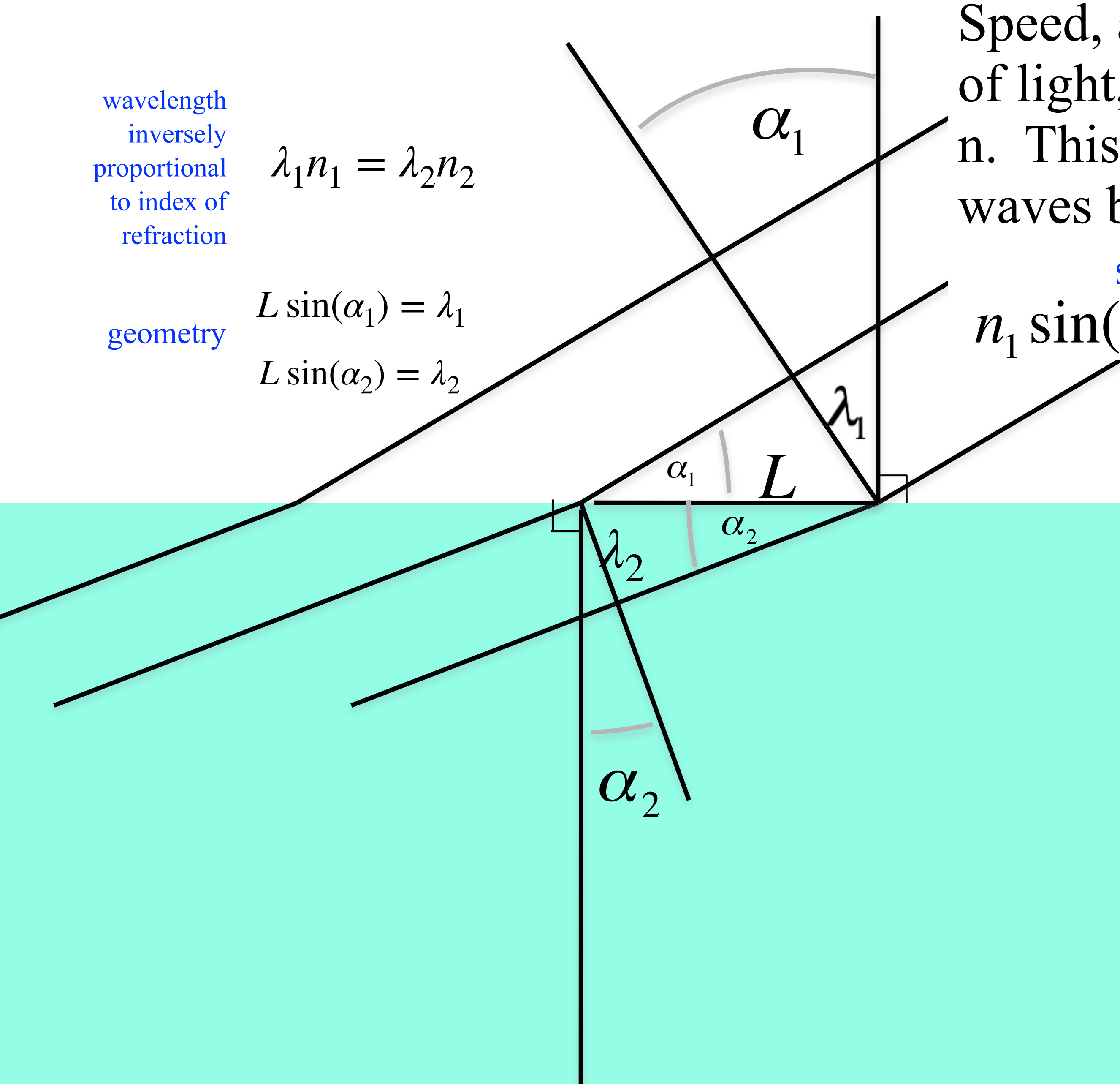
wavelength inversely proportional to index of refraction

$$\lambda_1 n_1 = \lambda_2 n_2$$

geometry

$$L \sin(\alpha_1) = \lambda_1$$

$$L \sin(\alpha_2) = \lambda_2$$



indices of refraction

n_1

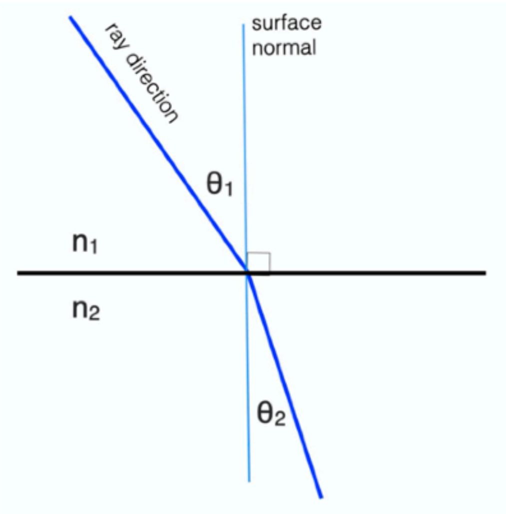
n_2

Snell’s law, for small angles

$$n_1 \sin(\alpha_1) = n_2 \sin(\alpha_2)$$

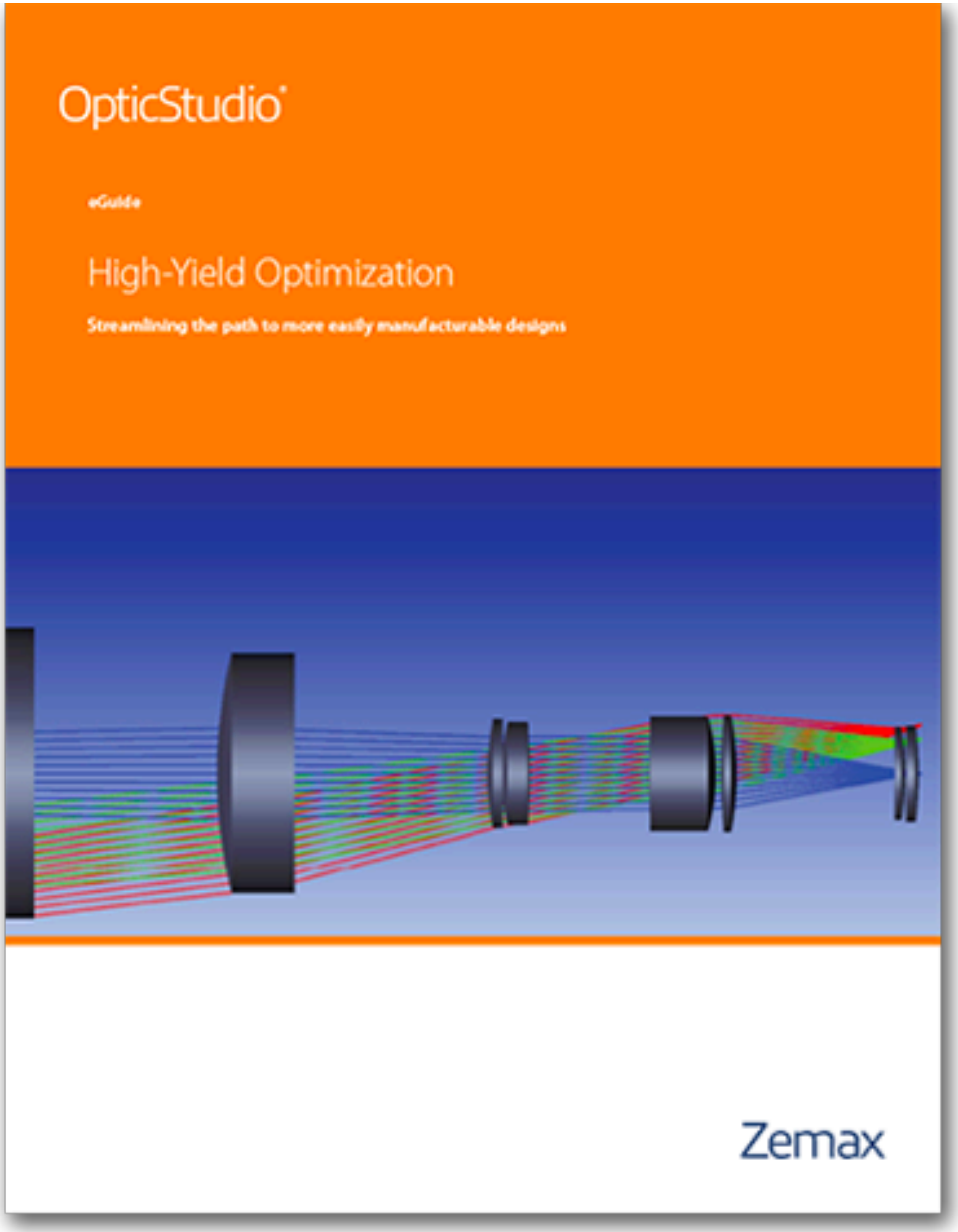
For small angles,

$$n_1 \alpha_1 = n_2 \alpha$$

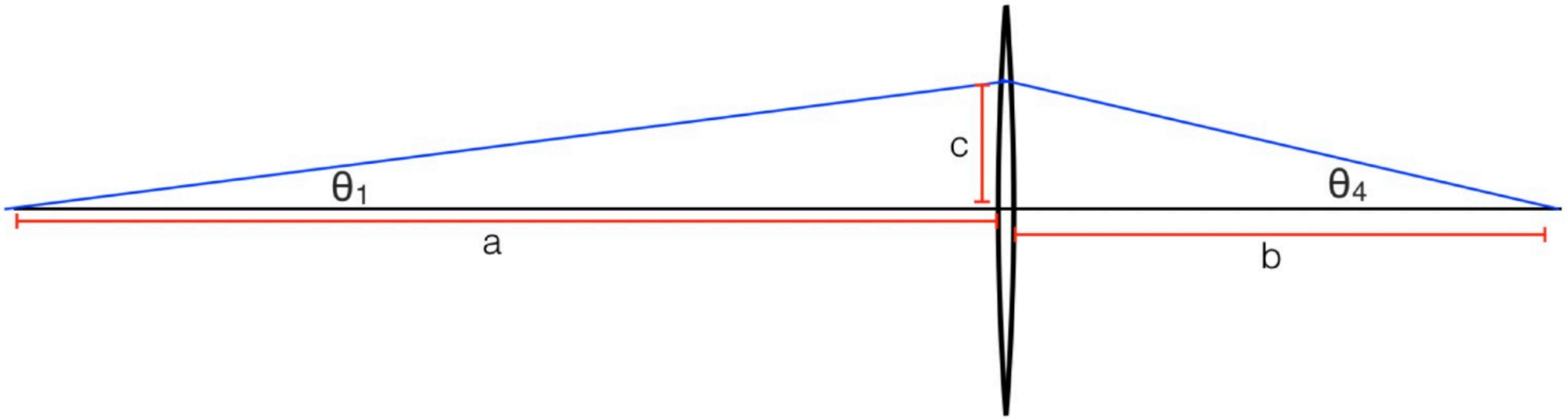


Modern camera lens systems are designed by computer, using commercial programs such as Zemax. (Max was the name of the original programmer's dog, but was taken as a trademarked name, so they went with Zemax)

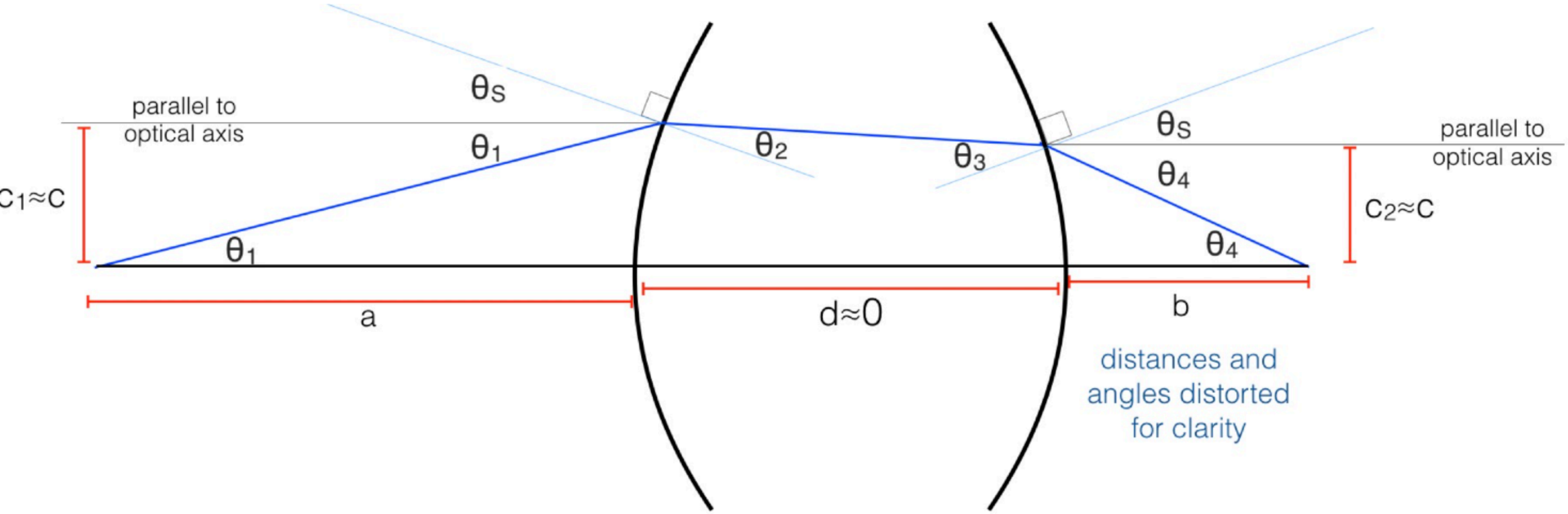
But let's design a very simple lens by hand…



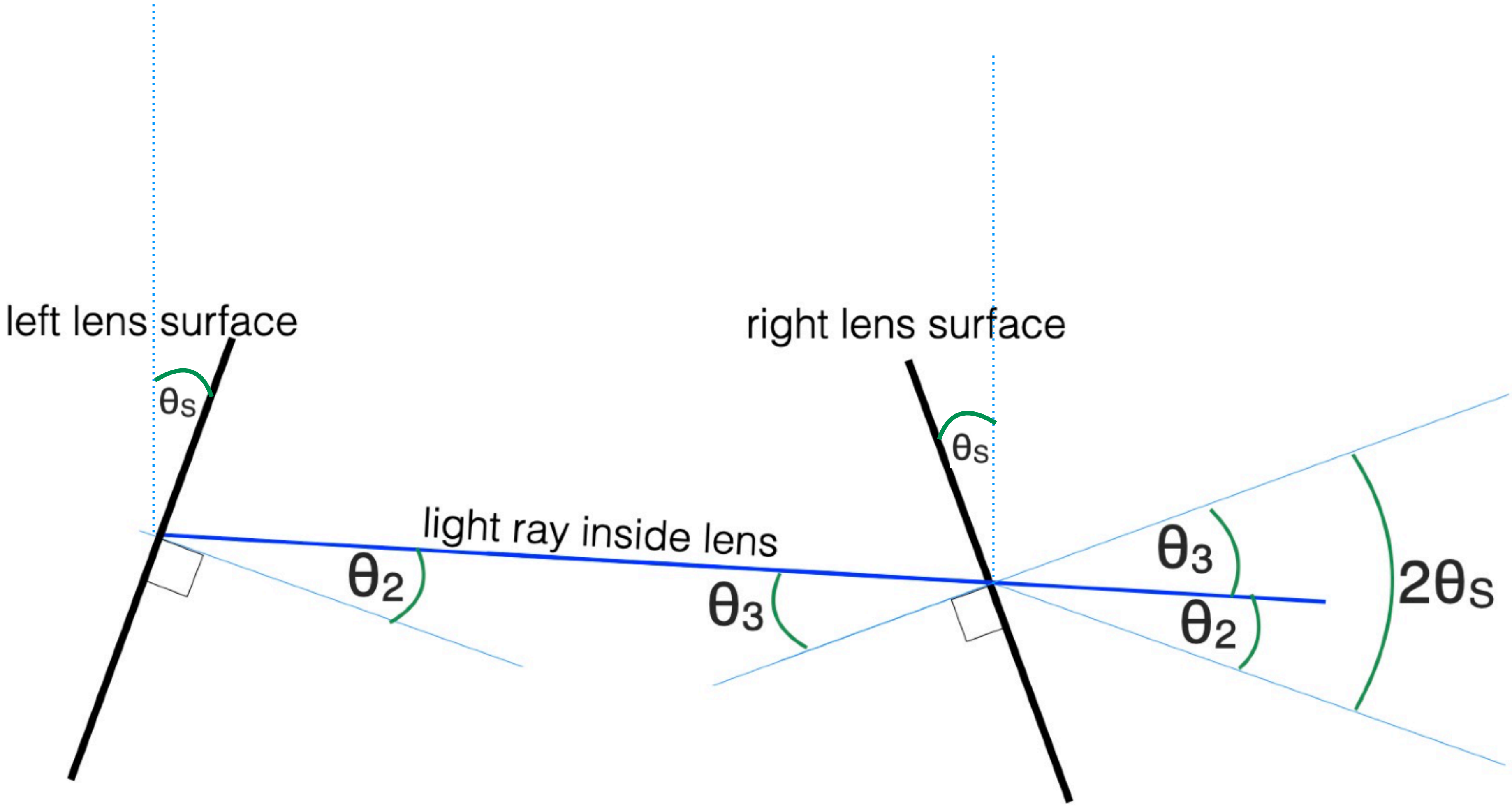
what shape should we make a thin lens so that it will focus light?



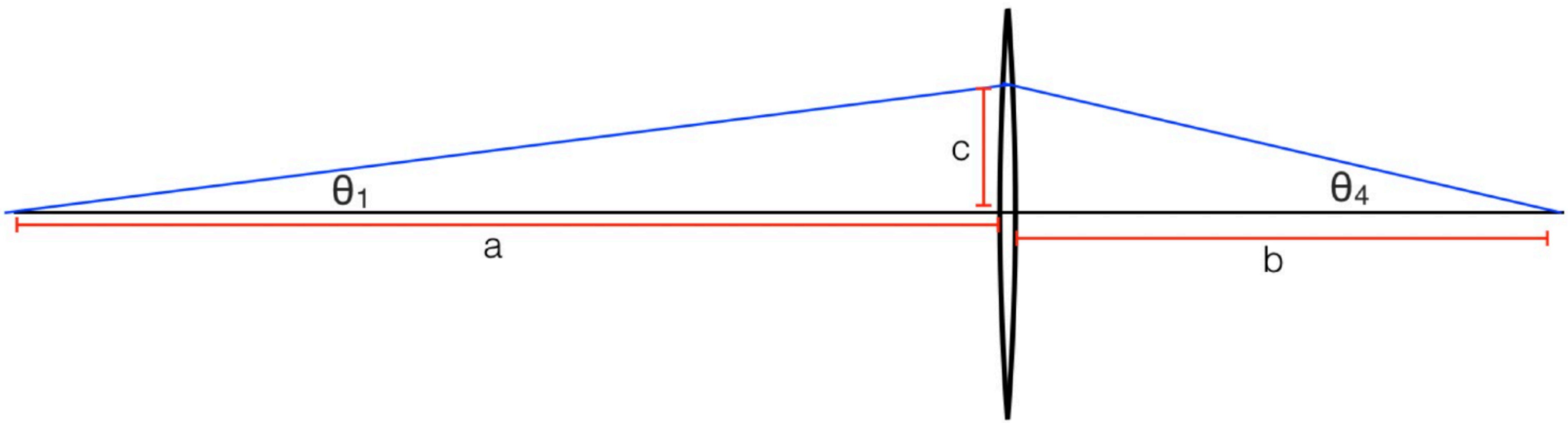
with angles distorted for labeling clarity



Angle	Description	Relation	Reason
θ_1	initial angle from optical axis	$\theta_1 = \frac{c}{a}$	small angle approx.
θ_2	angle of refracted ray wrt front surface normal	$n\theta_2 = \theta_1 + \theta_S$	Snell's law, small angle approx.
θ_3	angle of refracted ray wrt back surface normal	$2\theta_S = \theta_2 + \theta_3$	symmetry of lens, thin lens approx.
$\theta_4 + \theta_S$	angle of ray exiting lens wrt back surface normal	$n\theta_3 = \theta_4 + \theta_S$	Snell's law, small angle approx.
θ_4	final angle from optical axis	$\theta_4 = \frac{c}{b}$	small angle approx.

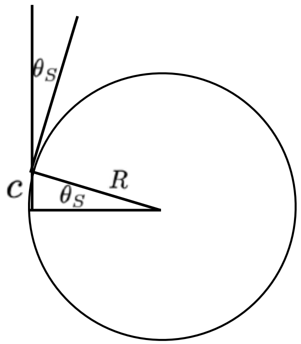

left lens surface
θs
right lens surface
θs
light ray inside lens
θ2
θ3
θ3
θ2
2θs

What shape should we make a lens so that it will focus light?



$$\theta_S = \frac{c}{2(n-1)}\left(\frac{1}{a} + \frac{1}{b}\right) \tag{4.10}$$

Lensmaker's equation



For thin lenses, both parabolic and spherical shapes satisfy that constraint. For a spherical lens surface, curving according to a radius R, we have $\sin(\theta_S) = \frac{c}{R}$. For small angles θ_S, this reduces to

$$\theta_S = \frac{c}{R}, \tag{4.11}$$

where R is the radius of the sphere, which has the desired property that $\theta_S \propto c$. Substituting Eq. (4.11) into the focusing condition, Eq. (4.10) yields the *Lensmaker's Formula*,

$$\theta_S = \frac{c}{2(n-1)}\left(\frac{1}{a} + \frac{1}{b}\right)$$

from previous slide

$$\frac{1}{R} = \frac{1}{2(n-1)}\left(\frac{1}{a} + \frac{1}{b}\right)$$

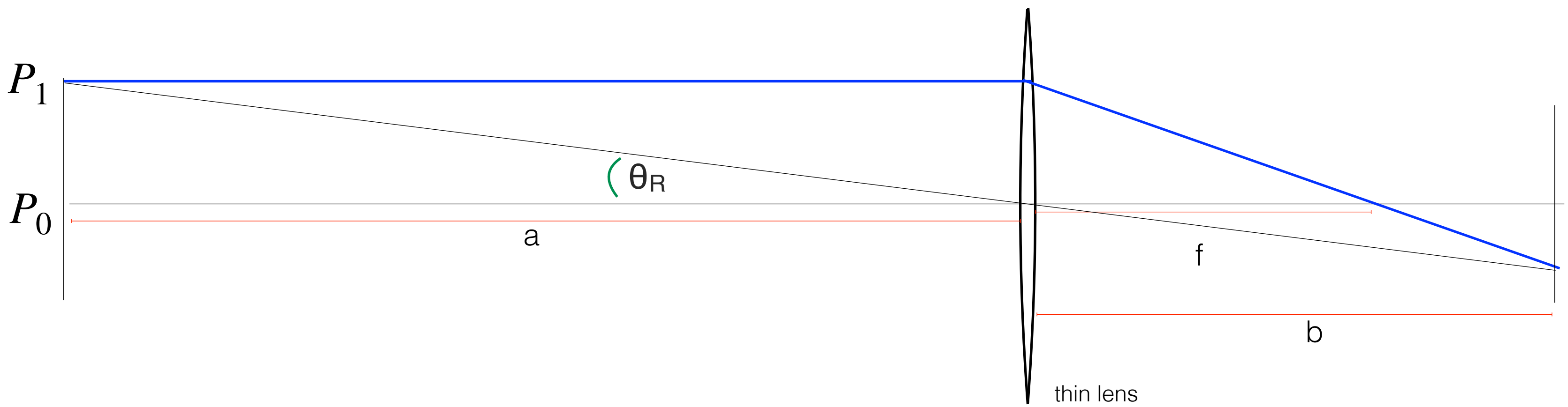
combine with 4.11

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}, \tag{4.12}$$

where the lens *focal length*, f is

$$f = \frac{R}{2(n-1)} \tag{4.13}$$

Note: (1) off-axis rays are focussed, too, and (2) rays from infinity focus at a distance f



$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}$$

Lens demonstration

- Verify:
  - Focusing property
  - Lens maker's equations

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}$$

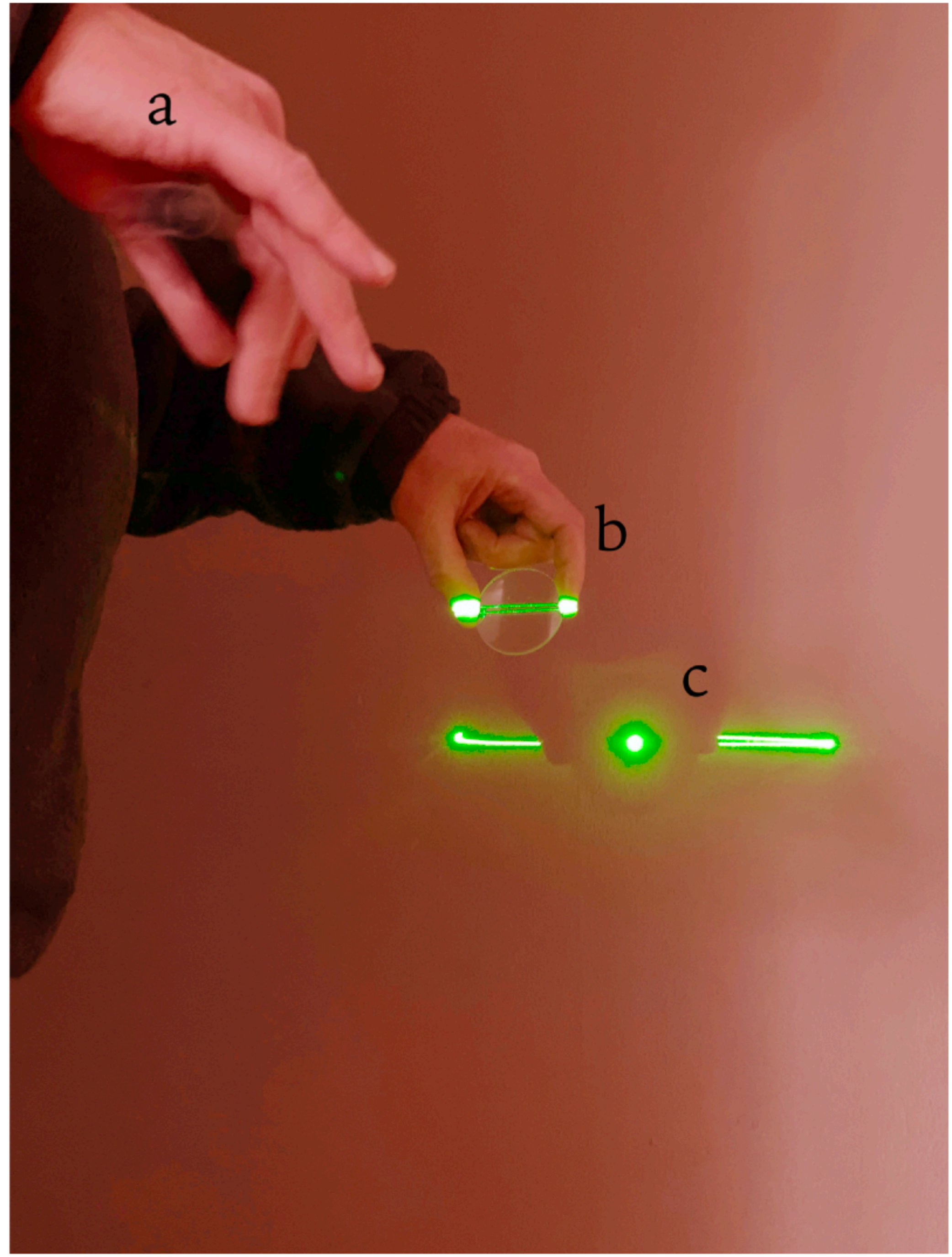
$$f = \frac{R}{2(n-1)}$$

lens focal length: 20cm

lens to laser pointer center of rotation
= 23.5 inches = 59.7 cm

lens to wall = 12.5 inches = 31.7 cm

1/59.7 + 1/31.7 = 1/20.7



Lens Demonstration

more general cameras

Photometric properties of general imagers

$$\vec{y} = \mathbf{A}\vec{x} \tag{1.9}$$

For the case of conventional cameras, where the observed intensities, $\vec{y}$ are an image of the reflected intensities in the scene, $\vec{x}$, then $\mathbf{A}$ is approximately an identity matrix.

For more general cameras, $\mathbf{A}$ may be very different from an identity matrix, and we will need to estimate $\vec{x}$ from $\vec{y}$. In the presence of noise, there may not be a solution $\vec{x}$ that exactly satisfies Eq. (1.9), so we often seek to satisfy it in a least squares sense. In most cases, $\mathbf{A}$ is either not invertable, or is poorly conditioned. It is often useful to introduce a regularizer, an additional term in the objective function to be minimized. If the regularization term favors small $\vec{x}$, then the objective term to minimize, E, could be

$$E = |\vec{y} - \mathbf{A}\vec{x}|^2 + \lambda|\vec{x}|^2 \tag{1.10}$$

Photometric properties of general imagers

Setting the derivative of Eq. (1.10) with respect to the elements of the vector $\vec{x}$ equal to zero, we have

$$\begin{aligned} 0 &= \nabla_x |\vec{y} - \boldsymbol{A}\vec{x}|^2 + \nabla_x \lambda |\vec{x}|^2 && (1.11) \\ &= \boldsymbol{A}^T \boldsymbol{A}\vec{x} - \boldsymbol{A}^T \vec{y} + \lambda \vec{x} && (1.12) \\ & && (1.13) \end{aligned}$$

or

$$\vec{x} = (\boldsymbol{A}^T \boldsymbol{A} + \lambda \boldsymbol{I})^{-1} \boldsymbol{A}^T \vec{y} \quad (1.14)$$

See, e.g.: https://en.wikipedia.org/wiki/Matrix_calculus

system matrix, A, for pinhole imager

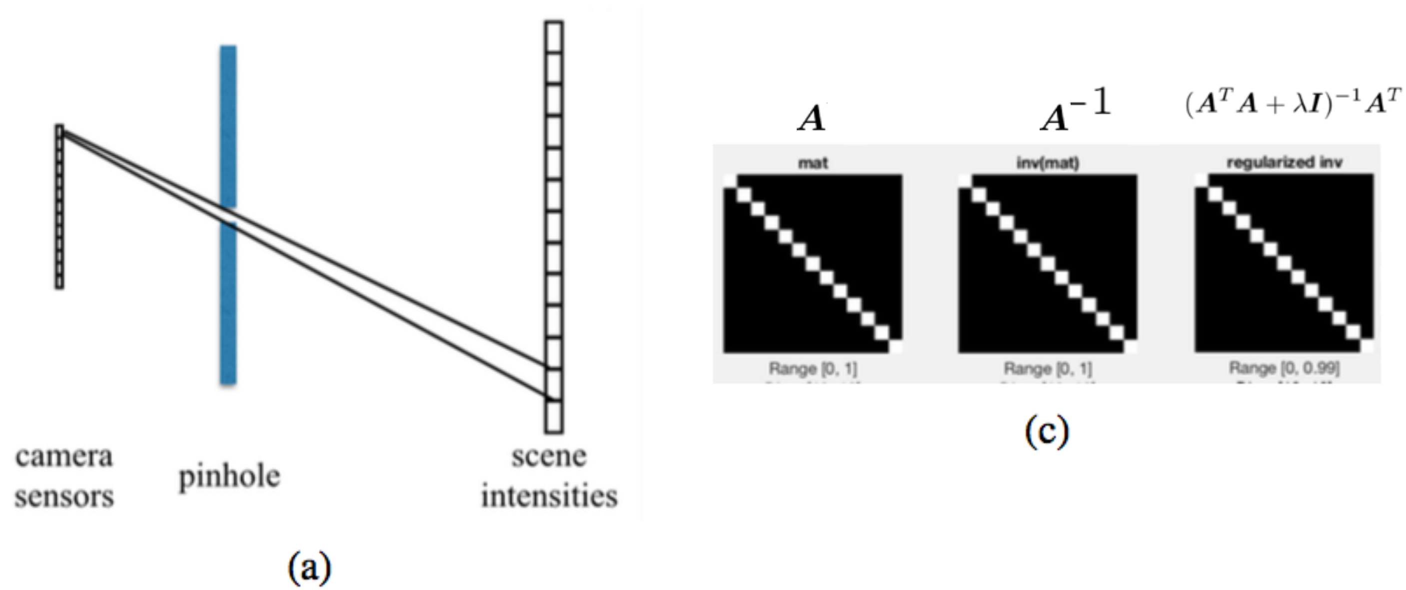


Figure 1.8
(a) Schematic drawing of a small-hole 1-d pinhole camera.(b) Visualization of imaging matrices: The imaging matrix relating scene intensities to sensor readings; the inverse of that matrix; the regularized inverse. For the small-pinhole imager, all three matrices are approximately identity matrices.

system matrix, A, for large aperture pinhole imager

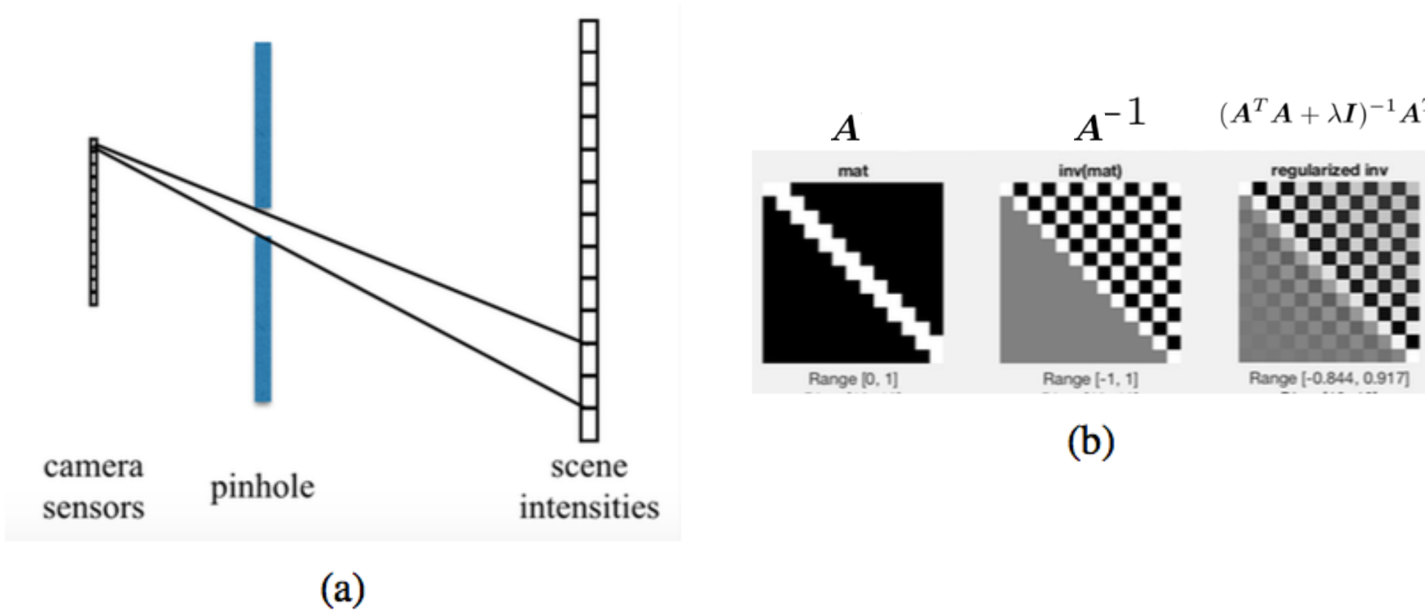
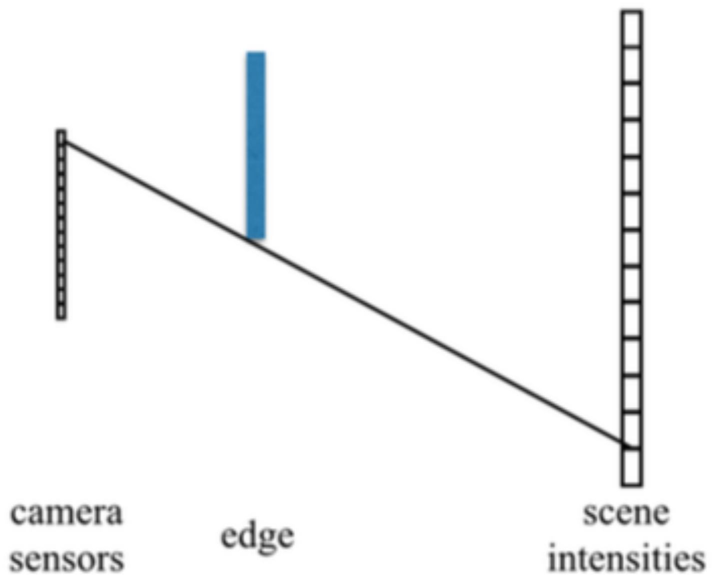
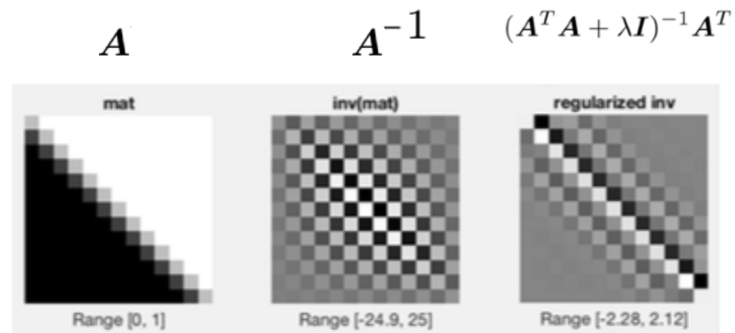


Figure 1.9
(a) Large-hole 1-d pinhole camera. (b) Visualization of imaging matrices: The imaging matrix relating scene intensities to sensor readings; the inverse of that matrix; the regularized inverse. For the small-pinhole imager, all three matrices are approximately identity matrices.

system matrix, A, for an edge

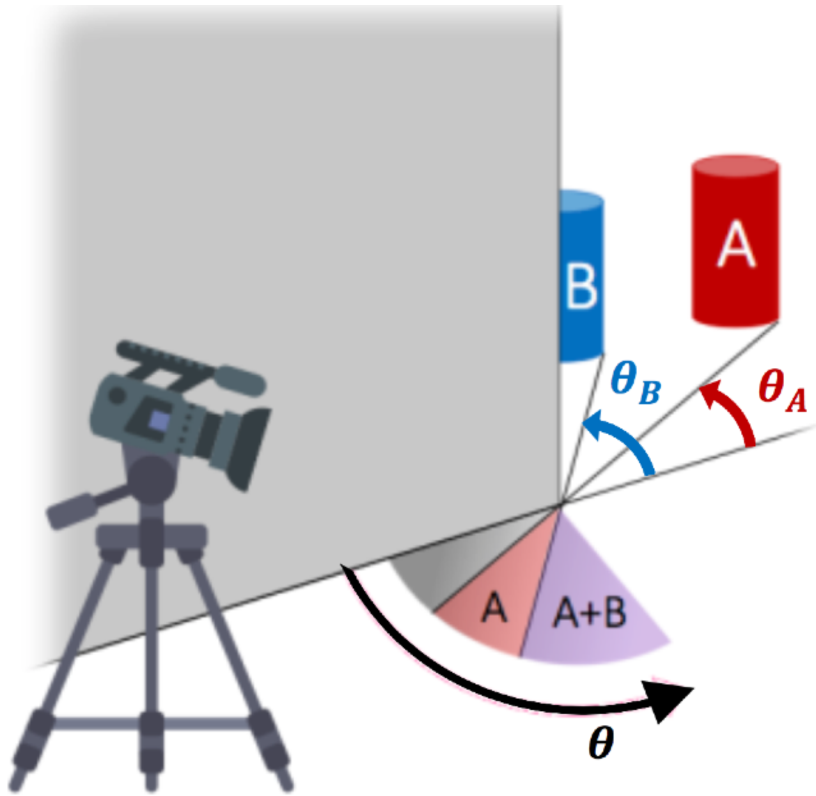


(a)



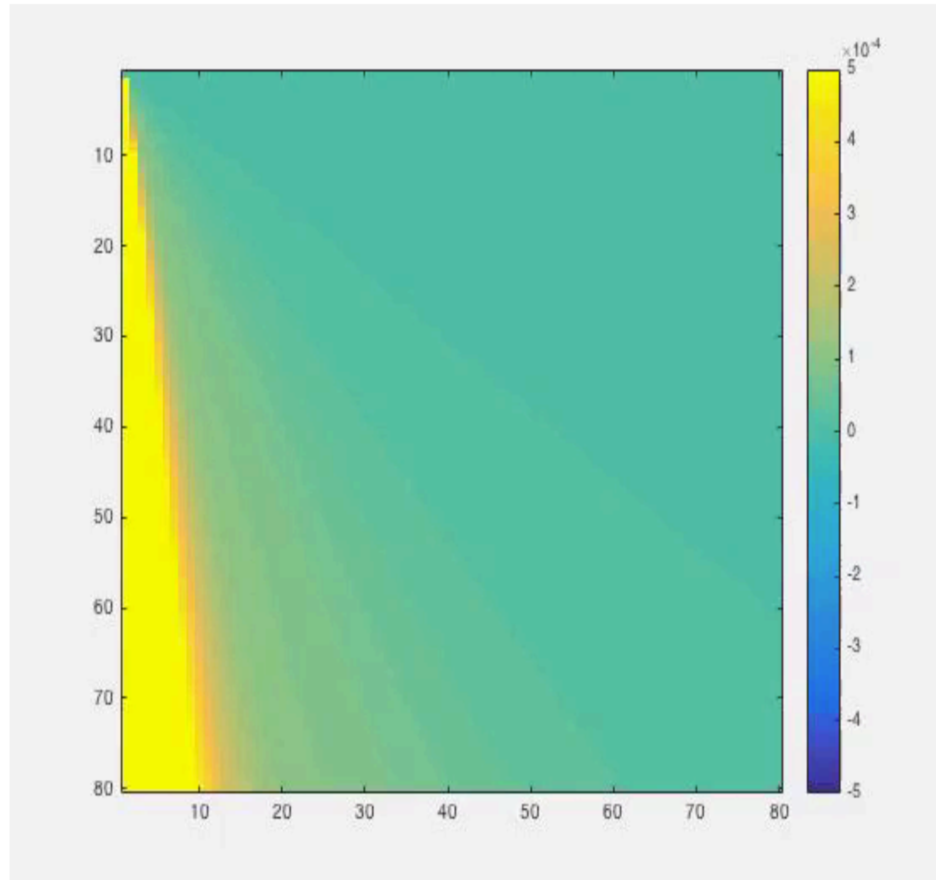
(c)

Another occlusion-based camera: edge camera

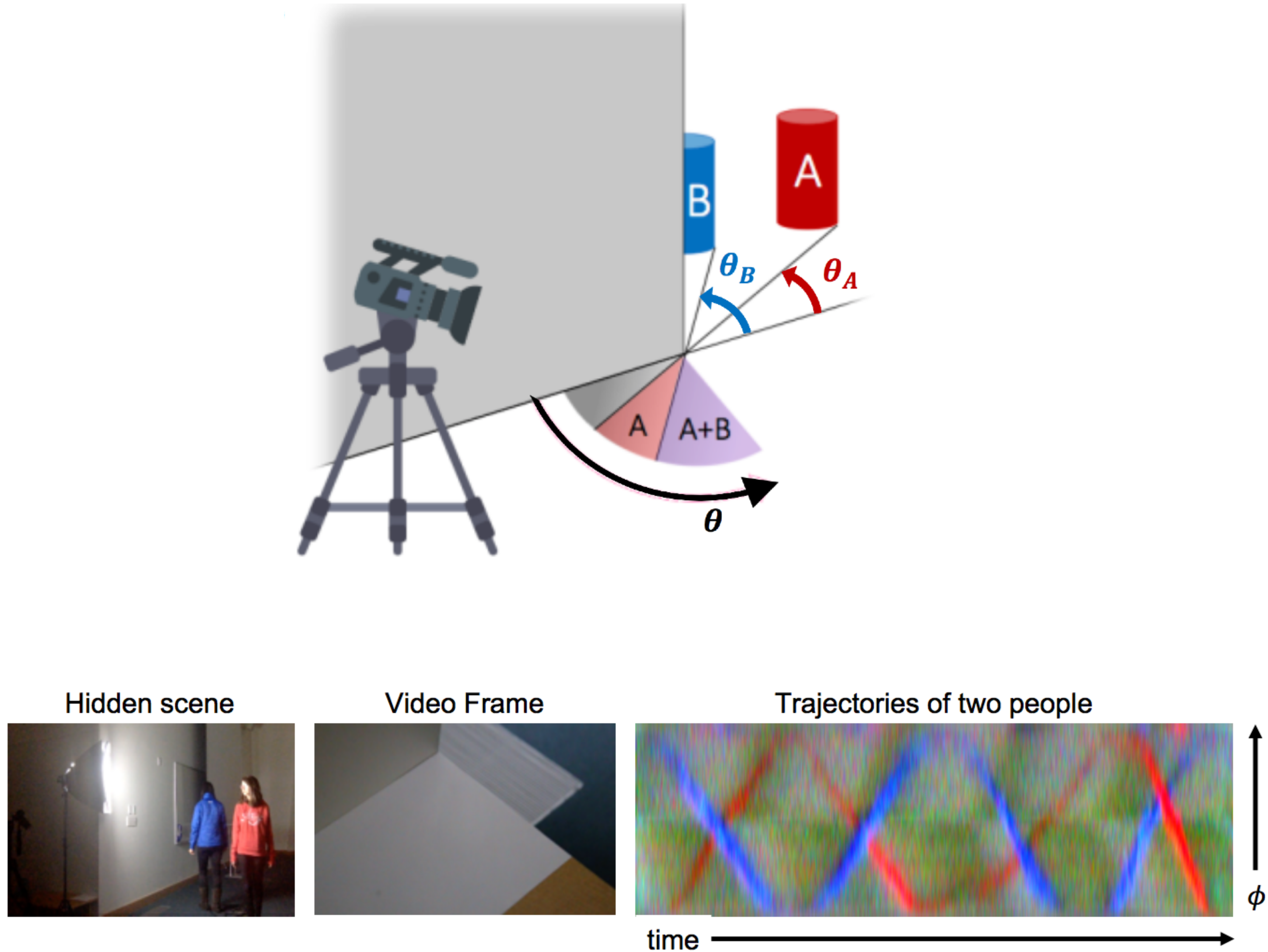


Corner Camera 1-D Image Computation

Rectified Image



Images you multiply the rectified ground plane images by to recover the input image around the corner (projected to 1d), for each different angle.


B
A
θ_B
θ_A
A
A+B
θ
Hidden scene
Video Frame
Trajectories of two people
time
ϕ

Experiment Proof of Concept

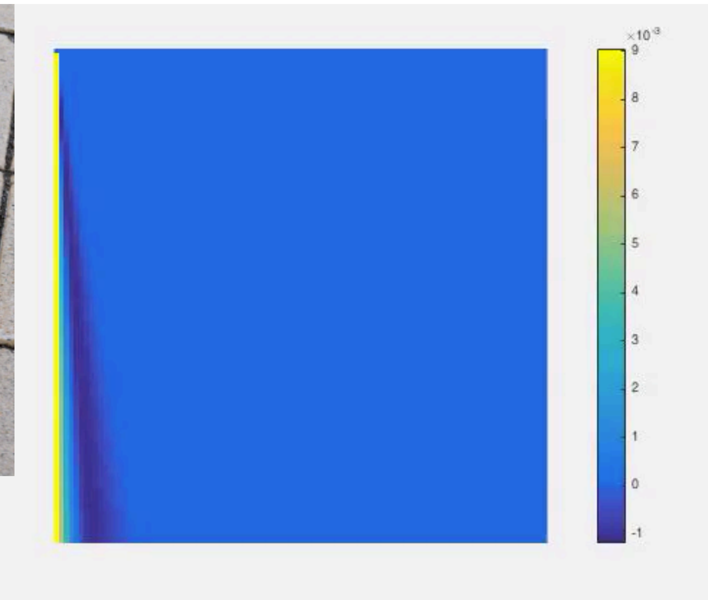
Experimental Proof of Concept

Experimental Proof of Concept

Experimental Proof of Concept

Video Corresponding to 1-D Camera

Corner camera 1-d image computation

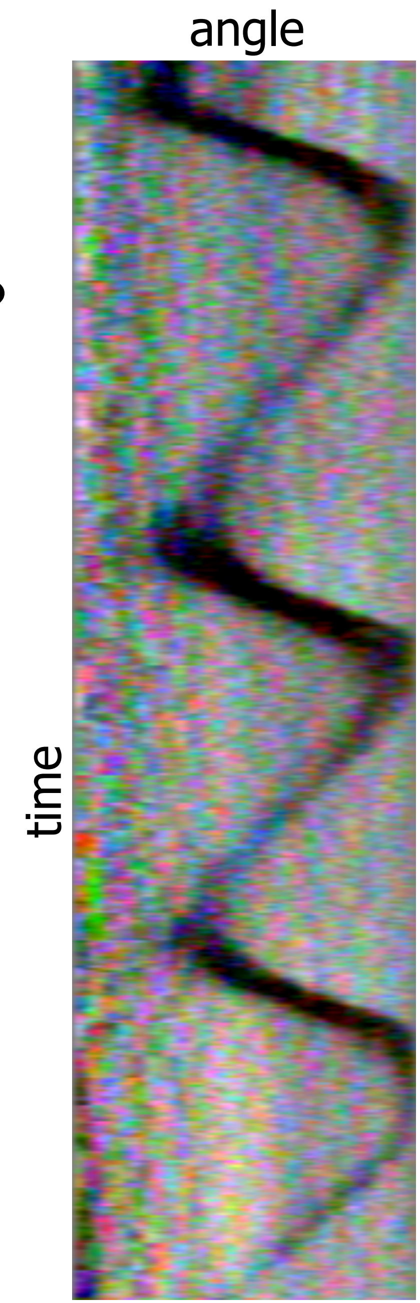


Input video image

mask images to read-out 1-d image of scene around the corner

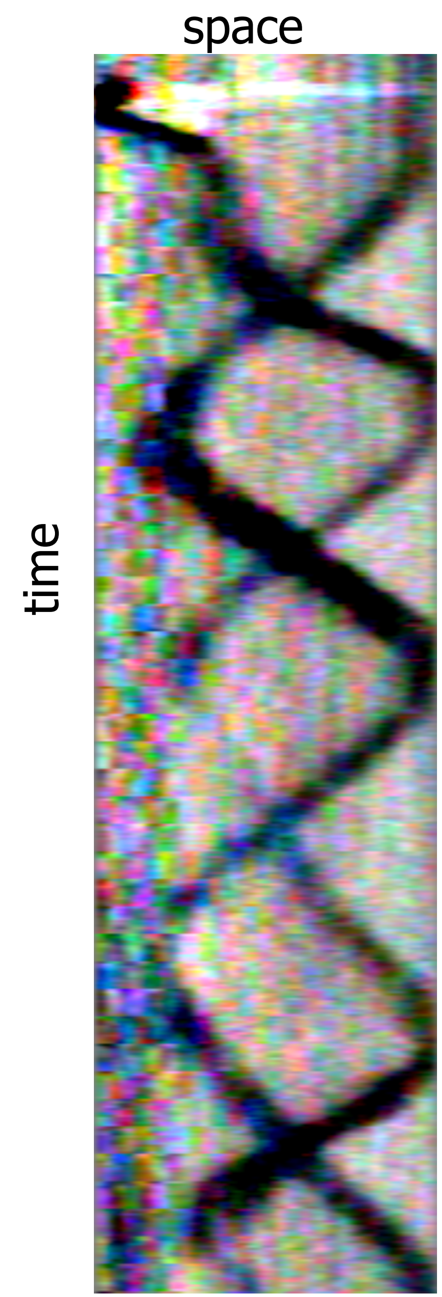
1-D Corner Camera Output

- How many people?
- Where slowed down, where moved quickly?



1-D Corner Camera Output

- How many people?
- How fast is each person moving?



Additional Results

Paper ID: 1983

Summary

- Pinhole camera models the geometry of perspective projection
- Lenses gather light and form images
- We designed a lens
  - Thin lens, spherical surfaces, first order optics
- Cameras as general linear systems.
  - specified by transfer matrix relating illumination in world to recorded data.
  - example: corner cameras