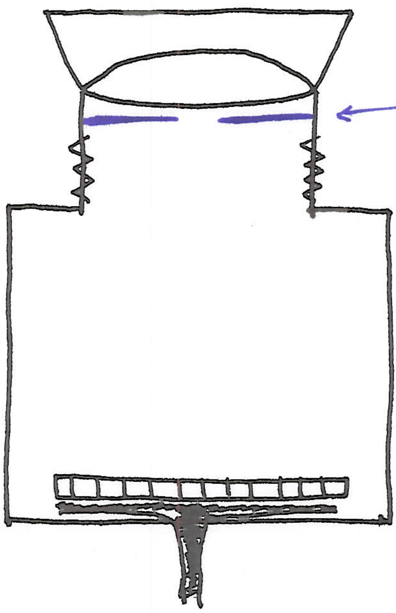
IMAGE FORMATION AND IMAGE SENSING

(CHAPTER 2)

In the synthetic world, we have a simpler & less complete picture. There's no eye, but there is a camera.



diaphragm / aperture

manual / automatic focal length adjustment (zoom).

CCD or other sensing device
- has timing mechanism

↓ analog or digital signal
- if analog, must be digitized for computer. usually by a framegrabber

functional aspects of sensing process (some may be fixed)

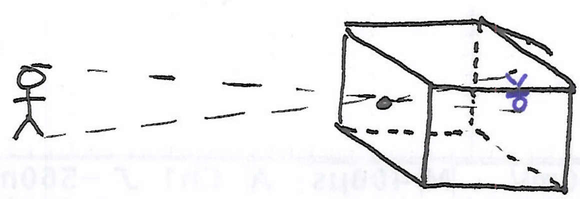
- data transmitted to computer via digital signal.
- lens curvature fixed, use focal length adjustments
  - move lens closer to / further from sensing device.
- have an aperture which can be opened / closed like iris.
- sensing device can control exposure timing (or gain).

- some manufacturers have designed foveating camera.
- if we want stereo, need to add another camera with appropriate geometry.

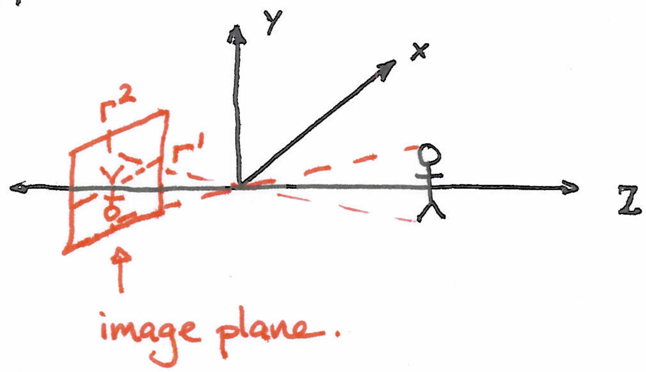
How does image formation work?

1. what light gets sensed?
2. how does camera design impact sensing?
3. how is light digitized?

Image formation: Projection Equations

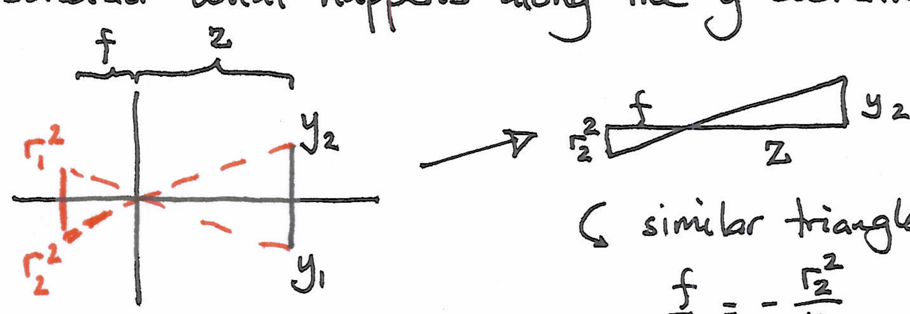


- rays of light pass through pinhole and form an inverted image on the surface.
- basically, this is like:



now, consider what happens along the y-coordinate

f - focal length



similar triangles

$$\frac{f}{Z} = -\frac{r_2^2}{y_2}$$

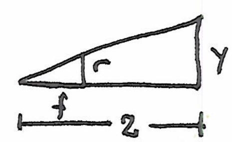
$\Rightarrow$

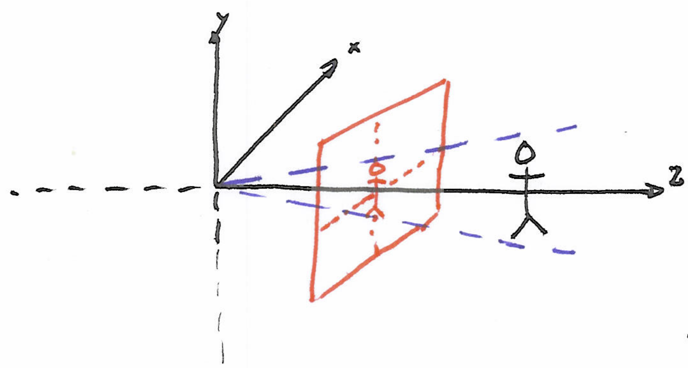
$$r_2^2 = -\frac{f y_2}{z} \qquad \text{likewise, (for fact)} \quad r_1^2 = -\frac{f y_1}{z}$$

A similar ~~ste~~ equation holds for the x/r^1 pairing. Thus, the point (x, y, z) projects onto

$$r^1 = -\frac{fx}{z} \qquad r^2 = -\frac{fy}{z} \tag{1}$$

If we move the location of the imaging plane from $z = -f$ to $z = f$, we get



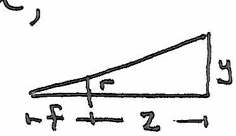


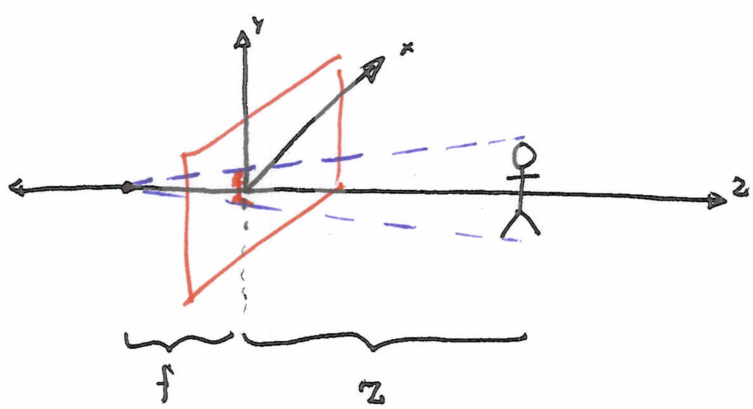
$$r^1 = \frac{fx}{z} \qquad r^2 = \frac{fy}{z} \tag{2}$$

- no need to worry about image reversal.

viewpoint centered coordinate system

Others consider an <u>image centered coordinate system</u>,





$$r^1 = \frac{fx}{f+z}, \quad r^2 = \frac{fy}{f+z} \tag{3}$$

- These different equations are called "perspective projection" equations. Equation (2) is the one typically used, but the others are equivalent.

- Location of the pinhole is called the center of projection.

- Note that if the focal length is much smaller than the distance, then (2) approximates (3).
  ($f \ll Z$)

* if the focal length is large compared to Z, or variation in Z is very small compared to f, then can use orthographic projection.

CASE: $f \gg Z$

$\Rightarrow$ (3)

$$r^1 = \frac{fx}{f+Z} \approx \frac{fx}{f} = x$$

$$r^2 = \frac{fy}{f+Z} \approx \frac{fy}{f} = y$$

CASE: $Z = Z_0 + \delta Z \wedge \delta Z \ll Z_0$

$\Rightarrow$ (2)

$$r^1 = \frac{fx}{Z_0 + \delta Z} \approx \frac{fx}{Z_0} = \frac{f}{Z_0} x = \mu x$$

$$r^2 = \frac{fy}{Z_0 + \delta Z} \approx \frac{fy}{Z_0} = \frac{f}{Z_0} y = \mu y$$

where $\mu = \dfrac{f}{Z_0}$

magnification factor.

$\Rightarrow$

$$r^1 = \mu x$$

$$r^2 = \mu y$$

- usually $\mu = 1$ in definition of orthographic projection, but need not be.

Pinhole has issues

- a point size hole is not possible, in practice one uses a very small circle.
- this lead to other imaging problems (diffraction).

Lenses used to overcome these issues.

What does the lens do?

- role of lens is to focus light rays
- ideal lens produces same projection as the pinhole for a given distance.

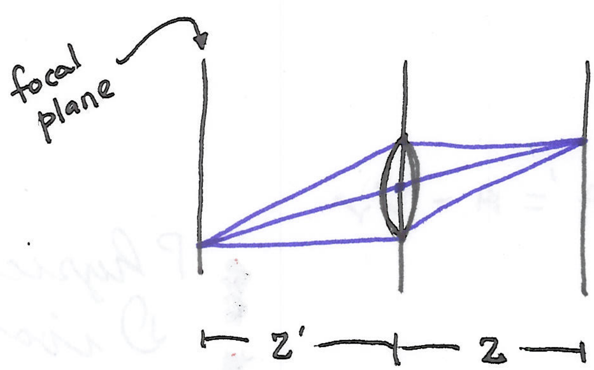
only focuses a part of the scene

all points satisfying $\frac{1}{z'} + \frac{1}{z} = \frac{1}{f}$

lens to image plane distance.

focal length of the lens.

(thin lens formula)

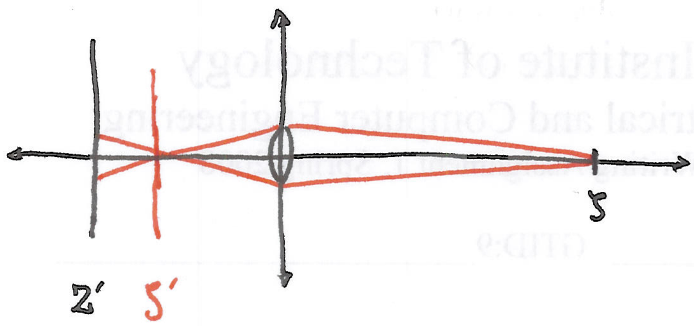


problem: focal distance z' varies with true distance z, but the focal plane is fixed.

⇒

points at other distances may not focus.

Let's consider a point located elsewhere.



⇒ instead of being at z', the imaged point is at ζ.

focal plane should be at ζ' not z'.

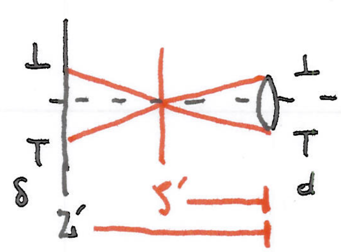
what happens if it is at z' then?

based on the image above, it does not project to a point, but across some region. can it be quantified?

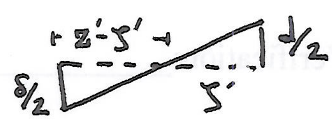
facts: ideal distance is z. focal plane at z'

true distance is ζ. focal distance is ζ'.

lens diameter is d.



⇒



using similar triangles,

$$\frac{d/2}{\zeta'} = \frac{\delta/2}{z'-\zeta'}$$

query: what is δ? (as a quantity)

this is degree of blurring.

$\Rightarrow$

$$\zeta'\delta = d(z' - \zeta')$$

$\Rightarrow$

$$\delta = \frac{d(z'-\zeta')}{\zeta'}$$

$\Rightarrow$ make independent of ζ too far or too close. only care about lengths, not sign.

$$\delta = \frac{d|z'-\zeta'|}{\zeta'}$$

$\Rightarrow$

$$\delta = d\left|\frac{z'}{\zeta'} - 1\right|$$

recall that $\frac{1}{z'} + \frac{1}{z} = \frac{1}{f}$

$\Rightarrow$

$$z' = \frac{fz}{z-f}$$

$\Rightarrow$

$$\delta = d\left|\frac{fz/(z-f)}{f\zeta/(\zeta-f)} - 1\right| = d\left|\frac{\cancel{f}z(\zeta-f)}{\cancel{f}\zeta(z-f)} - 1\right|$$

$$= d\left|\frac{z\zeta(1-f/\zeta)}{z\zeta(1-f/z)} - 1\right| = d\left|\frac{(1-f/\zeta)}{(1-f/z)} - 1\right|$$

$\Rightarrow$

$$\delta = d\left|\frac{1-f/\zeta}{1-f/z} - 1\right| \qquad (*)$$

wrong distance ζ gives rise to blurring of point, with blur radius of δ given by $(*)$.

- when dealing with a physical sensor, there is a finite region size within which a point light source is sensed (think of it as a light bucket).

  blurring up to the size of the bucket has no effect on the sensor's ability to detect.

  $\Rightarrow$

  for physical sensors, there is a range of distances below and above z for which an object is still in focus,

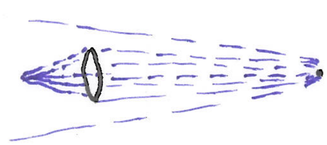
  $$\zeta \in [z_{min}, z_{max}] \Rightarrow \text{point at } \zeta \text{ in focus.}$$

- 'depth of field' = range of distances over which objects are focused sufficiently well.

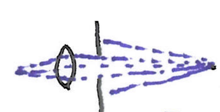
- the aperture affects depth of field.

  by restricting what light rays enter, it reduces the effective lens diameter. $\Rightarrow$ larger depth of field.

  also reduces quantity of incoming light $\Rightarrow$ dimmer image.



w/out aperture



some rays are blocked.

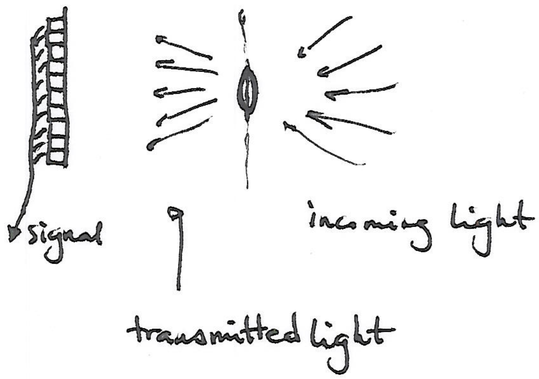
w/aperture

- larger lens $\Rightarrow$ less depth of field.

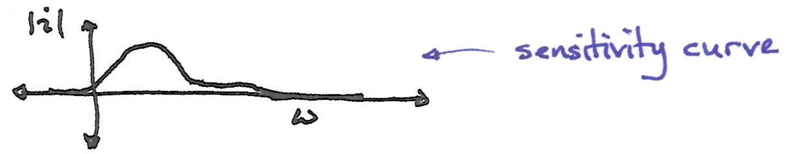
Image Sensing: Capturing the Light.

- light from world projects onto image plane.
- at the image plane, a gridded array of photosensitive cells exists.

  converts photons to voltages



- different sensors have different sensitivities in the electromagnetic spectrum. UV, IR, Visible, ...

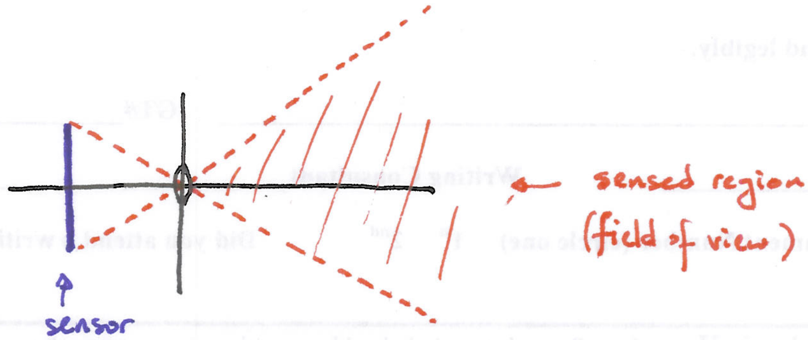


- gridded array results in quantization of image data into picture cells (pixels).
  - records/measures average irradiance over area of sensor element/cell.

- finite extent of sensor dimensions ⇒ can only sense part of the scene.

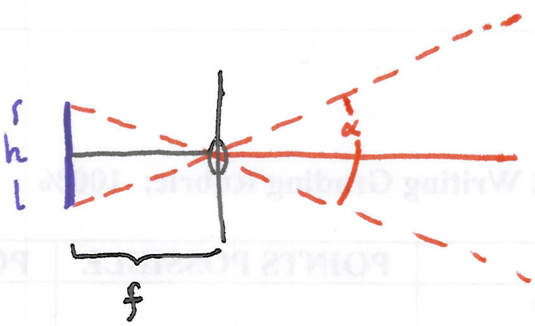
  - sensed portion of world is called "field of view".

field of view determined by camera geometry

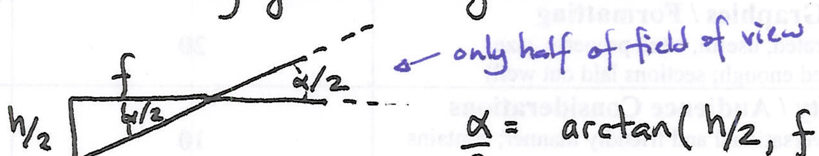


a light ray must land on the sensor to be sensed (the blue region)

$\Rightarrow$



field of view usually given in angular units.



$$\frac{\alpha}{2} = \arctan(h/2, f)$$

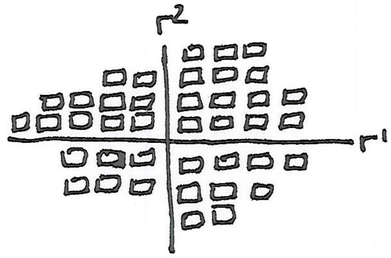
$$\Rightarrow \quad \alpha = 2\arctan(h/2, f)$$

for a rectangular sensor of dimensions $h \times w$, then

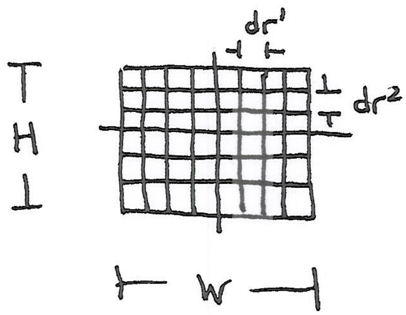
$\alpha' = 2\arctan(w/2, f)$ horizontal field of view.

$\alpha^2 = 2\arctan(h/2, f)$ vertical field of view.

the sensed image is a quantized version of the projected image.



if light focuses onto (r^1, r^2) then we need to convert it to pixels.



$$R^1 = \text{floor}(r^1/dr^1) + \frac{W}{2}$$

$$R^2 = \text{floor}(r^2/dr^2) + \frac{H}{2}$$

for H,W even

OR

(H,W) - size of pixel array called "resolution" of the image.

$$R^1 = \text{floor}(r^1/dr^1) + \frac{W-1}{2}$$

$$R^2 = \text{floor}(r^2/dr^2) + \frac{H-1}{2}$$

for H,W odd

- pixel captures light from angular extent

$$d\alpha^1 = 2\arctan\left(\frac{r^1 + dr^1}{2f}\right) - 2\arctan\left(\frac{r^1}{2f}\right)$$

$\Rightarrow$

$$d\alpha^1 = 2 \cdot \frac{dr^1}{2f} \quad \frac{\arctan\left(\frac{r^1}{2f} + \frac{dr^1}{2f}\right) - \arctan\left(\frac{r^1}{2f}\right)}{dr^1/2f}$$

$$\approx 2\frac{dr^1}{2f} \cdot \frac{1}{1 + \left(\frac{r^1}{2f}\right)^2} = \frac{2 \cdot dr^1 \cdot 2f}{2f + (r^1)^2}$$

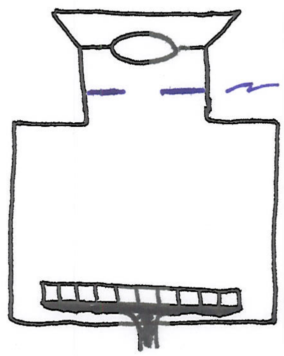
$\Rightarrow$

$$d\alpha^1 \approx \frac{4f\, dr^1}{2f + (r^1)^2} \qquad d\alpha^2 \approx \frac{4f\, dr^2}{2f + (r^2)^2}$$

- thus, for ~~a~~ fixed sensor dimensions, the pixel dimensions (W, H) directly determine (dr^1, dr^2) which affects the ability to see small angular extent, e.g., ability to resolve small objects/details.

Hence the name resolution.

... back to the sensor system ...



} affects light gathering ability.

sensor may need to adjust based on light.

sensing parameters

- can adjust timing to allow more/less light
- can adjust sensitivity (gain adjustment)

can have adverse effects, increased shot noise.

(see online reviews of cameras and effect of varying ASA/ISO)

alternatively, image processing can be used to adjust the image.

- histogram stretching/equalization
- contrast enhancement, sharpening
- brightness adjustments

- may still have sensitivity issues, but they may be controllable.

In the end, we get an image function $I: \mathbb{R}^2 \to D$

where $D = \mathbb{R}$ grayscale
$D = \mathbb{R}^3$ RGB color
$D = \mathbb{R}^n$ vector valued / multispectral

to emphasize discrete domain of pixels, $I: \mathbb{Z}^2 \to D$