ECE 1390/2390
Image Processing and Computer Vision – Fall 2021

Camera Models and *Perspective Imaging*

Ahmed Dallal
Assistant Professor of ECE
University of Pittsburgh

Reading

- FP, Chapter 1, 2.1, and 2.2

Cameras and images

What is an image?

- Up until now: a function – a 2D pattern of intensity values
- Today: a 2D projection of 3D points

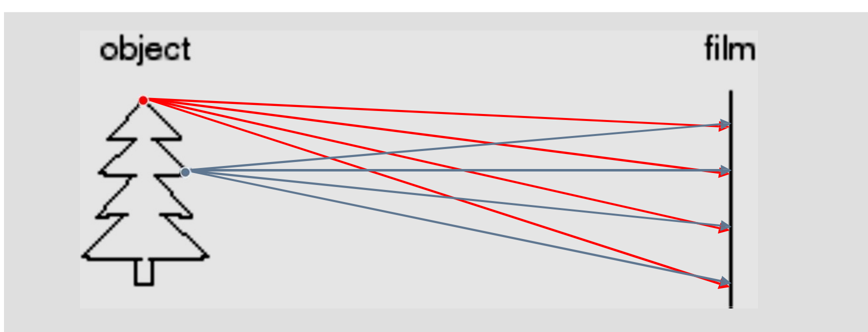
What is a camera/imaging system?

- Some device that allows the projection of light from 3D points to some "medium" that will record the light pattern.
- A key to this is "projection"…

Projection

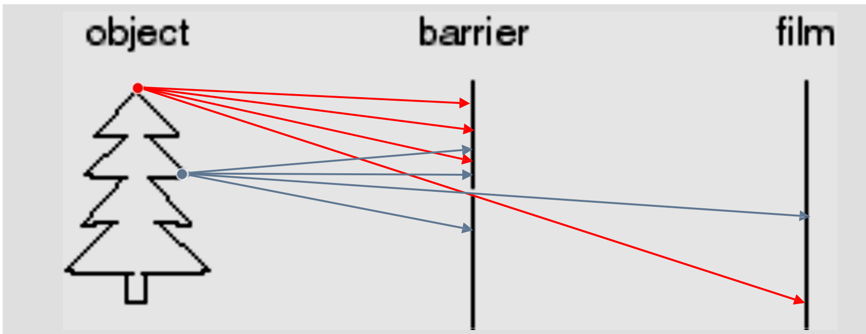
Projection

Image formation – (bad) method



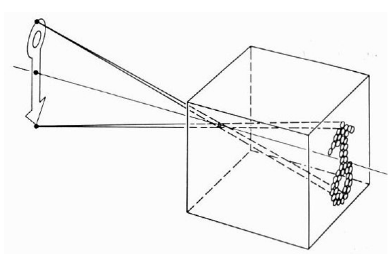
- Let’s design a camera
  - Idea 1: put a piece of film in front of an object
  - Do we get a reasonable image?

Pinhole camera



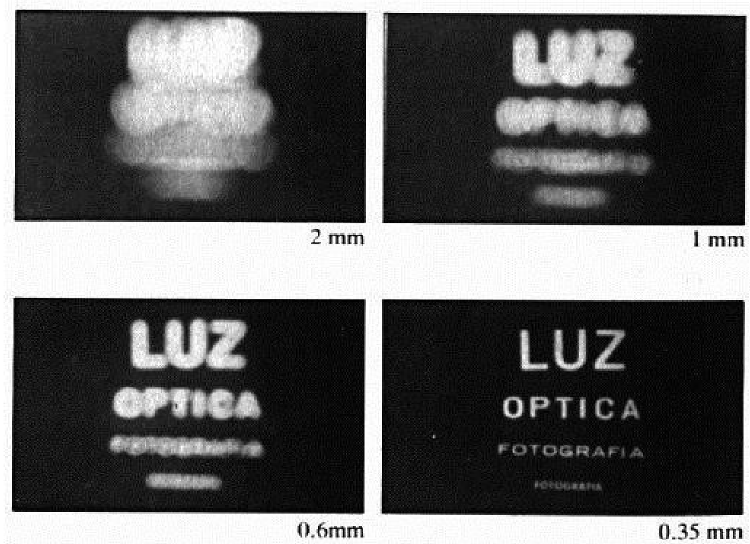
- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening known as the **aperture**
  - How does this transform the image?

Camera Obscura (Latin: Darkened Room)

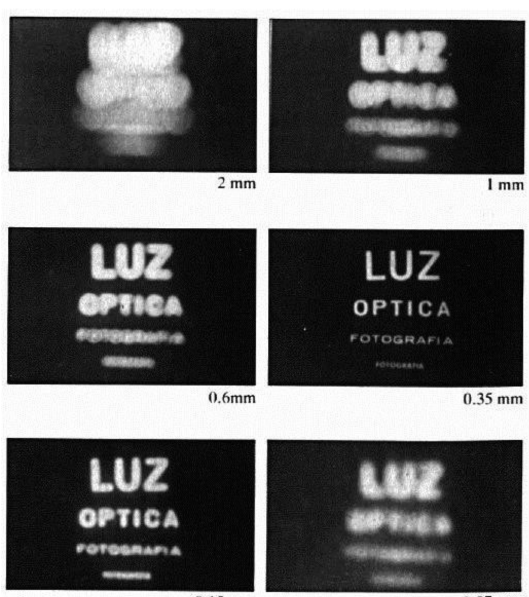


- The first camera
  - Known to Aristotle (384-322 BCE)
  - According to DaVinci "When images of illuminated objects ... penetrate through a small hole into a very dark room ... you will see [on the opposite wall] these objects in their proper form and color, reduced in size, in a reversed position, owing to the intersection of the rays".
  - Depth of the room is the "focal length"
  - How does the aperture size affect the image?

Shrinking the aperture

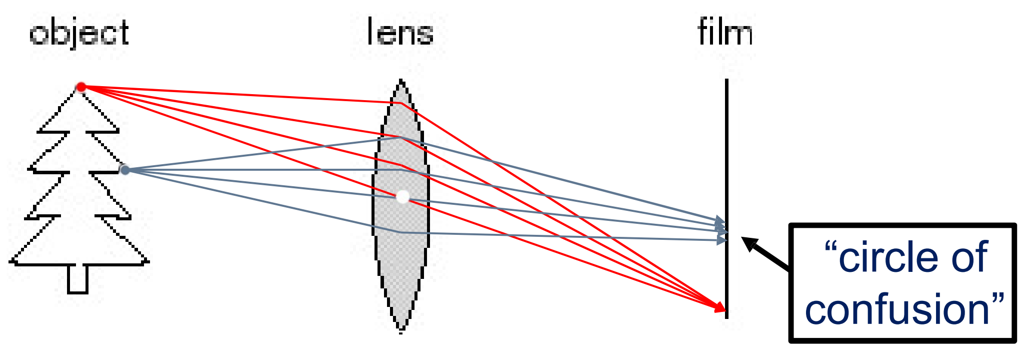


Shrinking the aperture



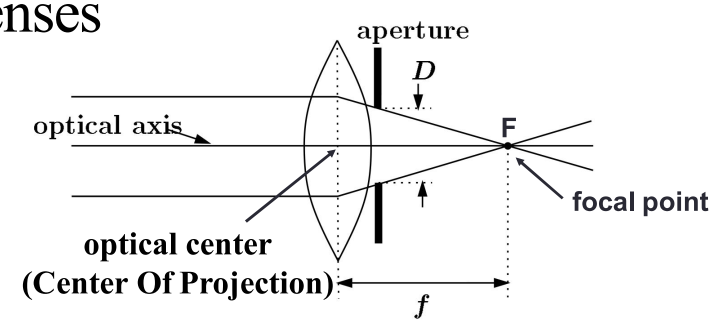
A little bit of computational photography

Adding a lens – and concept of focus



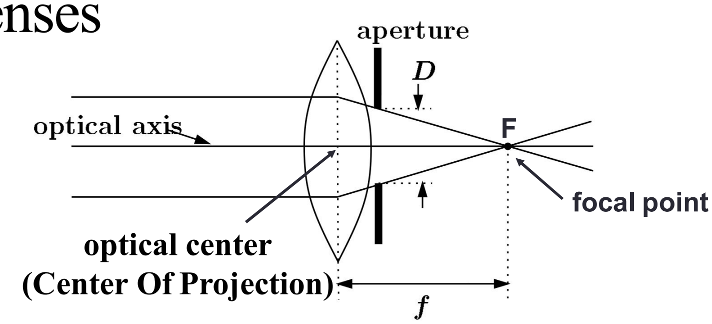
- A lens focuses light onto the film
  - There is a specific distance at which objects are "in focus"
    - other points project to a "circle of confusion" in the image
  - Changing the shape of the lens changes this distance

Lenses



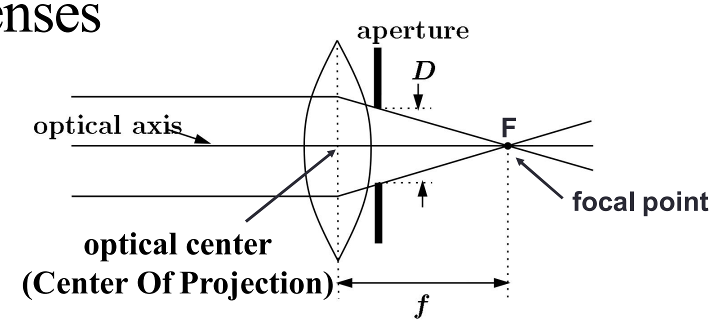
- A lens focuses parallel rays onto a single focal point
  - focal point at a distance *f* beyond the plane of the lens
    - *f* is a function of the shape and index of refraction of the lens

Lenses



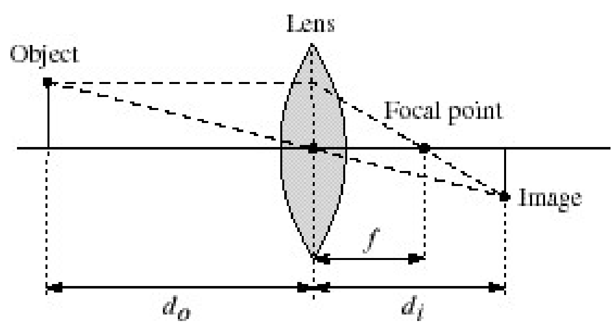
- A lens focuses parallel rays onto a single focal point
  - Aperture of diameter D restricts the range of rays
    - aperture may be on either side of the lens

Lenses



- A lens focuses parallel rays onto a single focal point
  - Lenses used to be typically spherical (easier to produce) but now many “aspherical” elements – designed to improve variety of “aberrations”.

Thin lenses



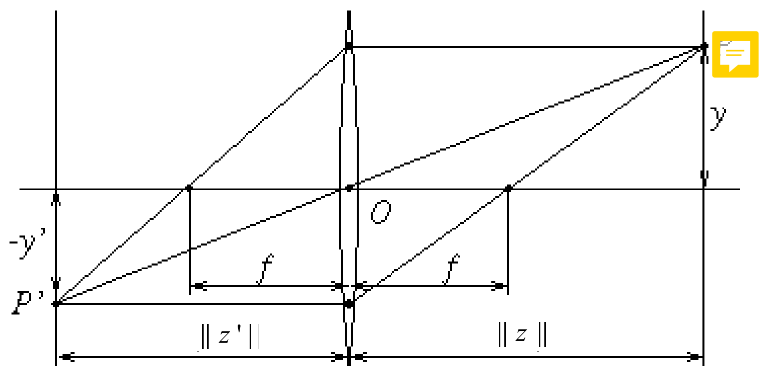
Thin lens model assumes thickness is small compared to curvature:

1. Any ray parallel to the axis on one side of the lens passes through the focal point on the other side.
2. Any ray that passes through the center of the lens will not change its direction.

This gives rise to the “thin lens equation”…

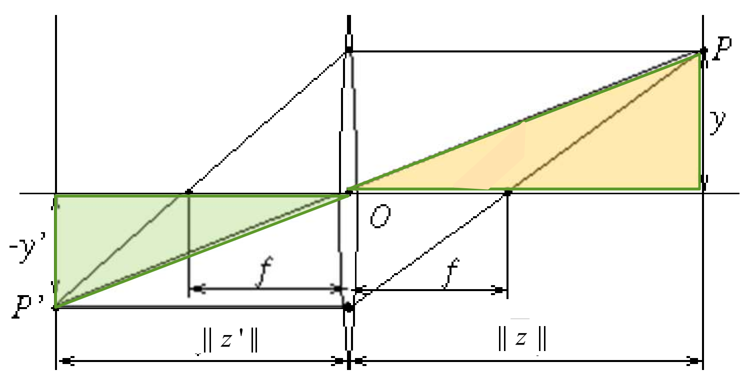
Slide by Steve Seitz

The thin lens



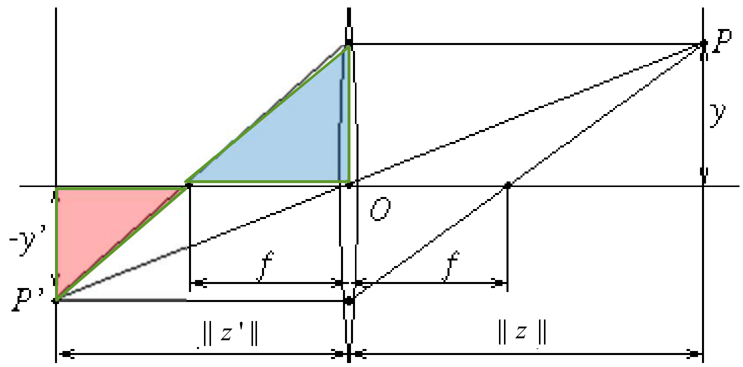
Computer Vision - A Modern Approach
Slides by D.A. Forsyth

The thin lens



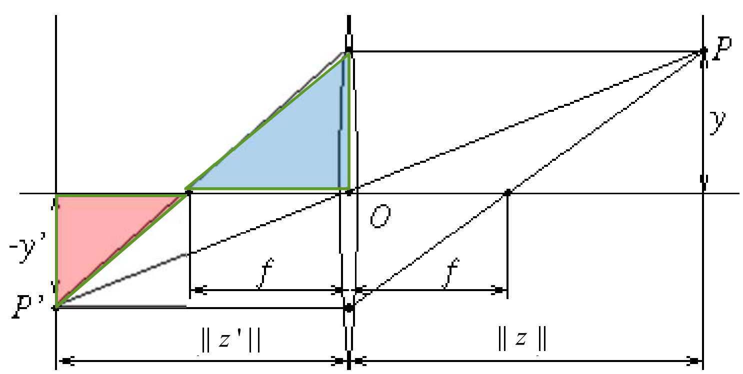
$$\frac{-y'}{y} = \frac{||z'||}{||z||}$$

The thin lens



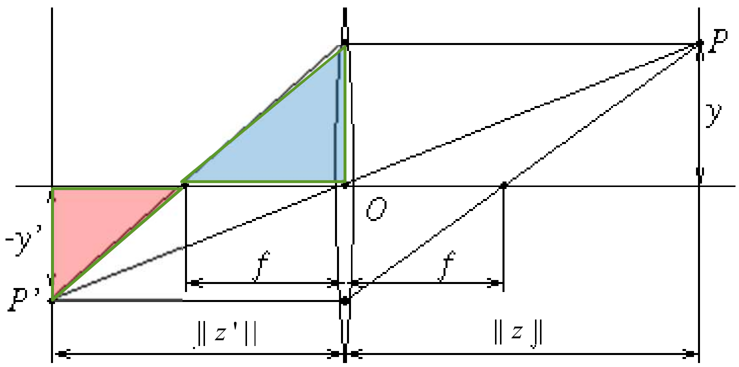
$$\frac{-y'}{y} = \frac{\| z' \|}{\| z \|} \qquad \frac{-y'}{y} = \frac{\| z' \| - f}{f} \qquad \boxed{\rightarrow \frac{\| z' \|}{\| z \|} = \frac{\| z' \| - f}{f}}$$

The thin lens equation



$$\frac{\| z' \|}{\| z \|} = \frac{\| z' \| - f}{f} \qquad \rightarrow \frac{1}{\| z \|} = \frac{1}{f} - \frac{1}{\| z' \|} \qquad \boxed{\rightarrow \frac{1}{\| z' \|} + \frac{1}{\| z \|} = \frac{1}{f}}$$

The thin lens equation



Any object point satisfying this equation is in focus.

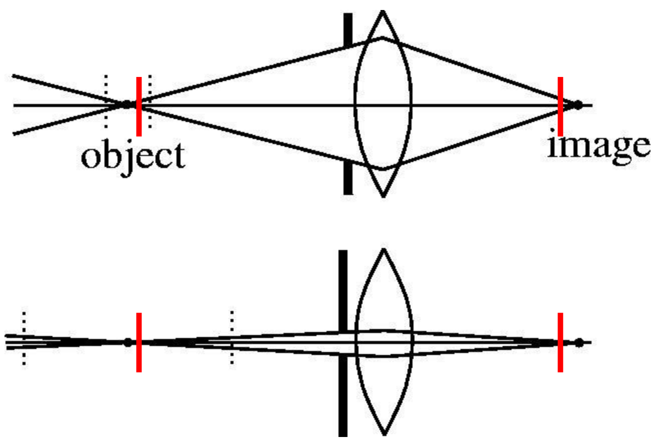
$$\rightarrow \boxed{\frac{1}{\|z'\|} + \frac{1}{\|z\|} = \frac{1}{f}}$$

Thin lenses

http://www.phy.ntnu.edu.tw/java/Lens/lens_e.html

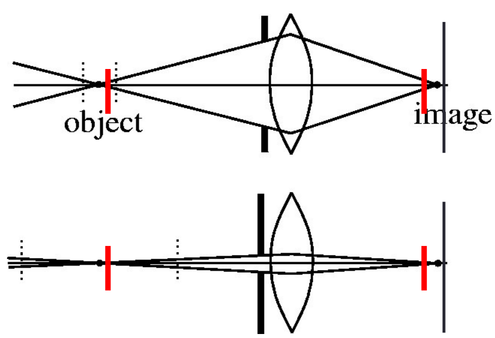
(by Fu-Kwun Hwang)

Depth of field



http://en.wikipedia.org/wiki/Depth_of_field

Depth of field



f / 5.6

f / 32

- Changing the aperture size affects depth of field
  - A smaller aperture increases the range in which the object is approximately in focus
  - But small aperture reduces amount of light – need to increase exposure

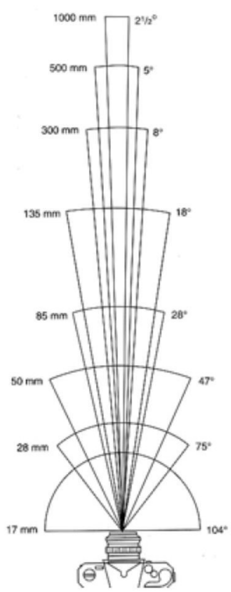
Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field

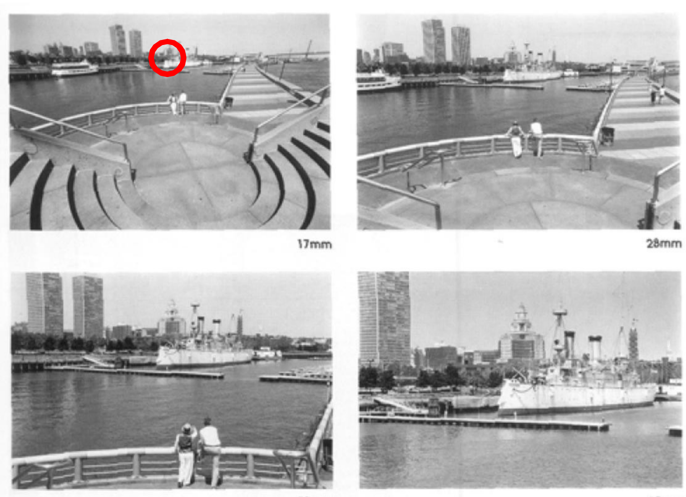
Varying the aperture

Large apeture = small DOF

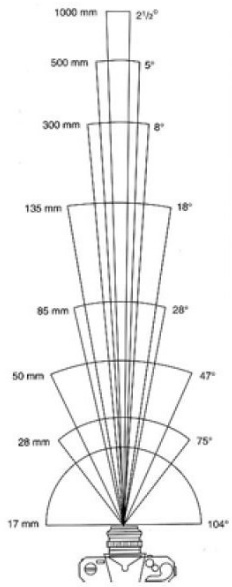
Small apeture = large DOF

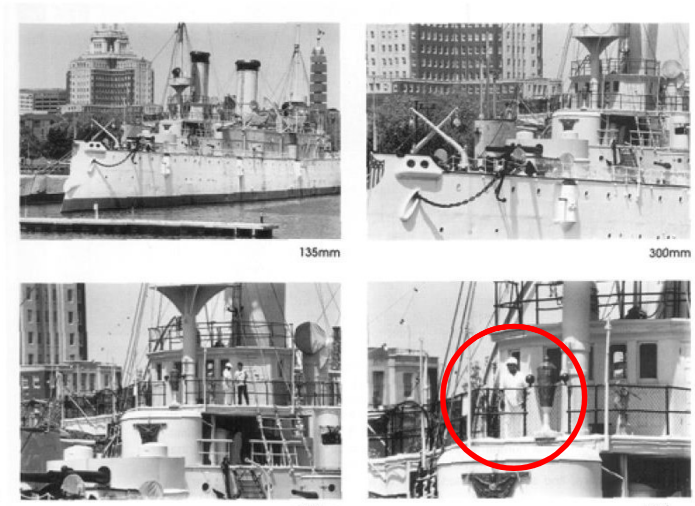
Field of View (Zoom)



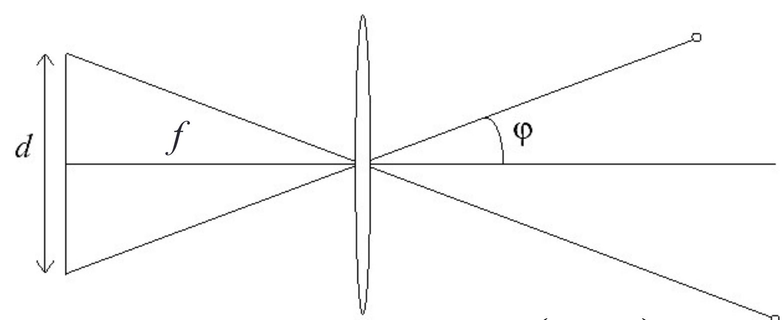


Field of View(Zoom)





FOV depends on Focal Length



d is the “retina” or sensor size

$$\phi = 2\tan^{-1}\left(\frac{d/2}{f}\right)$$

Larger Focal Length => Smaller FOV

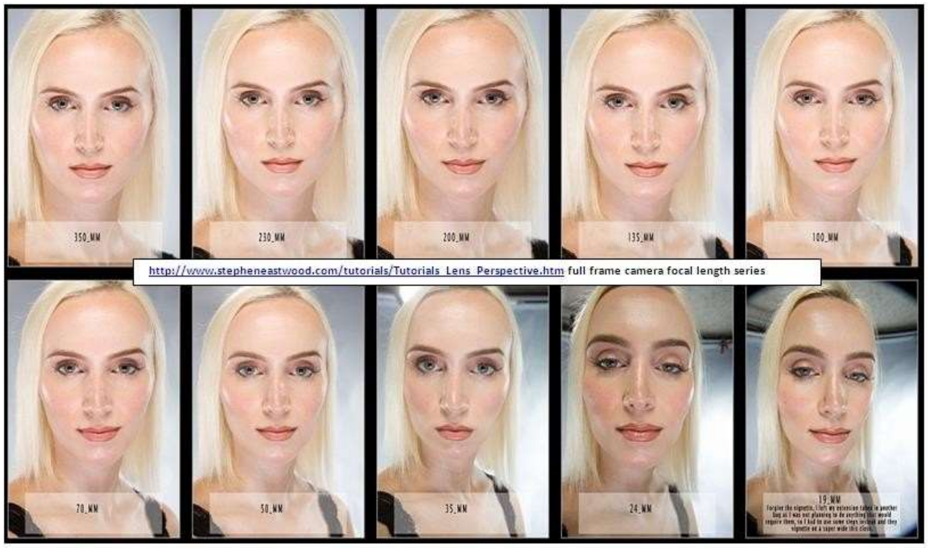
Zooming and Moving are not the same…

Field of View / Focal Length

Large FOV, small *f*
Camera close to car

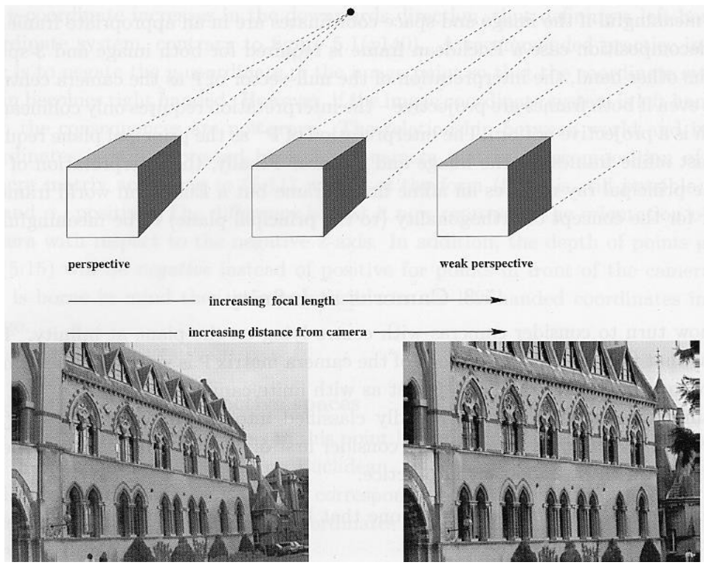
Small FOV, large *f*
Camera far from the car

Perspective and Portraits



Perspective and Portraits

Effect of focal length on perspective effect



Dolly Zoom

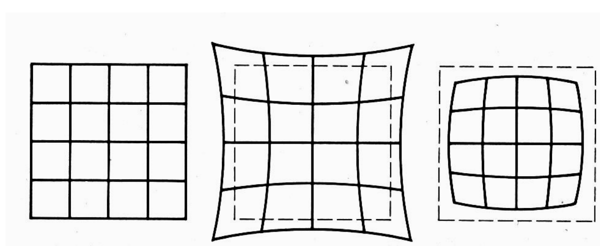
- Cinematic techniques

https://www.studiobinder.com/blog/best-dolly-zoom-vertigo-effect/

But reality can be a problem…

- Lenses are not thin
- Lenses are not perfect
- Sensing arrays are almost perfect
- Photographers are not perfect

Geometric Distortion



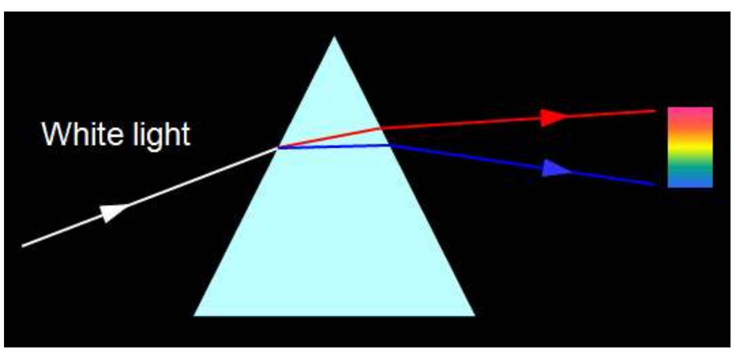
No distortion Pin cushion Barrel

- Radial distortion of the image
  - Caused by imperfect lenses
  - Deviations are most noticeable for rays that pass through the edge of the lens

Correcting radial distortion

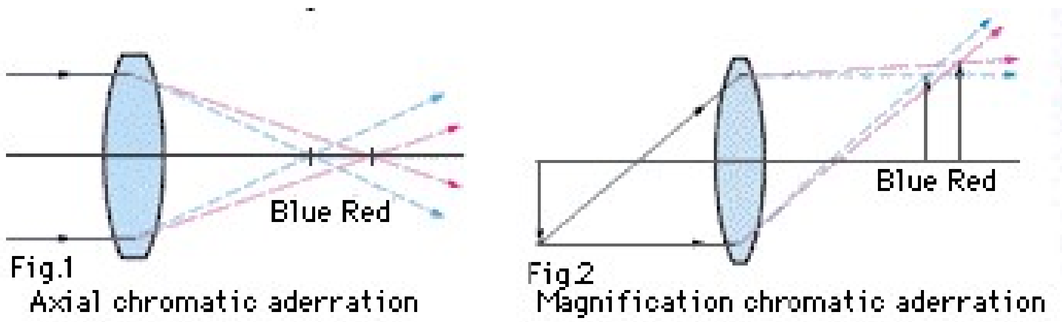
from Helmut Dersch

Chromatic Aberration



Rays of different wavelength
focus in different planes

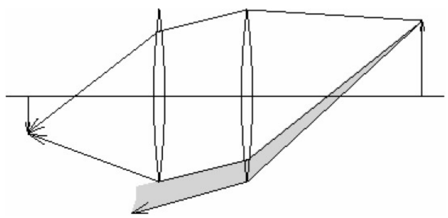
Chromatic Aberration



Rays of different wavelength
focus in different planes

Chromatic Aberration

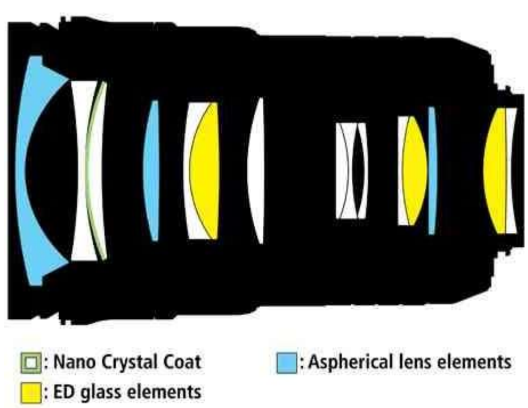
Vignetting



- Some light misses the lens or is otherwise blocked by parts of the lens

Lens systems

Nikon 24-70mm zoom

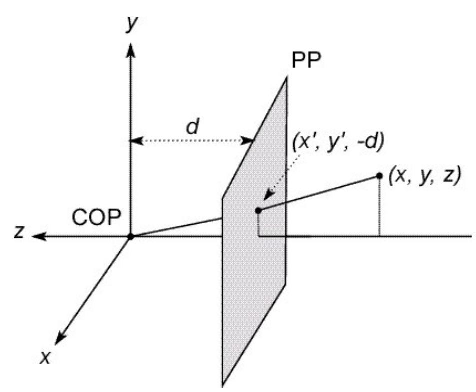


- Real lenses combat these effects with multiple elements.
- Computer modeling has made lenses lighter and better.
- Special glass, aspherical elements, etc.

Perspective imaging

Modeling projection – coordinate system

- We will use the pin-hole model as an approximation
- Put the optical center (**C**enter **O**f **P**rojection) at the origin
- STANDARD (x,y) COORDINATE SYSTEM
- Put the image plane (**P**rojection **P**lane) *in front* of the COP
- The camera looks down the *negative* z axis
  - we need this if we want right- handed-coordinates



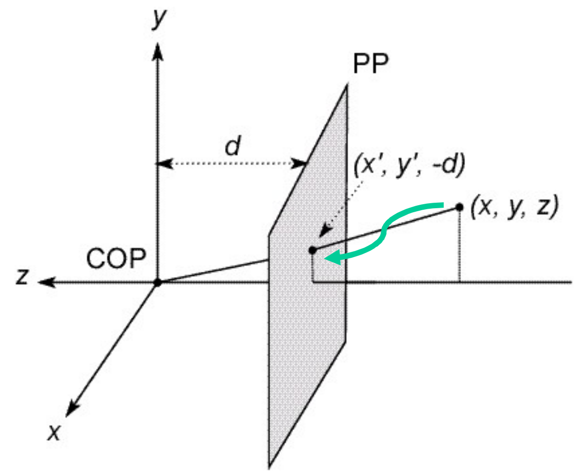
Modeling projection

Projection equations

- Compute intersection with Perspective Projection of ray from (x,y,z) to COP
- Derived using similar triangles

$$(X, Y, Z) \rightarrow (-d\frac{X}{Z}, -d\frac{Y}{Z}, -d)$$

(assumes normal Z negative – we'll change later)



Modeling projection

Projection equations

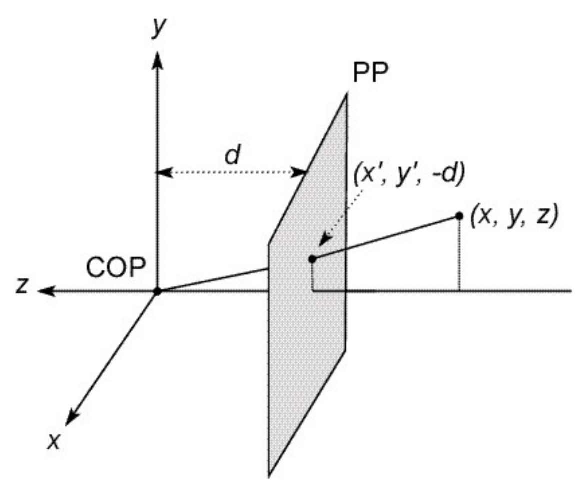
$$(X, Y, Z) \rightarrow (-d\frac{X}{Z}, -d\frac{Y}{Z}, -d)$$

We get the projection by throwing out the last coordinate:

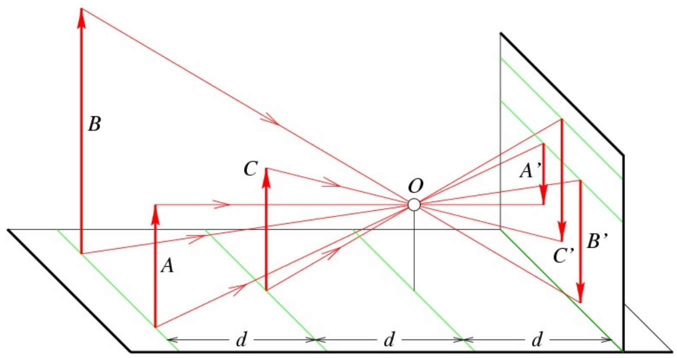
$$(x, y, z) \rightarrow (-d\frac{x}{z}, \ -d\frac{y}{z})$$

$$(x', y') = (-d\frac{X}{Z}, -d\frac{Y}{Z})$$

Distant objects are smaller



Distant objects appear smaller



Quiz

- When objects are very far away, the real X and real Y can be huge. If I move the camera (the origin) those numbers hardly change. This explains:

a) Why the moon follows you.

b) Why the North Star is always North.

c) Why you can tell time from the Sun regardless of where you are?

d) All of the above.

Homogeneous coordinates

Is this a linear transformation?

No – division by the (not constant) Z is non-linear

Trick: add one more coordinate:

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous image (2D) coordinates

homogeneous scene (3D) coordinates

Homogeneous coordinates

Converting *from* homogeneous coordinates:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w) \qquad \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

(this makes homogenous coordinates invariant under scale)

Perspective Projection

- Projection is a matrix multiply using homogeneous coordinates:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow \left(f\frac{x}{z}, f\frac{y}{z} \right)$$

$$\Rightarrow (u, v)$$

S. Seitz

Perspective Projection

- Projection is a matrix multiply using homogeneous coordinates:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow \left(f\frac{x}{z}, f\frac{y}{z} \right)$$

$$\Rightarrow (u, v)$$

This is known as perspective projection

- The matrix is the projection matrix
- The matrix is only defined up to a scale
- *f* is for “focal length – used to be *d*

S. Seitz

Perspective Projection

- How does scaling the projection matrix change the transformation?

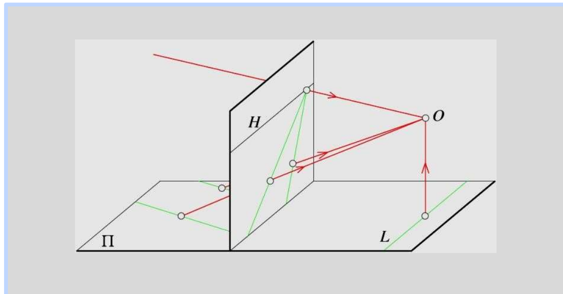
$$\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} fx \\ fy \\ z \end{bmatrix} \Rightarrow \left(f\frac{x}{z}, f\frac{y}{z} \right)$$

Perspective Projection

```
% Project a point from 3D to 2D using a matrix operation

%% Given: Point p in 3-space [x y z], and focal length f
%% Return: Location of the projected point on 2D image plane [u v]
function p_img = project_point(p, f)
    %% Define projection matrix (size: 3x4)
    A = [f 0 0 0;
         0 f 0 0;
         0 0 1 0];

    %% Convert p to homogeneous coordinates and transpose (size: 4x1)
    p_hom = [p 1]'; % Note: Here single quote (') is the transpose operator

    %% Apply projection transformation
    p_proj = A * p_hom; % Note: Use * not .* (NOT element-wise multiplication)

    %% Convert back to non-homogeneous coordinates and return
    p_img = [(p_proj(1) / p_proj(3)) (p_proj(2) / p_proj(3))];
endfunction
```

Geometric properties of projection

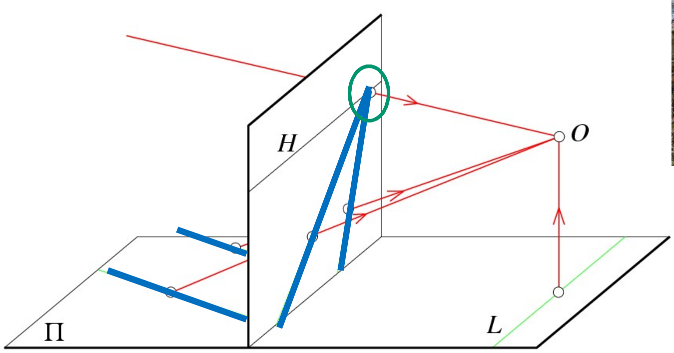
- Points go to **points**
- Lines go to **lines**
- Polygons go to **polygons**



- Degenerate case:
  - line in the world through focal point yields **point**

Parallel lines in the world meet in the image

"**Vanishing**" point



Parallel lines in the world meet in the image

Parallel lines converge in math too…

Image plan

Line in 3-space

$$x(t) = x_0 + at$$
$$y(t) = y_0 + bt$$
$$z(t) = z_0 + ct$$

Perspective projection of the line

$$x'(t) = \frac{fx}{z} = \frac{f(x_0 + at)}{z_0 + ct}$$

$$y'(t) = \frac{fy}{z} = \frac{f(y_0 + bt)}{z_0 + ct}$$

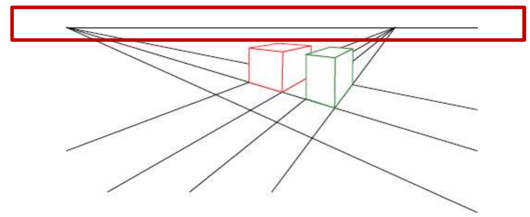
In the limit as $t \rightarrow \pm\infty$ we have (for $c \neq 0$):

$$x'(t) \rightarrow \frac{fa}{c}, \quad y'(t) \rightarrow \frac{fb}{c}$$

All points will converge to x' and y'

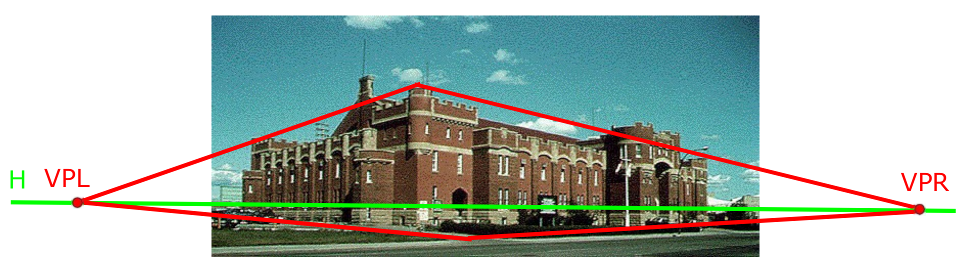
Vanishing points

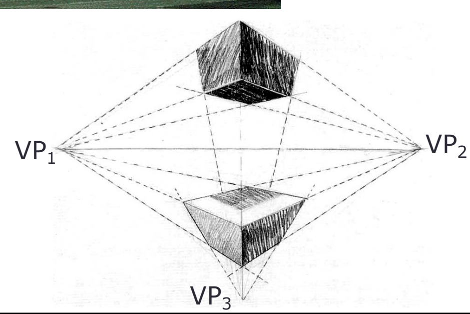
- Each set of parallel lines (=direction) meets at a different point
  - The *vanishing point* for this direction
- Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  - The line is called the *horizon* for that plane



- Good ways to spot faked images
  - scale and perspective don’t work
  - vanishing points behave badly

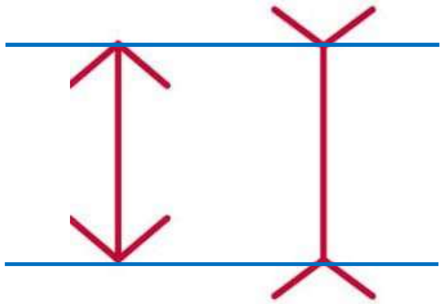
Vanishing points

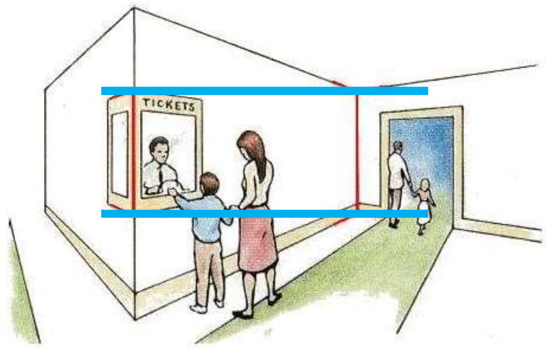




Different directions correspond to different vanishing points

Human vision: Müller-Lyer Illusion

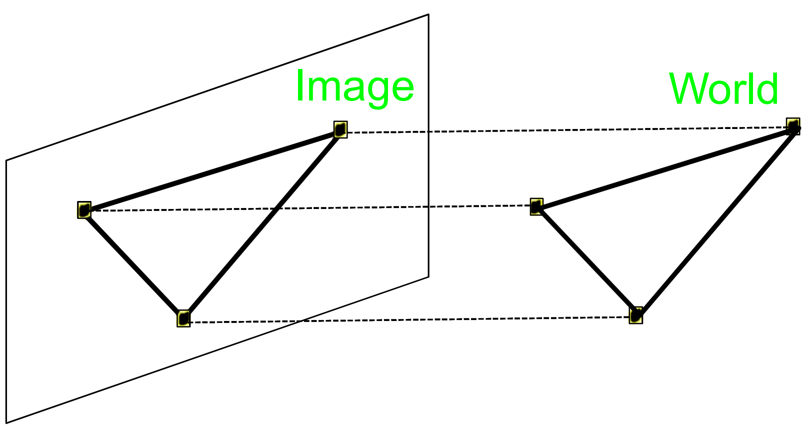




Which line is longer?

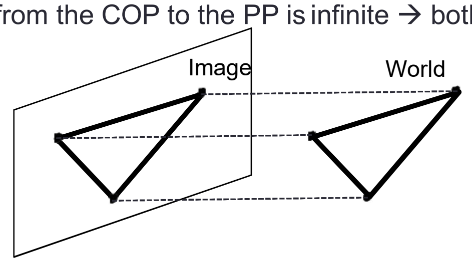
http://www.michaelbach.de/ot/sze_muelue/index.html

Other models: Orthographic projection



Orthographic projection

- Special case of perspective projection
  - Distance from the COP to the PP is infinite → both f and Z are very large

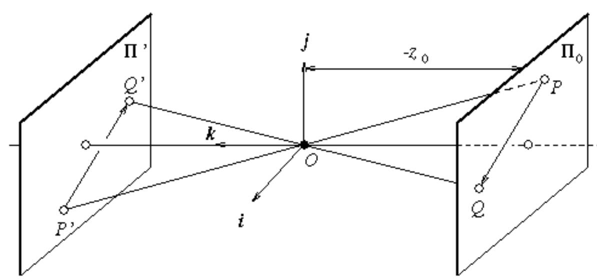


  - Good approximation for telephoto optics
  - Also called "parallel projection": (x, y, z) → (x, y)
  - What's the projection matrix?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow (x, y)$$

Other projection models: Weak perspective

- *Perspective* effects, but not over the scale of *individual* objects
- Collect object points into a group at about the same depth, then divide each point by the depth of its group
- Advantage: easy
- Disadvantage : only approximate



$$(x, y, z) \rightarrow \left(\frac{fx}{z_0}, \frac{fy}{z_0} \right)$$

Other projection models: Weak perspective

- *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
- Advantage: easy
- Disadvantage : only approximate

$$(x, y, z) \rightarrow \left(\frac{fx}{z_0}, \frac{fy}{z_0} \right)$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/s \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1/s \end{bmatrix} \Rightarrow (sx, sy)$$

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: $(x, y, z) \rightarrow (x, y)$

Fun with Perspective

- www.streetpainting3d.com
- oozandoz.com
- www.instructables.com