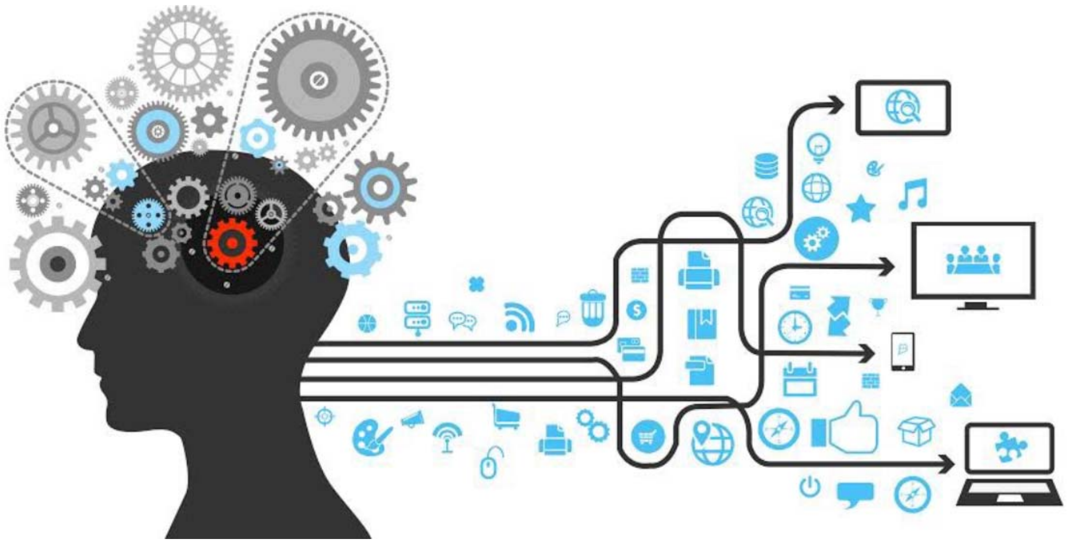
Computer Vision

Image formation

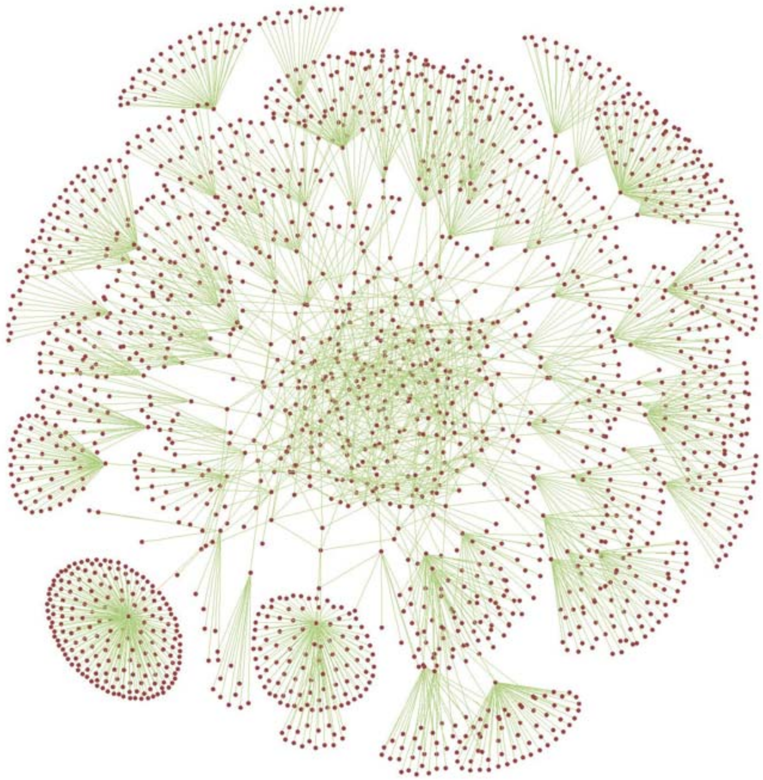


School of Electronic & Electrical Engineering

Sungkyunkwan University

Hyunjin Park

MIP Lab.
Medical Image Processing



Geometric Camera Models

Factors in image formation

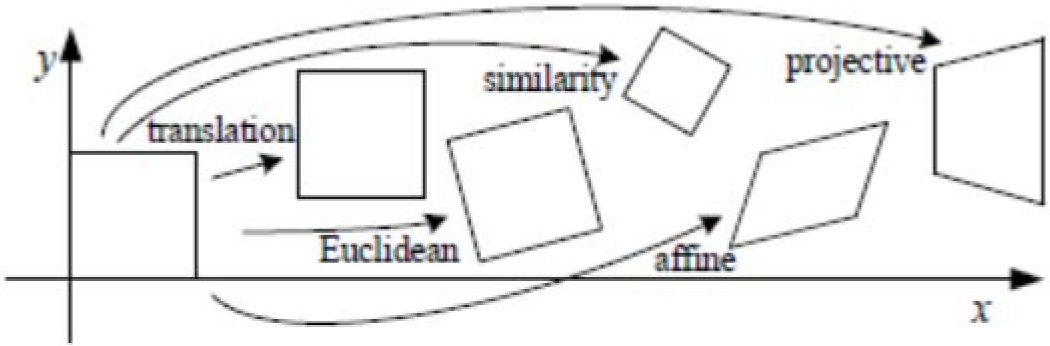
- ***Geometry***
  - Relationship between points in the 3D world and their images
- ***Radiometry***
  - Relationship between the amount of light
    radiating from a surface and the amount of incident at its image
  - Measurement of optical radiation (ultraviolet, visible and infrared)
- ***Photometry***
  - Ways of measuring the intensity of light
  - Measurement of light that is detectable by human eye
- ***Digitization***
  - Ways of converting continuous signals to digital approximations

Basic: Geometric primitives

- ***2D points***
  - $\boldsymbol{x} = (x, y) \in R^2$
- ***2D lines***
  - $\bar{\boldsymbol{x}} \cdot \bar{\boldsymbol{l}} = ax + by + c = 0$
  - $\bar{\boldsymbol{l}} = (a, b, c)$
- ***3D points***
  - $\bar{\boldsymbol{x}} = (\bar{x}, \bar{y}, \bar{z}, \bar{w})$
- ***3D planes***
  - $\bar{\boldsymbol{x}} \cdot \bar{\boldsymbol{m}} = ax + by + cz + d = 0$
  - $\bar{\boldsymbol{m}} = (a, b, c, d)$
- ***3D lines***
  - $\boldsymbol{r} = (1 - \lambda)\boldsymbol{p} + \lambda \boldsymbol{q}$

Basic: Transformations

변환	동차 행렬 $\dot{\mathbf{H}}$	설명
이동	$T(t_y,t_x)=\begin{pmatrix}1&0&0\\0&1&0\\t_y&t_x&1\end{pmatrix}$	y방향으로 t_y, x방향으로 t_x만큼 이동
회전	$R(\theta)=\begin{pmatrix}\cos\theta&-\sin\theta&0\\\sin\theta&\cos\theta&0\\0&0&1\end{pmatrix}$	원점을 중심으로 시계방향으로 θ만큼 회전
크기	$S(s_y,s_x)=\begin{pmatrix}s_y&0&0\\0&s_x&0\\0&0&1\end{pmatrix}$	y방향으로 s_y, x방향으로 s_x만큼 확대
기울임	$Sh_y(h_y)=\begin{pmatrix}1&0&0\\h_y&1&0\\0&0&1\end{pmatrix}$, $Sh_x(h_x)=\begin{pmatrix}1&h_x&0\\0&1&0\\0&0&1\end{pmatrix}$	Sh_y : y방향으로 h_y만큼 기울임 Sh_x : x방향으로 h_x만큼 기울임



Basic: Transformations

- 2D transformations

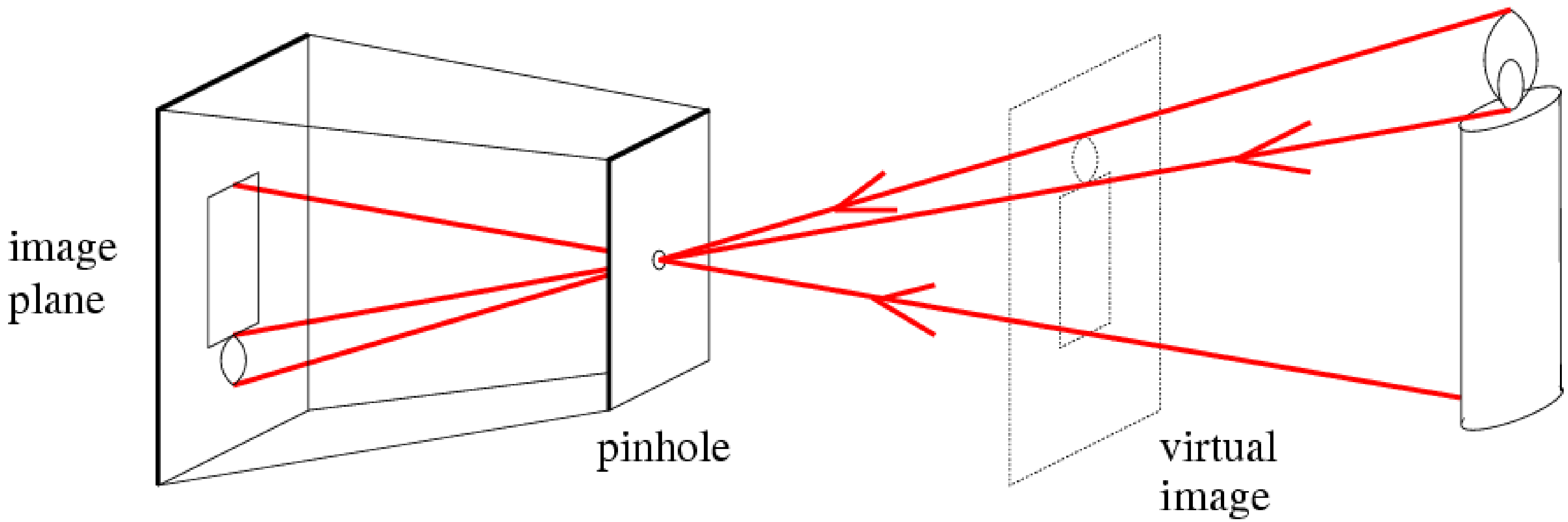
Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	3	lengths	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	4	angles	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{2\times 3}$	6	parallelism	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{3\times 3}$	8	straight lines	

- 3D transformations

Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{3\times 4}$	3	orientation	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{3\times 4}$	6	lengths	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{3\times 4}$	7	angles	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{3\times 4}$	12	parallelism	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{4\times 4}$	15	straight lines	

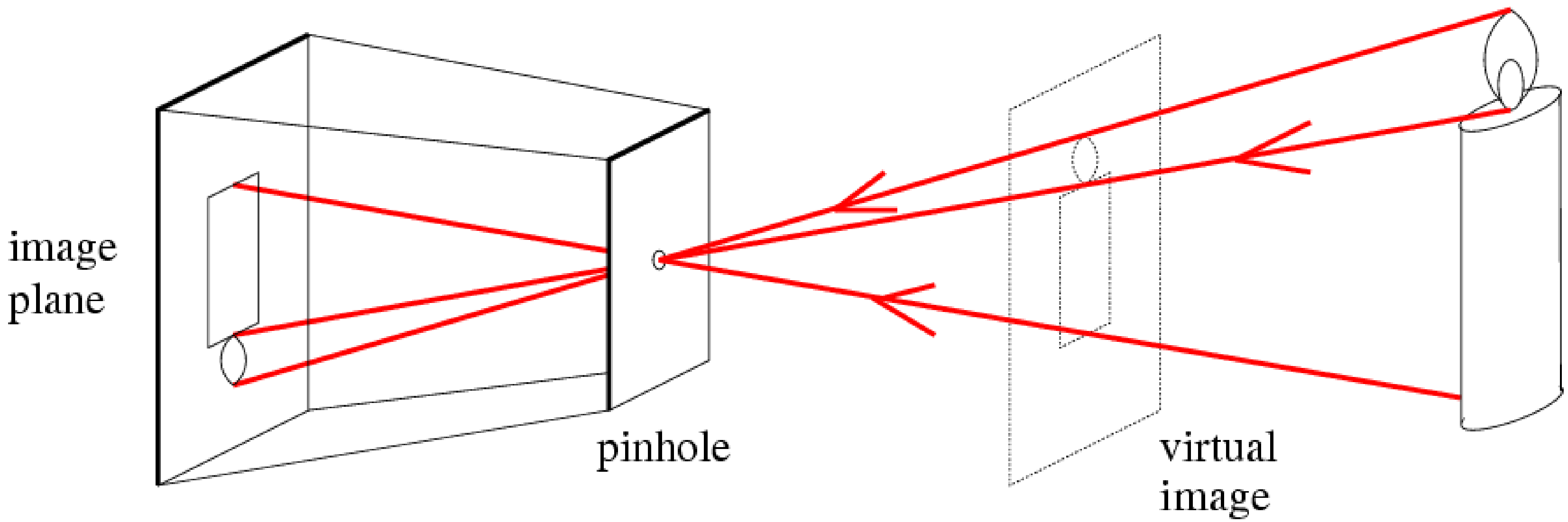
Pinhole camera model

- Pinhole camera
  - A simple model to approximate imaging process, ***perspective projection***



Pinhole camera model

- Pinhole camera
  - Box with a small hole in it
  - Only one ray from any given point can enter the camera
  - When 3D world is projected to 2D, image is inverted

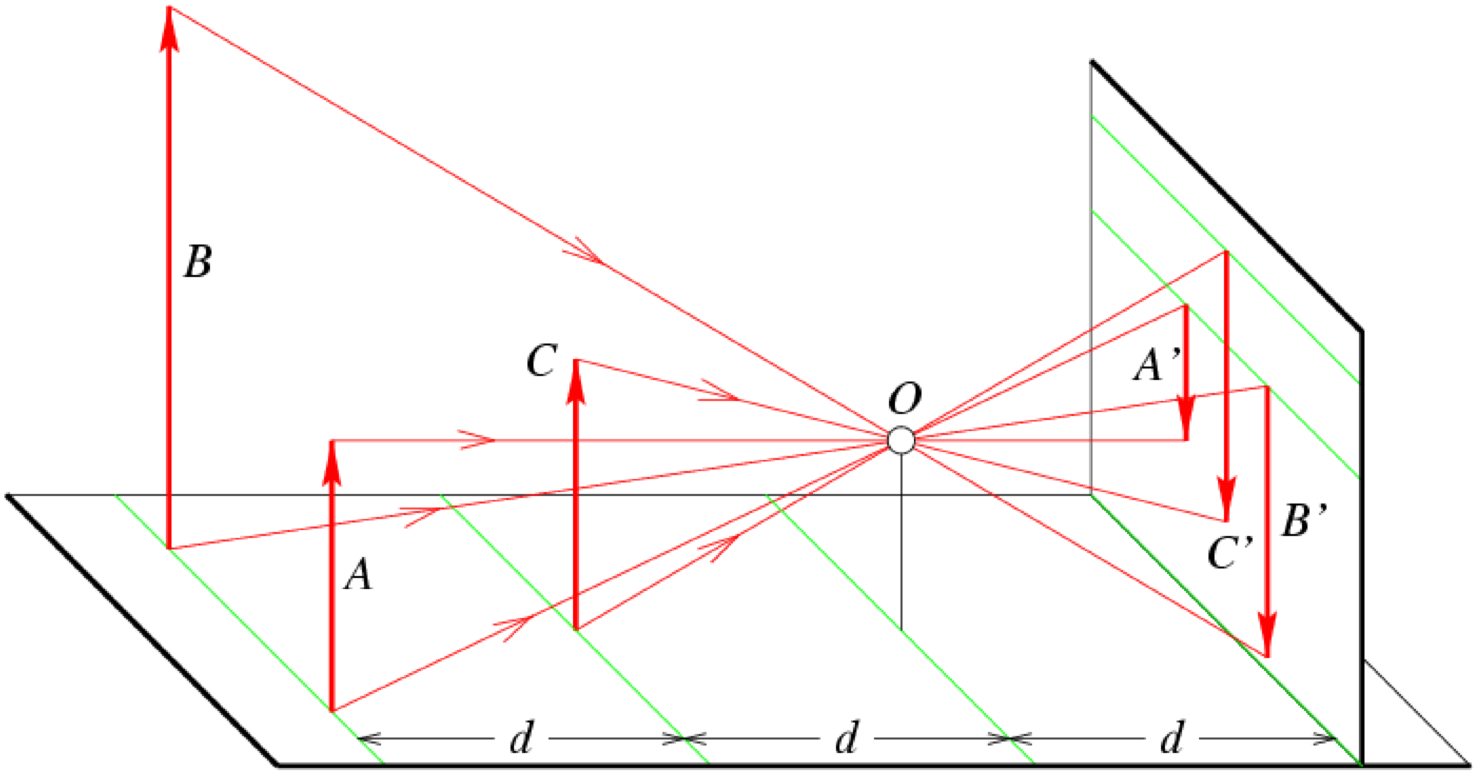


Perspective effects

- Distant objects are smaller

Perspective effects

- Distant objects are smaller

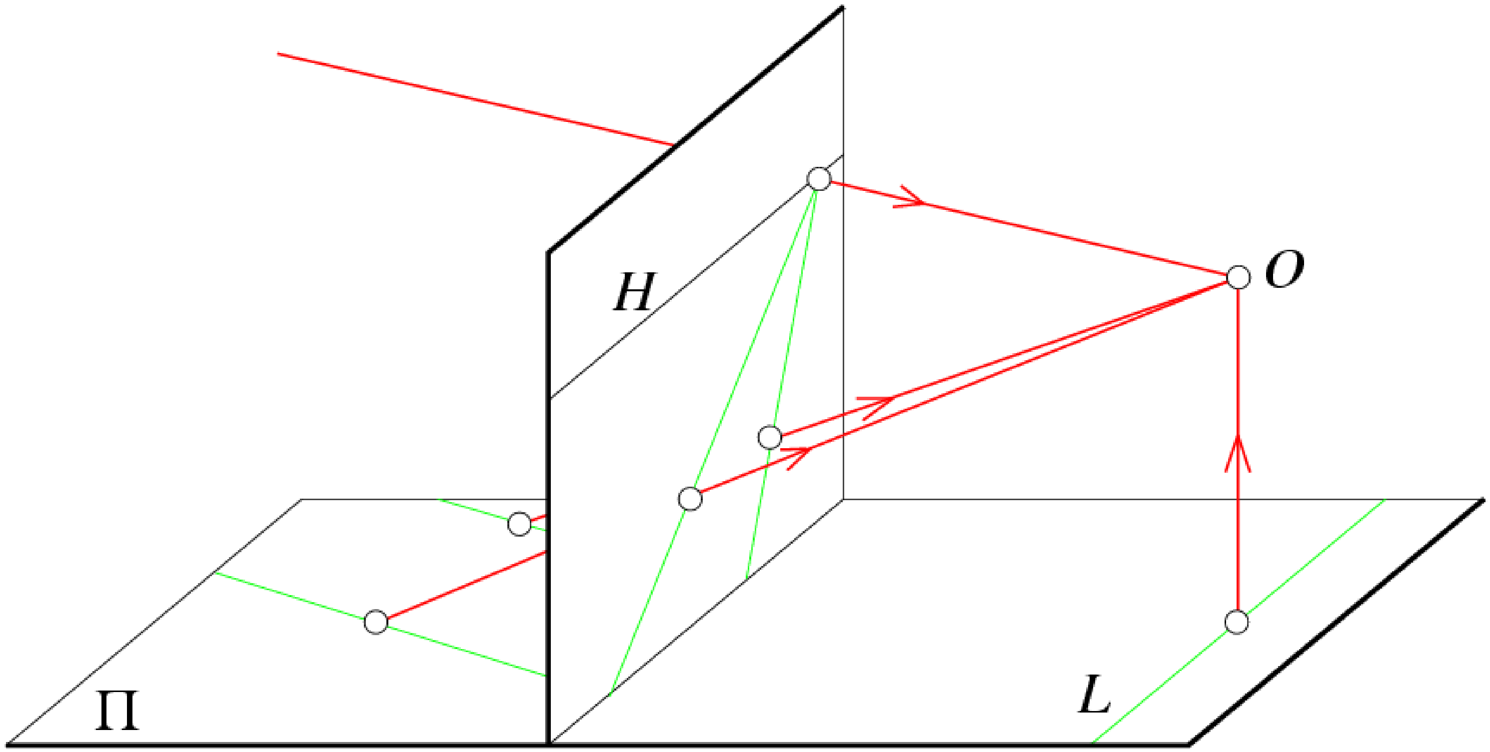


Perspective effects

- Parallel lines intersect in the image: Vanishing points

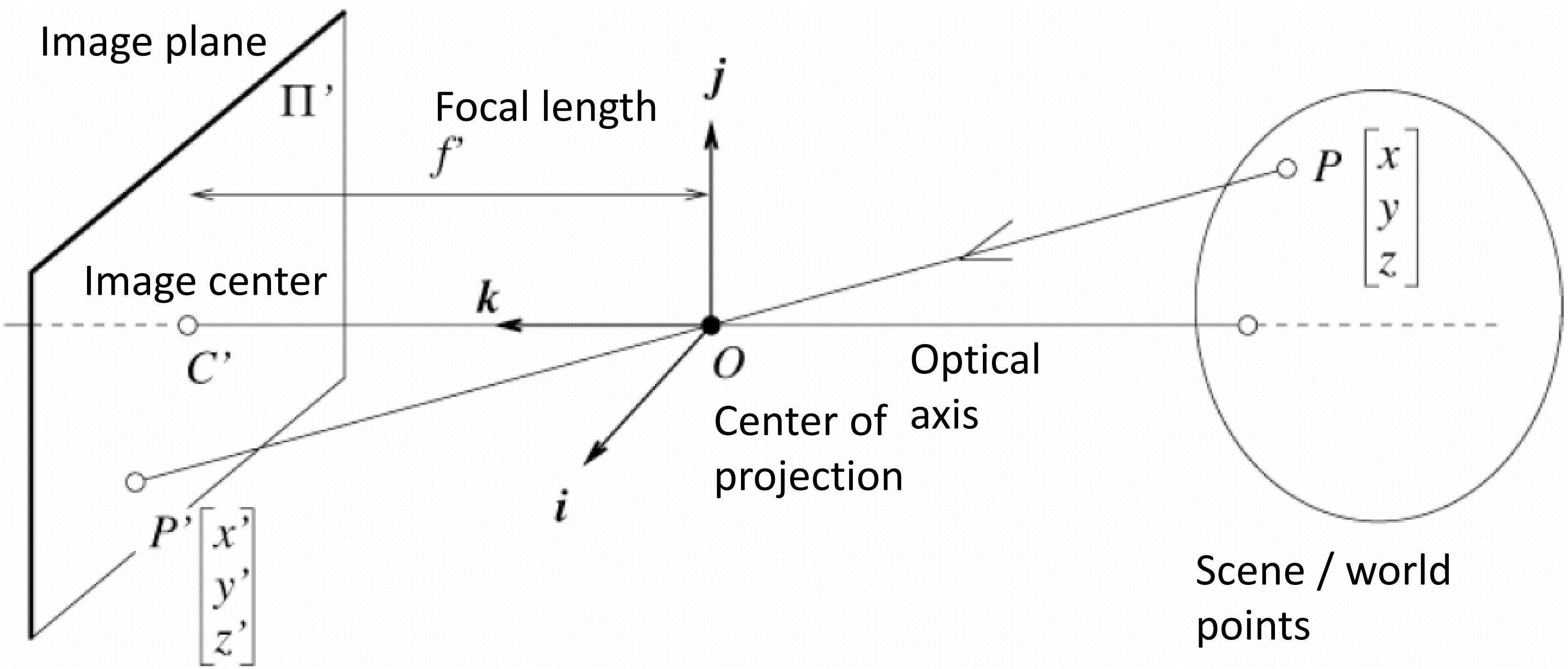
Perspective effects

- Parallel lines intersect in the image: Vanishing points



Perspective equations

- 3D world is mapped to 2D image

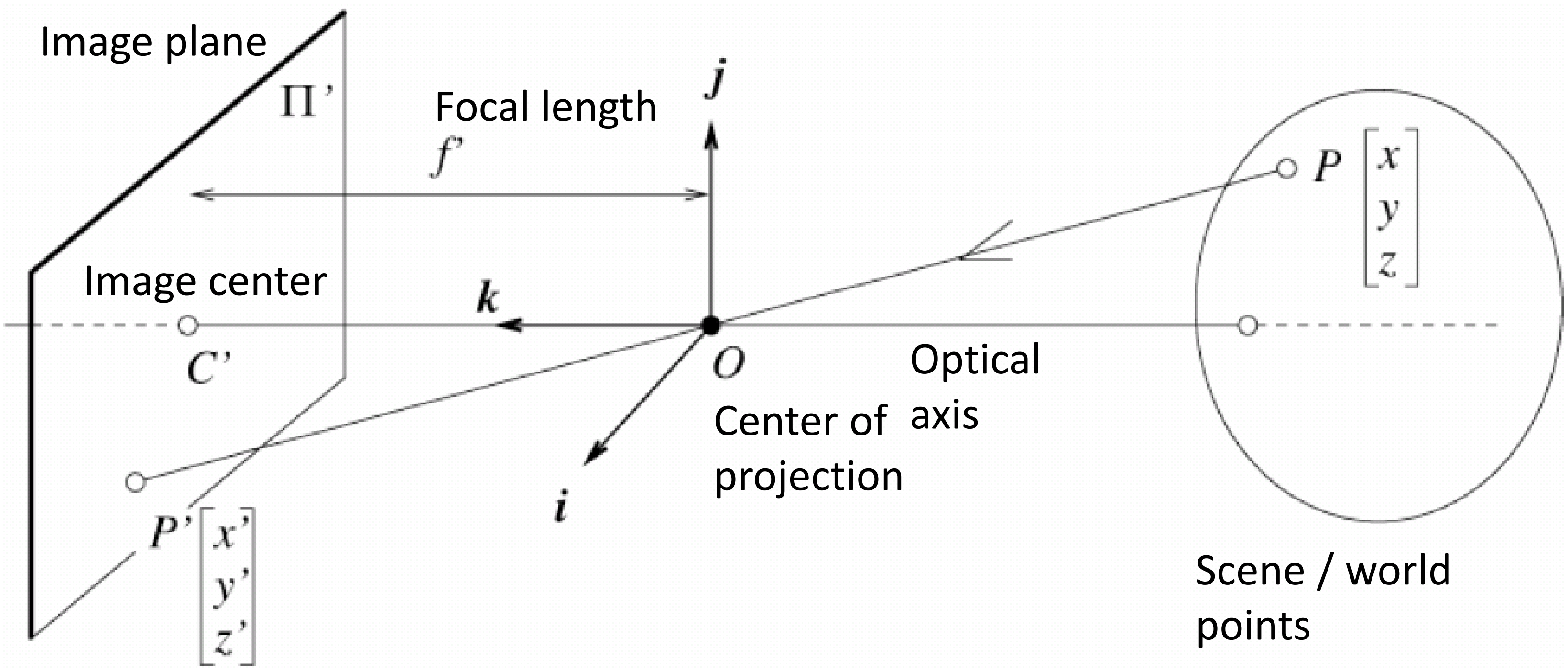


$$(x, y, z) \rightarrow \left(f'\frac{x}{z}, f'\frac{y}{z}\right)$$

Scene point → Image coordinates

Perspective equations

- 3D world is mapped to 2D image



$$\overline{OP'} = \lambda \overline{OP} \Leftrightarrow \begin{array}{c} x' = \lambda x \\ y' = \lambda y \\ z' = \lambda z \end{array} \Leftrightarrow \lambda = \frac{x'}{x} = \frac{y'}{y} = \frac{z'}{z}$$

$$(x', y') = \left(f' \frac{x}{z}, f' \frac{y}{z}\right)$$

i.e., $P' = \left(f' \frac{x}{z}, f' \frac{y}{z}, f'\right)$

Homogeneous coordinates

$$(x', y') = \left(f'\frac{x}{z}, f'\frac{y}{z}\right)$$

$$\text{i.e., } P' = \left(f'\frac{x}{z}, f'\frac{y}{z}, f'\right)$$

- Is this a linear transformation?
  - NO!
  - It is a non-linear transformation because it was divided by z

Homogeneous coordinates

- Introduce an ***extra coordinate representing scale***

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

Homogeneous
image coordinates

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

Homogeneous
scene coordinates

- Converting from homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow \left(\frac{x}{w}, \frac{y}{w}\right)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow \left(\frac{x}{w}, \frac{y}{w}, \frac{z}{w}\right)$$

Homogeneous coordinates

- Turn previous expression into homogeneous coordinates

$$(x', y') = \left(f'\frac{x}{z}, f'\frac{y}{z}\right)$$

$$\Rightarrow (x', y', 1) = \left(f'\frac{x}{z}, f'\frac{y}{z}, 1\right) = \left(x, y, \frac{z}{f'}\right)$$

Perspective projection matrix

- Perspective projection matrix

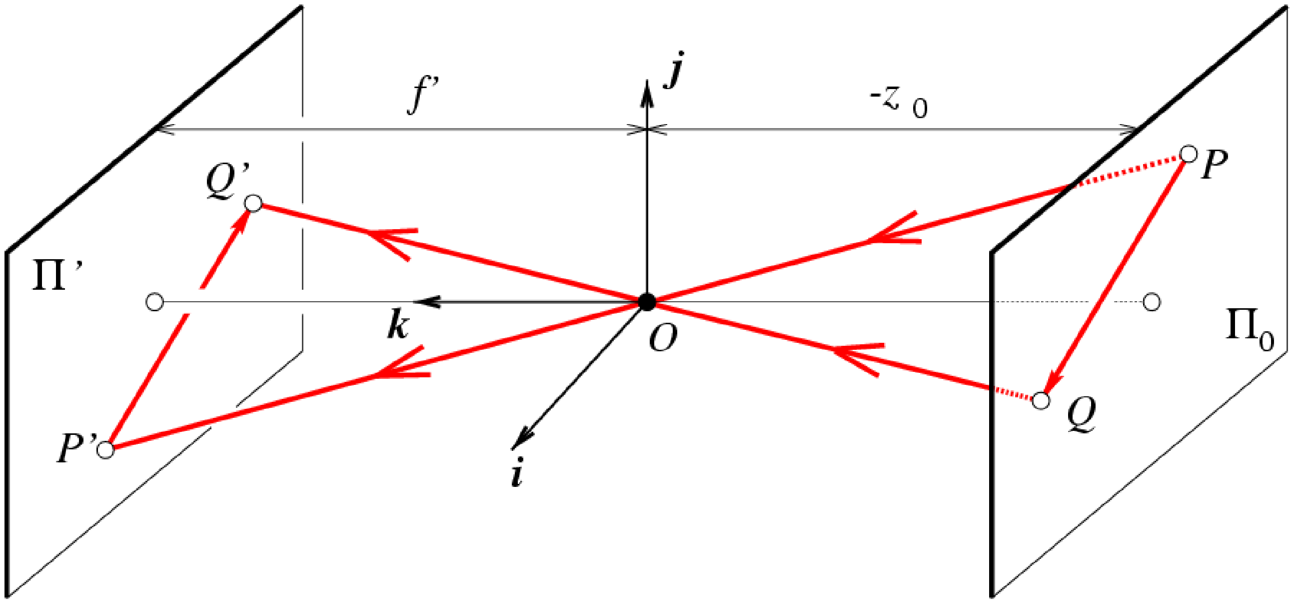
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- Projection: A matrix multiplication using homogeneous coordinates

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

Weak perspective

- ***Approximation:*** Treat magnification as constant
- Assumption: Scene depth << Average distance to camera



$$\begin{cases} x' = -mx \\ y' = -my \end{cases}, \text{ where } m = -\frac{f'}{z_0} \text{ is the magnification}$$

Weak perspective

- The error in image position

$$E = P_{perspective} - P_{weak\ perspective}$$

$$= \frac{f'}{z_0+\Delta z}\begin{pmatrix} x \\ y \end{pmatrix} - \frac{f'}{z_0}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \frac{f' z_0 - f'(z_0+\Delta z)}{(z_0+\Delta z) z_0}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \frac{f'}{z_0} \cdot \frac{\Delta z}{z_0+\Delta z}\begin{pmatrix} x \\ y \end{pmatrix}$$

- ***Factors contribute to the validity of the model***
  - Small focal length (f')
  - Small field of view (x/z_0 and y/z_0)
  - Small depth variation of the object (Δz)

Weak perspective

- Projection matrix for weak perspective projection

$$\begin{cases} x' = -mx = \dfrac{f'}{z_0} x \\ y = -my = \dfrac{f'}{z_0} y \end{cases}$$

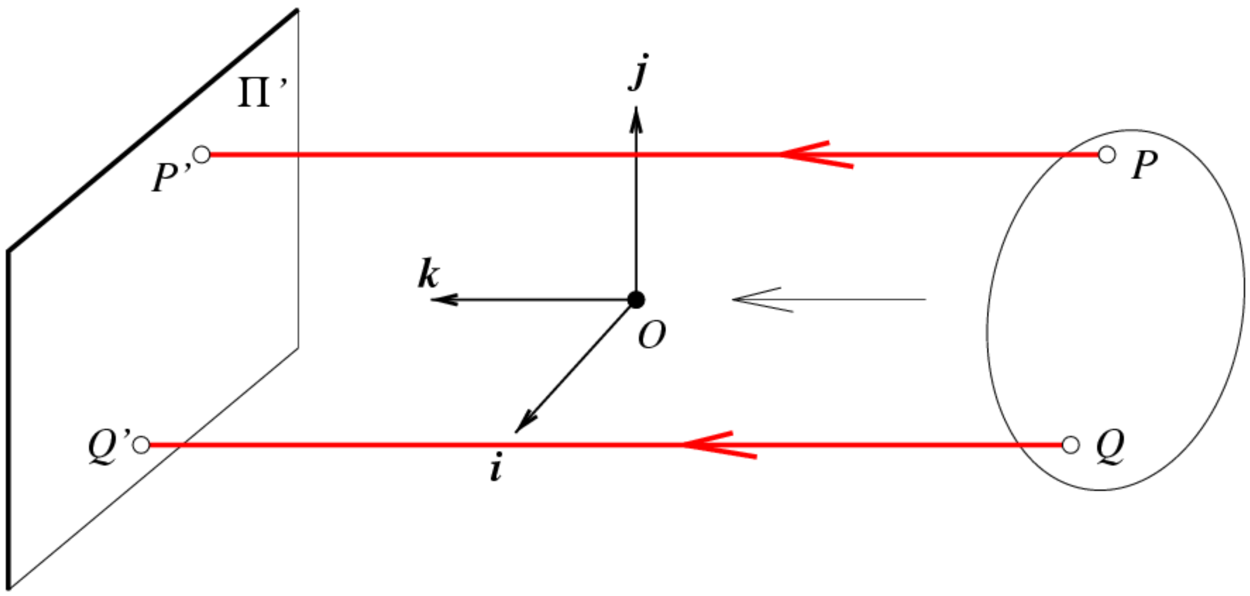
- Homogeneous coordinates

$$(x', y', 1) = \left(\frac{f'}{z_0} x, \frac{f'}{z_0} y, 1\right) = \left(x, y, \frac{z_0}{f'}\right)$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & z_0/f' \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Orthographic projection

- Given camera at constant distance from the scene ($m = -1$)
- World points projected along rays ***parallel*** to optical access



$$\begin{cases} x' = x \\ y' = y \end{cases}$$

Orthographic projection

- Projection matrix for orthographic projection

$$\begin{cases} x' = x \\ y' = y \end{cases}$$

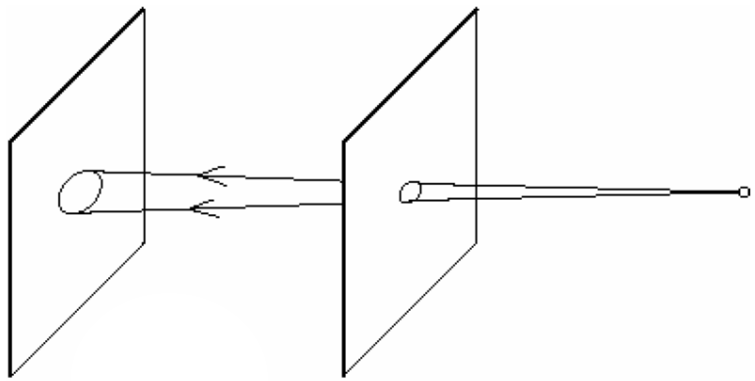
- Homogeneous coordinates

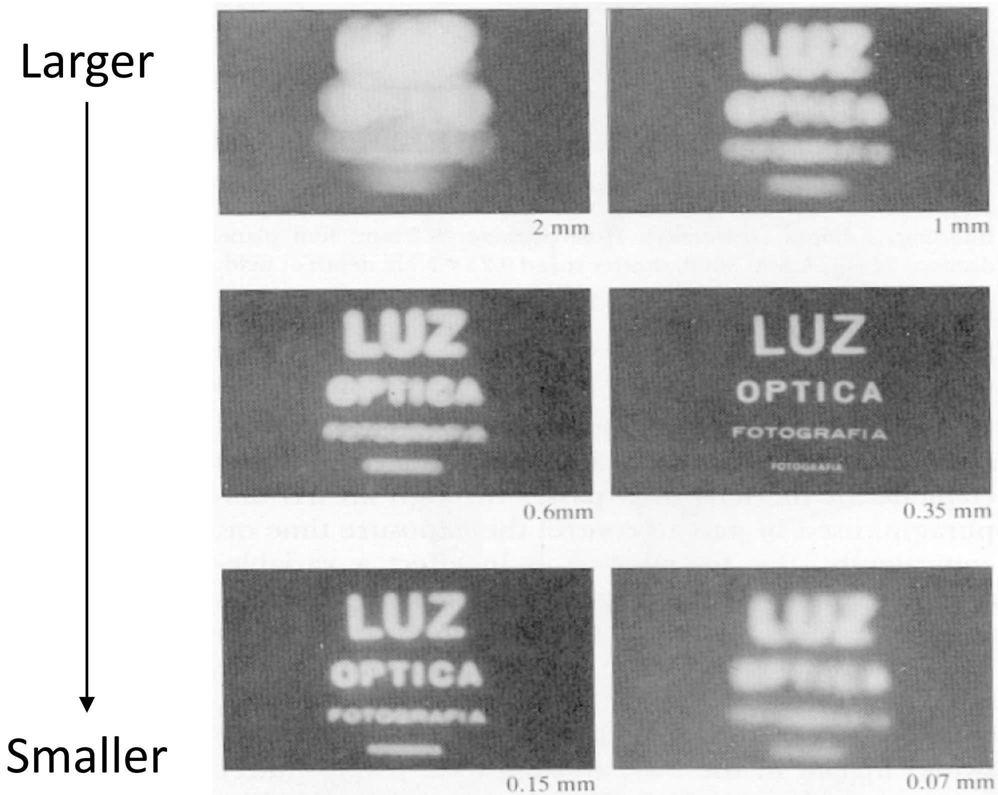
$$(x', y', 1) = (x, y, 1)$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Pinhole size (aperture)

- How does the size of the aperture affect the image we'd get?



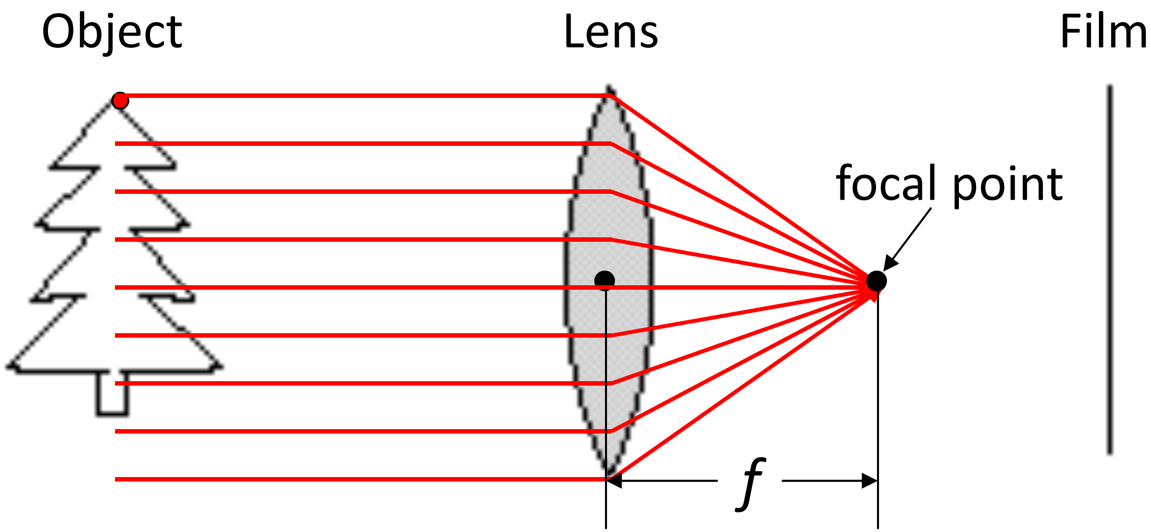


Pinhole size (aperture)

- Pinhole too big
  - Many directions are averaged
  - Blurring the image
- Pinhole too small
  - Diffraction effects blur the image
- Generally, pinhole cameras are ***dark***
  - Because very small set of rays from a particular point hits the screen

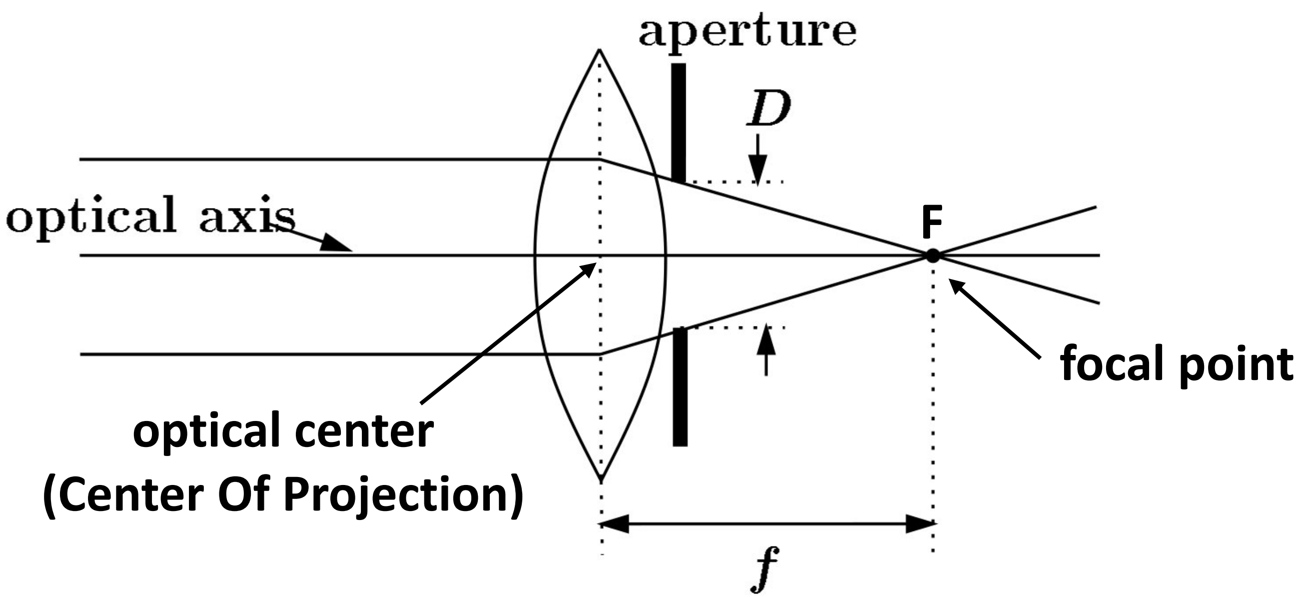
Lens

- A lens ***focuses light*** onto the film
  - Rays passing through the center are not deviated
  - All parallel rays converge to one point on a plane located at the focal length f

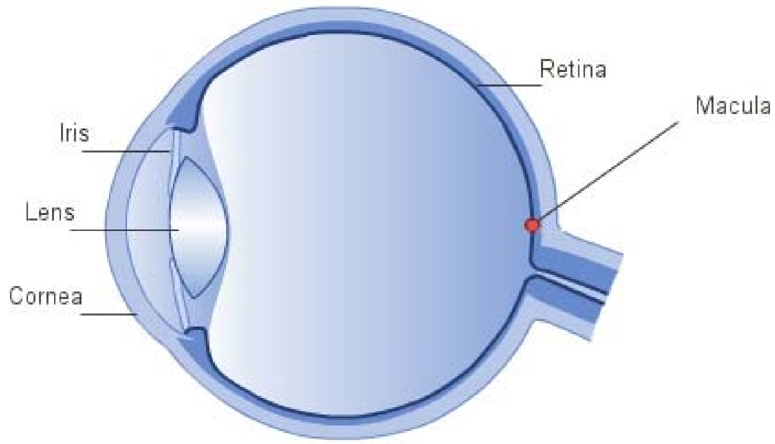


Cameras with lenses

- A lens focuses parallel rays onto a single focal point
- Gather more light, while keeping focus
  - Make pinhole perspective projection practical



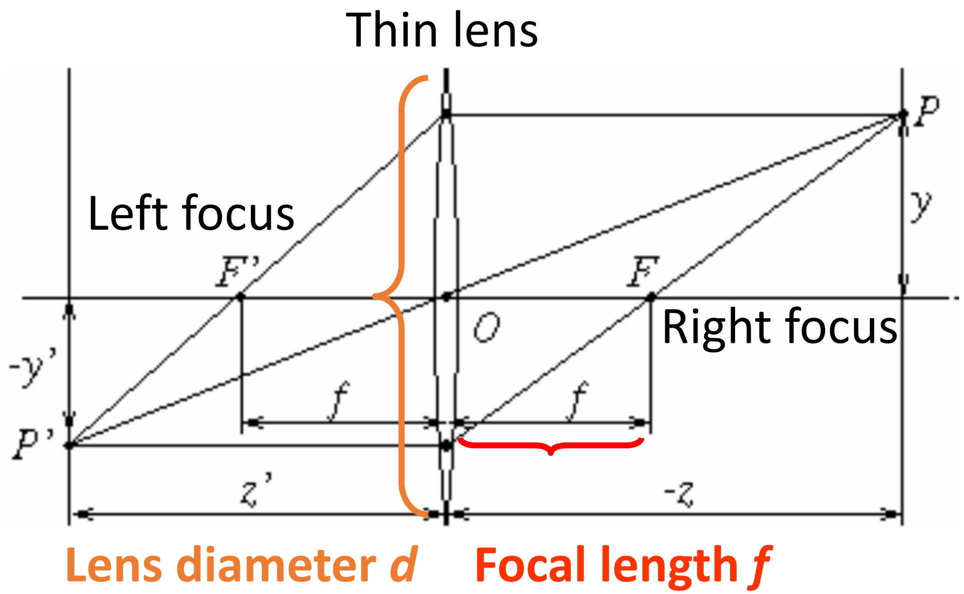
Human eye



- Human eye is a lens
- Iris
  - Control amount of light passing through lens
- Retina
  - Contains sensor cells, where image is formed
- Fovea
  - Highest concentration of cones

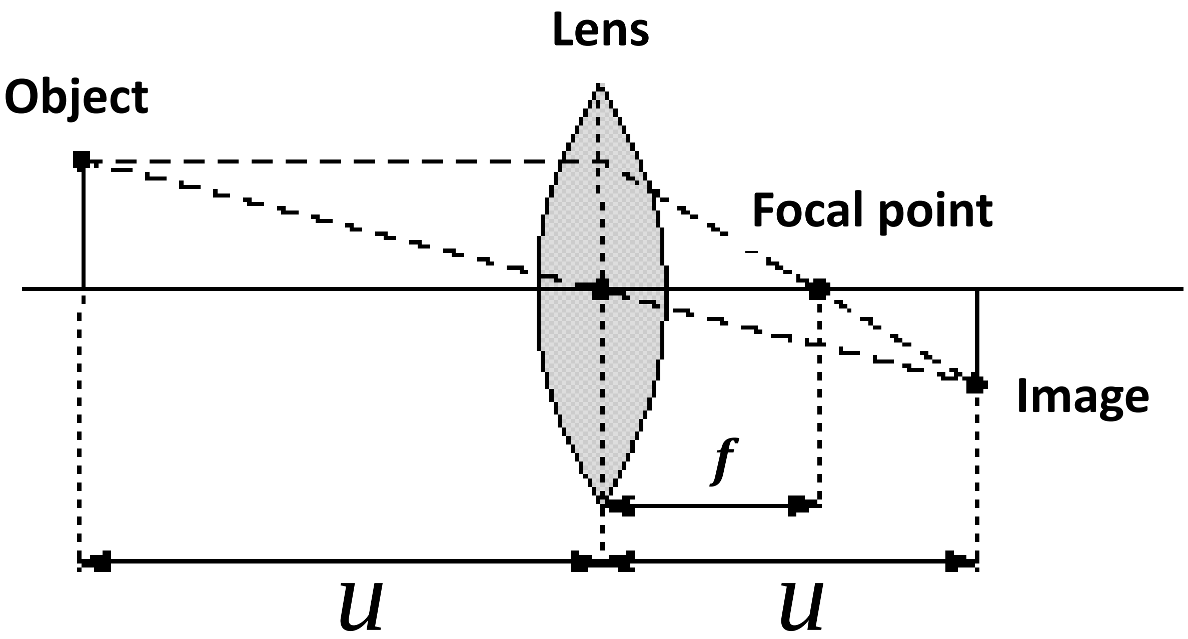
Thin lens

- Thin lens
  - Rays entering parallel on one side go through focus on other, and vice versa
  - In ideal case, all rays from P imaged at P'



Thin lens

- Thin lens equation

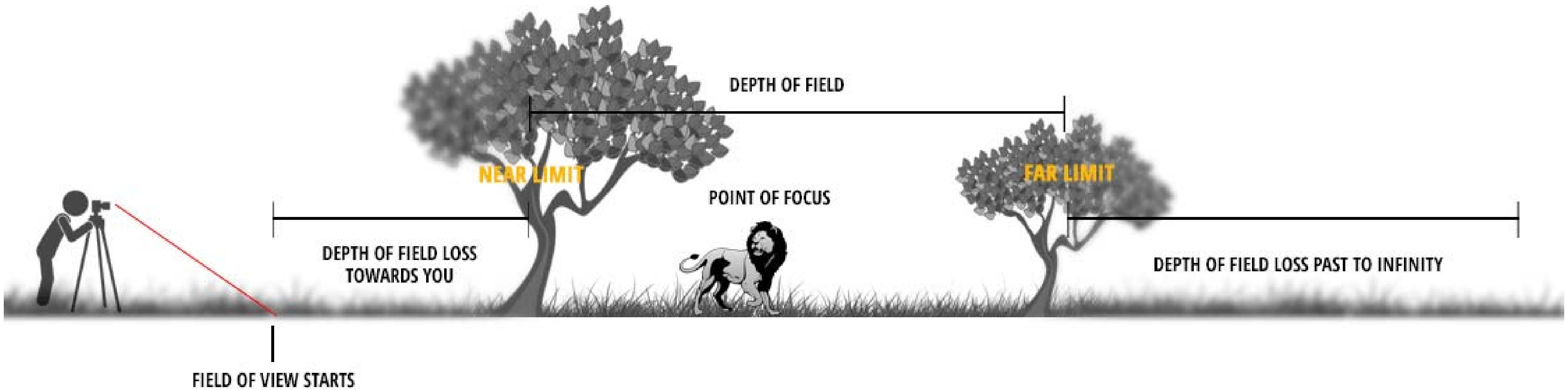


$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

- Any object point satisfying the thin lens equation is in focus

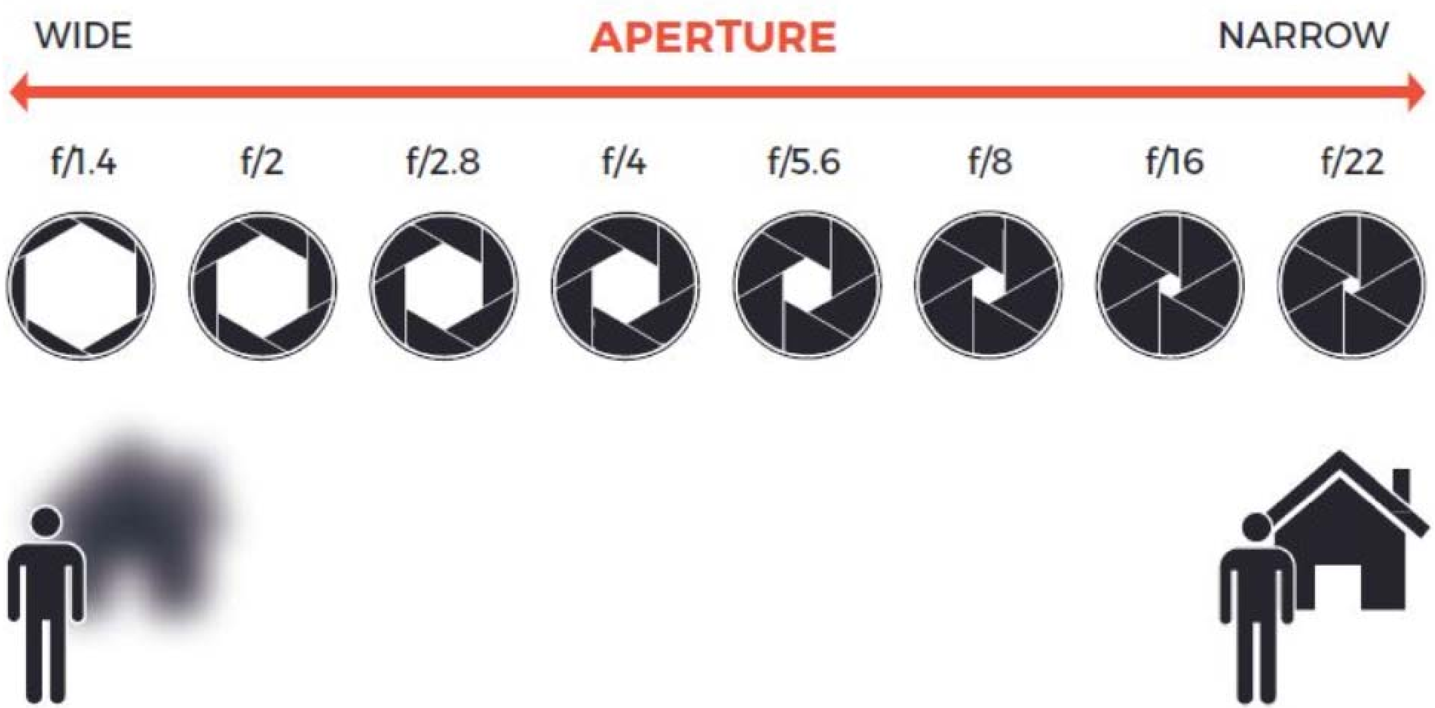
Depth of field

- Depth of field
  - Distance between image planes where blur is tolerable



Aperture

- Aperture
  - Smaller aperture increases the range that the object is approximately in focus



Field of view

- Field of view
  - Angular measure of portion of 3D space seen by the camera

28 mm lens, 65.5° × 46.4°

50 mm lens, 39.6° × 27.0°

70 mm lens, 28.9° × 19.5°

210 mm lens, 9.8° × 6.5°

Field of view

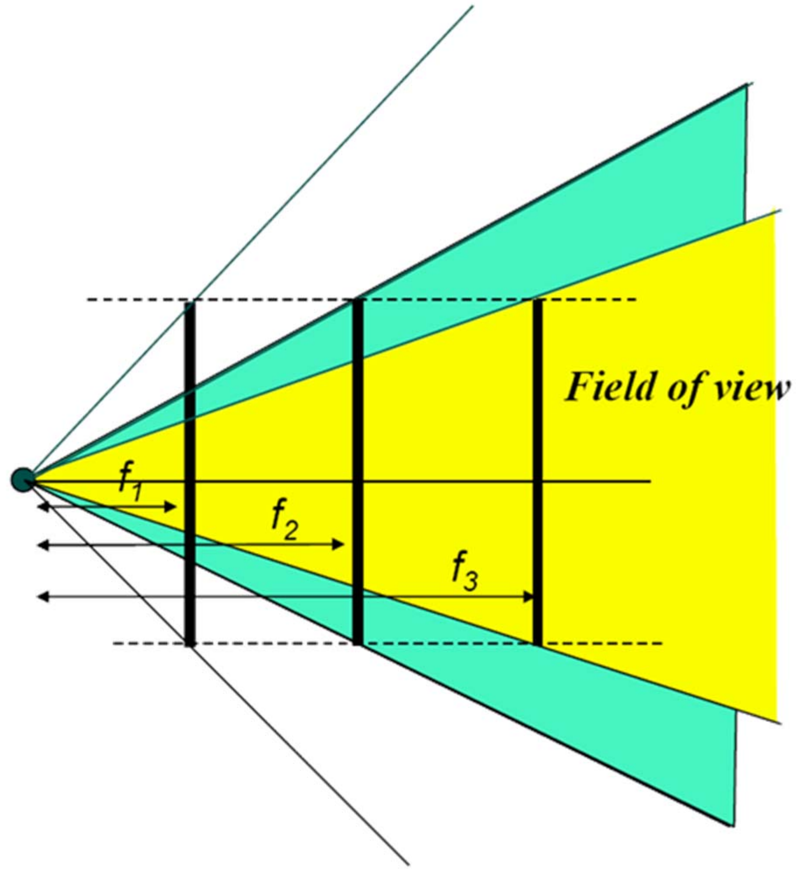
- Field of view depends on focal length (f)

- As f gets smaller,

  image becomes wider angle
  - More world points
    project onto finite image plane

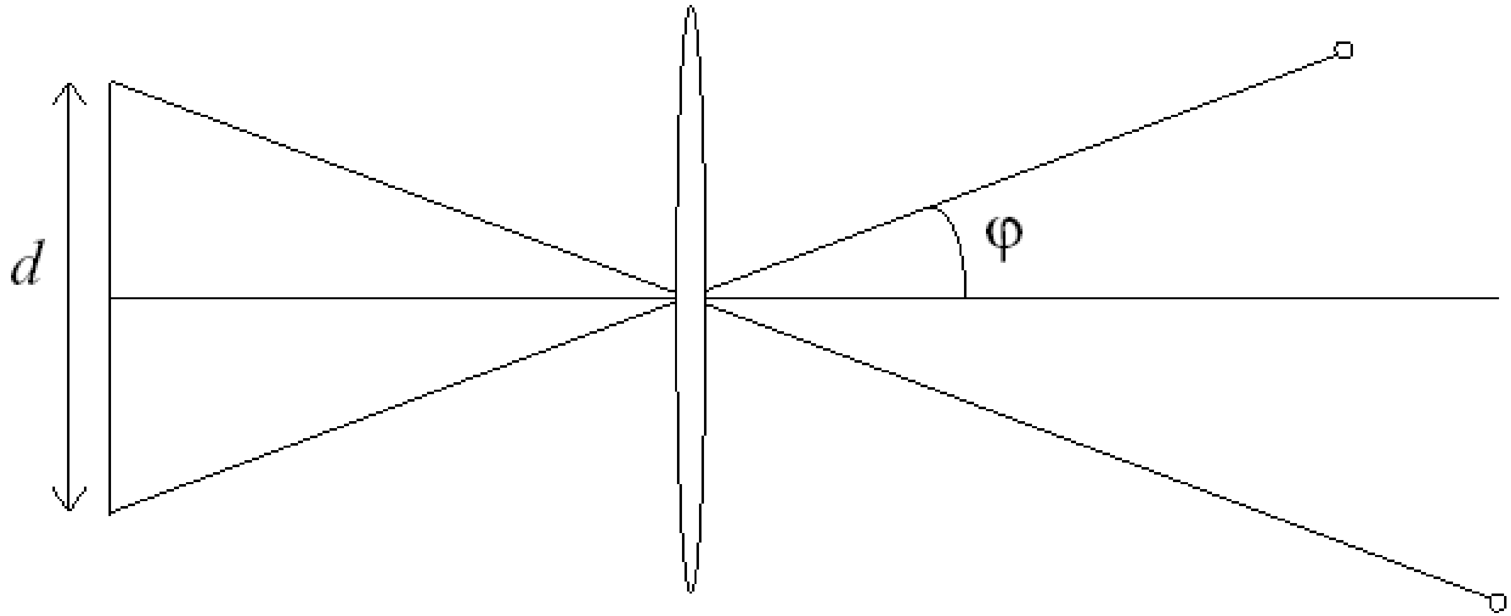
- As f gets larger,

  image becomes more telescopic
  - Small part of the world
    projects onto the finite image plane



Field of view

- Size of field of view governed by size of the camera retina



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

Smaller field of view = Larger focal length

Vignetting

- Tendency for the brightness of the image to fall off towards the edge of the image

Vignetting

- The amount of light (E) hitting a pixel surface area δ_i

$$E = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4 \alpha$$