Computer Vision I - Algorithms and Applications: *Image Formation Process*

Carsten Rother

13/11/2013

COMPUTER VISION LAB DRESDEN

TECHNISCHE UNIVERSITÄT DRESDEN

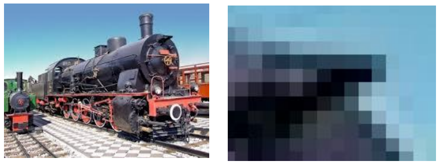
Slide credits

Stefan Roth, Konrad Schindler, Svetlana Lazebnik, Steve Seitz, Fredo Durand, Alyosha Efros, Dimitri Schlesinger, and potentially others

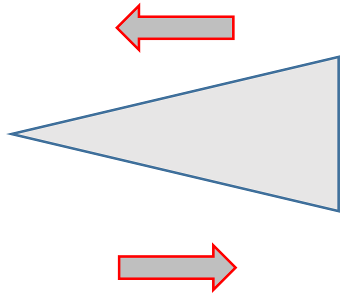
Reminder from first lecture

- Computer Vision is an inverse Problem

Computer Graphics



2D pixel representation



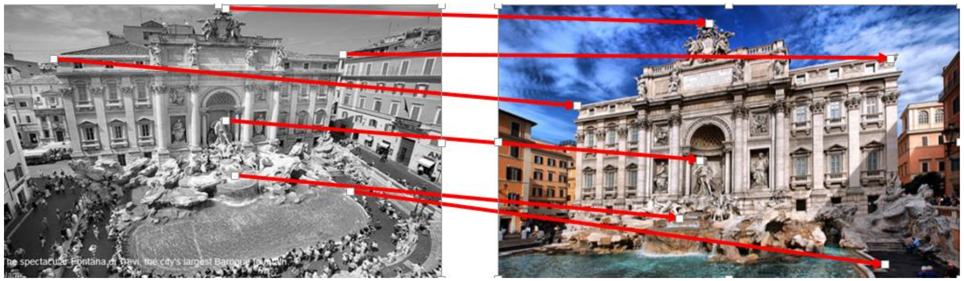
3D Rich Representation,

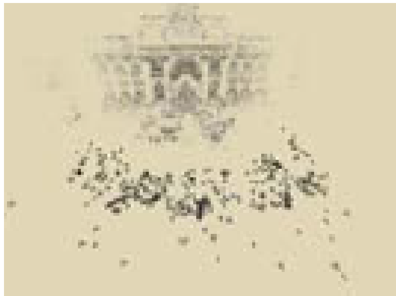
Script = {Camera, Light, Geometry, Material, Objects, Scene, Attributes, Others}

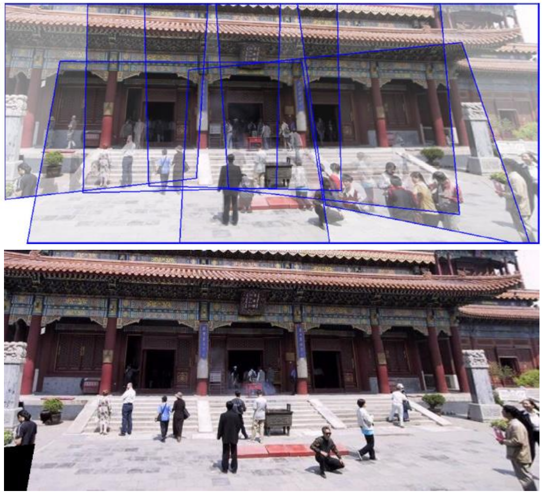
Computer Vision

- What general (prior) knowledge of the world (not necessarily visual) can be exploit?
- What properties / cues from the image can be used?

Reminder: Sparse versus Dense Matching: Tasks and Applications







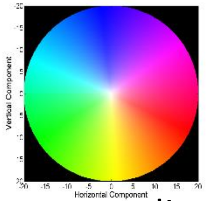
Tasks:

- Find places where we could match features (points, lines, regions, etc)
- Extract appearance - features descriptors
- Find all possible (putative) appearance matches between images
- Verify with geometry

For what applications is sparse matching enough:

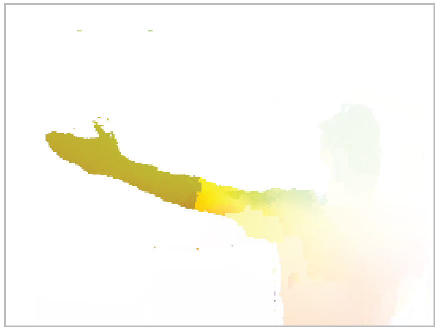
- Sparse 3D reconstruction of a rigid scene
- Panoramic stitching of a rotating / translating camera
- Augemted Reality / Video

Reminder: Sparse versus Dense Matching



Flow encoding

Kinect RGB and Depth data input



Dense flow:
frame 1->2

Dense flow:
frame 2->1

3D view interpolation

Reminder: Using multiple Images: Define Challenges

A road map for the next five lectures

- L4: Geometry of a Single Camera
  and Image Formation Process
- L5: Sparse Matching two images: Appearance
- L6: Sparse Matching two images: Geometry
- L7: Sparse Reconstructing the world (Geometry of n-views)
- L8: Dense Geometry estimation
  (stereo, flow and scene flow, registration)

Roadmap this lecture (image formation process)

- Geometric primitives and transformations (sec. 2.1.1-2.1.4)

- Geometric image formation process (sec. 2.1.5, 2.1.6)
  - Pinhole camera
  - Lens effects

- The Human eye

- Photometric image formation process (sec. 2.2)

- Camera Types and Hardware (sec 2.3)

Roadmap this lecture (image formation process)

- Geometric primitives and transformations (sec. 2.1.1-2.1.4)
- Geometric image formation process (sec. 2.1.5, 2.1.6)
  - Pinhole camera
  - Lens effects
- The Human eye
- Photometric image formation process (sec. 2.2)
- Camera Types and Hardware (sec 2.3)

Some Basics

- Real coordinate space R^2 example: $\binom{1}{2}$
- Real coordinate space R^3 example: $\begin{pmatrix}1\\2\\3\end{pmatrix}$

- Euclidean Space R^3 where angles and length are defined. Operations we need are

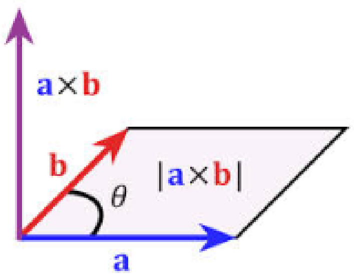
<u>scalar product:</u>

$$\boldsymbol{x}\,\boldsymbol{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 \quad \text{where } x = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

<u>cross/vector product:</u>

$$\boldsymbol{x} \times \boldsymbol{y} = [\boldsymbol{x}]_\times \, \boldsymbol{y}$$

$$[\boldsymbol{x}]_\times = \begin{bmatrix}0 & -x_3 & x_2\\ x_3 & 0 & -x_1\\ -x_2 & x_1 & 0\end{bmatrix}$$



Euclidean Space

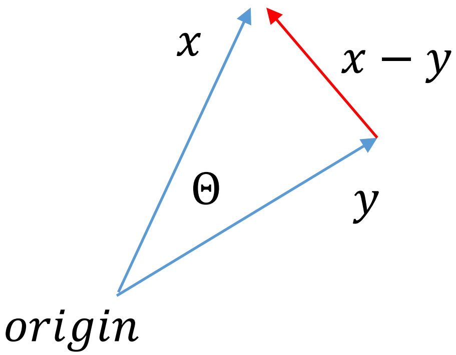
- Euclidean Space R^2 or R^3 has angles and distances defined

Angle defined as: $\theta = \arccos\left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\right)$

$$\|\mathbf{x}\| = \sqrt{x\,x}$$

Length of the vector x

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$



Projective Space

- 2D Point in a real coordinate space:
  $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \in R^2$ has 2 DoF (degrees of freedom)

- 2D Point in a real coordinate space:
  $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in R^3$ has 3 DoF

- Definition: A point in 2-dimensional projective space P^2 is defined as
  $p = \begin{pmatrix} x \\ y \\ w \end{pmatrix} \in P^2$, such that all vectors $\begin{pmatrix} kx \\ ky \\ kw \end{pmatrix}$ $(\forall\, k \neq\, 0)$
  define the same point p in P^2 (equivalent classes)

- Writing: sometimes $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \sim \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}$; we write $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}$

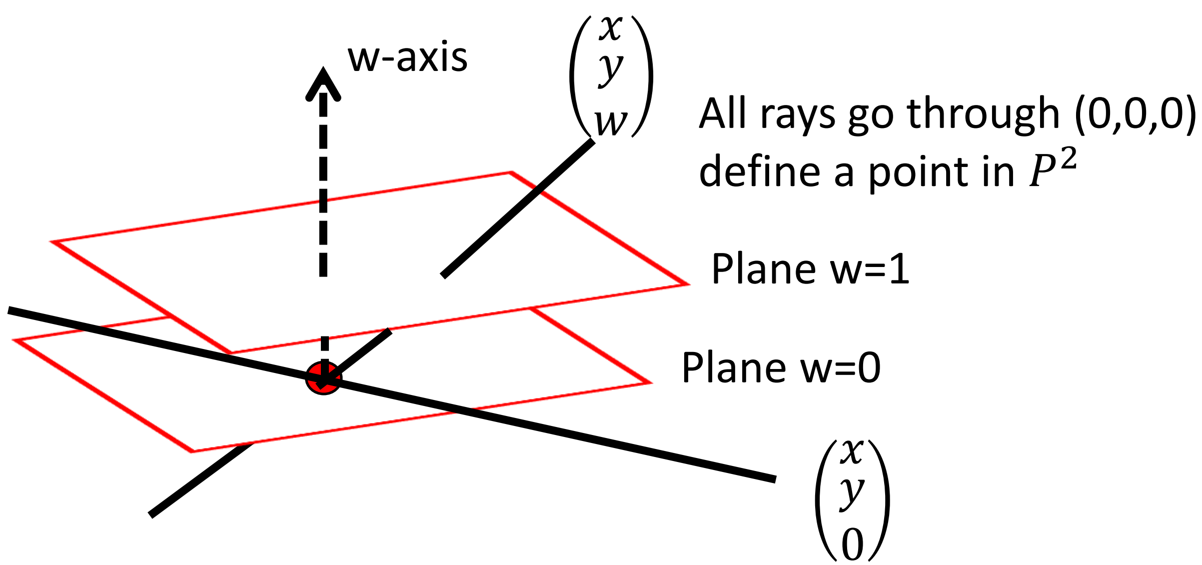
Projective Space - visualization

Definition: A point in 2-dimensional projective space P^2 is defined as

$$p = \begin{pmatrix} x \\ y \\ w \end{pmatrix} \in P^2, \text{ such that all vectors } \begin{pmatrix} kx \\ ky \\ kw \end{pmatrix} \quad (\forall\, k \neq 0)$$

define the same point p in P^2 (equivalent classes)

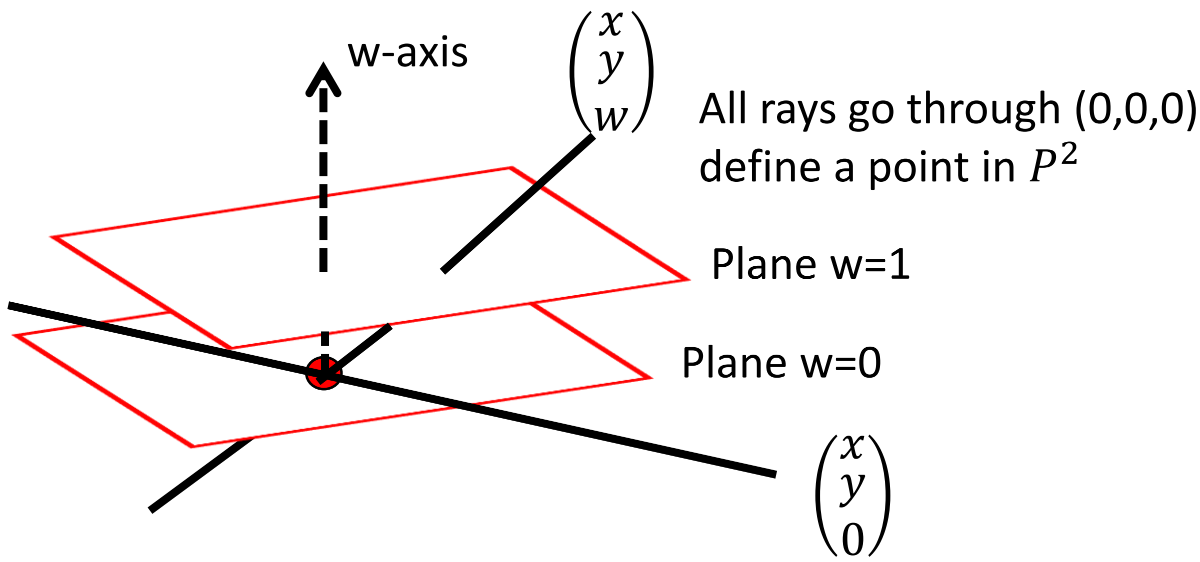
A point in P^2 is a ray in R^3 that goes through the origin:



Projective Space

- All points in P^2 are given by: $R^3 \setminus \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

- A point $\begin{pmatrix} x \\ y \\ w \end{pmatrix} \in P^2$ has 2 DoF (3 elements but norm of vector can be set to 1)



Real Coordinate Space versus Projective Space

Real coordinate space R^2/R^3

Primitives:

- Points
- Lines
- Conics (Quadric in 3D)
- (Planes in 3D)

Operations with Primitives:

- intersection
- tangent

Transformations:

- Rotation
- Translation
- projective
- ….

Projective space P^2/P^3

Primitives:

- Points
- Lines
- Conics (Quadric in 3D)
- (Planes in 3D)

Operations with Primitives:

- intersection
- tangent

Transformations:

- Rotation
- Translation
- projective
- ….

From R^2 to P^2 and back

- From R^2 to P^2:

$$p = \begin{pmatrix} x \\ y \end{pmatrix} \in R^2 \quad \Rightarrow \quad p = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \in P^2$$

- a point in inhomogeneous coordinates
- we soemtimes write $\tilde{p}$ for inhomogeneous coordinates

- a point in homogeneous coordinates

- From P^2 to R^2:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \in P^2 \quad \Rightarrow \quad \begin{pmatrix} x/w \\ y/w \end{pmatrix} \in R\hat{}2$$

$$for\ w \neq 0$$

what does it mean if w=0?

We can do this transformation with all primitives

From R^2 to P^2 and back: Example

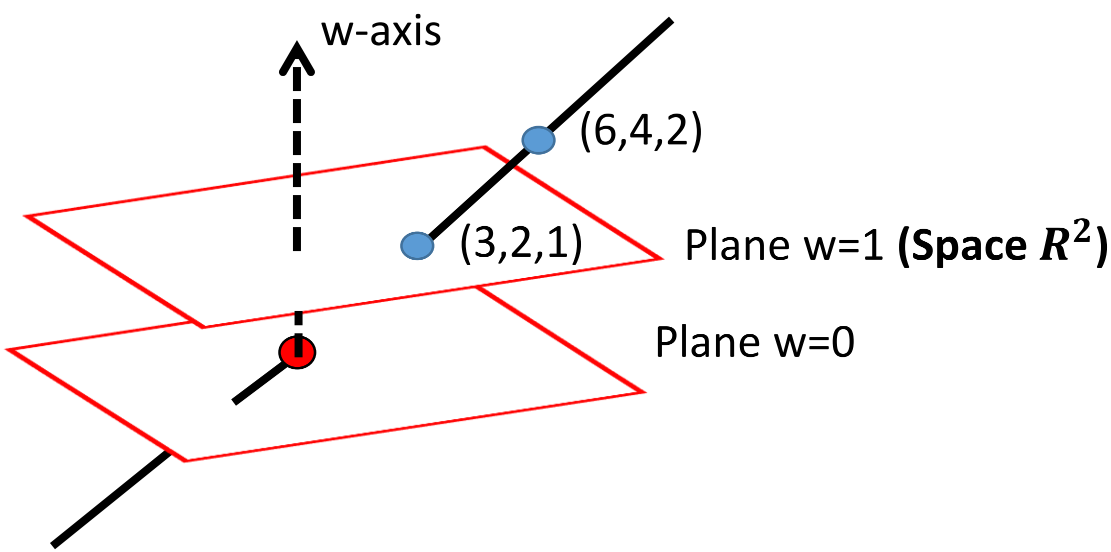
- From R^2 to P^2:

$$p = \begin{pmatrix} x \\ y \end{pmatrix} \in R^2 \quad \Longrightarrow \quad p = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \in P^2$$

- a point in inhomogeneous coordinates

- a point in homogeneous coordinates

$$p = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \in R^2 \quad \Longrightarrow \quad p = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4.5 \\ 3 \\ 1.5 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix} \in P^2 \quad \Longrightarrow \quad p = \begin{pmatrix} 6/2 \\ 2/2 \end{pmatrix} \in R^2$$

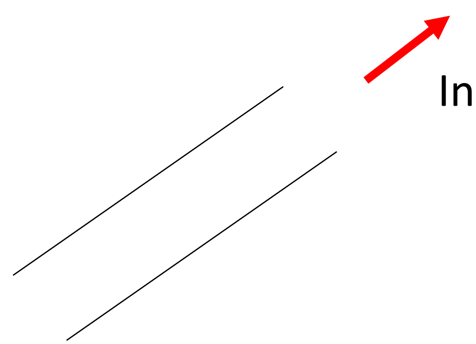


Why bother about P^2

- All Primitives, operations and transformations are defined in R^2 and P^2

- <u>Advantage of P^2:</u>
  - Many transformation and operations are written more compactly (e.g. linear transformations)
  - We will introduce new special “primitives” that are useful when dealing with “parallelism”

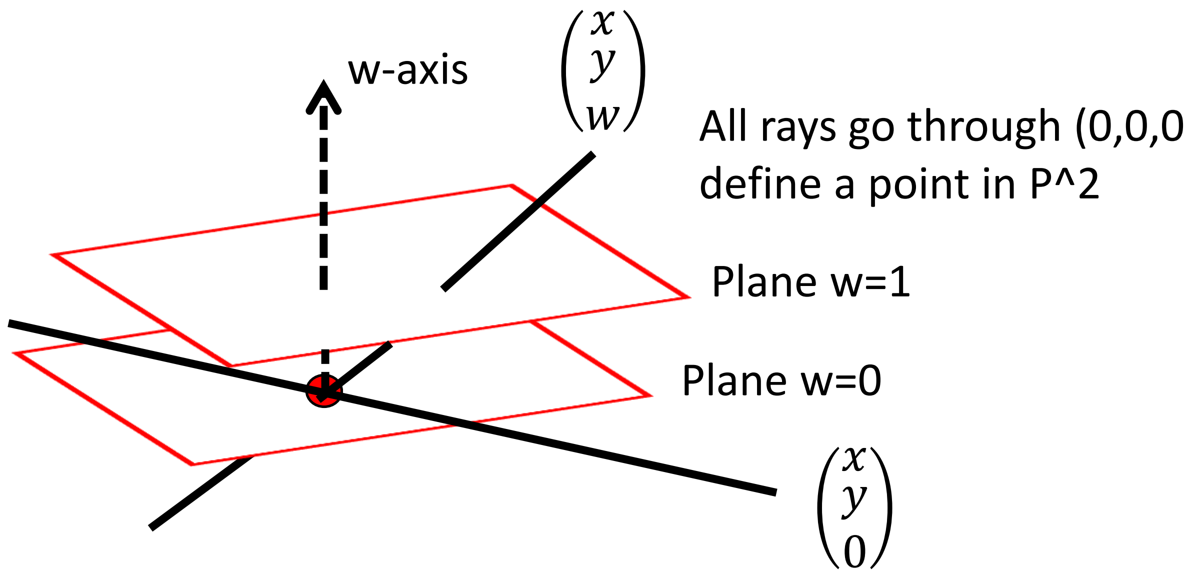
Example will come later



In P^2 they meet in a „point at inifinty“

Two parallel lines in R^2

Points at infinity



Points with coordinate $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ are ideal points or points at infinity

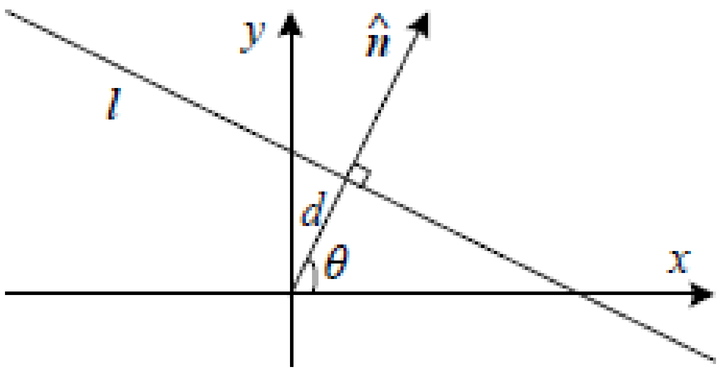
$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in P^2$ ⇒ Not defined in R^2 $since\ w = 0$

Lines in R^2

- For Lines in coordinate space R^2 we can write

  $l = (n_x, n_y, d)$ with $n = \left(n_x, n_y\right)^t$ is normal vector and $||n|| = 1$

- A line has 2 DoF



- A point (x, y) lies on l if:

  $n_x\, x + n_y\, y + d = 0$

- Normal can also be encoded

  with an angle θ:
  $n = (\cos\theta, \sin\theta)^t$

Lines in P^2

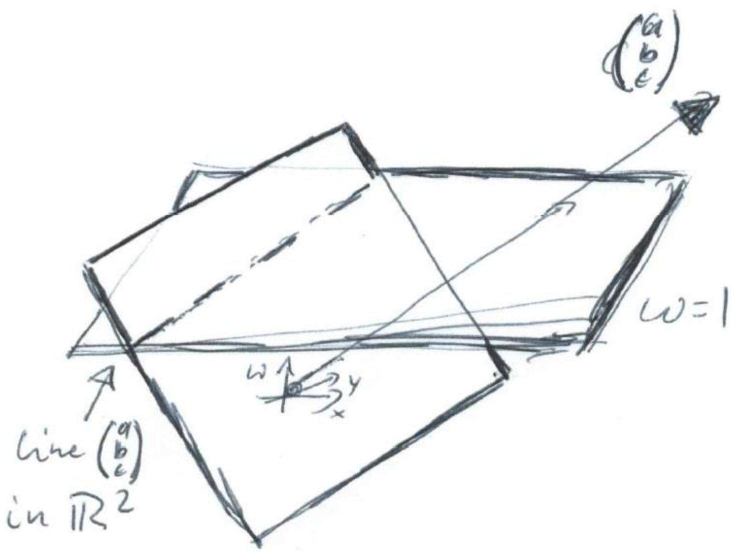
- Points in P^2: $\boldsymbol{x} = (x, y, w)$
- Lines in P^2: $\boldsymbol{l} = (a, b, c)$

  (again equivalent class: $(a, b, c) = (ka, kb, kc)\ \forall\ k \neq 0$)

  Hence also 2 DoF

- All points (x, y, w) on the line (a, b, c) satisfy: $ax + by + cz = 0$

this is the equation of a plane in R^3 with normal (a,b,c) going through (0,0,0)

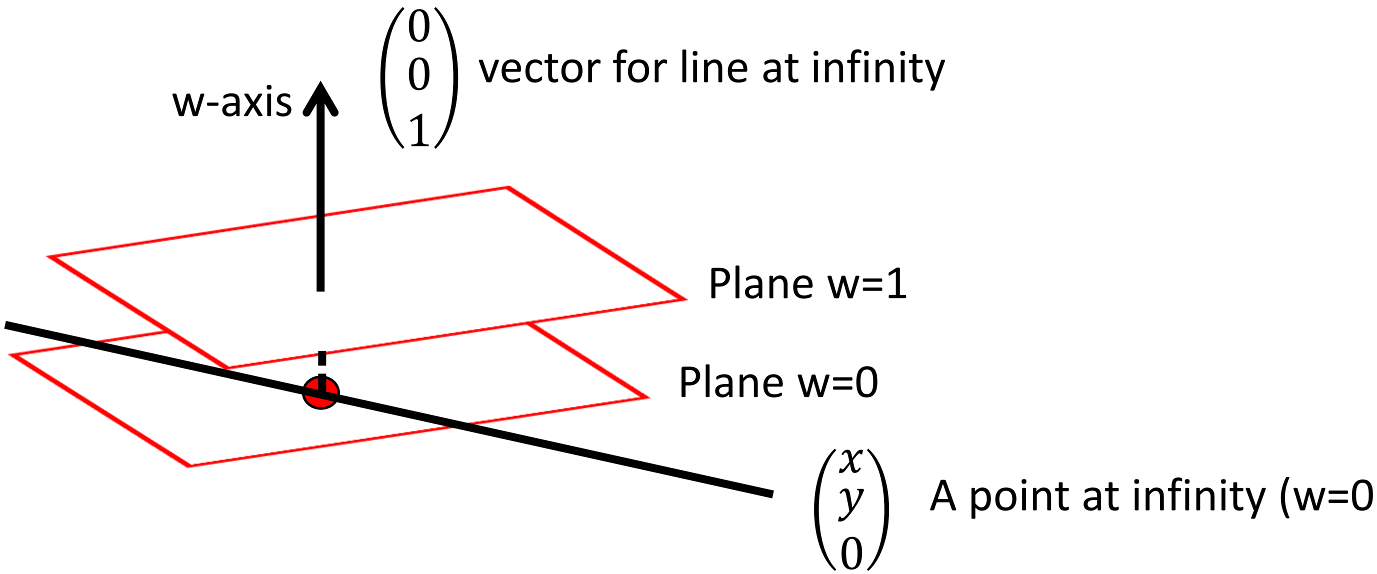


Line at Infininty

- There is a “special” line, called line at infinity: (0,0,1)

- All points at infinity (x,y,0) lie on the line at infinity (0,0,1):

  x*0 + y*0 + 0*1 = 0

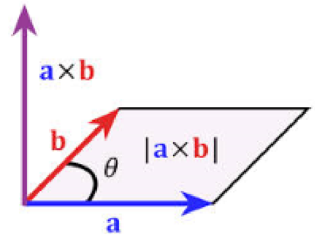


A Line defined by two points in P^2

- The line through two points $\boldsymbol{x}$ and $\boldsymbol{x}'$ is given as: $\boldsymbol{l} = \boldsymbol{x} \times \boldsymbol{x}'$
- Proof:

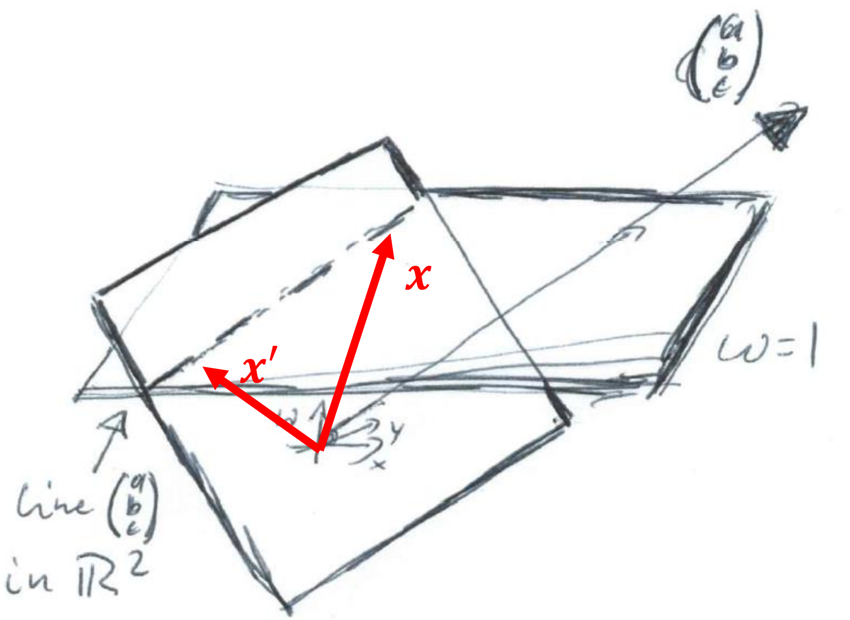
$$\boldsymbol{x}(\boldsymbol{x} \times \boldsymbol{x}') = \boldsymbol{x}'(\boldsymbol{x} \times \boldsymbol{x}') = \mathbf{0}$$

vectors are orthogonal



$$\boldsymbol{x}\,\boldsymbol{l} = \boldsymbol{x}'\boldsymbol{l} = \mathbf{0}$$

The line $\boldsymbol{l}$ goes through points $\boldsymbol{x}$ and $\boldsymbol{x}'$

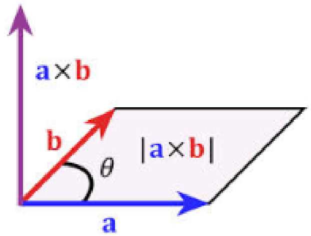


The Intersection of two lines in P^2

- Intersection of two lines $\boldsymbol{l}$ and $\boldsymbol{l}'$ is the point $\boldsymbol{x} = \boldsymbol{l} \times \boldsymbol{l}'$
- Proof:

$$\boldsymbol{l}(\boldsymbol{l} \times \boldsymbol{l}') = \boldsymbol{l}'(\boldsymbol{l} \times \boldsymbol{l}') = \boldsymbol{0}$$

vectors are orthogonal



$$\boldsymbol{l}\boldsymbol{x} = \boldsymbol{l}'\boldsymbol{x} = \boldsymbol{0}$$

The point $\boldsymbol{x}$ lies on the lines $\boldsymbol{l}$ and $\boldsymbol{l}'$

Note the „Theorem“ and Proofs have been very similiar, we only interchanged points and lines

Duality of points and lines

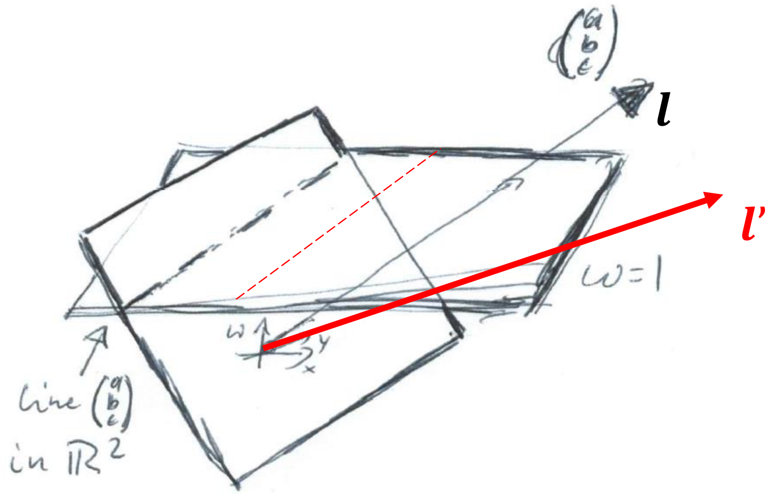
- Note $\boldsymbol{lx} = \boldsymbol{xl} = \mathbf{0}$ ($\boldsymbol{x}$ and $\boldsymbol{l}$ are "interchangeable")

- **Duality theorem:** Two any theorem of 2D projective geometry there corresponds a dual theorem, which may be derived by interchanging the roles of points and lines in the original theorem.

The intersection of two lines $\boldsymbol{l}$ and $\boldsymbol{l}'$ is the point $\boldsymbol{x} = \boldsymbol{l} \times \boldsymbol{l}'$

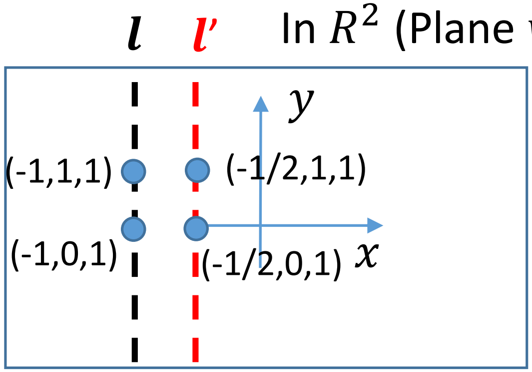
The line through two points $\boldsymbol{x}$ and $\boldsymbol{x}'$ is the line $\boldsymbol{l} = \boldsymbol{x} \times \boldsymbol{x}'$

Parallel lines meet at a point in Infinty



$$\boldsymbol{l} = \begin{pmatrix}1\\0\\1\end{pmatrix};\ \boldsymbol{l'} = \begin{pmatrix}2\\0\\1\end{pmatrix}$$

In R^2 (Plane $w = 1$)



intersection

$$\boldsymbol{l} \times \boldsymbol{l'} = \begin{bmatrix}0 & -1 & 0\\1 & 0 & -1\\0 & 1 & 0\end{bmatrix}\begin{pmatrix}2\\0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\end{pmatrix}$$

Point at infinty

Points at infinty in 3D

- Parallel lines in 3D meet at a point at infinty
- Points at infinty can be real points in a camera

$$\begin{pmatrix} b \\ f \\ 1 \end{pmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \end{bmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

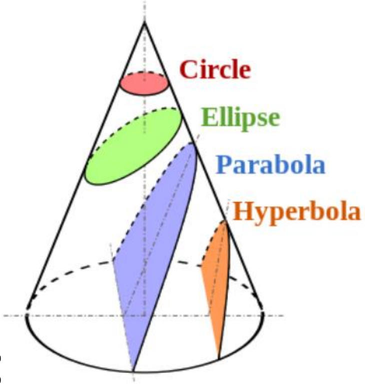
Real point in the image

3x4 Camera Matrix 3D->2D projection

3D Point at infinty

2D conic "Kegelschnitt"

- Conics are shapes that arise when a plane intersects a cone
- In compact form: $\boldsymbol{x}^t \boldsymbol{C}\,\boldsymbol{x} \ = \ \boldsymbol{0}$ where **C** has the form:



$$\mathbf{C} = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}$$

- This can be written as in in-homogenous coordinates:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

where $\boldsymbol{x} \ =$

- $\boldsymbol{C}$ has 5DoF since unique up to scale:

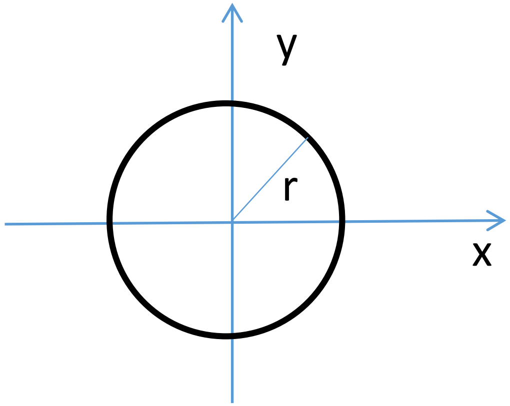
$$\boldsymbol{x}^t \boldsymbol{C}\,\boldsymbol{x} \ = k\boldsymbol{x}^t \boldsymbol{C}\,\boldsymbol{x} = \boldsymbol{x}^t k \boldsymbol{C}\,\boldsymbol{x} = \boldsymbol{0}$$

- <u>Properties:</u> $\boldsymbol{l}$ is tangent to $\boldsymbol{C}$ at a point $\boldsymbol{x}$ if $\boldsymbol{l} \ = \ \boldsymbol{C}\boldsymbol{x}$

Example: 2D Conic

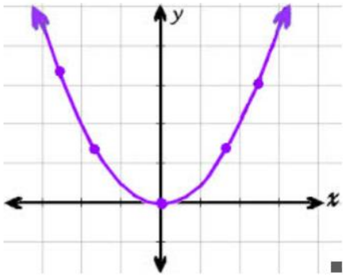
A circle:

$$x^2 + y^2 - r^2 = 0$$



Parabola:

$$y^2 = 0$$



Define a conic with five points

Given 5 points $(x_i, y_i, 1)$ we can write:

$$a x_i^2 + b x_i y_i + c y_i^2 + d x_i + e y_i + f = 0.$$

as $\begin{pmatrix} x_i^2 & x_i y_i & y_i^2 & x_i & y_i & 1 \end{pmatrix} \mathbf{c} = 0$ with $\mathbf{c} = (a, b, c, d, e, f)^\mathsf{T}$

That gives:

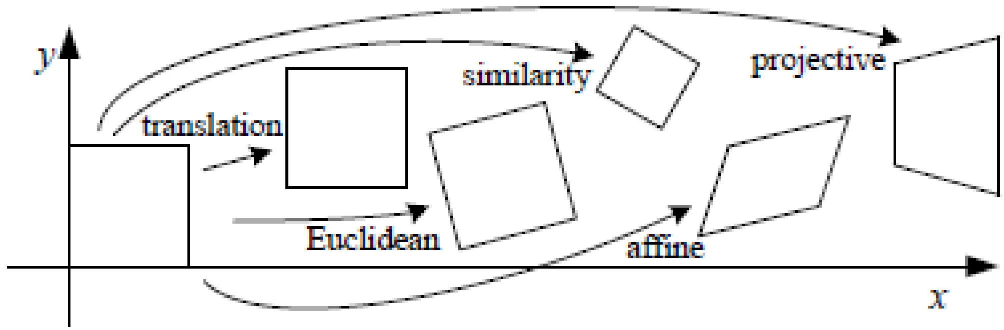
$$\begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4 y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5 y_5 & y_5^2 & x_5 & y_5 & 1 \end{bmatrix} \mathbf{c} = \mathbf{0}$$

This is a $5 \times\ 6$ matrix. The 1D null-space gives the conic up to scale.

Compute Nullspace with Gaussian elimination or SVD

2D transformations

2D Transformations in R^2



Definition:

- Euclidean: translation + rotation
- Similarity (rigid body transform): Euclidean + scaling
- Affine: Similarity + shearing
- Projective: arbitrary linear transform in homogenous coordinates

2D Transformations of points

- 2D Transformations in homogenous coordinates:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} a & b & d \\ e & f & h \\ i & j & l \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Transformation matrix

- Example: translation

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

homogeneous coordinates

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

inhomogeneous coordinates

Advantage of homogeneous coordinates (going into P^2)

2D transformations of points

Group	Matrix	Distortion	Invariant properties
Projective 8 dof	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$		Concurrency, collinearity, **order of contact**: intersection (1 pt contact); tangency (2 pt contact); inflections (3 pt contact with line); tangent discontinuities and cusps. cross ratio (ratio of ratio of lengths).
Affine 6 dof	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Parallelism, ratio of areas, ratio of lengths on collinear or parallel lines (e.g. midpoints), linear combinations of vectors (e.g. centroids). The line at infinity, $\mathbf{l}_\infty$.
Similarity 4 dof	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Ratio of lengths, angle. The circular points, **I**, **J** (see section 2.7.3). *(two special points on the line at infity)*
Euclidean 3 dof	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Length, area

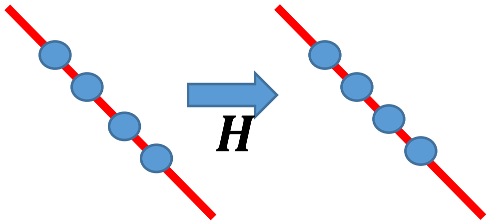
Here R is a 2 x 2 rotation matrix with 1 DoF which can be written as: $\begin{bmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{bmatrix}$

[from Hartley Zisserman Page 44]

2D transformations of lines and conics

All points move: $\boldsymbol{x}' = \boldsymbol{H}\boldsymbol{x}$ then:

1) Line (defined by points) moves:
$\boldsymbol{l}' = (\boldsymbol{H}^{-1})^{\mathrm{t}}\ \boldsymbol{l}$



2) conic (defined by points) moves:
$\boldsymbol{C}' = (\boldsymbol{H}^{-1})^{\mathrm{t}}\ \boldsymbol{C}\,\boldsymbol{H}^{-1}$

Example: Projective Transformation

Picture from top

1. Circles on the floor are circles in the image
2. Squares on the floor are squares in teh image

Affine transformation

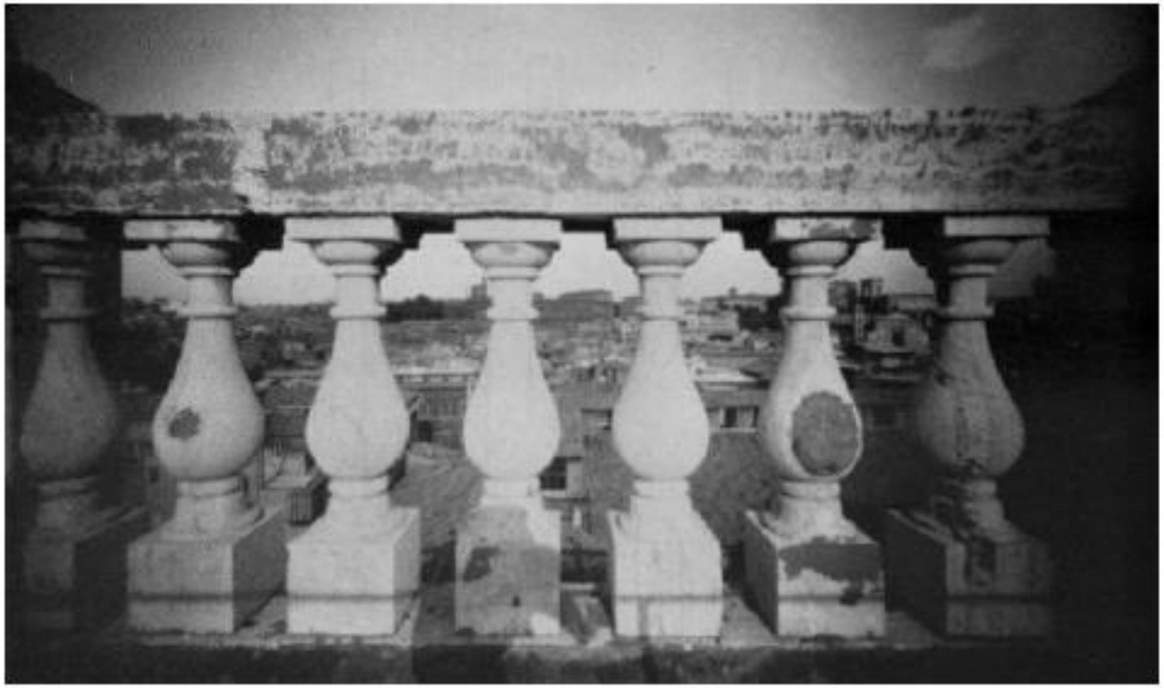
1. Circles on the floor are ellipse in the image
2. Squares on the floor are sheared in the image
3. Lines are still parallel

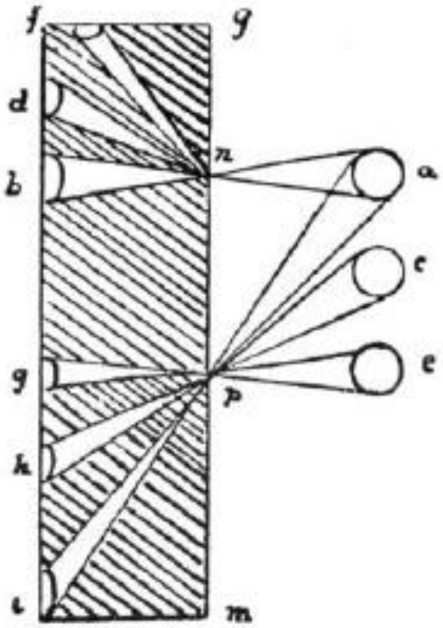
Picture from the side (projective transformation)

1. Lines converge to a vanishing point not at infinity

Persepcitive Distortion

- The exterior columns appear bigger
- The distortion is not due to lens flaws
- Problem was pointed out by Da Vinci



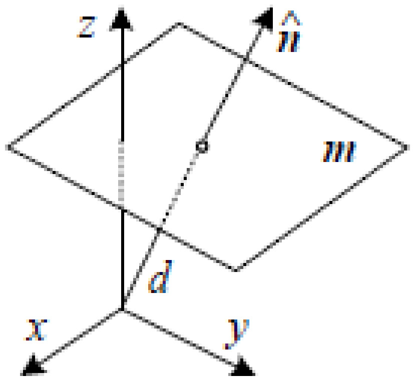


Now in 3D: Points

- $\boldsymbol{x} = (x, y, z) \in R^3$ has 3 DoF
- In homogeneous coordinates: $(x, y, z, 1) \in P^3$
- P^3 is defined as the space $R^3 \setminus (0,0,0,0)$ such that points $(x, y, z, w)$and (kx, ky, kz, kw) are the same for all $k \neq 0$
- Points: $(x, y, z, 0) \in P^3$ are called points at infinity

Now in 3D: Planes

- Planes in R^3 are defined as by 4 paramters (3 DoF):
  - Normal: $n = (n_x, n_y, n_z)$
  - Offset: d



- All points (x, y, z) lie on the plane if:

  $x\, n_x + y\, n_y + z\, n_z + d = 0$

- In homogeous coordinates:

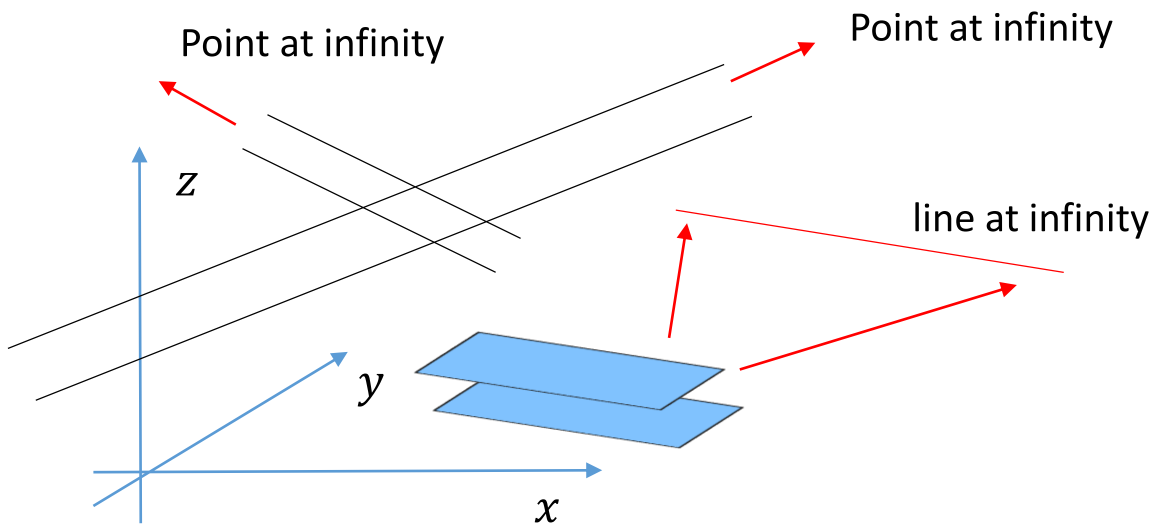
  $\boldsymbol{x}\, \pi = 0$, where $x = (x, y, z, 1)$ and $\pi = (n_x, n_y, n_z, d)$

- Planes in P^3 are written as: $\boldsymbol{x}\, \pi = 0$

- Points and planes are dual in P^3 (as points and lines have been in P^2)

- Plane at infinity is π = (0,0,0,1) since all points at infinity (x,y,z,0) lie on it.

Plane at infinity

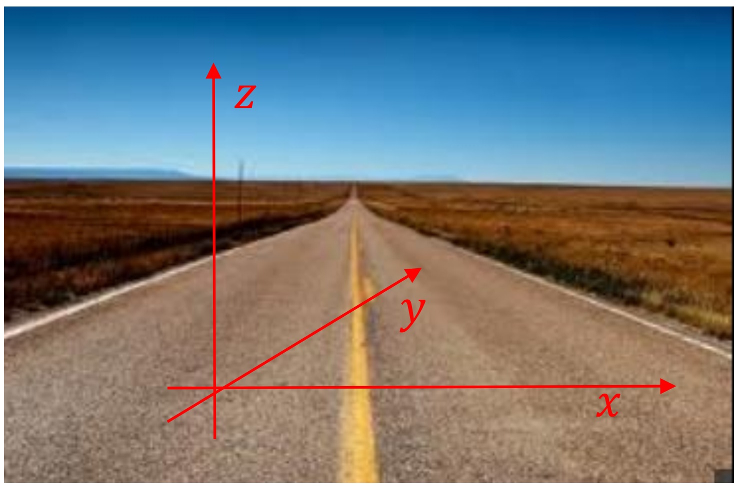


All of these elements at infinity lie on the plane at infinity

What is the horizon?

The ground plane is special (we/things stand on it)

Horizon is a line at infinity where plane at infinity intersects ground plane

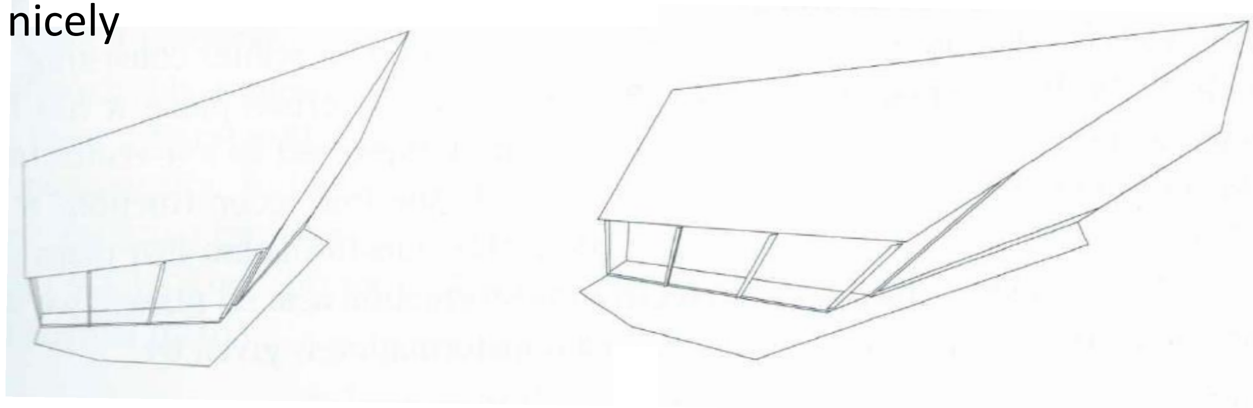


Ground plane: (0,0,1,0)
Plane at infinity: (0,0,0,1)

Many lines and planes in our real world meet at the horizon (since parallel to ground plane)

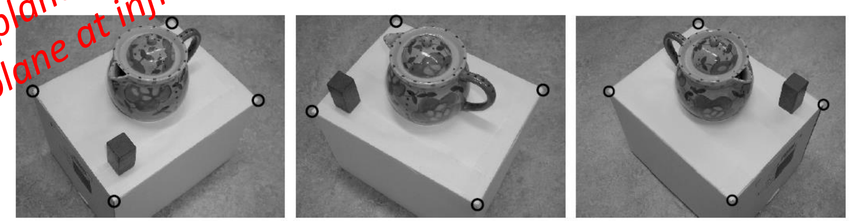
Why plane at infinity is important (we do later)

Plane at infinity is important to visualize 3D reconstructions nicely

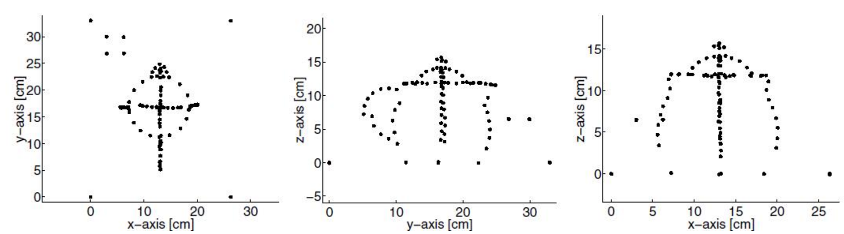


Plane at infinity can be used to simplify 3D reconstruction

Real plane defined as plane at infinity

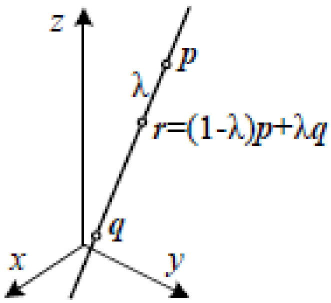


(a) (b) (c)



Now in 3D: Lines

- Unfortunately not a compact form (as for points)



- A simple representation in R^3.
  Define a line via two points $p, q \in R^3$:

$$\boldsymbol{r} = (1-\lambda)\boldsymbol{p} + \lambda\,\boldsymbol{q}$$

- A line has 4 DoF (both points $\boldsymbol{p}, \boldsymbol{q}$ can move arbitrary on the line)

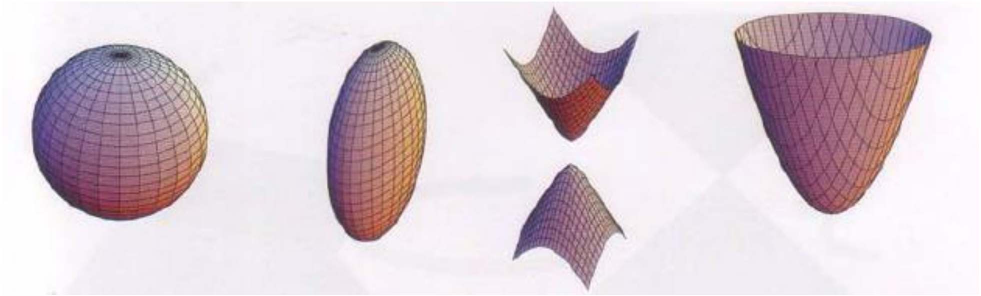
- A more compact, but more complex, way two define a 3D Line is to use Plücker coordinates:

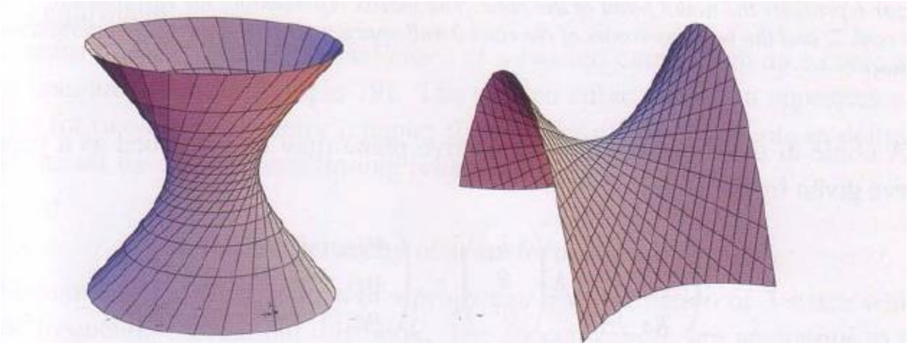
$$\boldsymbol{L} = \boldsymbol{p}\boldsymbol{q}^{t} - \boldsymbol{q}\boldsymbol{p}^{t} \text{ where } \det(\boldsymbol{L}) = 0$$

here $\boldsymbol{L}, \boldsymbol{p}, \boldsymbol{q}$ are in homogenous coordinates

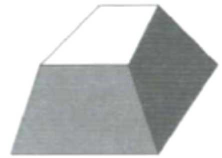
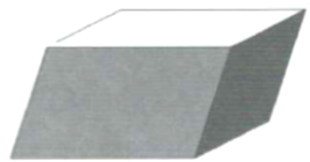
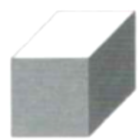
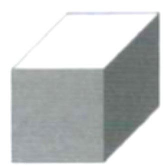
Now in 3D: Quadrics

- Points $\boldsymbol{X}$ on the quadric if: $\boldsymbol{X}^T \boldsymbol{Q} \boldsymbol{X} = 0$
- A quadric $\boldsymbol{Q}$ is a surface in P^3
- A quadric is a symmetric 4×4 matrix with 9 DoF





3D Transformation

Group	Matrix	Distortion	Invariant properties
Projective 15 dof	$\begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{v}^{\mathsf{T}} & v \end{bmatrix}$		Intersection and tangency of surfaces in contact. Sign of Gaussian curvature.
Affine 12 dof	$\begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix}$		Parallelism of planes, volume ratios, centroids. The plane at infinity, π_∞, (see section 3.5).
Similarity 7 dof	$\begin{bmatrix} s\mathtt{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix}$		The absolute conic, Ω_∞, (see section 3.6).
Euclidean 6 dof	$\begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix}$		Volume.

3D Rotations

Rotation $\boldsymbol{R}$ in 3D has 3 DoF. It is slightly more complex, and several options exist:

1) Euler angles: rotate around, x, y, z-axis in order
   (depends on order, not smooth in parameter space)

2) Axis/angle formulation:
   $$\boldsymbol{R}(\boldsymbol{n}, \Theta) = \boldsymbol{I} + \sin\Theta\,[\boldsymbol{n}]_\times + (1 - \cos\Theta)[\boldsymbol{n}]_\times^2$$
   $\boldsymbol{n}$ is the normal vector (2 DoF) and Θ the angle (1 DoF)

3) Another option is unit quaternions (see book page 40)

Roadmap this lecture (image formation process)

- Geometric primitives and transformations (sec. 2.1.1-2.1.4)

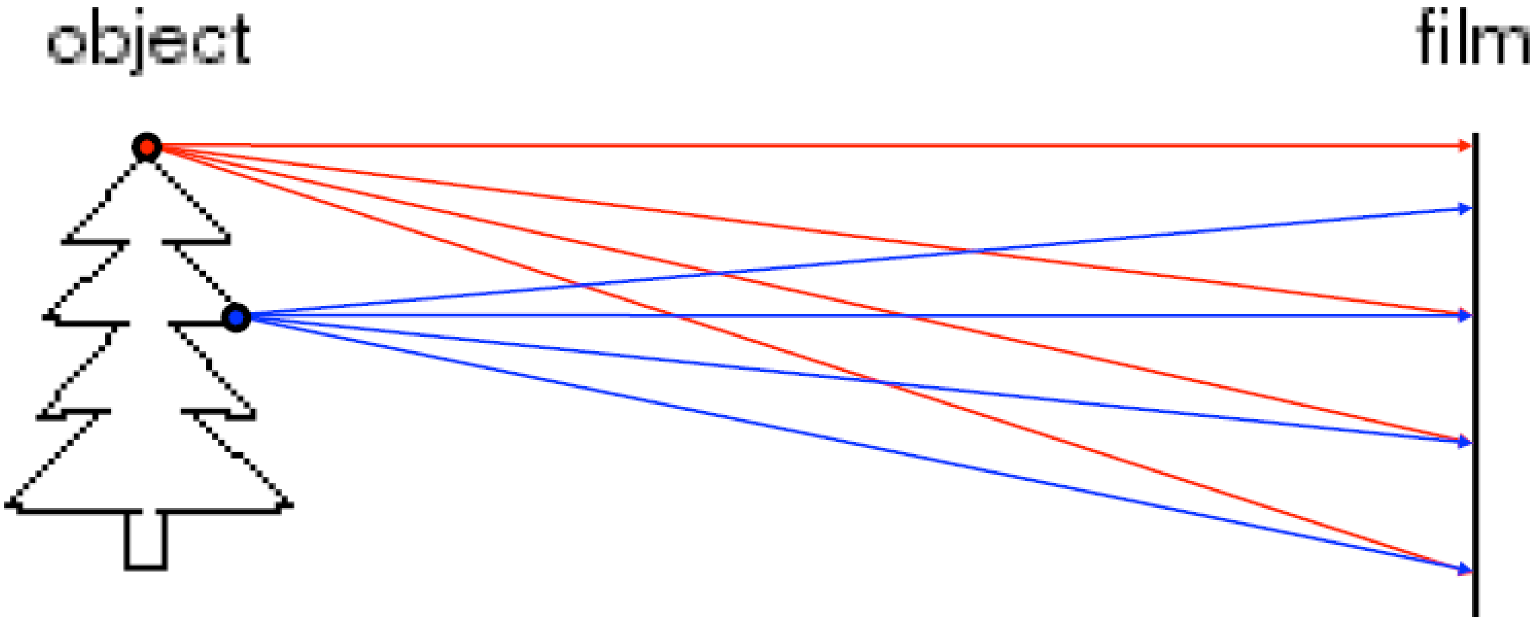
- Geometric image formation process (sec. 2.1.5, 2.1.6)
  - Pinhole camera
  - Lens effects

- The Human eye

- Photometric image formation process (sec. 2.2)

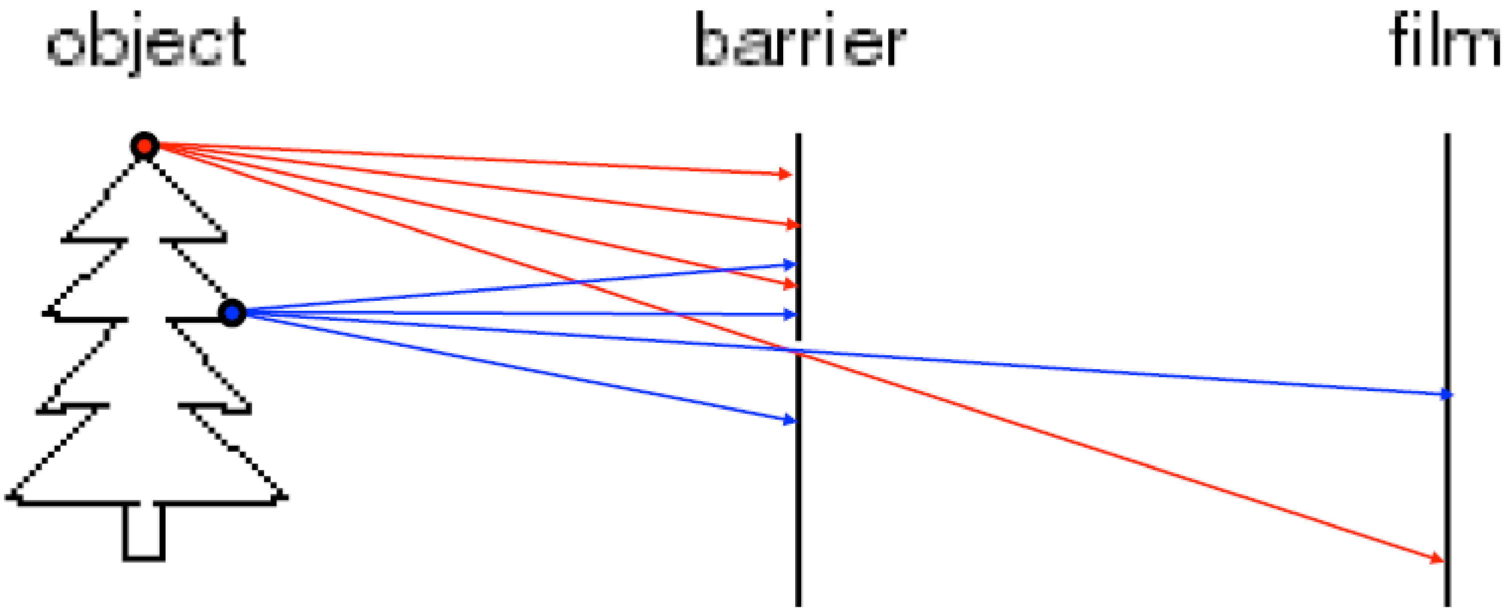
- Camera Types and Hardware (sec 2.3)

How can we capture the world



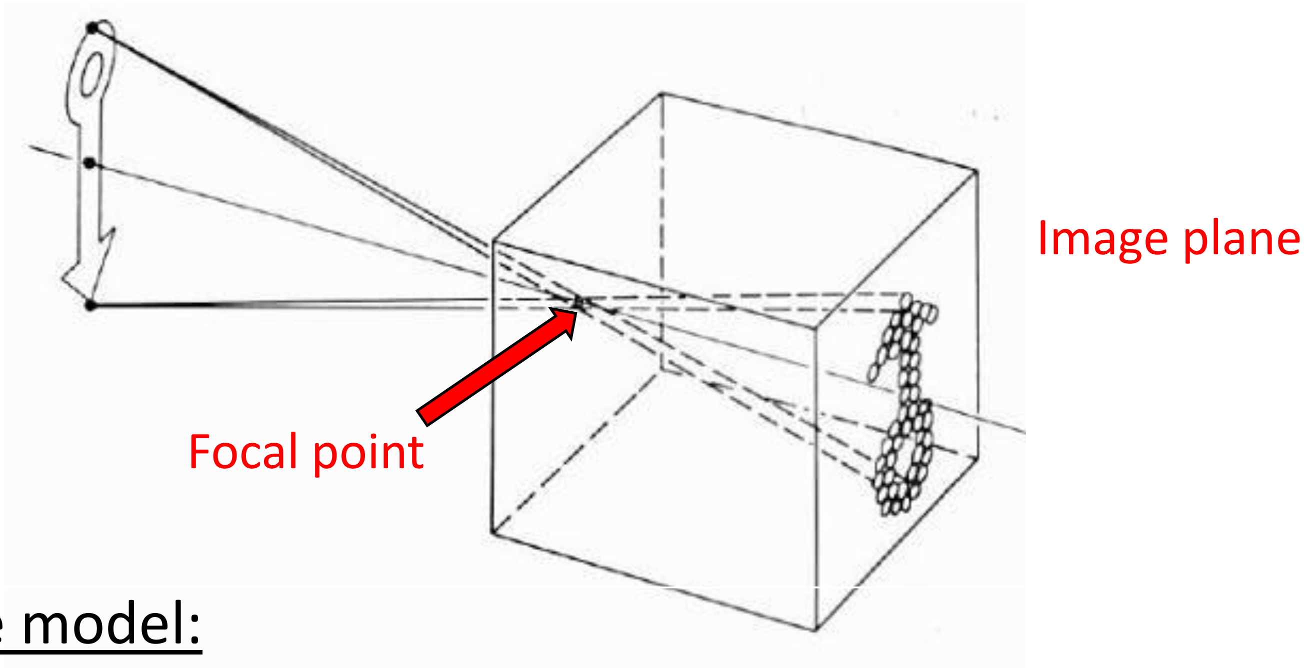
- Let’s design a camera
  - Idea 1: put a piece of film (or a CCD) in front of an object
  - Do we get a reasonable image?

Pinhole Camera



- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the aperture (“Blende”)

Pinhole camera model

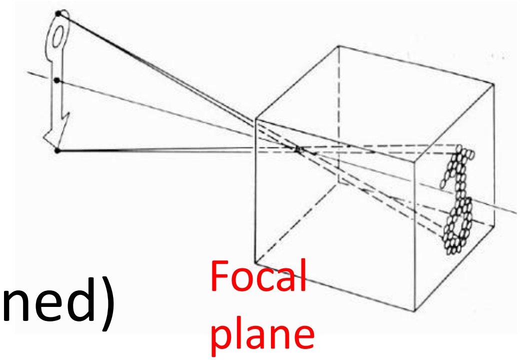


Pinhole model:

- Captures pencil of rays – all rays through a single point
- Projected rays are straight lines
- The point where all rays meet is called center of projection (focal point)
- The image is formed on the image plane

Pinhole camera – Properties

- Many-to-one: any point along the same ray maps to the same point in the image



- Points map to points

(But projection of points on focal plane is undefined)

- Lines map to lines (collinearity is preserved)
  (But line through focal point projects to a point)

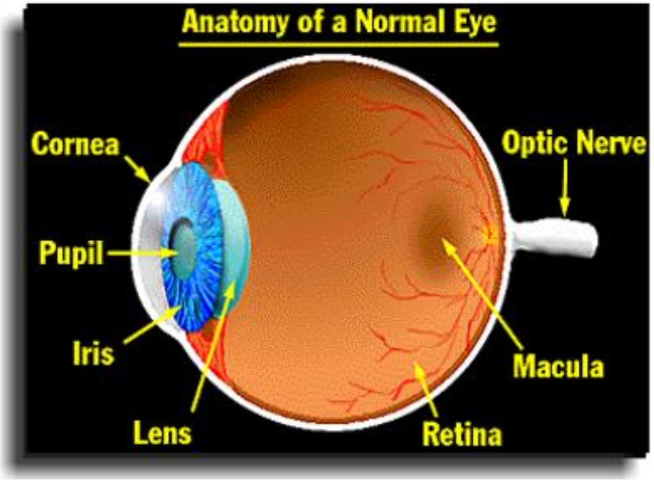
- Planes map to planes (or half-plane)

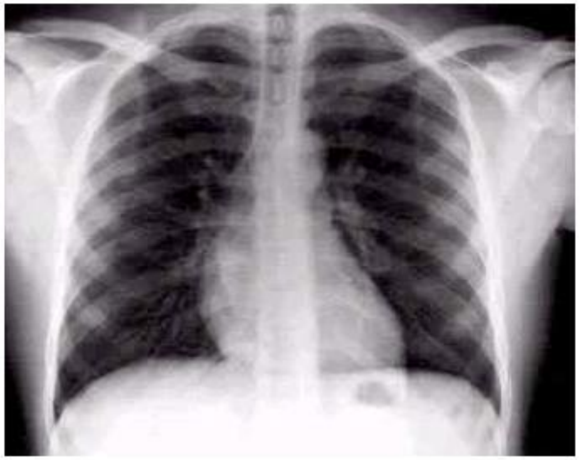
(But plane through focal point projects to line)

Pinhole Camera

A model for many common sensors:

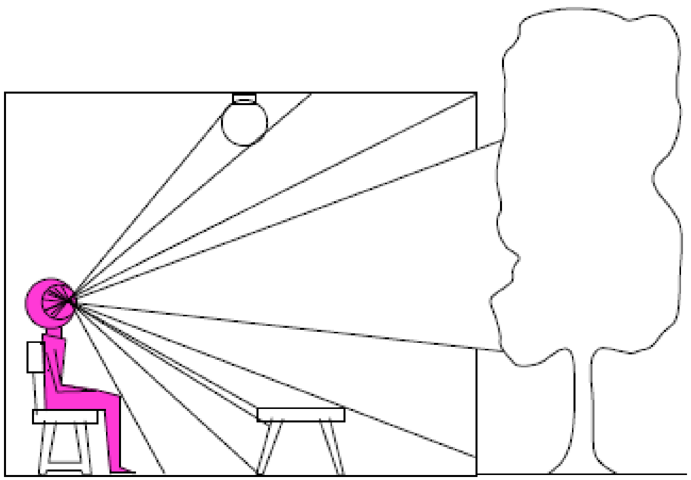
- Photographic cameras
- human eye
- X-ray machines





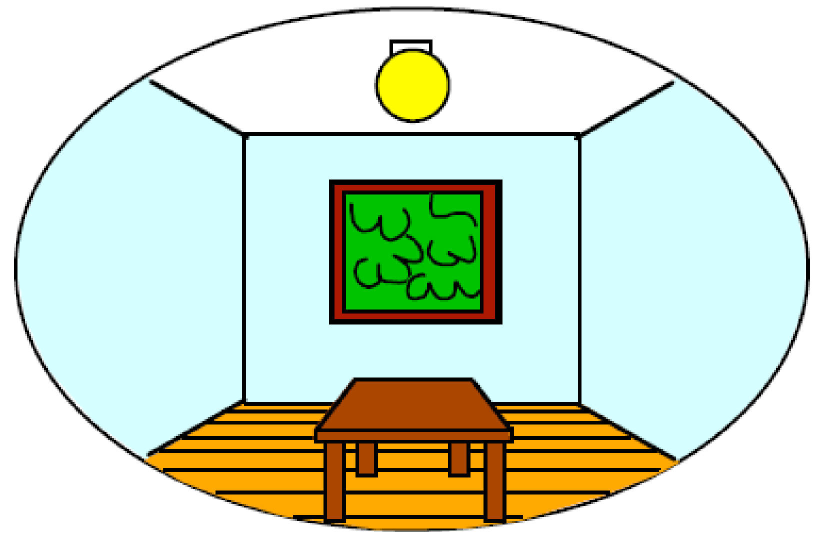
Dimensionality Reduction Machine (3D to 2D)

3D world



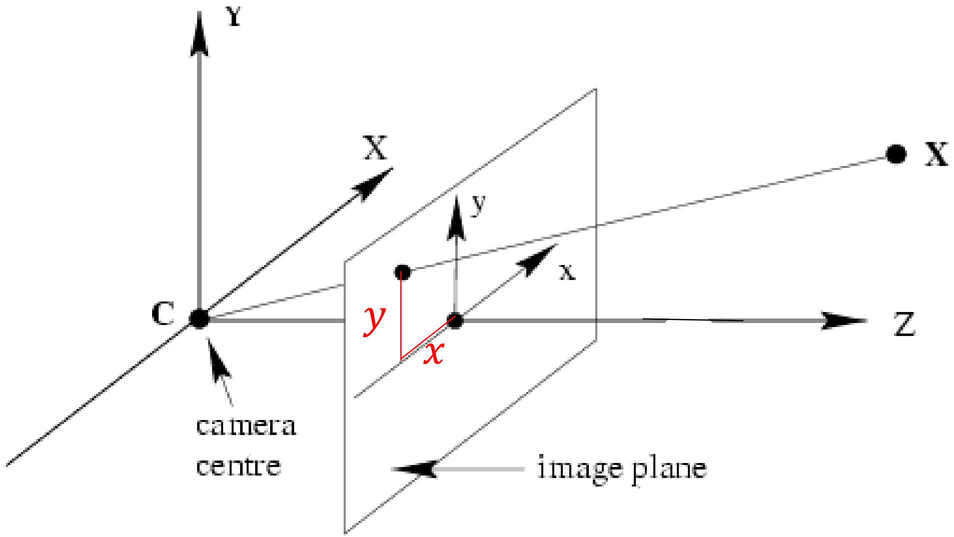
Point of observation

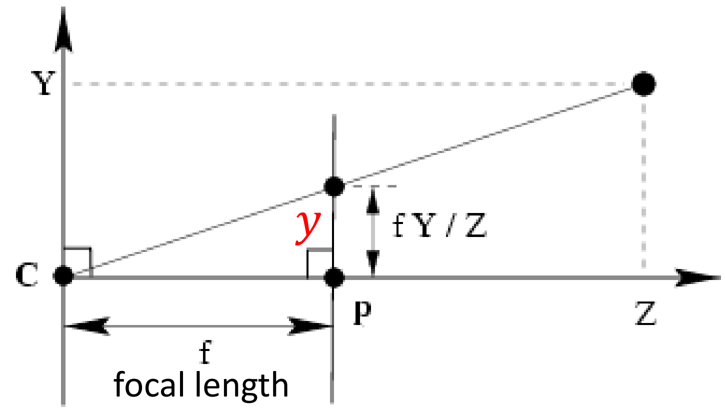
2D image



- What have we lost?
  - Angles
  - Distances (lengths)

Pinhole camera model – in maths





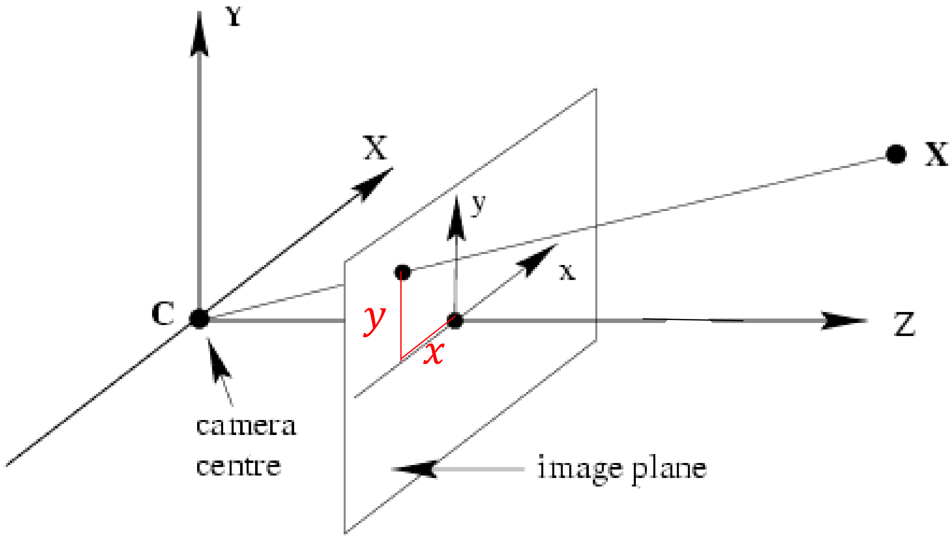
- Similar trinagles: $\frac{y}{f} = \frac{Y}{Z}$
- That gives: $y = f\frac{Y}{Z}$ and $x = f\frac{X}{Z}$

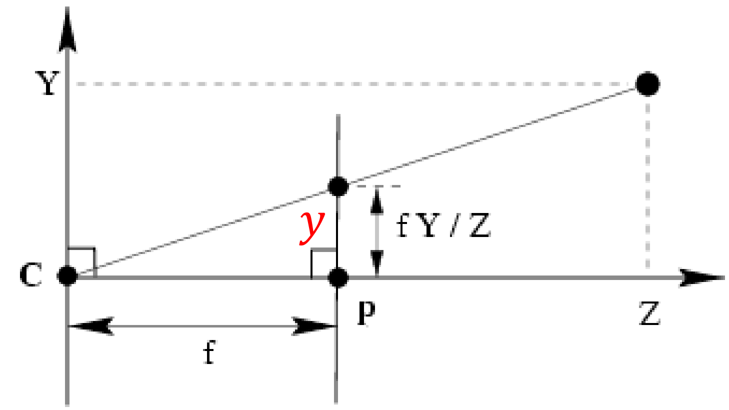
- That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ (remeber “=“ means equal up to scale)

2D homogenous coordinate

3D inhomogenous coordinate

Pinhole camera model – in maths





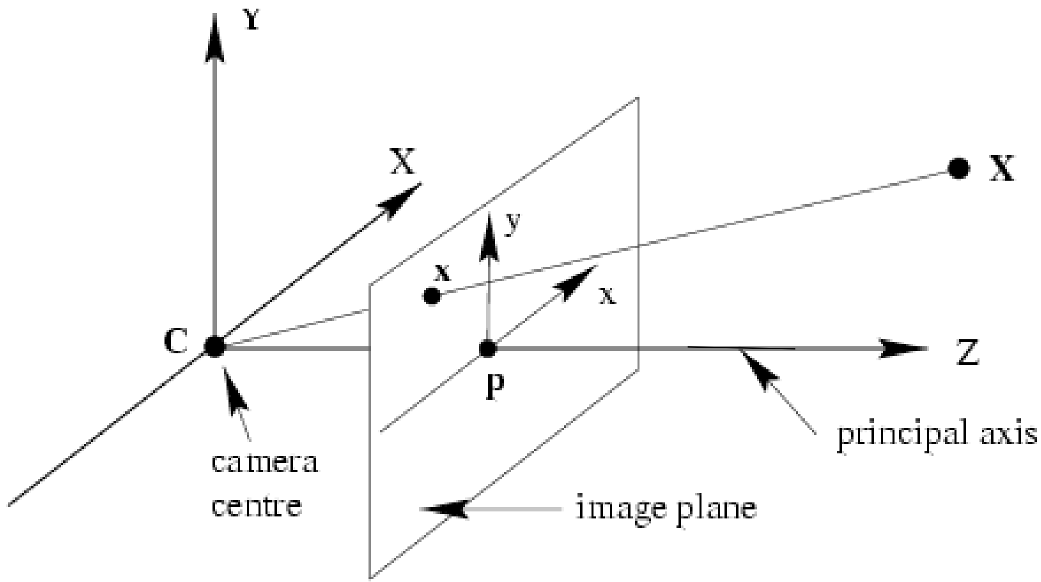
That gives: $$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Calibration matrix } \boldsymbol{K}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

In short $\boldsymbol{x} = \boldsymbol{K}\,\tilde{\boldsymbol{X}}$ (here $\tilde{\boldsymbol{X}}$ means inhomogeneous coordinates)

Intrinsic Camera Calibration means we know $\boldsymbol{K}$ (we do that later)

We can go from image points into the 3D world: $\tilde{\boldsymbol{X}} = \boldsymbol{K}^{-1}\,\boldsymbol{x}$

Pinhole camera - definitions



- Principal axis: line from the camera center perpendicular to the image plane
- Normalized (camera) coordinate system: camera center is at the origin and the principal axis is the z-axis
- Principal point (p): point where principal axis intersects the image plane (origin of normalized coordinate system)

Principal Point

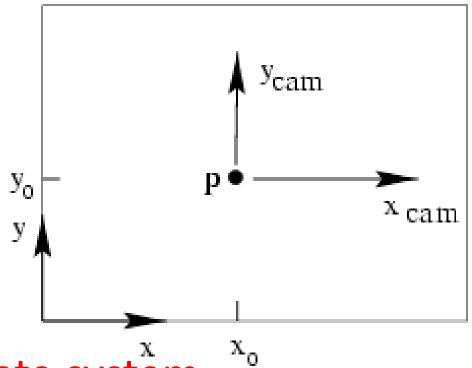


Image
coordinate system

Principal point (p_x, p_y)

- Camera coordinate system: origin is at the principal point
- Image coordinate system: origin is in the corner
  In practice: principal point in center of the image

Adding principal point into $\boldsymbol{K}$

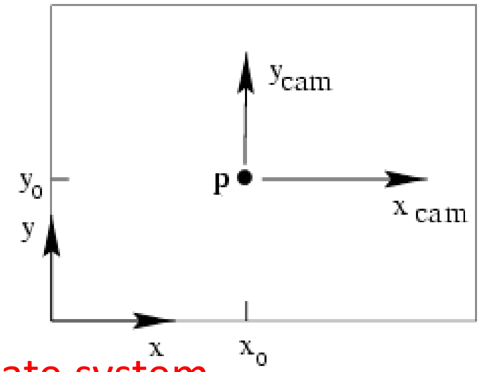


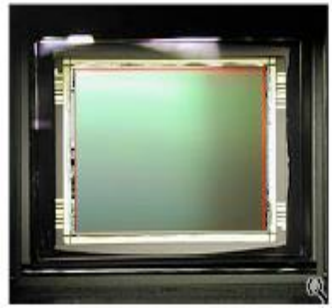
Image
coordinate system

Principal point (p_x, p_y)

Projection with principal point : $y = f\frac{Y}{Z} + p_y = \frac{fY + Zp_y}{Z}$ and $x = f\frac{X}{Z} + p_x = \frac{fX + Zp_x}{Z}$

That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

Pixel Size and Shape



- m_x pixels per unit (m,mm,inch,...) in horizontal direction
- m_y pixels per unit (m,mm,inch,...) in vertical direction
- s' skew of a pixel
- *In practice (close to): m=1 s = 0*

That gives: $$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} m_x & s' & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Simplified to: $$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Final calibration matrix } \boldsymbol{K}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

f now in units of pixels

Camera intrinsic parameters - Summary

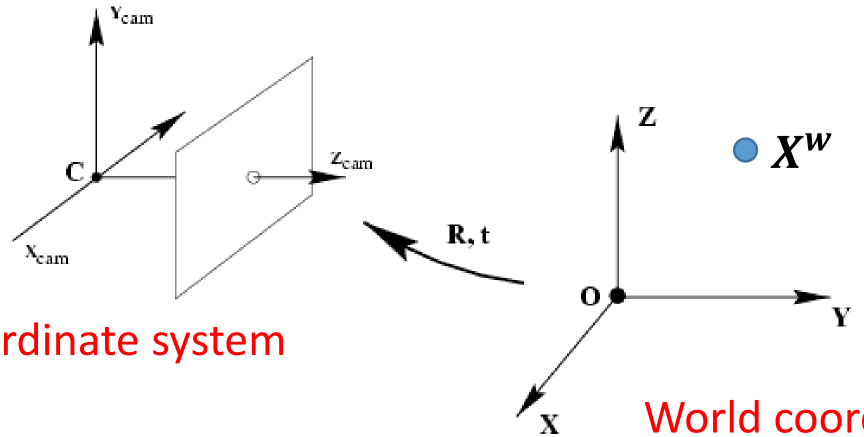
- Intrinsic parameters
  - Principal point coordinates (p_x, p_y)
  - Focal length f
  - Pixel magnification factors m
  - Skew (non-rectangular pixels) s

$$\boldsymbol{K} = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

For later:
We sometimes have to only guess these values and then they are optimized over (bundle adjustment):

- p in image center,
- $s = 0, m = 1$
- f= EXIF tag (or guess, e.g. two times image size)

Putting the camera into the world



camera coordinate system

World coordinate system

Given a 3D homogenous point $\boldsymbol{X^w}$ in world coordinate system

1) Translate from world to camera coordinate system:

$$\tilde{\boldsymbol{X}}^{c'} = \tilde{\boldsymbol{X}}^{w} - \tilde{\boldsymbol{C}}$$

$$\tilde{\boldsymbol{X}}^{c'} = \underbrace{(\boldsymbol{I}_{3\times3} \mid -\tilde{\boldsymbol{C}})}_{3\times4 \text{ matrix}} \boldsymbol{X}^{w}$$ where $\boldsymbol{I}_{3\times3}$ is 3x3 identity matrix

2) Rotate world coordinate system into camera coordinate system

$$\tilde{\boldsymbol{X}}^{c} = \boldsymbol{R}\,(\boldsymbol{I}_{3\times3} \mid -\tilde{\boldsymbol{C}})\,\boldsymbol{X}^{w}$$

3) Apply camera matrix

$$\boldsymbol{x} = \boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times3} \mid -\tilde{\boldsymbol{C}})\,\boldsymbol{X}$$

Camera matrix

- Camera matrix $\boldsymbol{P}$ is defined as:

$$\boldsymbol{x} = \underbrace{\boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times 3} \mid -\tilde{\boldsymbol{C}})}_{\boldsymbol{P}\,(3\times 4)\text{ camera matrix has 11 DoF}}\,\boldsymbol{X}$$

- In short we write: $\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X}$

- The camera center is the (right) nullspace of P

$\boldsymbol{P}\,\boldsymbol{C} = \boldsymbol{K}\,\boldsymbol{R}\,(\tilde{\boldsymbol{C}} - \tilde{\boldsymbol{C}}) = 0$

Camera parameters - Summary

- Camera matrix P has 11 DoF

$$\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X}$$

$$\boldsymbol{x} = \boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times 3}\,|\,-\tilde{\boldsymbol{C}})\,\boldsymbol{X}$$

$$\boldsymbol{K} = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

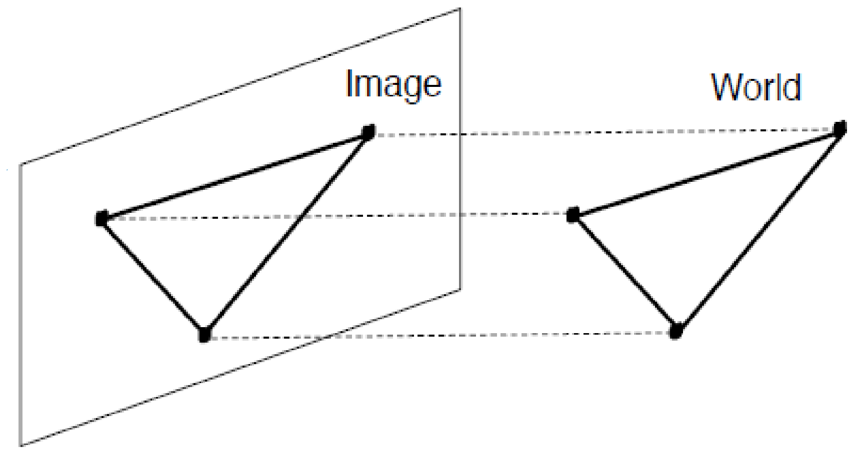
- <u>Intrinsic parameters</u>
  - Principal point coordinates (p_x, p_y)
  - Focal length f
  - Pixel magnification factors m
  - Skew (non-rectangular pixels) s

- <u>Extrinsic parameters</u>
  - Rotation $\boldsymbol{R}$ (3DoF) and translation $\tilde{\mathbf{C}}$ (3DoF) relative to world coordinate system

Orthographic Projection

Special case of perspective projection

- Distance from center of projection to image plane is infinite (infinite focal length)
- Also called "parallel projection"



- Most simple from of projection

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Affine cameras

- Most general camera that does parallel projection are:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ w \end{bmatrix}$$

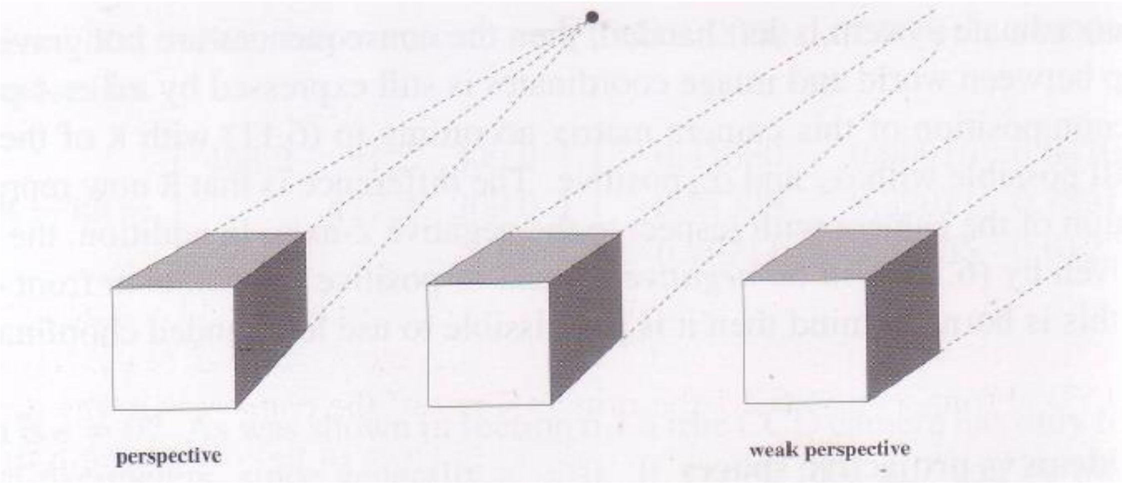
- Parallel lines map to parallel lines (since points at infinity stay)

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 0 \end{bmatrix}$$

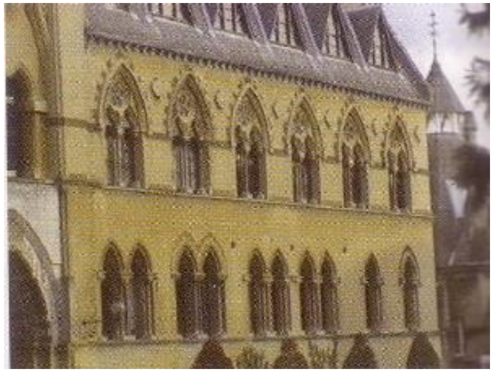
- Affine cameras simplify 3D reconstruction task, hence good to get an approximate solution

- See more details on: HZ (Hartley, Zissermann) chapter 6.3

Going from perspective to orthographic




(normal focal length)


(very large focal length)

Roadmap this lecture (image formation process)

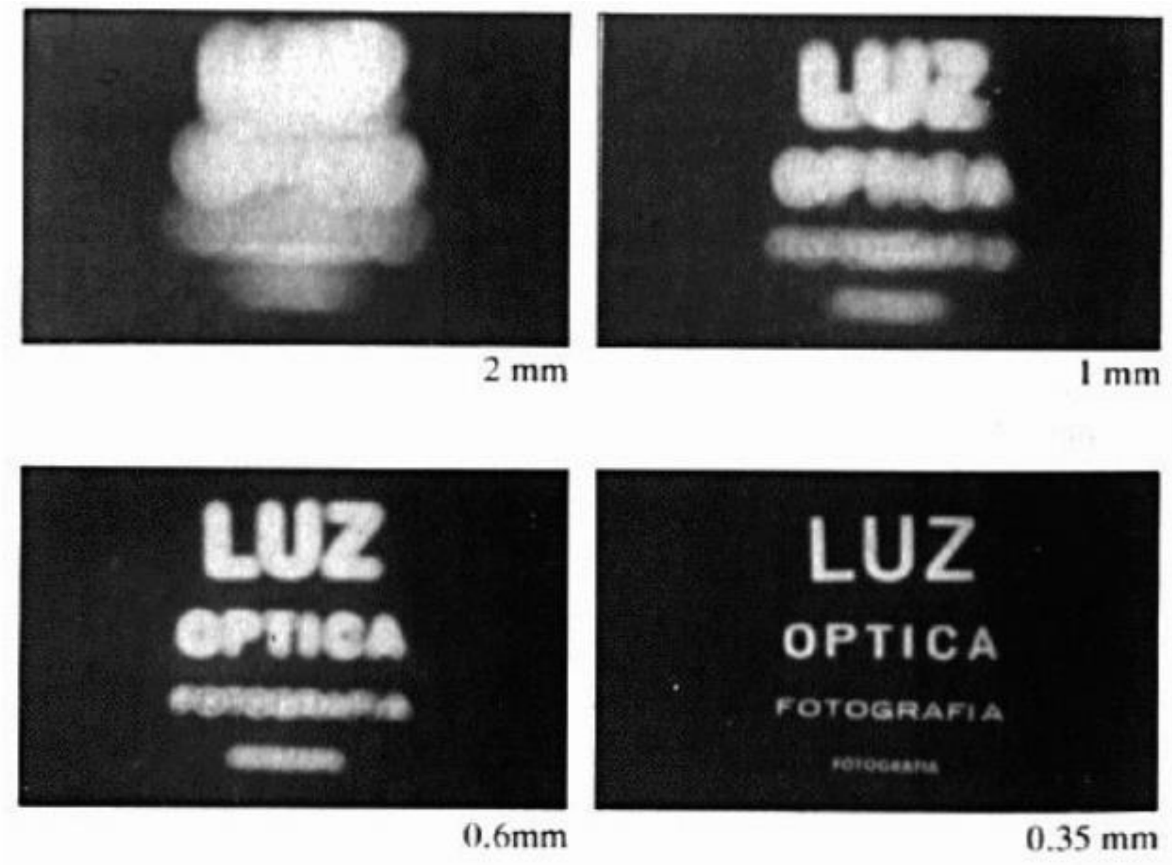
- Geometric primitives and transformations (sec. 2.1.1-2.1.4)
- Geometric image formation process (sec. 2.1.5, 2.1.6)
  - Pinhole camera
  - Lens effects
- The Human eye
- Photometric image formation process (sec. 2.2)
- Camera Types and Hardware (sec 2.3)

Home-made pinhole camera

Why so blurry?

http://www.debevec.org/Pinhole/

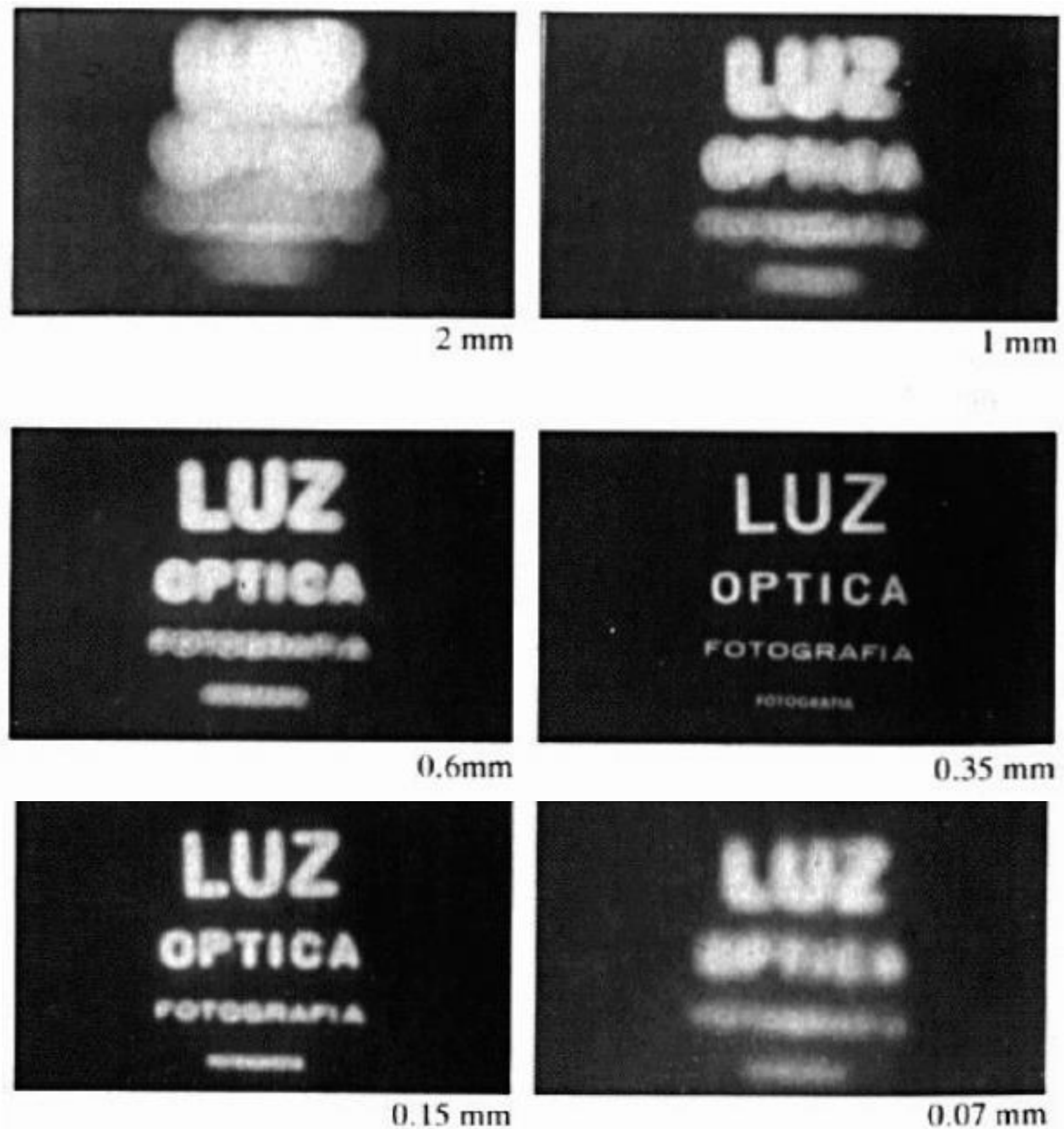
Shrinking the aperture



Why not make the aperture as small as possible?

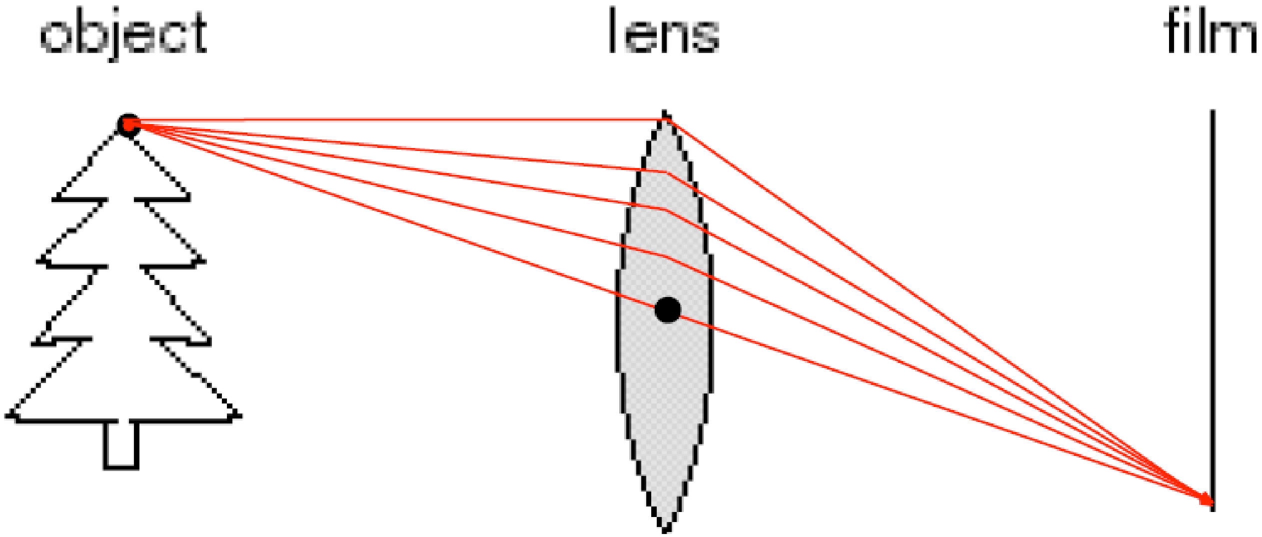
- Less light gets through
- Diffraction effects…

Shrinking the aperture



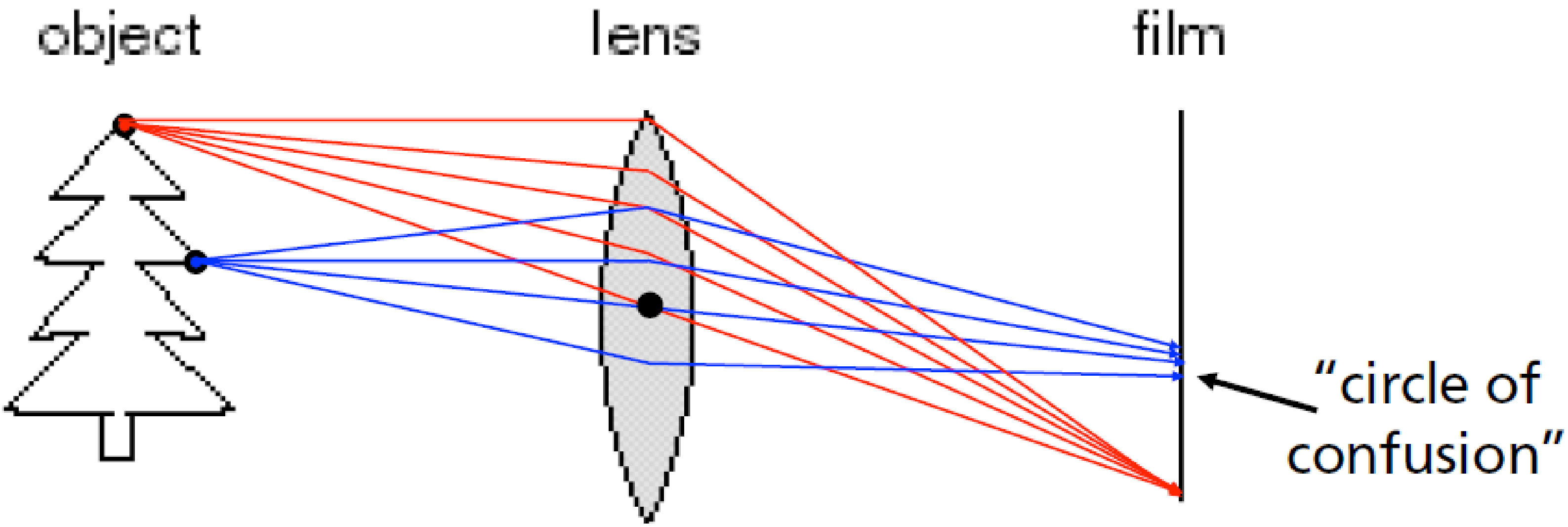
Diffraction effects. Noise due to long exposure

Adding a lens



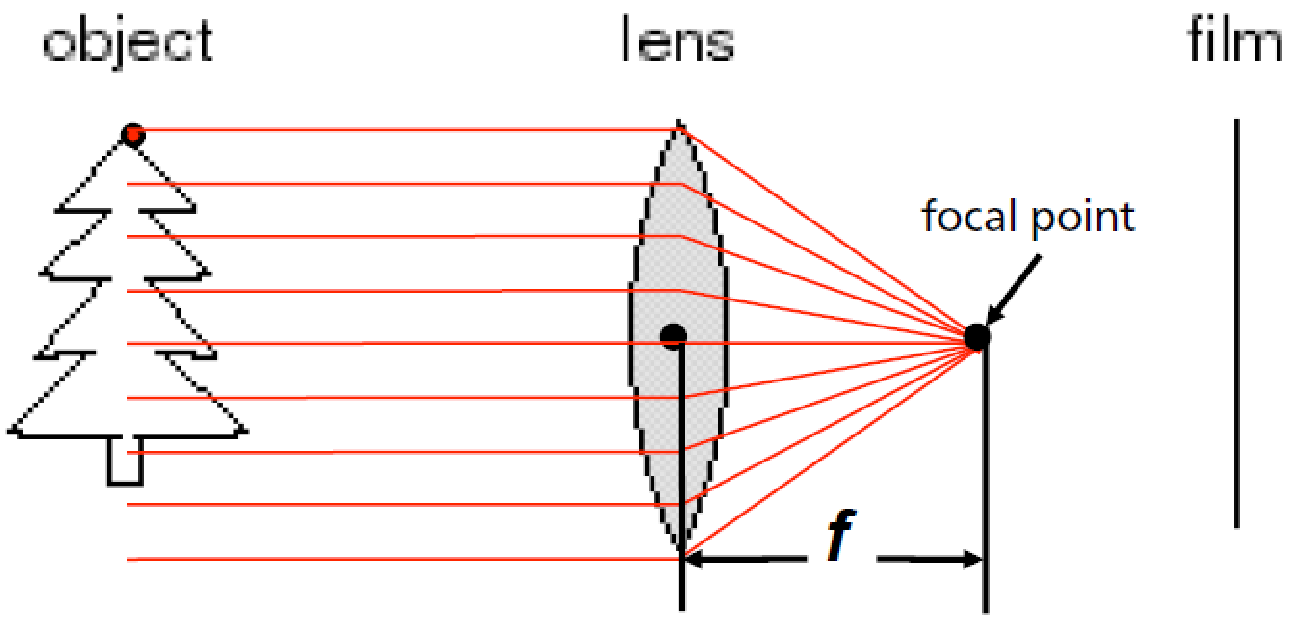
- A lens focuses light onto the film
- Lets enough ligh through
- Rays passing through the center are not deviated

In Focus



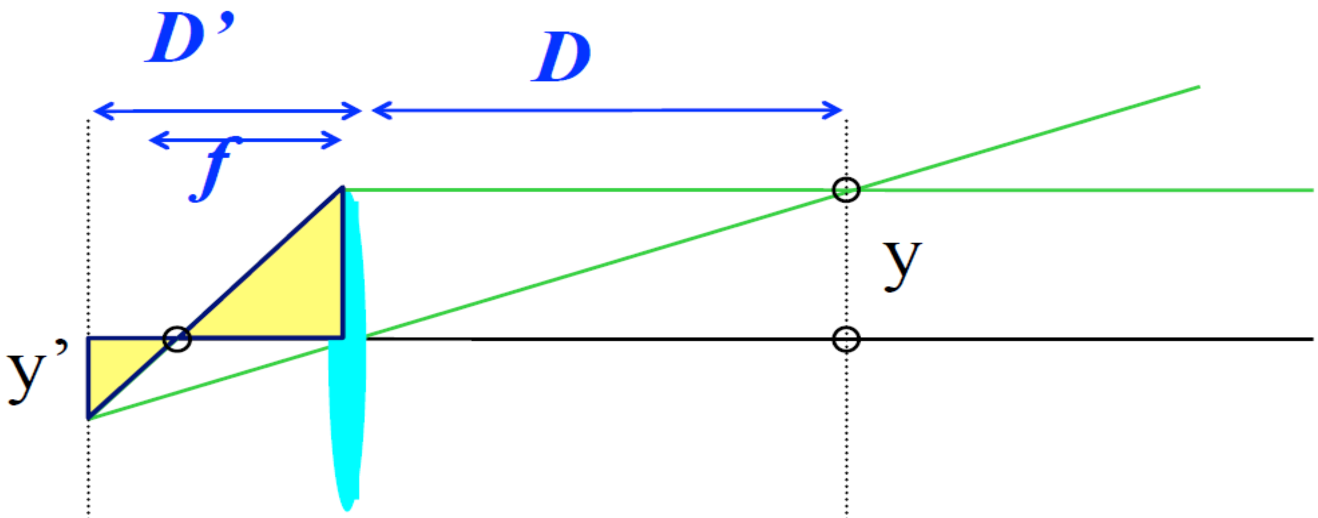
- There is a specific distance at which objects are "in focus"
- other points project to a "circle of confusion" in the image

Focal point



- All parallel rays converge to one point on a plane located at the *focal length f*

Thin lens formula



Green similiar triangles:

$$\frac{y'}{y} = \frac{D'}{D}$$

Yellow similiar triangles:

$$\frac{y'}{y} = \frac{D' - f}{f}$$

$$\frac{D'}{D} = \frac{D' - f}{f} \Rightarrow \frac{D'}{D} = \frac{D'}{f} - 1 \Rightarrow \boxed{\frac{1}{D} + \frac{1}{D'} = \frac{1}{f}}$$

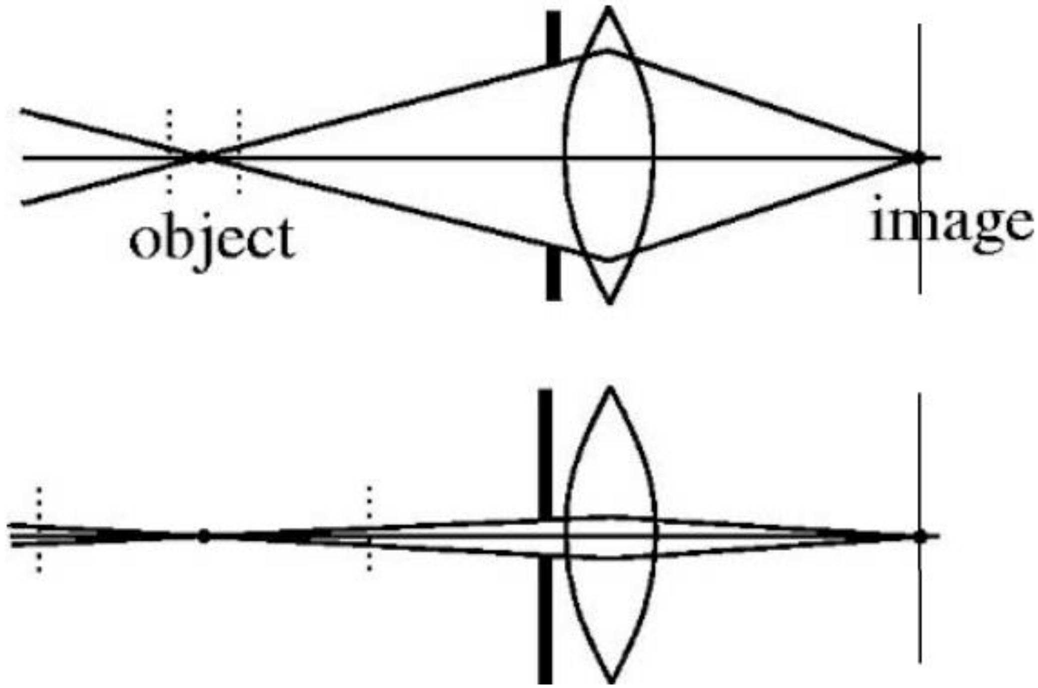
Thin lens formula

Any point satisfying the thin lens equation is in focus.

Depth of Field

http://www.cambridgeincolour.com/tutorials/depth-of-field.htm

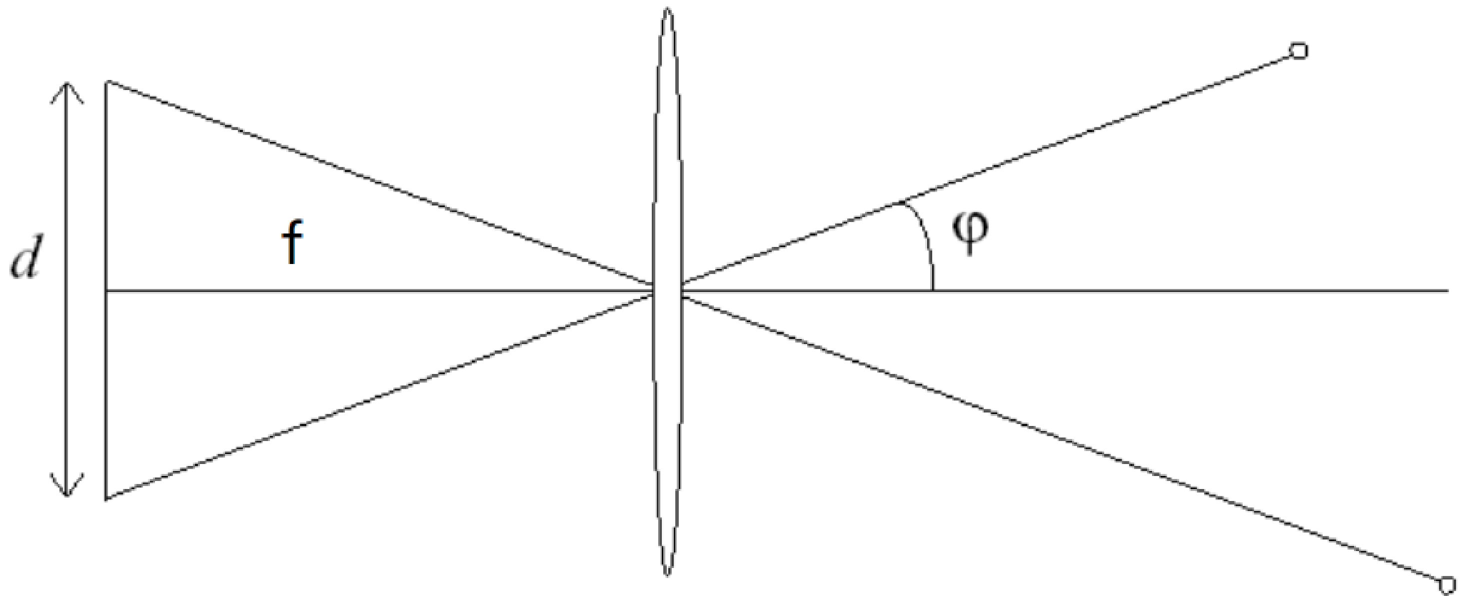
Control the depth of field



Changing the aperture size affects depth of field

- A smaller aperture increases the range in which the object is approximately in focus (but longer exposure needed)

Fields of View

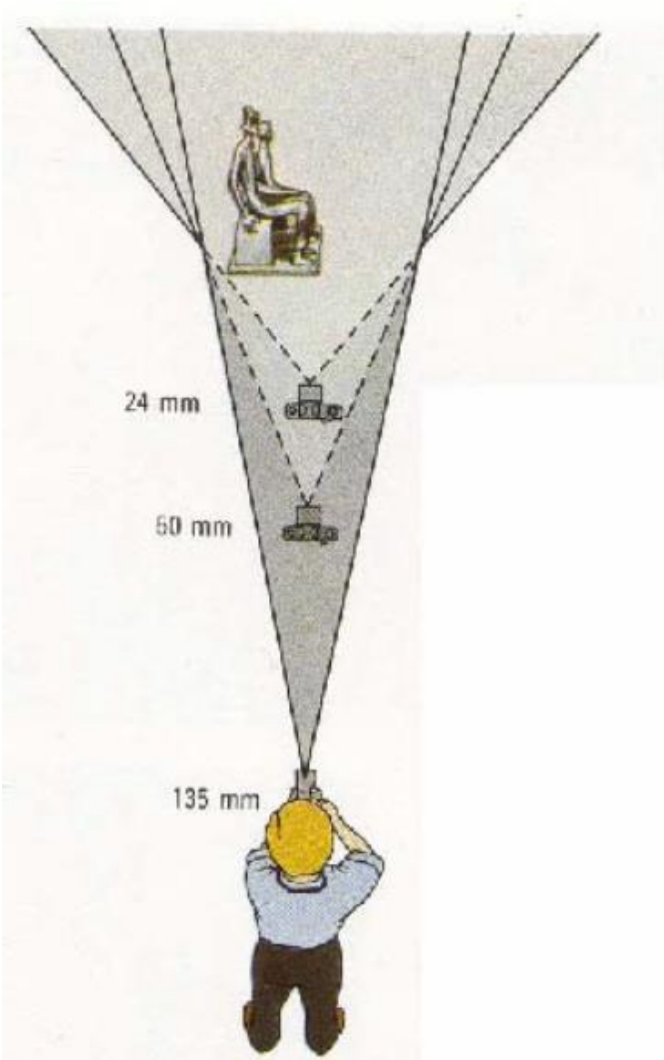


FOV depends on focal length and size of the camera retina

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

Smaller FOV = larger Focal Length

Field of View / Focal Length



Large FOV, small f
Camera close to car

Small FOV, large f
Camera far from the car

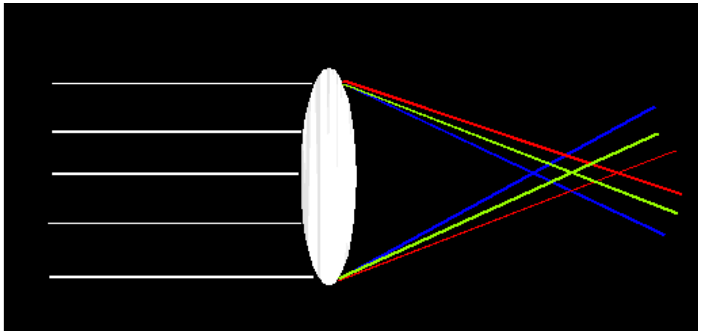
Close to affine camera (look at front light)

Same effect for faces

Small f Large f

Lens Flaws: Chromatic Aberration

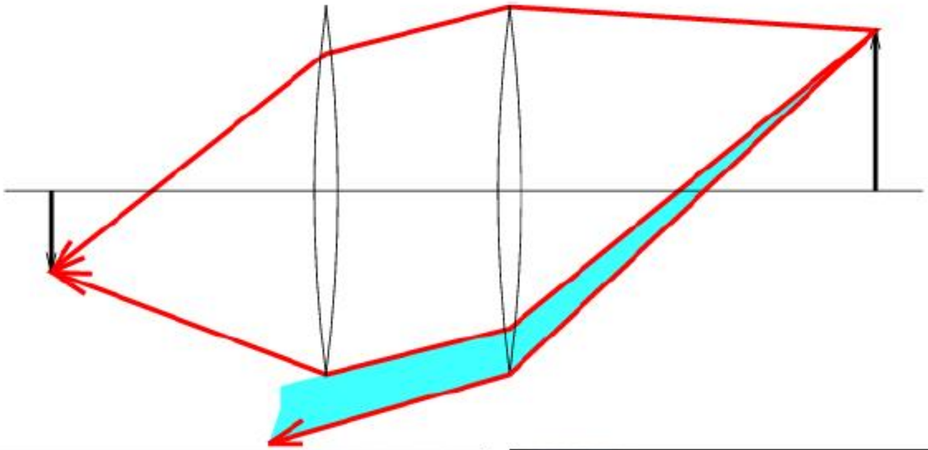
Lens has different refractive indices for different wavelengths: causes color fringing



High quality lens (top)
low quality lens (bottom)
blur + green edges

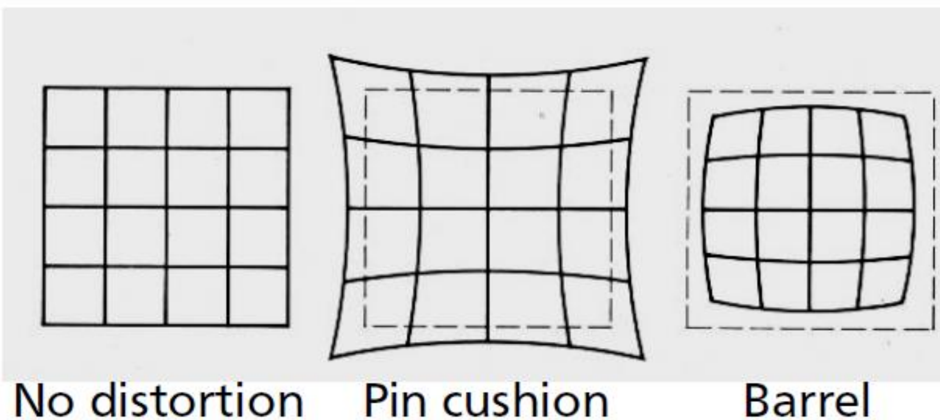
Purple fringing

Lens flaws: Vignetting



Lens distortion

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



Reading for next class

This lecture:

- Geometric primitives and transformations (sec. 2.1.1-2.1.4)
- Geometric image formation (sec 2.1.5, 2.1.6, HZ ?)

Next lecture:

- Photometric image formation (sec 2.2)
- Camera Types and Hardware (sec 2.3)
- Appearance matching: (sec. 4.1.2-4.1.3)