Image Formation and Cameras

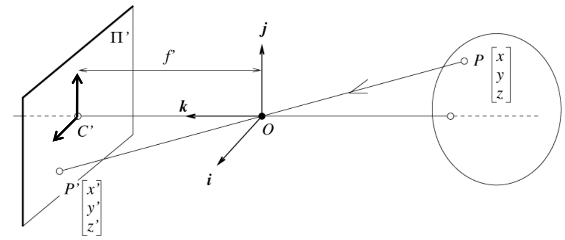
Guest Professor: Dr. Ana Murillo

Computer Vision I

CSE 252A

Lecture 4

Equation of Perspective Projection



Cartesian coordinates:

- We have, by similar triangles, that (x, y, z) -> (f' x/z, f' y/z, f')
- Establishing an image plane coordinate system at C' aligned with i and j, we get $(x,y,z) \rightarrow (f'\frac{x}{z}, f'\frac{y}{z})$

Projective geometry provides an elegant means for handling these different situations in a unified way and **homogenous coordinates** are a way to represent entities (points & lines) in projective spaces.

Projective Geometry

- Axioms of Projective Plane
  1. Every two distinct points define a line
  2. Every two distinct lines define a point (intersect at a point)
  3. There exists three points, A,B,C such that C does not lie on the line defined by A and B.
- Different than Euclidean (affine) geometry
- Projective plane is "bigger" than affine plane – includes "line at infinity"

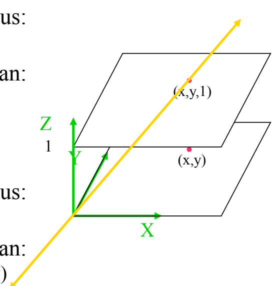
Conversion
Euclidean -> Homogenous -> Euclidean

In 2-D

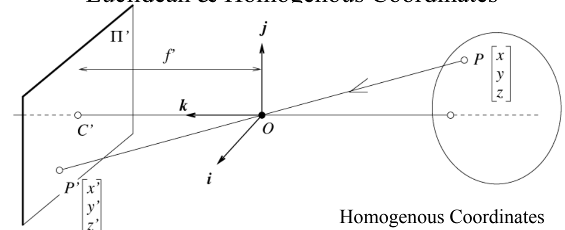
- Euclidean -> Homogenous:
  (x, y) -> k (x,y,1)
- Homogenous -> Euclidean:
  (x, y, z) -> (x/z, y/z)

In 3-D

- Euclidean -> Homogenous:
  (x, y, z) -> k (x,y,z,1)
- Homogenous -> Euclidean:
  (x, y, z, w) -> (x/w, y/w, z/w)



The equation of projection: Euclidean & Homogenous Coordinates



Cartesian coordinates:

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

Homogenous Coordinates and Camera matrix

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Projective transformation

- Also called a homography
- This is a mapping from 2-D to 2-D in homogenous coordinates
- 3 x 3 linear transformation of homogenous coordinates

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{21} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- Points map to points
- Lines map to lines

Mapping from a Plane to a Plane under Perspective is given by a projective transform H

x' = Hx   H is a 3x3 matrix,
x is a 3x1 vector of homogenous coordinates

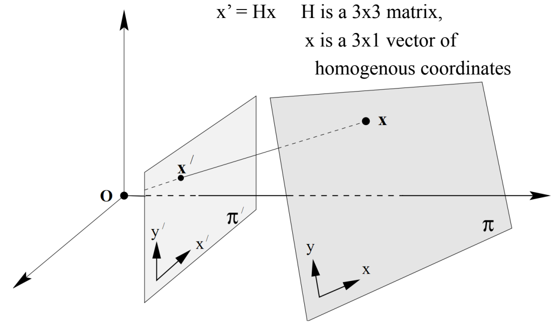
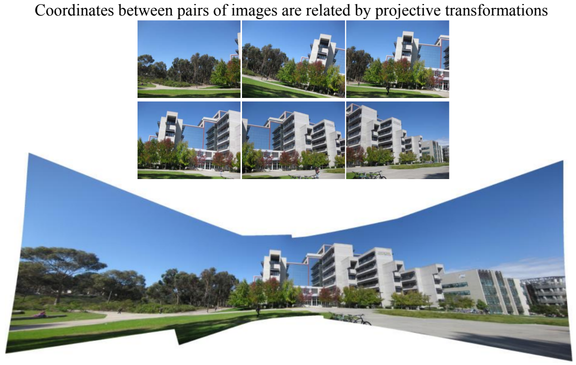


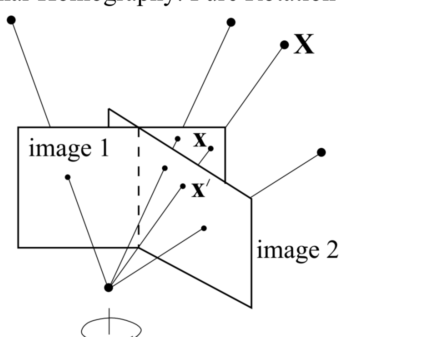
Figure borrowed from Hartley and Zisserman "Multiple View Geometry in computer vision"

Application: Panoramas

Coordinates between pairs of images are related by projective transformations



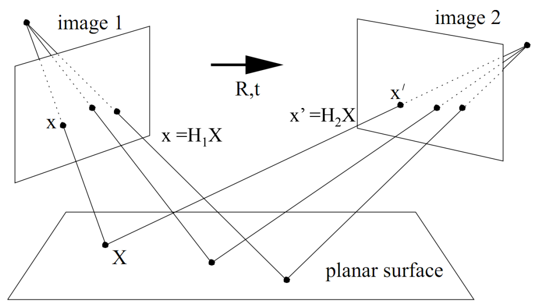
Planar Homography: Pure Rotation



$x' = H_2X = H_2(H_1^{-1} x) = (H_2H_1^{-1})x$

Figure borrowed from Hartley and Zisserman "Multiple View Geometry in computer vision"

Planar Homography



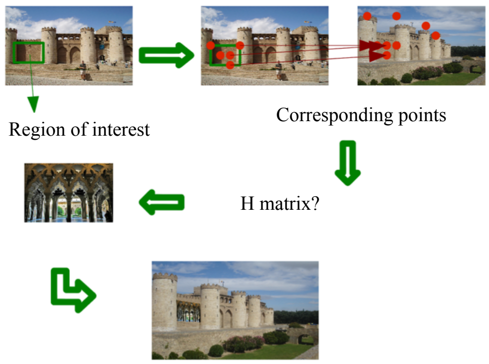
$x' = H_2X = H_2(H_1^{-1} x) = (H_2H_1^{-1})x$

Figure borrowed from Hartley and Zisserman "Multiple View Geometry in computer vision"

More applications: OCRs, scan,…

$\mathbf{x}' = \mathbf{H}\,\mathbf{x}$

Augmented reality



Vanishing Point

- In the **projective space**, parallel lines meet at a point at infinity.
- The vanishing point is the perspective projection of that point at infinity, resulting from multiplication by the camera matrix.

Some applications …

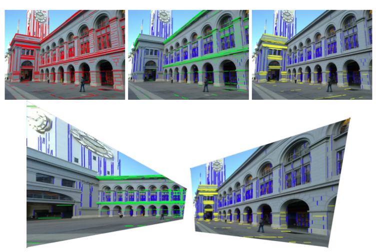
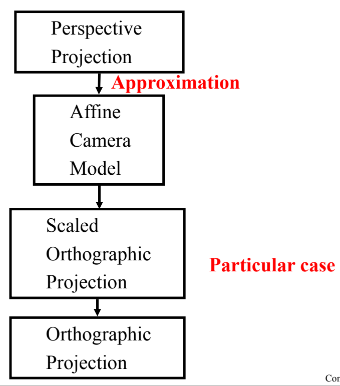


Fig. 4. Top row. Left: Query image with detected line segments. Middle and right: Lines belonging to the same vanishing point have been given the same color. Each image shows only the lines corresponding to one pair of orthogonal vanishing points. Bottom row: Two rectifications of the query image according to the two chosen pairs of vanishing points.

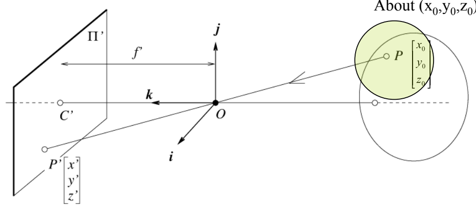
Figure from "Handling Urban Location Recognition as a 2D Homothetic Problem" G. Baatz, K. Koserl, D. Chen, R. Grzeszczuk, M. Pollefeys ECCV 2010.

Simplified Camera Models



Affine Camera Model

Appropriate in Neighborhood About (x_0, y_0, z_0)



- Take perspective projection equation, and perform Taylor series expansion about some point P= (x_0, y_0, z_0).
- Drop terms that are higher order than linear.
- Resulting expression is affine camera model

- Perspective

$$\begin{bmatrix}u\\v\end{bmatrix} = \frac{f}{z}\begin{bmatrix}x\\y\end{bmatrix}$$

- Assume that $f=1$, and perform a Taylor series expansion about (x_0, y_0, z_0)

$$\begin{bmatrix}u\\v\end{bmatrix} = \frac{1}{z_0}\begin{bmatrix}x_0\\y_0\end{bmatrix} - \frac{1}{z_0^2}\begin{bmatrix}x_0\\y_0\end{bmatrix}(z-z_0) + \frac{1}{z_0}\begin{bmatrix}1\\0\end{bmatrix}(x-x_0) + \frac{1}{z_0}\begin{bmatrix}0\\1\end{bmatrix}(y-y_0) + \frac{1}{2}\frac{2}{z_0^3}\begin{bmatrix}x_0\\y_0\end{bmatrix}(z-z_0)^2 + \cdots$$

- Dropping higher order terms and regrouping.

$$\begin{bmatrix}u\\v\end{bmatrix} \approx \frac{1}{z_0}\begin{bmatrix}x_0\\y_0\end{bmatrix} + \begin{bmatrix}1/z_0 & 0 & -x_0/z_0^2\\0 & 1/z_0 & -y_0/z_0^2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \mathbf{Ap}+\mathbf{b}$$

Affine camera model in Euclidean Coordinates

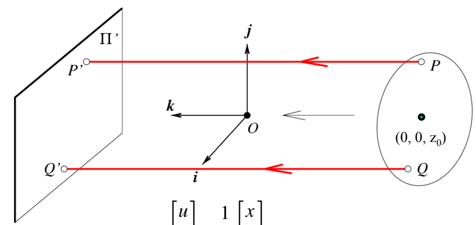
$$\begin{bmatrix} u \\ v \end{bmatrix} \approx \frac{1}{z_0}\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} 1/z_0 & 0 & -x_0/z_0^2 \\ 0 & 1/z_0 & -y_0/z_0^2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{Ap}+\mathbf{b}$$

Rewrite affine camera model
in terms of Homogenous Coordinates

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \approx \begin{bmatrix} 1/z_0 & 0 & -x_0/z_0^2 & x_0/z_0 \\ 0 & 1/z_0 & -y_0/z_0^2 & y_0/z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Scaled orthographic projection

Starting with Affine Camera Model, take Taylor series about $(x_o, y_0, z_0) = (0, 0, z_0)$ – a point on the optical axis



$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{z_0}\begin{bmatrix} x \\ y \end{bmatrix}$$

– That is the z coordinate is dropped, and the image a scaling of the x and y coordinates, where the **scale is $1/z_0$**, the depth of the point of the expansion.

The projection matrix for scaled orthographic projection

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1/z_0 & 0 & 0 & 0 \\ 0 & 1/z_0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

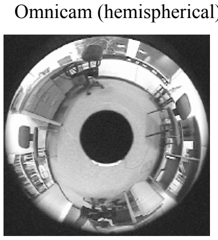
- Parallel lines project to parallel lines
- Ratios of distances are preserved under orthographic projection

For all cameras?

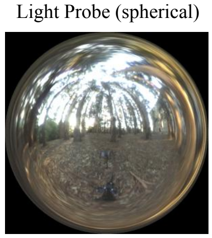
Other camera models

- Generalized camera – maps points lying on rays and maps them to points on the image plane.

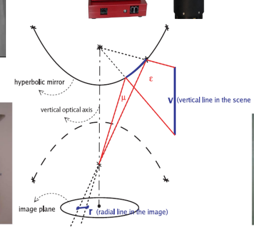
Omnicam (hemispherical)

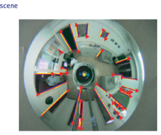


Light Probe (spherical)



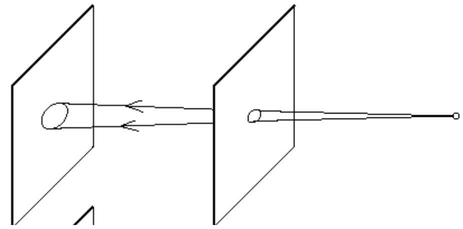
Some Alternative "Cameras"





Beyond the pinhole Camera

Beyond the pinhole Camera
Getting more light – Bigger Aperture

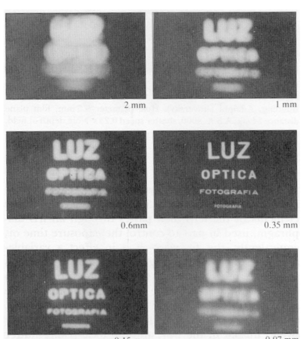


Pinhole Camera Images with Variable Aperture

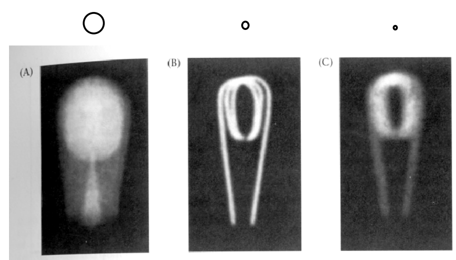
2 mm **1mm**

.6 mm **.35 mm**

.15 mm **.07 mm**

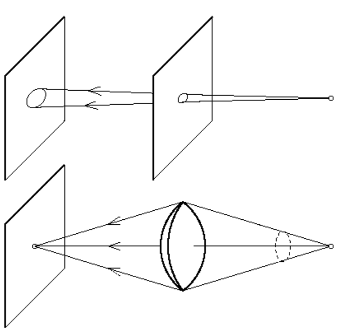


Limits for pinhole cameras



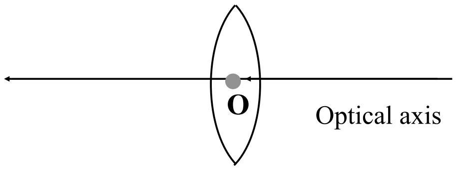
2.18 DIFFRACTION LIMITS THE QUALITY OF PINHOLE OPTICS. These three images of a bulb filament were made using pinholes with decreasing size. (A) When the pinhole is relatively large, the image rays are not properly converged, and the image is blurred. (B) Reducing the size of the pinhole improves the focus. (C) Reducing the size of the pinhole further worsens the focus, due to diffraction. From Ruechardt, 1958.

The reason for lenses
We need light, but big pinholes cause blur.



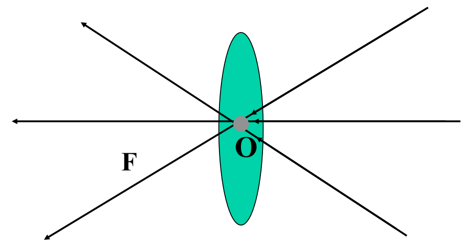
Lenses

Thin Lens



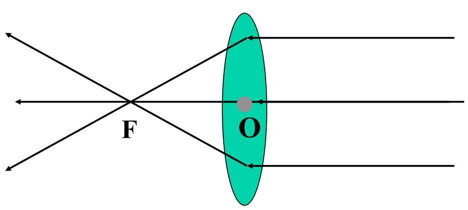
- Rotationally symmetric about optical axis.
- Spherical interfaces.

Thin Lens: Center



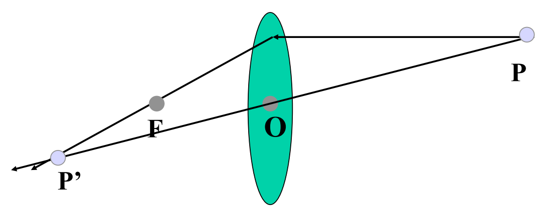
- All rays that enter lens along line pointing at **O** emerge in same direction.

Thin Lens: Focus



Parallel lines pass through the focus, F

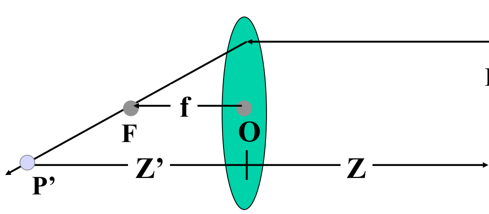
Thin Lens: Image of Point



– All rays passing through lens and starting at **P** converge upon **P'**

– So light gather capability of lens is given the area of the lens and all the rays focus on P' instead of become blurred like a pinhole

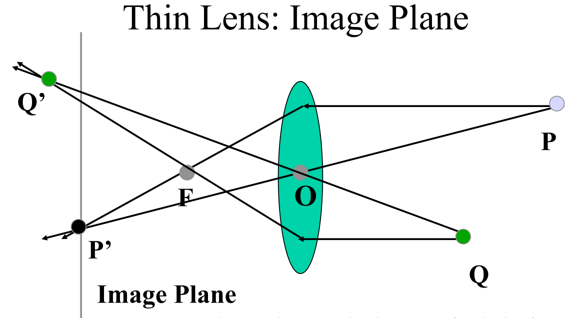
Thin Lens: Image of Point



$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

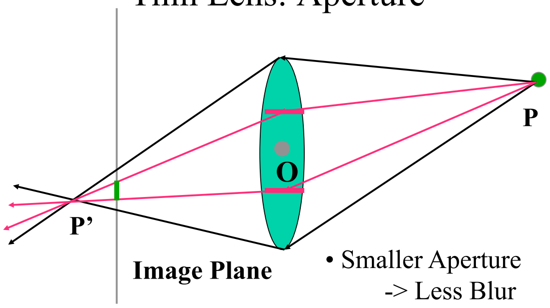
Relation between depth of Point (Z) and the depth where it focuses (Z')

Thin Lens: Image Plane



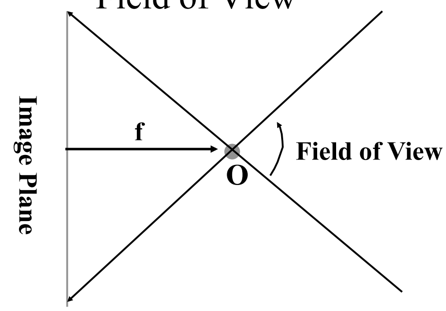
A price: Whereas the image of **P** is in focus, the image of **Q** isn't.

Thin Lens: Aperture



- Smaller Aperture -> Less Blur
- Pinhole -> No Blur

Field of View



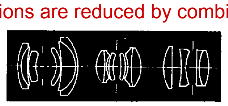
–Field of view is a function of **f** and size of image plane.

Deviations from the lens model

Deviations from this ideal are *aberrations*
Two types

1. geometrical
   - spherical aberration
   - astigmatism
   - distortion
   - coma
2. chromatic

Aberrations are reduced by combining lenses

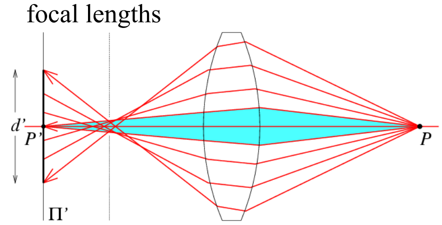


Compound lenses

Spherical aberration

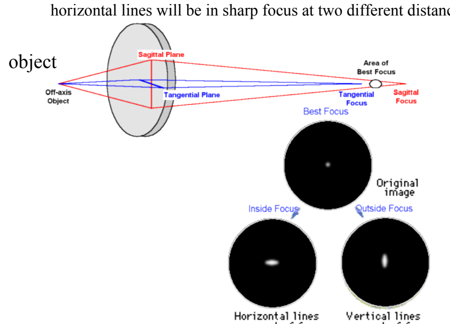
Rays parallel to the axis do not converge

Outer portions of the lens yield smaller focal lengths



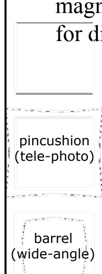
Astigmatism

An optical system with astigmatism is one where rays that propagate in two perpendicular planes have different focus. If an optical system with astigmatism is used to form an image of a cross, the vertical and horizontal lines will be in sharp focus at two different distances.



Distortion

magnification/focal length different for different angles of inclination



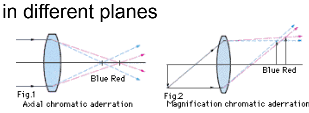
Can be corrected! (if parameters are know)

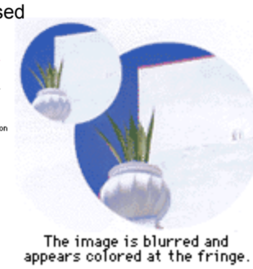
Chromatic aberration

(great for prisms, bad for lenses)

Chromatic aberration

rays of different wavelengths focused in different planes





cannot be removed completely

Vignetting

– Only part of the light reaches the sensor
– Periphery of the image is dimmer

Human eye