Image Formation:
Geometric Camera Models

Introduction to Computer Vision
CSE 152
Lecture 2

Announcements

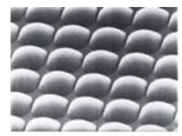
- Course website
  - http://cseweb.ucsd.edu/classes/sp16/cse152-a/
- Homework 0 will be assigned today
  - Piazza and MATLAB
  - Due Wed, Apr 6, 11:59 PM
- Wait list
- Reading:
  - Chapters 1: Geometric camera models

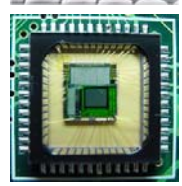
Earliest Surviving Photograph

- First photograph on record, “la table service” by Nicephore Niepce in 1822.
- Note: First photograph by Niepce was in 1816.

How Cameras Produce Images

- Basic process:
  - photons hit a detector
  - the detector becomes charged
  - the charge is read out as brightness
- Sensor types:
  - CCD (charge-coupled device)
    - high sensitivity
    - high power
    - cannot be individually addressed
    - blooming
  - CMOS
    - simple to fabricate (cheap)
    - lower sensitivity, lower power
    - can be individually addressed





Images are two-dimensional patterns of brightness values.

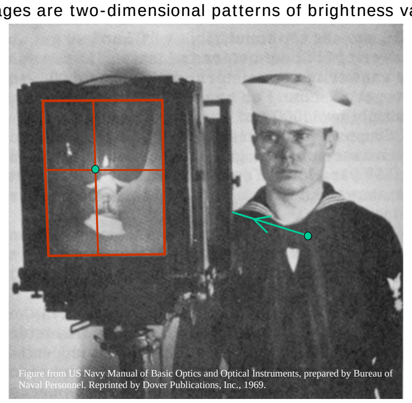


Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

They are formed by the projection of 3D objects.

Effect of Lighting: Monet

Change of Viewpoint: Monet

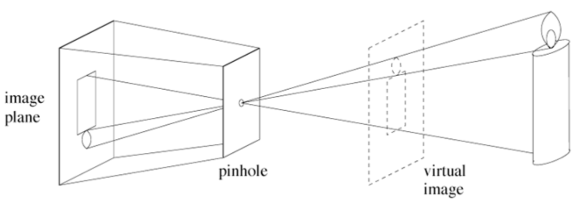
Haystack at Chailly at sunrise (1865)

Image Formation: Outline

- Geometric camera models
- Light and shading
- Color

Pinhole Camera: Perspective projection

- Abstract camera model - box with a small hole in it



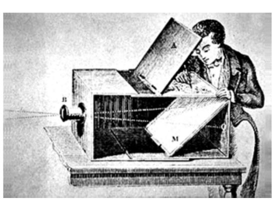
 Forsyth&Ponce

Camera Obscura

"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". --- Leonardo *Da Vinci*

http://www.acmi.net.au/AIC/CAMERA_OBSCURA.html (Russell Naughton)

Camera Obscura



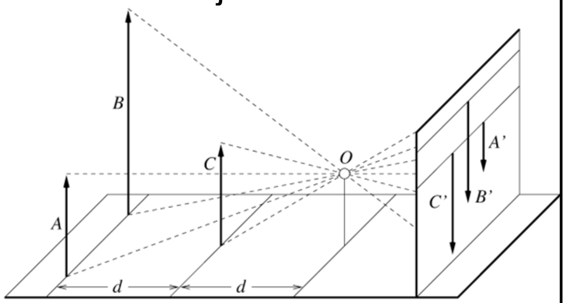
- Used to observe eclipses (e.g., Bacon, 1214-1294)
- By artists (e.g., Vermeer).

Camera Obscura

Jetty at Margate England, 1898.

http://brightbytes.com/cosite/collection2.html (Jack and Beverly Wilgus)

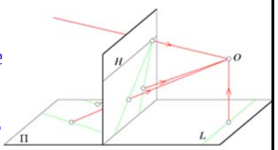
Distant objects are smaller



(Forsyth & Ponce)

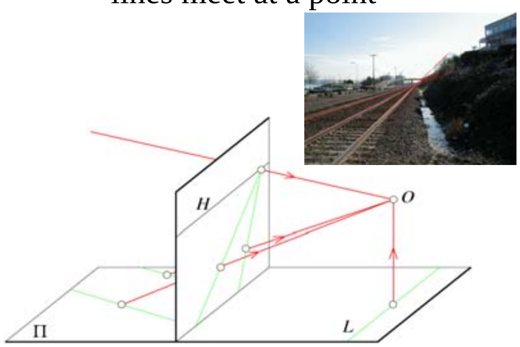
Geometric properties of projection

- 3-D points map to points
- 3-D lines map to lines
- Planes map to whole image or half-plane
- Polygons map to polygons

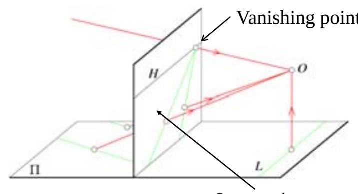


- Important point to note: Angles & distances not preserved, nor are inequalities of angles & distances.
- Degenerate cases:
  - line through focal point project to point
  - plane through focal point projects to a line

In the perspective image, two parallel lines meet at a point



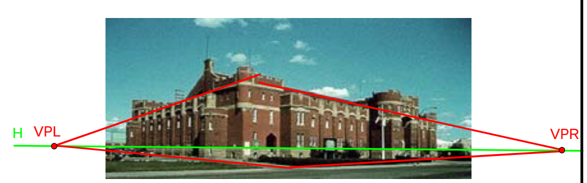
Parallel lines meet in the image



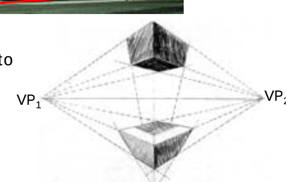
- Formed by line through O
- Parallel to the given line(s)
- A single line can have a vanishing point

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.

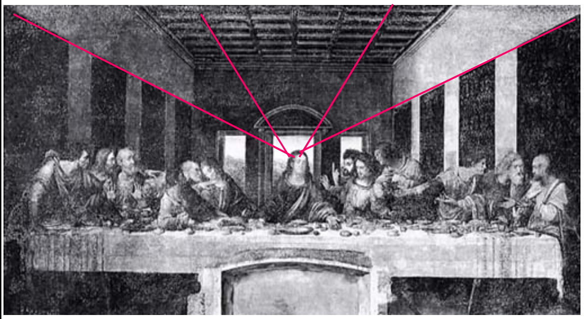
Vanishing points



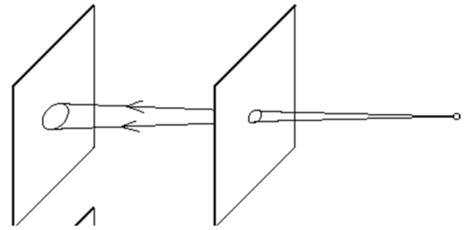
Different directions correspond to different vanishing points



Vanishing Points



Beyond the pinhole Camera
Getting more light – Bigger Aperture

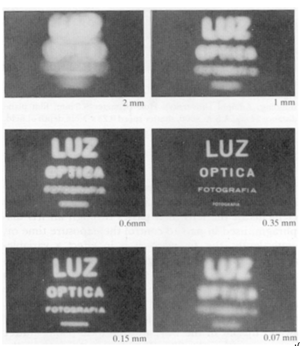


Pinhole Camera Images with Variable Aperture

2 mm **1mm**

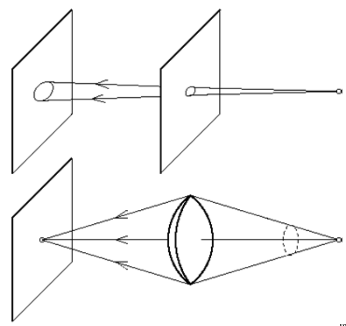
.6 mm **.35 mm**

.15 mm **.07 mm**

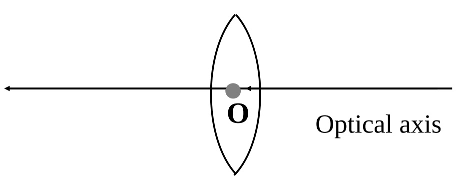


The reason for lenses

We need light, but big pinholes cause blur.

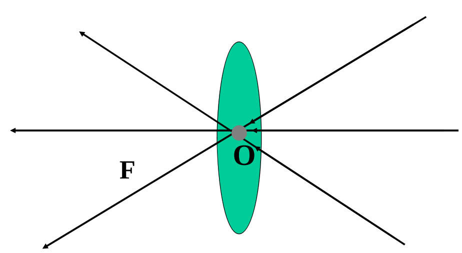


Thin Lens



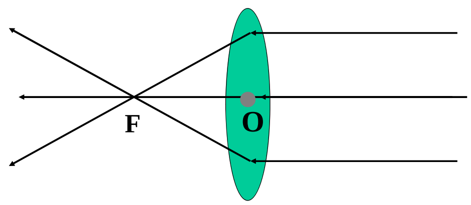
- Rotationally symmetric about optical axis.
- Spherical interfaces.

Thin Lens: Center



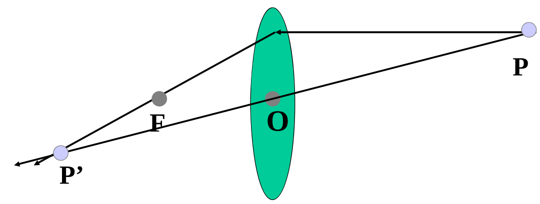
- All rays that enter lens along line pointing at **O** emerge in same direction.

Thin Lens: Focus



Parallel lines pass through the focus, F

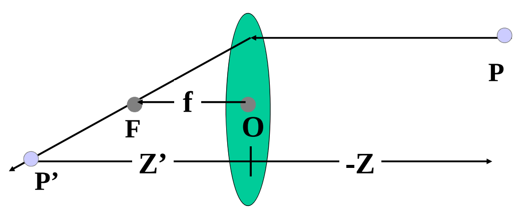
Thin Lens: Image of Point



– 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

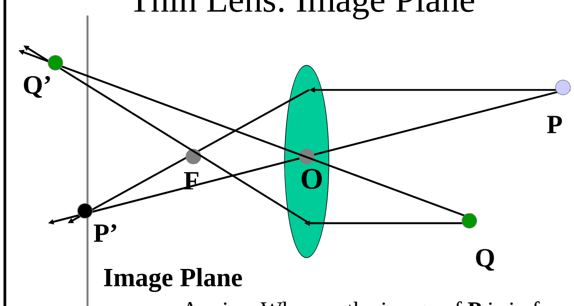
Thin Lens: Image of Point



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

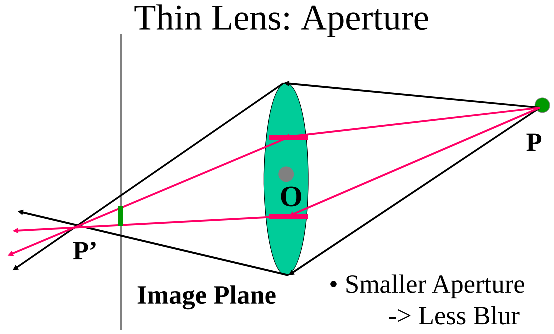
Relation between depth of Point (-Z) and the depth where it focuses (Z')

Thin Lens: Image Plane



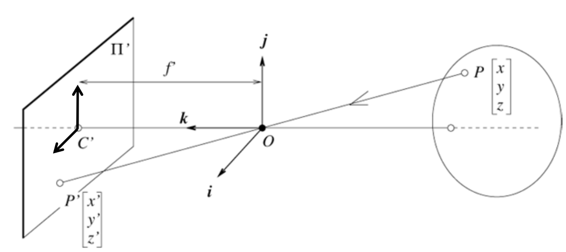
A price: Whereas the image of **P** is in focus, the image of **Q** isn't.

Thin Lens: Aperture



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

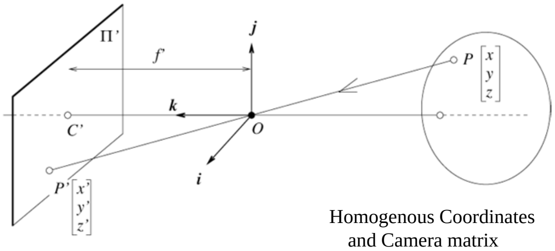
Equation of Perspective Projection



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, image coordinates of the projection of P are $(x,y,z) \rightarrow (f'\frac{x}{z}, f'\frac{y}{z})$

The equation of projection



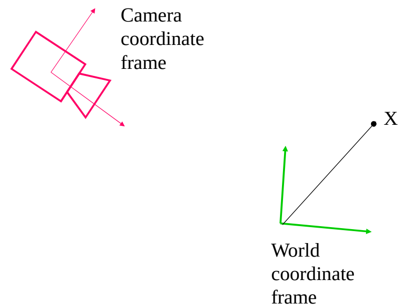
Homogenous Coordinates and Camera matrix

Cartesian coordinates:

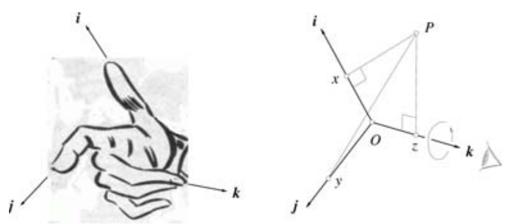
$$(X,Y,Z) \rightarrow (f\frac{X}{Z}, f\frac{Y}{Z}) = (x,y)$$

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

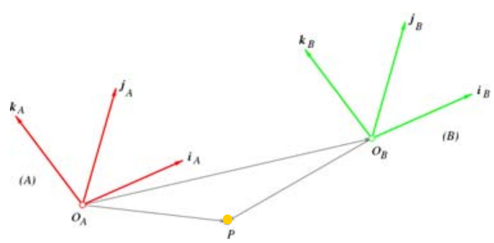
What if camera coordinate system differs from world coordinate system?



Euclidean Coordinate Systems

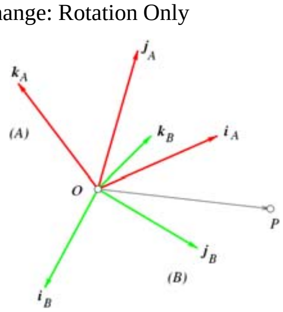


Coordinate Change: Translation Only



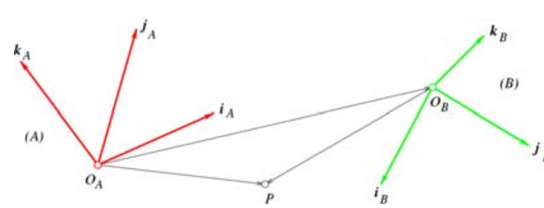
$$X' = X + t$$

Coordinate Change: Rotation Only



$$X' = R X$$

Coordinate Changes: Rotation and Translation



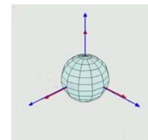
$$X' = R X + t$$

Some points about SO(n)

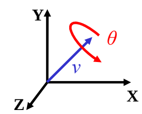
- SO(n) = { $R \in \Re^{nxn}$: $R^TR = I$, det(R) = 1}
  - SO(2): rotation matrices in plane $\Re^2$
  - SO(3): rotation matrices in 3-space $\Re^3$
- Forms a Group under matrix product operation:
  - Identity
  - Inverse
  - Associative
  - Closure
- Closed (finite intersection of closed sets)
- Bounded $R_{i,j} \in [-1, +1]$
- Does not form a vector space.
- Manifold of dimension n(n-1)/2
  - Dim(SO(2)) = 1
  - Dim(SO(3)) = 3

Parameterizations of SO(3)

–Even though a rotation matrix is 3x3 with nine numbers, it only has three degrees of freedom, it can be parameterized with three numbers. There are many parameterizations.

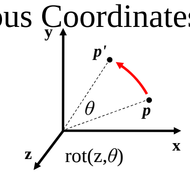


- Other common parameterizations
  - Euler Angles
  - Axis Angle
  - Quaternions
    - four parameters; homogeneous



Rotation: Homogenous Coordinates

- About z axis



$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Rotation

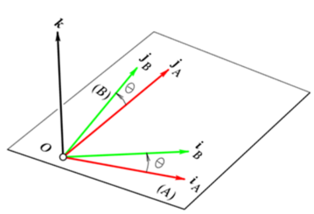
- About x axis:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

- About y axis:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

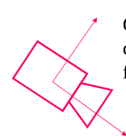
Euler Angles: Roll-Pitch-Yaw



–Composition of rotations

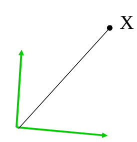
$$R = rot(\hat{k}'', \alpha) rot(\hat{j}', \beta) rot(\hat{k}, \varphi)$$

What if camera coordinate system differs from world coordinate system?



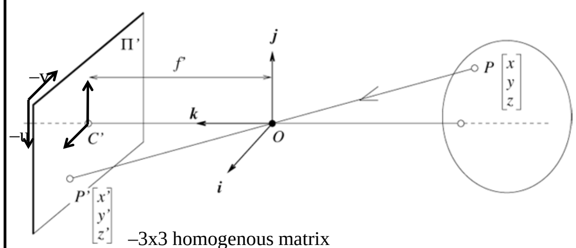
Camera coordinate frame

$$X_{Camera} = R\, X_{World} + t$$



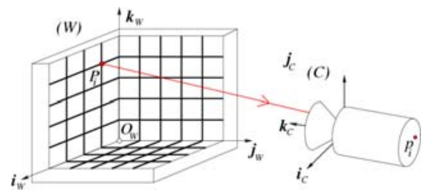
World coordinate frame

Intrinsic parameters



–3x3 homogenous matrix
–Focal length
–Principal Point
–Units (e.g. pixels)
–Pixel Aspect ratio

Camera Calibration



Given n points $P_1, \ldots, P_n$ with *known* positions and their images $p_1, \ldots, p_n$ –, estimate intrinsic and extrinsic camera parameters

- See Text book for how to do it.
- Camera Calibration Toolbox for Matlab (Bouguet)
  – http://www.vision.caltech.edu/bouguetj/calib_doc/

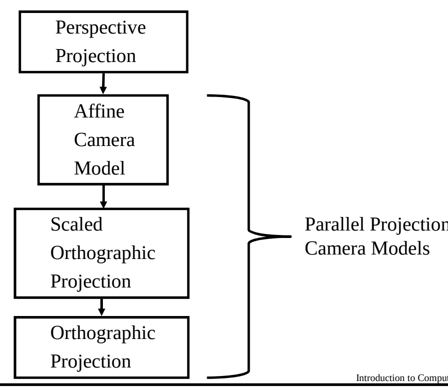
Camera parameters

- Extrinsic Parameters: Since camera may not be at the origin, there is a rigid transformation between the world coordinates and the camera coordinates
- Intrinsic parameters: Since scene units (e.g., cm) differ image units (e.g., pixels) and coordinate system may not be centered in image, we capture that with a 3x3 transformation comprised of focal length, principal point, pixel aspect ratio, and skew

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} \text{Transformation} \\ \text{represented by} \\ \text{intrinsic parameters} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \text{Rigid Transformation} \\ \text{represented by} \\ \text{extrinsic parameters} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

3 x 3 4 x 4

Camera Models

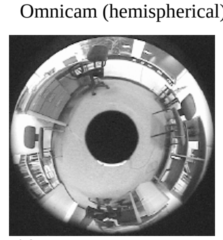


For all cameras?

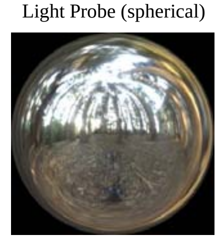
Other camera models

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

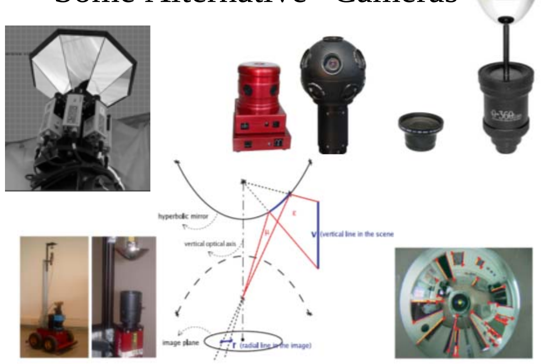
Omnicam (hemispherical)



Light Probe (spherical)



Some Alternative "Cameras"



Next Lecture

- Image Formation: Light and Shading
- Reading:
  - – Chapter 2: Light and Shading