Perspective Projection

CS-477 Computer Vision

Dr. Mohsin Kamal
Associate Professor
dr.mohsinkamal@seecs.edu.pk

School of Electrical Engineering and Computer Science (SEECS)
National University of Sciences and Technology (NUST), Pakistan

- Image points are 2-dimensional
$$\mathbf{x} = (x, y)^T \in \mathbb{R}^2$$
- Homogeneous Coordinates
  - Vectors that differ only by scale are equivalent
$$\tilde{x} = (\tilde{x}, \tilde{y}, 1)^T \in \mathbb{P}^2$$
  - $\mathbb{P}^2$ is the 2D projective space
  - $(10, 20, 1)^T \equiv (30, 60, 3)^T \equiv (5, 10, \frac{1}{2})^T$
  - Every point has infinite representations
- A homogenous vector $\tilde{x}$ can be converted to non-homogeneous coordinates by dividing by $\tilde{w}$
$$\tilde{x} = (\tilde{x}, \tilde{y}, \tilde{w})^T = \tilde{w}(x, y, 1) = \tilde{w}\bar{x}, \quad x = \frac{\tilde{x}}{\tilde{w}}, y = \frac{\tilde{y}}{\tilde{w}}$$

Ideal Points

- Ideal points: when $\tilde{w} = 0$
- They define a direction from origin
- They do not have a non-homogeneous equivalent
- The ideal points in homogeneous coordinates are used to represent directions or points at infinity in projective geometry.

Projective space $\mathbb{P}^2$

- Every coordinate is defined by a 3D-vector
- The first two elements of the vector define its direction only (outward from origin)
- Each point in non-homogeneous coordinates has a whole equivalent class of points in homogeneous coordinates
- Point $(0,0,0)^T$ does not define a direction, hence is excluded from $\mathbb{P}^2$

$$\mathbb{P}^2 = \mathbb{R}^3 - (0,0,0)^T$$

How projective space is different from Euclidean space?
Projective space points at infinity whereas Eucalidean space is a flat space

Why Projective Geometry?

There are four reason:

- Camera is a projective engine
- Points at infinity are handled
- Algebra is simpler than usual
- It is the most general framework to work in

- Projective space contains Affine space
- Affine or Euclidean upgrades can be made if required

You can consider affine space as the "finite" part of projective space, while projective space includes both finite and infinite points

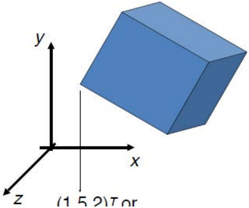
3D Points

- World points are 3-dimensional
- In homogeneous coordinates

$$\mathbf{x} = (x, y, z)^T \in \mathbb{R}^3$$

- Homogeneous coordinates for 3D representation
  - Vectors that differ only by scale are equivalent

$$\tilde{x} = (\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w})^T \in \mathbb{P}^3$$
$$\mathbb{P}^3 = \mathbb{R}^4 - (0, 0, 0, 0)^T$$



Lines in 2D

- Equation of line in 2D

$$ax + by + c = 0$$

- Thus, a line can be represented by vector $(a, b, c)^T$
- $(a, b, c)^T$ and $k(a, b, c)^T$ mean the same line for $k \neq 0$
- Thus lines can be represented by equivalence classes of vectors in $\mathbb{R}^3 - (0, 0, 0)^T$ i.e., Projective space $\mathbb{P}^2$

Point on a line

- 2D point $\mathbf{x} = (x, y)^T$
- Point will lie on line iff $ax + by + c = 0$
- This can be written as inner product[1]

$$(x, y, 1)(a, b, c)^T = 0$$
$$(x, y, 1)l = 0$$

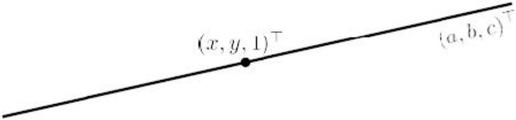
- Any non-zero k can be multiplied to the point, without loss of generality
- Hence points can also be represented as homogeneous vectors

[1] Inner product measures the projection of one vector onto another

Point on a line

- Point **x** lies on line *l* iff

$$\mathbf{x}^\top \mathbf{l} = 0$$

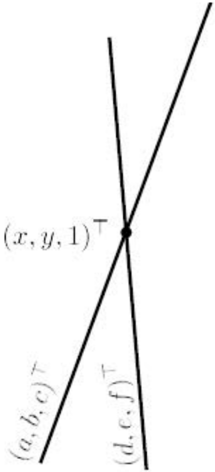


- Even though **x** and *l* are 3D-vectors, they have 2 degrees of freedom[2] each

[2] a line is specified by two parameters (the two independent ratios $\{a : b : c\}$) and so has two degrees of freedom

Intersection of Two Lines-Projective Space

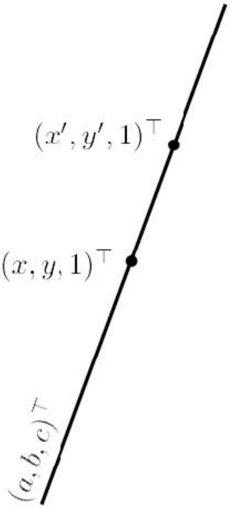
- Two lines will intersect at a point
- Let l and l' intersect at point, $\mathbf{x}$
- Then $l \times l' = \mathbf{x}$
- Proof:
  - The point $\mathbf{x}$ lies on both l and l'. Therefore
    $$\mathbf{x}^T l = 0$$
    $$\mathbf{x}^T l' = 0$$
  - This is non-trivially possible only when $\mathbf{x}$ is orthogonal to both l and l'.



The cross product of two vectors equal to zero implies that the two vectors are either parallel or antiparallel to each other.

Line joining two points

- Two points lie on a line
- Let **x** and **x'** lie on line *l*
- Then $\mathbf{x} \times \mathbf{x}' = l$
- Proof:
  - The line *l* passes through both **x** and **x'** Therefore
  $$l^T\mathbf{x} = 0$$
  $$l^T\mathbf{x}' = 0$$
  - This is non-trivially possible only when *l* is orthogonal to both **x** and **x'**.



Intersection of Parallel Lines

- Consider two parallel lines[3]

$$l : ax + by + c = 0$$
$$l' : ax + by + c' = 0$$

- Computing intersection (as before)

$$l \times l' = (c' - c)(b, -a, 0)^T$$

- Thus, point of intersection

$$(b, -a, 0)^T$$

- Converting to non-homogeneous coordinates:

$$\left(\frac{b}{0}, -\frac{a}{0}\right)^T$$

- Hence Parallel lines intersect at ideal points

[3]For parallel lines, the slope of both lines will be the same, but the intercept values may differ.

Ideal Points lie on a line

- Recall that all parallel lines intersect at an ideal point or point at infinity, of the form $(x, y, 0)^T$
- Consider two such ideal points:

$$\mathbf{x} = (x, y, 0)^T$$
$$\mathbf{x'} = (x', y', 0)^T$$

- The line joining them is given by:

$$l = \mathbf{x} \times \mathbf{x'}$$

or

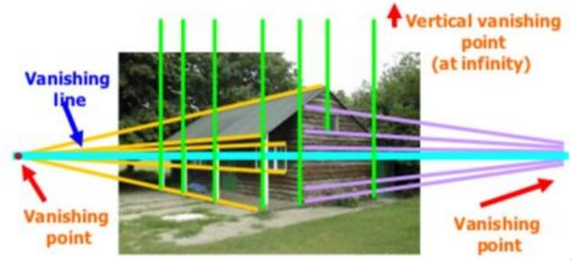
$$l = (0, 0, xy' - yx')^T = (0, 0, 1)^T$$

- Thus, all points at infinity lie on a single line, the **line at infinity**

$$l_\infty = (0, 0, 1)^T$$

Line at infinity

- Any line $l : (a, b, c)^T$ intersects l_∞ at $(b, -a, 0)^T$
- Any line parallel to l, i.e., $l' : (a, b, c')^T$ will intersects l_∞ also at $(b, -a, 0)^T$
- In non-homogeneous coordinates, $(b, -a)^T$ represents line direction
- Hence, as line direction varies, its intersection with l_∞ varies.
- Line at infinity is the set of directions for lines in a plane



Duality

$$\mathbf{x} \longleftrightarrow \mathbf{l}$$

$$\mathbf{x}^\top \mathbf{l} = 0 \longleftrightarrow \mathbf{l}^\top \mathbf{x} = 0$$

$$\mathbf{x} \times \mathbf{x}' = \mathbf{l} \longleftrightarrow \mathbf{l} \times \mathbf{l}' = \mathbf{x}$$

Duality theorem: To any theorem of 2-dimensional projective geometry, there corresponds a dual theorem, which may be derived by interchanging the role of points and lines in the original theorem

Example

Consider the two lines are parallel, and consequently intersect "at infinity". In homogeneous notation the lines are $l = (-1, 0, 1)^T, l = (-1, 0, 2)^T$ and find their intersection point[4].

[4] **Result:** The point at infinity in the direction of the y-axis.

Example

Example

Example

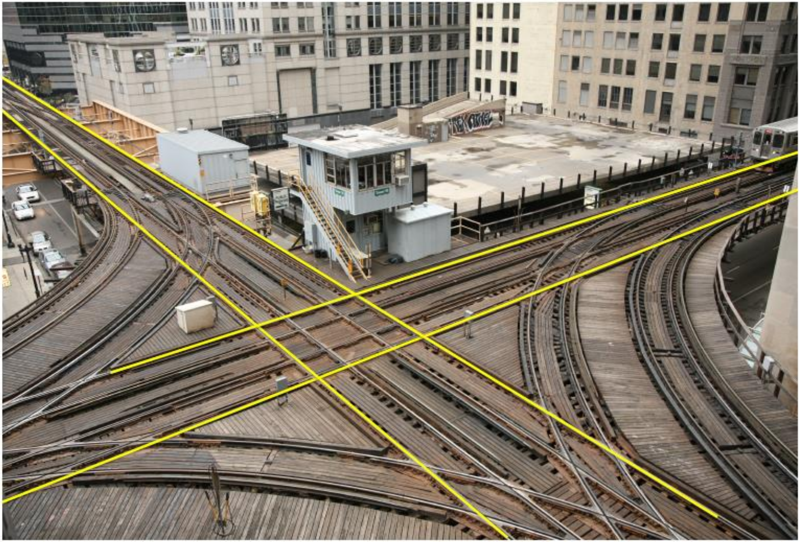
Example

Example

Example

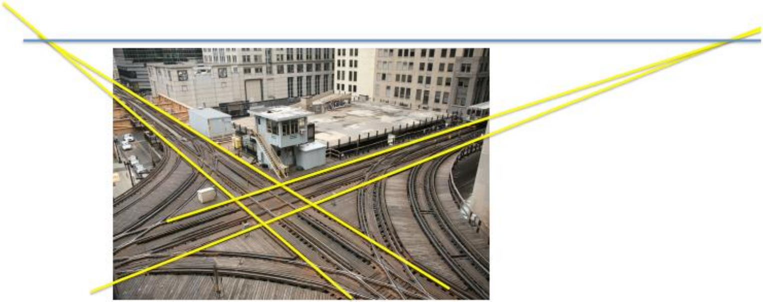
Example

Example



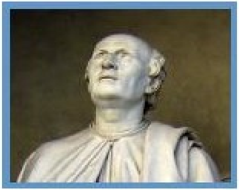
Example

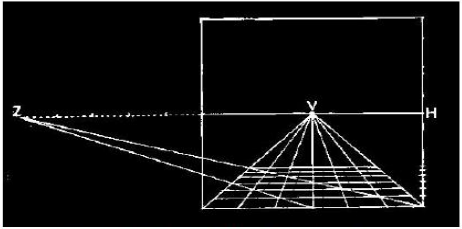
Reading: Section 2.2.1: Multi View Geometry by Andrew Zissserman



Fillippo Brunelleschi

- Architect, 1377-1446 AD
- Founding father of Renaissance
- Discovered the vanishing point/line
- Formulated the notion that linear perspective governed the pictorial representation of space

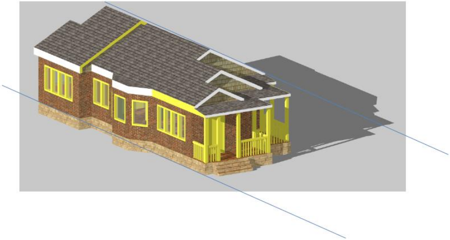




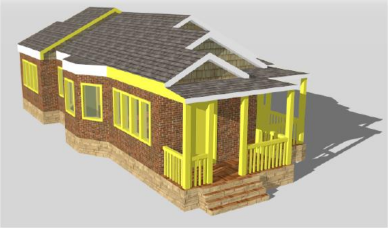
ide adapted from: A Brief History of Computer Vision,
esentation by Prof Yaser Sheikh, CMU

Types of projection

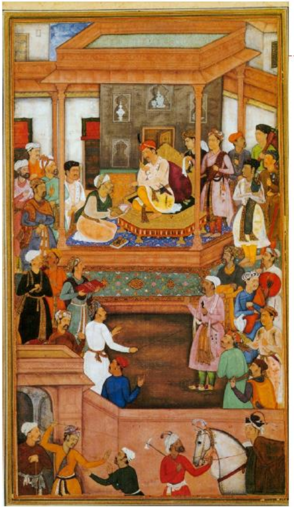
Orthographic



Perspective



Perspective



Without perspective

With perspective

Changing the Focal Length of the Lens

- Wide-angle portraits can look wonky
- Nose is nearest to camera and looks bigger
- Ears and hair are further away and look smaller

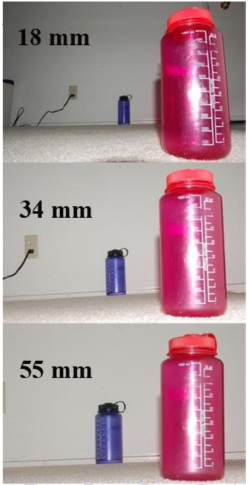
Wide-angle

Standard

Telephoto

Perspective Distortion

- **Perspective distortion:** The effect that further away objects appear smaller in size
- As focal length increases (more zoom), perspective distortion becomes less
- Orthographic camera can be considered as being very far away (so no variation in Z) and having very long focal length ($\frac{f}{Z} = 1$), and hence zero perspective distortion



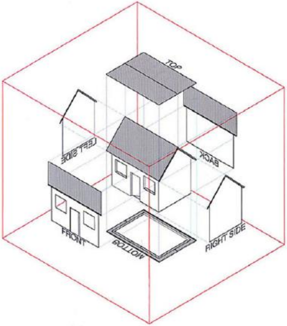
Changing the Focal Length of the Lens

- If focal length is small then perspective projection (wider field of view)
- If focal length reaches infinity then all rays become parallel resulting in orthographic projections

Top: Wide-angle ($f = 24mm$)
Bottom: Telephoto ($f = 392mm$)
(sensor: 35mm film)

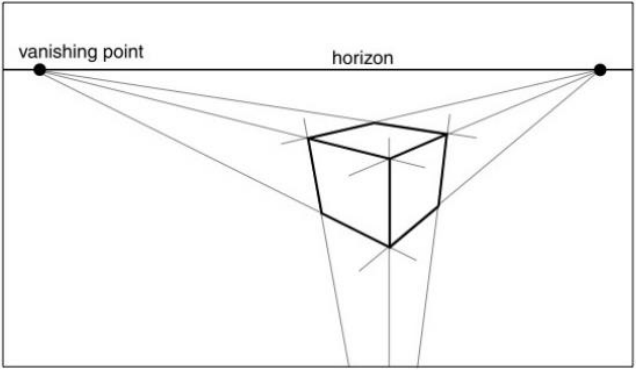
Orthographic camera

Parallel Lines remain parallel and do not converge (also termed parallel projection)



Three-point Perspective

In perspective view parallel lines will meet at a point called **vanishing point.** Every set of parallel lines has its own vanishing point. The lines parallel to ground will meet at **horizon**.

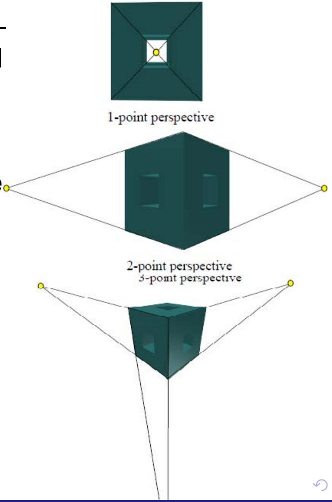


Conditions for 1,2 and 3 point perspective projections

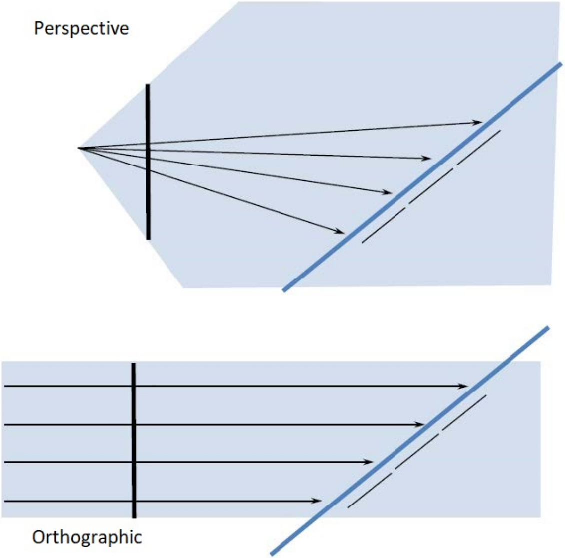
1 Point: Image plane or film should be parallel to one of the faces (2-axis are parallel to object axis)

2 Point: Vertical axis of the film is same as the vertical axis of the object

3 Point: None of the axis is parallel



Perspective vs orthographic camera



Camera obscura: dark room

Known during classical period in China and Greece (e.g., Mozi, China, 470BC to 390BC)

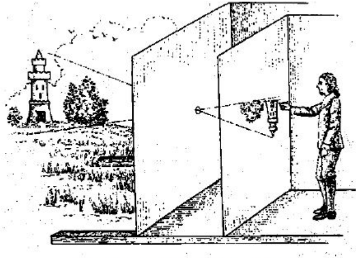
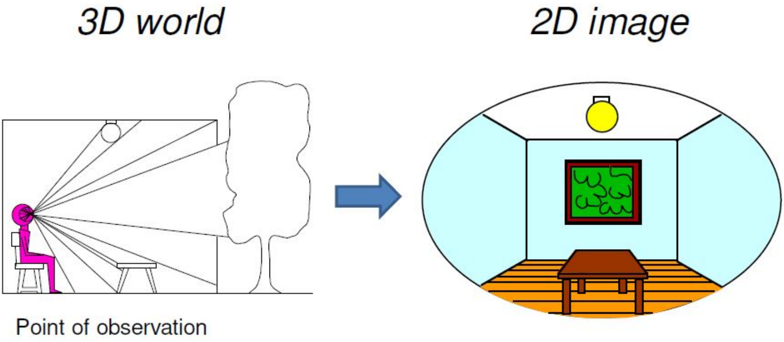


Illustration of Camera Obscura

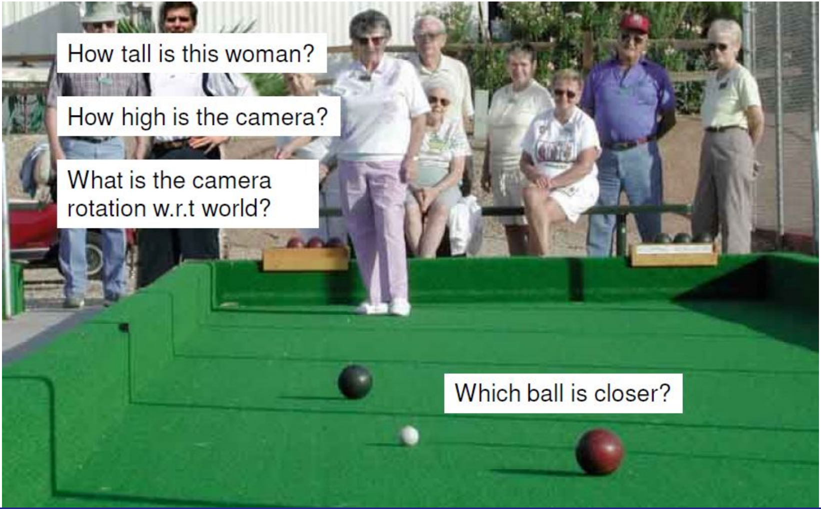
Freestanding camera obscura at UNC Chapel Hill

Photo by Seth Ilys

Camera: A Dimensionality Reduction Machine (3D to 2D)

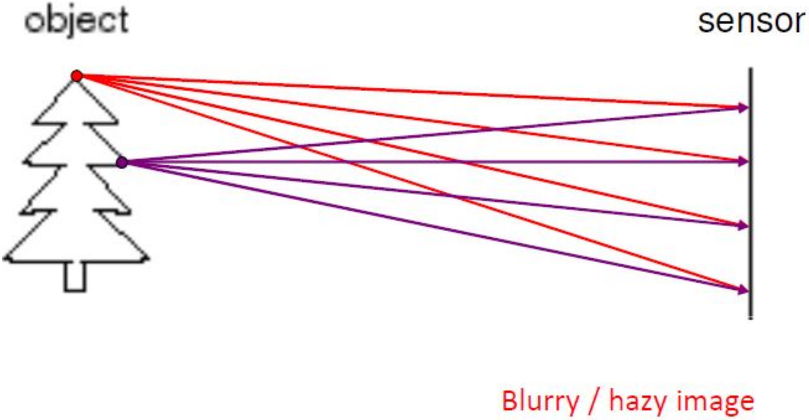


Cameras and World Geometry



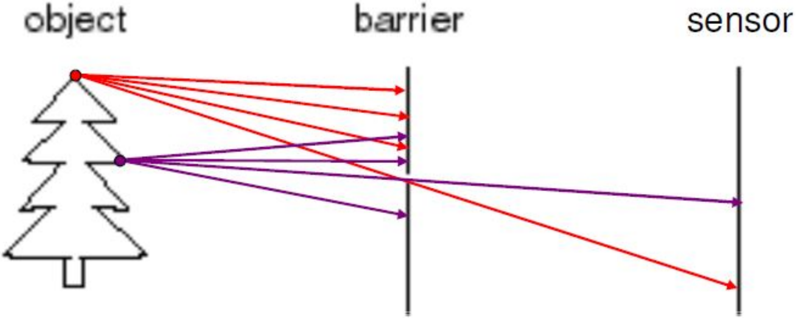
Let's design a camera

Let's design a camera
Do we get a reasonable image?

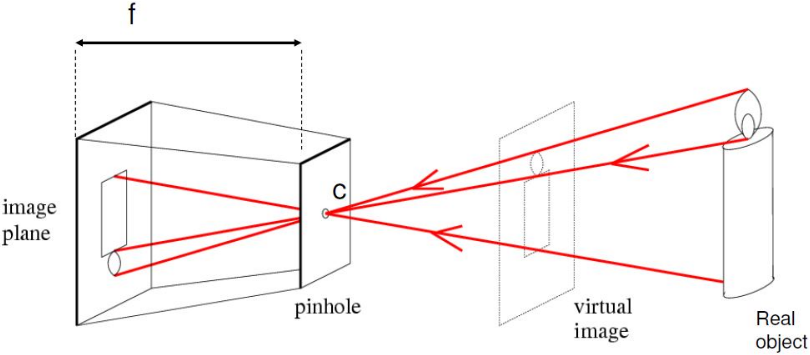


Let's design a camera

- Idea 2: Add a barrier to block most rays
  - Pinhole in barrier
  - Only sense light from one direction.
  - Reduces blurring.
  - In most cameras, this aperture can vary in size.



Pinhole camera model



f = Focal length

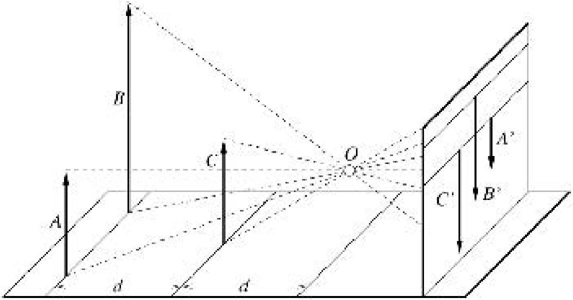
c = Optical center of the camera

Projective geometry

Length (and so area) is lost.

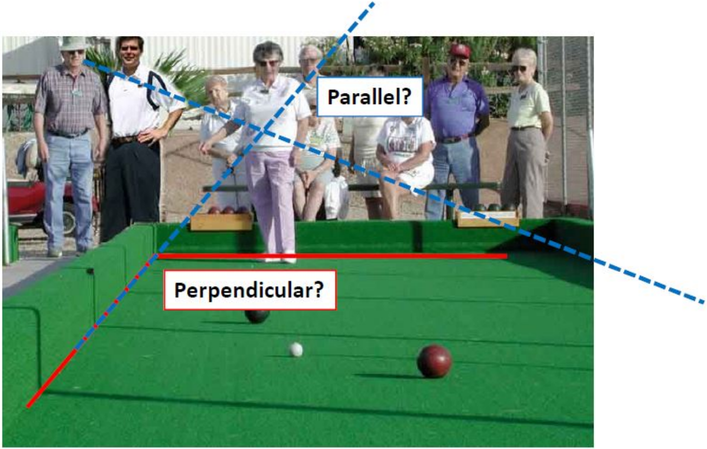
Length and area are not preserved

- Nonlinear transformation: an object with greater height but farther from the camera compared to object nearer the camera will be scaled down more



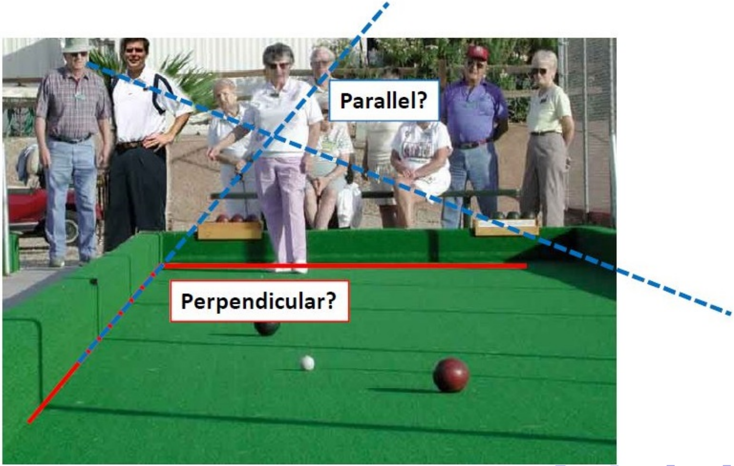
Projective geometry

Angles are lost.



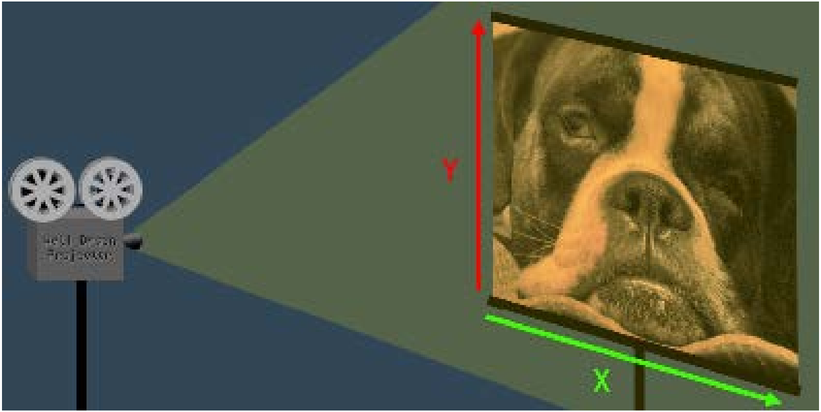
Projective geometry

What is preserved?
- Straight lines are still straight.



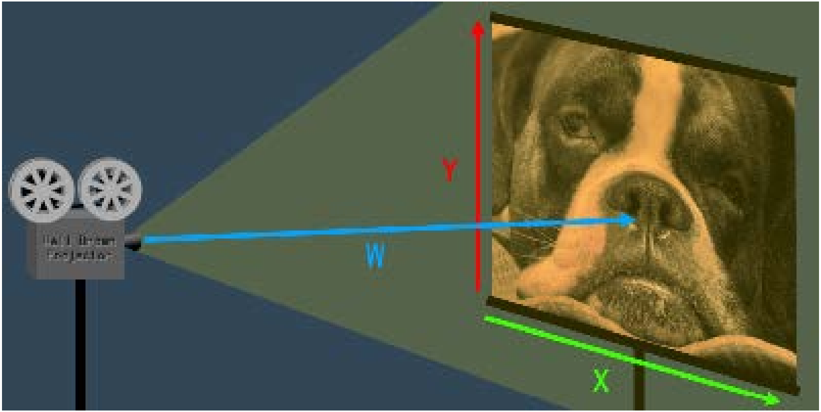
Projective geometry

- 2D point in Euclidean plane = (x,y) coordinates
- 2D point in projective plane = (x,y,w) coordinates



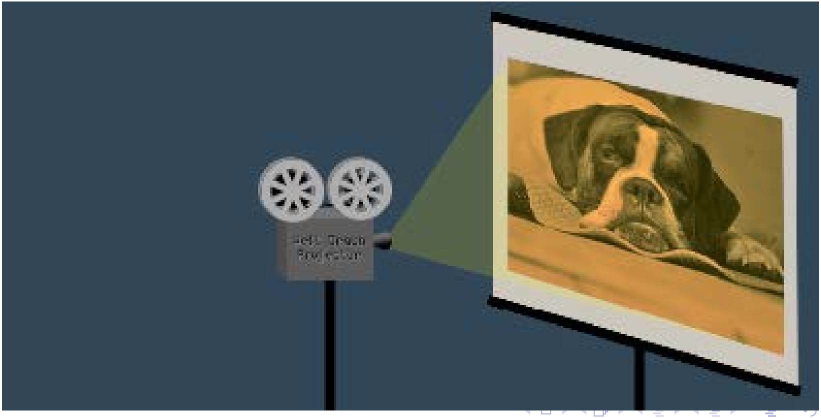
Projective geometry

- 2D point in Euclidean plane = (x,y) coordinates
- 2D point in projective plane = (x,y,w) coordinates (one additional dimension / parameter)



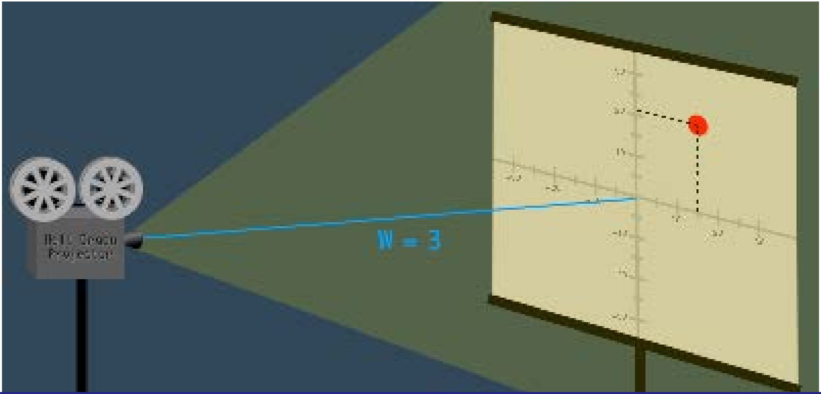
Varying w

As w becomes smaller, projected image becomes smaller (i.e., w affects scale of the image).



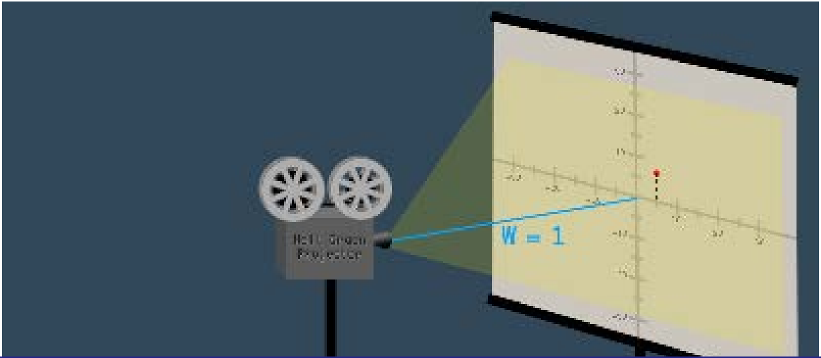
Varying w

Assume the projector is 3 meters away from the screen, and there is a dot on the 2D image at the coordinate (15,21). This gives us the projective coordinate vector (X,Y,W)=(15,21,3)



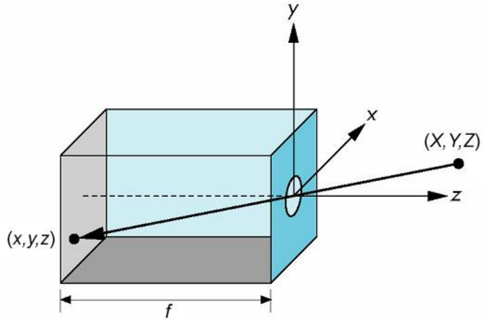
Varying w

Now, the projector is pushed closer to the screen so that the distance is 1 meter. The location of the point is then: (15/3,21/3,3/3)=(5,7,1), i.e, the dot is now at coordinate (5,7) in the Euclidean plane.



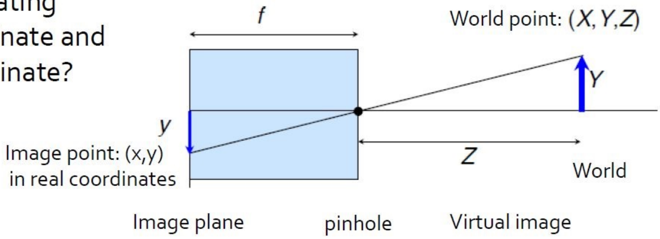
Projective geometry

- Orient along z-axis
- World point (X, Y, Z) [in world coordinates]
- Image point at (x, y, z) [in real world coordinates]



Perspective transform

Equation relating world coordinate and image coordinate?



$$\frac{-y}{Y} = \frac{f}{Z}$$

$$y = -\frac{fY}{Z} \qquad x = -\frac{fX}{Z}$$

It is customary to use a negative sign to indicate that the image is always formed upside down

Representation in homogeneous coordinates

We can write this as a matrix using the homogeneous coordinates

$$\begin{bmatrix} hx \\ hy \\ h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$hx = X, hy = Y, h = -\frac{Z}{f}$$

$$x = -\frac{fX}{Z}, y = -\frac{fY}{Z}$$

Representation in homogeneous coordinates

Any scaling of a homogeneous transform is equivalent

$$\begin{bmatrix} hx \\ hy \\ h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

is equivalent to

$$\begin{bmatrix} hx \\ hy \\ h \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$x = PX$$

Central projection

- The camera can be more compactly written as

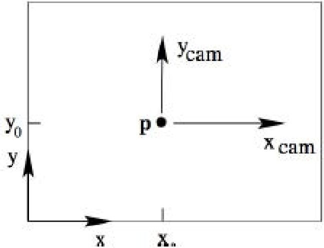
$$x = PX$$

- where P is a 3x4 matrix that maps from $\mathbb{P}^3 \rightarrow \mathbb{P}^2$
- P may also be written as:

$$P = diag(f, f, 1)[I|0]$$

Principal point offset

- The expression assumes that $P = diag(f, f, 1)[I|0]$ image origin is at the principal point.
- This may not be the case in general. For example:



- If the image coordinates of the principal point are $(p_x, p_y)^T$, then the camera mapping will be

$$(X, Y, Z)^T \rightarrow \left(\frac{fX}{Z} + p_x, \frac{fY}{Z} + p_y\right)^T$$

More General Perspective Camera Model

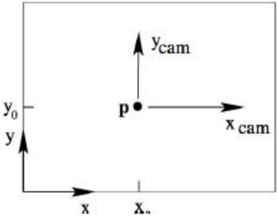
- General Perspective Transform

$$\begin{bmatrix} hx \\ hy \\ h \end{bmatrix} = \begin{bmatrix} m_x f & 0 & p_x & 0 \\ 0 & m_y f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- m_x and m_y are scaling, to convert to pixels
  - $m_x =$ number of pixels in x direction i.e., pixels per unit length in x direction
  - $m_y =$ number of pixels in y direction, i.e., pixels per unit length in y direction
- p_x and p_y are principal point offset

Principal point offset

We want the principal point to map to (p_x, p_y) instead of $(0, 0)$



$$(X, Y, Z)^T \rightarrow \left(\frac{fX}{Z} + p_x, \frac{fY}{Z} + p_y\right)^T$$

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Note: f will be positive in canonical representation of the camera model where the projection of the world does not invert

Example

Perspective transform: This relates the camera frame to the real image frame

Example: I take the image of a person (2m tall) standing 4m away from the camera, with a 35mm camera using the geometry shown previously. How high will be the image?
Answer: $y = -(35)(2000)/4000 = -17.5mm$
i.e., the image will be formed inverted of the length 17.5mm

How to convert to pixel frame (i.e. what will be the coordinates of the head of the person in the image)?

Example

- Suppose I know that the size of the film is $8cm \times 6cm$, and that the resolution of the camera is 640×480 pixels
- Implies, the center of the image is at $4cm \times 3cm$ from the corner, and is at location (320,240) which represents the principal point offset
- 17.5mm out of 60mm is 140 out of 480 pixels
- Hence the coordinates of the head will be

$$(same\ in\ x, 240 - 140\ in\ y) = (320, 100)$$