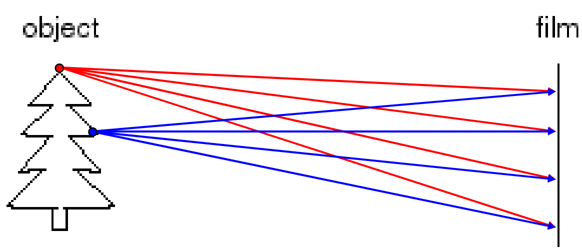
Image formation
Camera model

Oct 1. 2009

Jaechul Kim, UT-Austin

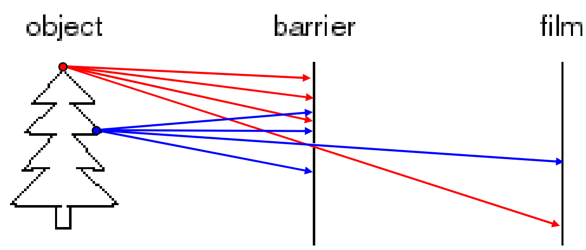
Image formation



- Let's design a camera
  - Idea 1: put a piece of film in front of an object
  - Do we get a reasonable image?

Slide by Steve Seitz

Pinhole camera

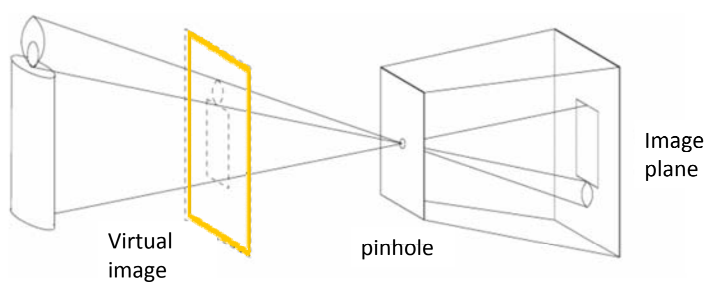


- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**
  - How does this transform the image?

Slide by Steve Seitz

Pinhole camera

- Pinhole camera is a simple model to approximate imaging process, perspective **projection**.



If we treat pinhole as a point, only one ray from any given point can enter the camera.

Fig from Forsyth and Ponce

Pinhole size / aperture

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

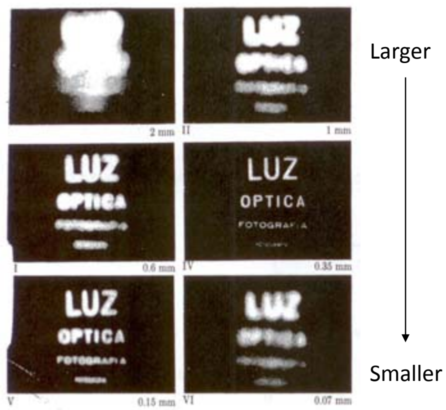
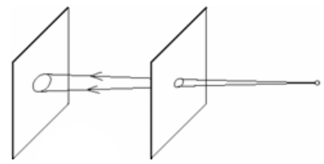
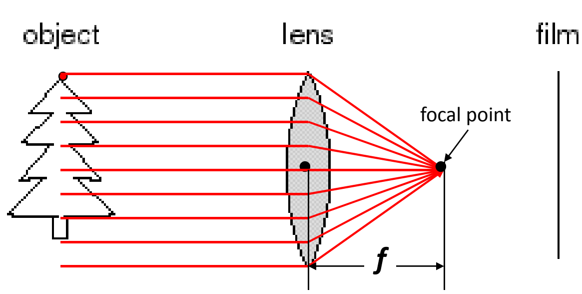


Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]



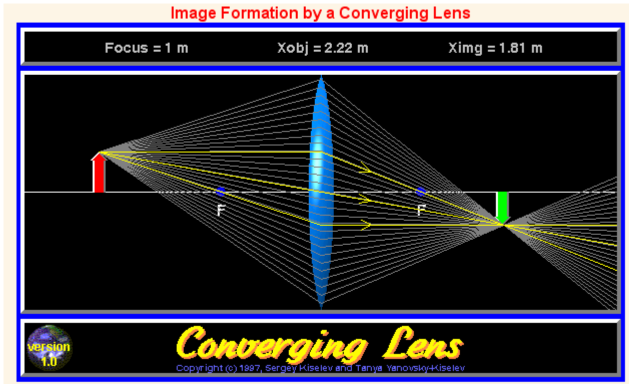
Adding a lens



- A lens focuses light onto the film
  - All parallel rays converge to one point on a plane located at the *focal length f*

Slide by Steve Seitz

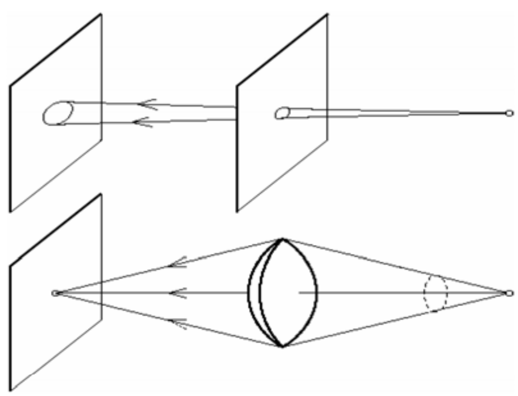
Adding a lens



- A lens focuses light onto the film
  - All rays radiating from an object point converge to one point on a film plane.

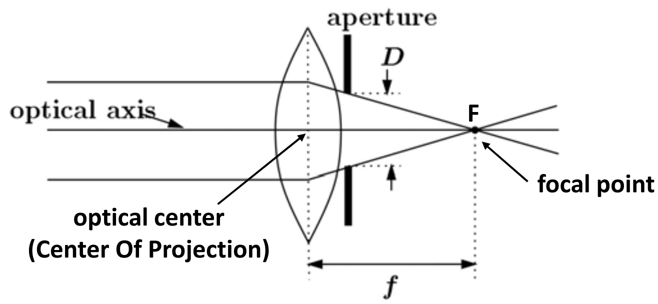
Image source:
http://www.physics.uoguelph.ca/applets/Intro_physics/kisalev/java/clens/index.html

Pinhole vs. lens

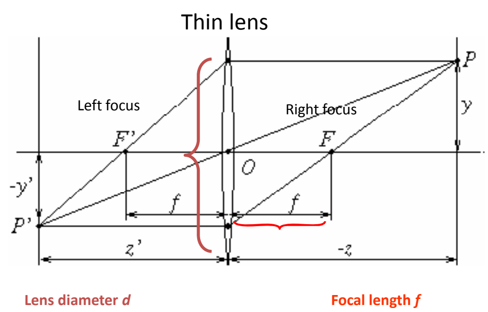


- A lens focuses rays radiating from an object point onto a single point on a film plane
- Gather more light, while keeping focus; make pinhole perspective projection practical

Cameras with lenses



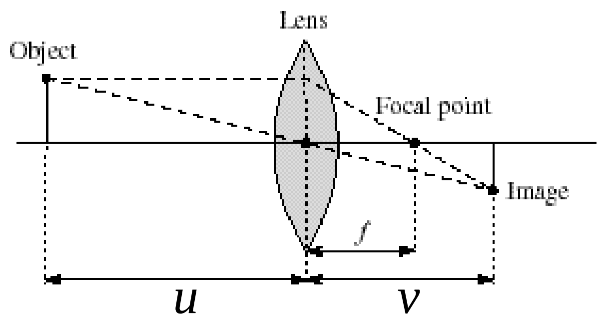
Thin lens



Rays entering parallel on one side go through focus on other, and vice versa.

In ideal case – all rays from P imaged at P'.

Thin lens equation



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

- Any object point satisfying this equation is in focus

Zoom lens

- A assembly of several lens
- By changing the lens formation, it varies its effective focal length.

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

For fixed v,

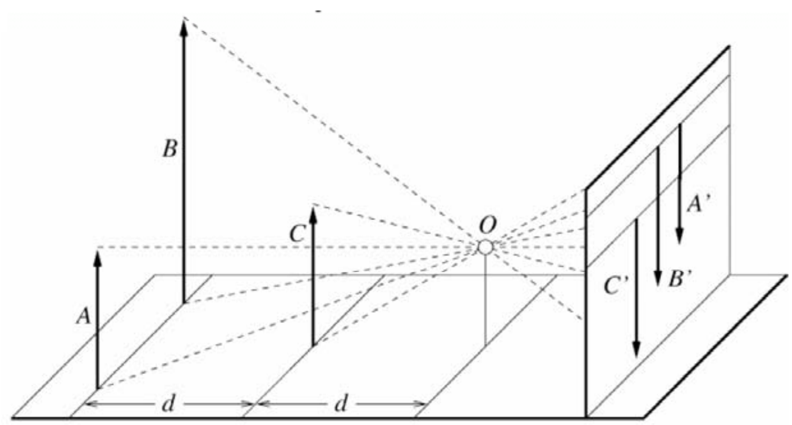
Large f $\longrightarrow$ Large u $\longrightarrow$ Far-away object is in focus. (Zoom out)

Small f $\longrightarrow$ Small u $\longrightarrow$ Near object is in focus. (Zoom in)

Perspective effects

Perspective effects

- Far away objects appear smaller



Forsyth and Ponce

Perspective effects

Perspective effects

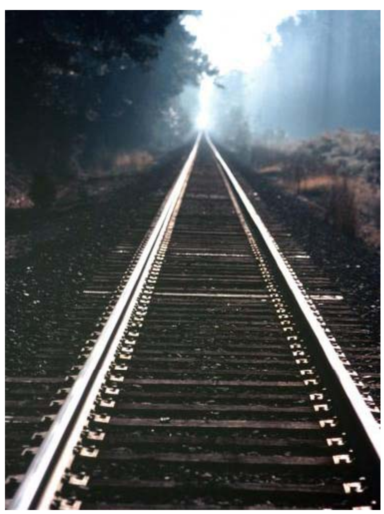


Image source: http://share.triangle.com/sites/share-uda.triangle.com/files/images/RailRoadTrackVanishingPoint_0.preview.jpg

Perspective effects

- Parallel planes in the scene intersect in a line in the image
- Parallel lines in the scene intersect in the image

Parallelism is “not” preserved under the perspective projection through camera.

Perspective effects

Perspective effects by camera projection can be thought as projective transformation between an object and its image.

A quadrangle in the image　Projective transformation　A rectangle in the scene

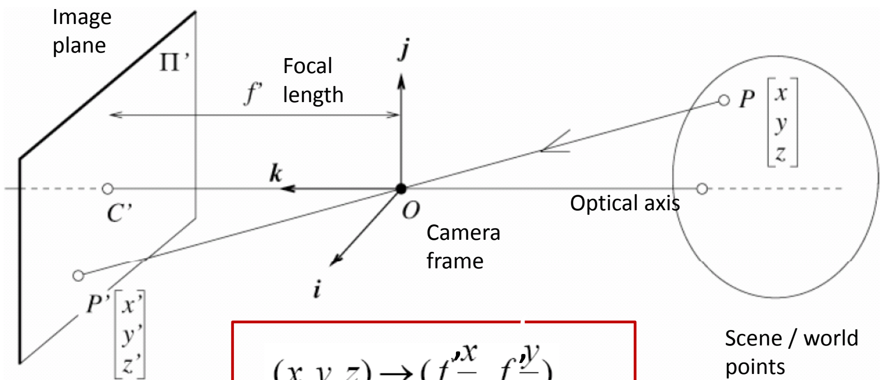
Both angle and length are not preserved via camera projection.

Image source: http://i.i.com.com/cnwk.1d/sc/30732122-2-440-camera+off-5.gif

Perspective projection model

(Pin-hole model revisited)

- 3d world mapped to 2d projection in image plane



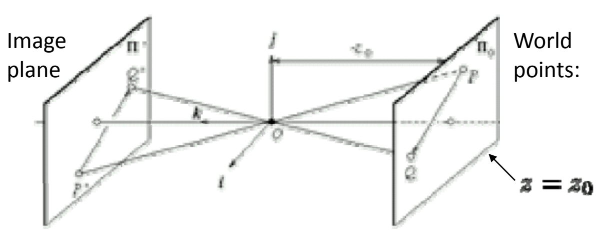
$$(x, y, z) \rightarrow (f' \frac{x}{z}, f' \frac{y}{z})$$

Scene point → Image coordinates

Forsyth and Ponce

Weak perspective

- Approximation: treat magnification as constant
- Assumes scene depth << average distance to camera

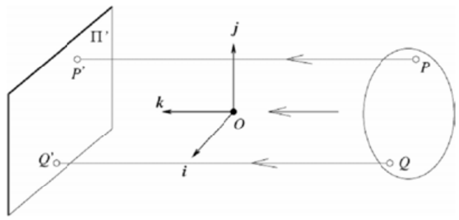


$$x' = f\frac{x}{z} \approx \frac{f}{z_0}x$$

$$y' = f\frac{y}{z} \approx \frac{f}{z_0}y$$

Orthographic projection

- World points projected along rays parallel to optical axis



$$x' = x$$
$$y' = y$$

Projective transformation (2D case)

- Hierarchy of transformations

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 1.7.3).
Euclidean 3 dof	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Length, area

General Imaging

(Full perspective camera)

Weak perspective camera

Scaled orthographic camera

Orthographic camera

Increasing focal, increasing distance

Multiple View Geometry in Computer Vision Second Edition
Richard Hartley and Andrew Zisserman,

Homogeneous coordinates

Trick: add one more coordinate:

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

homogeneous image coordinates

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

homogeneous scene coordinates

Converting *from* homogeneous coordinates

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

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

Slide by Steve Seitz

Homogeneous coordinates

Why do we use a homogeneous coordinates instead of Euclidean coordinates for describing camera model?

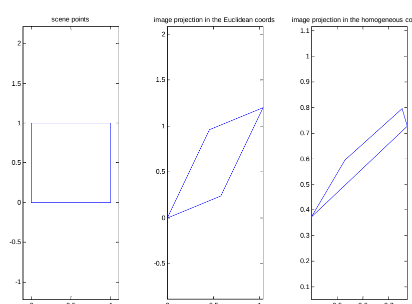
1. Euclidean cannot represent a (full) projective transformation in a linear matrix-vector form (i.e., y = Ax). It can only represent transformations up to affine.

```
scene_points = [0,0;0,1;1,1;1,0;0,0];
projection_matrix = rand(2,2);
image_points = projection_matrix*scene_points';
image_points = image_points';

scene_points_in_homogeneous = cat(2, scene_points, ones(5,1));
projection_matrix_homogeneous = rand(3,3);
image_points_homogeneous =
projection_matrix_homogeneous*scene_points_in_homogeneous';
image_points_homogeneous = image_points_homogeneous';
% back to the euclidean to display
for i = 1 : 5
   image_points_homogeneous(i,:) =
image_points_homogeneous(i,:)./image_points_homogeneous(i,3);
end

subplot(1,3,1); plot(scene_points(:,1), scene_points(:,2)); axis([-0.1 1.1 -0.1 1.1]);
axis equal;
subplot(1,3,2); plot(image_points(:,1), image_points(:,2));  axis equal;
subplot(1,3,3); plot(image_points_homogeneous(:,1),
image_points_homogeneous(:,2)); axis equal;
```

Matlab script

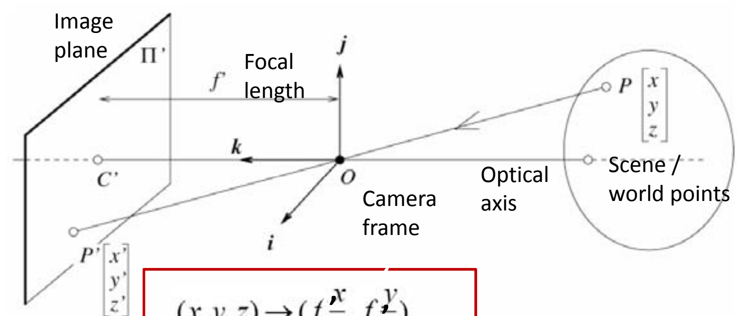


Scene points | **Projective transform under Euclidean coords** | **Projective transform under Homogeneous coords**

Homogeneous coordinates

Why do we use a homogeneous coordinates instead of Euclidean coordinates for describing camera model?

1. It converts the non-linear projection equation in the Euclidean coordinates into the linear form



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

Scene point → Image coordinates

→ division by z is nonlinear

Perspective Projection Matrix

- Projection becomes a linear matrix-vector multiplication using homogeneous coordinates:

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

divide by the third coordinate to convert back to non-homogeneous coordinates

Slide by Steve Seitz

Summary

- Pin-hole vs. Lens
  - What advantages can we obtain from using lens?
- Lens properties and thin lens equation
- Perspective effects by camera projection
  - Parallelism is not preserved.
- Various camera models and related projective transformations
- Homogeneous coordinates
  - Why we use it instead of Euclidean coordinates?
- Perspective projection matrix
  - This will be used later for camera calibration.