Back to the Beginning

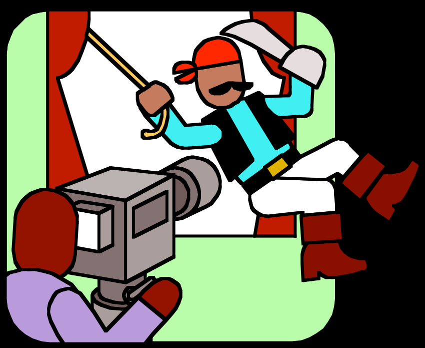
- Most classes on computer vision:
  - First optics and image formation
  - Then problem solving
- This class
  - First work on problems (e.g. classification)
  - Then back to optics and image formation
- Why?

- In the old days:
  - Computer Vision was about trying to learn the physical properties of the world
    - Exact shape of objects
    - Exact color of objects
    - Exact motion of objects
    - etc.
  - Learning the physics of image formation was *critical* in solving these problems:
    - Camera calibration
    - Shape from shading
    - Triangulation

- Today, it is generally accepted that many problems we care about...
  - Classification
  - Object avoidance
  - Navigation
- ...may be solved without exact geometry.
- However, knowledge of image formation may help us build models, predict problems, etc.

- Why so different?

Image Formation



Lecture Outline

- Light and Optics
  - Pinhole camera model
  - Perspective projection
  - Thin lens model
  - Fundamental equation
  - Distortion: spherical & chromatic aberration, radial distortion
  - Reflection and Illumination: color, Lambertian and specular surfaces, Phong, BDRF
- Sensing Light
- Conversion to Digital Images
- Sampling Theorem
- Other Sensors: frequency, type, ….

Abstract Image

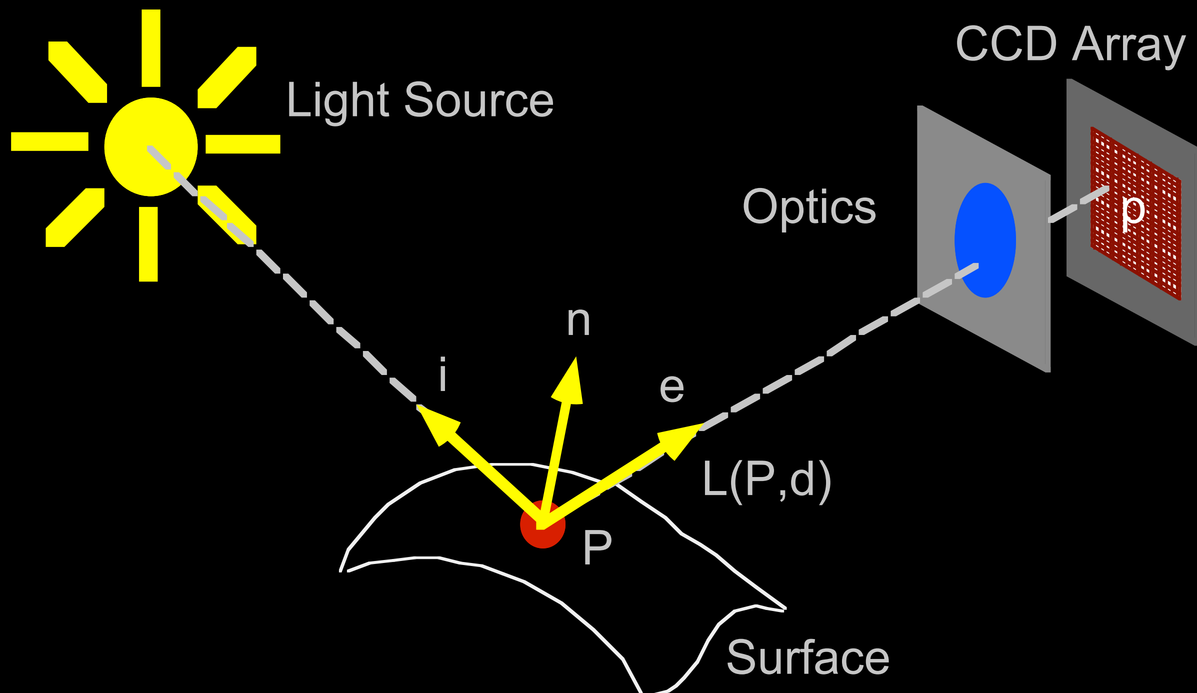
- An image can be represented by an image function whose general form is **f(x,y)**.
- **f(x,y)** is a vector-valued function whose argument represents a pixel location.
- The value of **f(x,y)** can have different interpretations in different kinds of images.

Examples

Intensity Image	- f(x,y) = intensity of the scene
Range Image	- f(x,y) = depth of the scene from imaging system
Color Image	- $\mathbf{f}(x,y) = \{f_r(x,y), f_g(x,y), f_b(x,y)\}$
Video	- f(x,y,t) = temporal image sequence

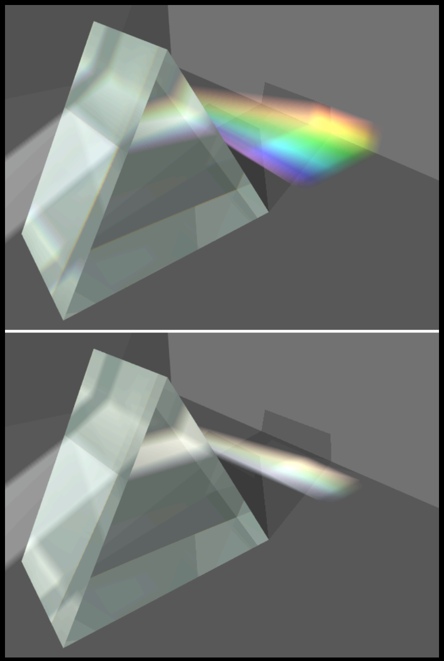
Basic Radiometry

- Radiometry is the part of image formation concerned with the relation among the amounts of light energy emitted from light sources, reflected from surfaces, and registered by sensors.



Light and Matter

- The interaction between light and matter can take many forms:
  - Reflection
  - Refraction
  - Diffraction
  - Absorption
  - Scattering
  - Emission



Lecture Assumptions

- Typical imaging scenario:
  - visible light
  - ideal lenses
  - standard sensor (e.g. TV camera)
  - opaque objects
- Goal

> To create 'digital' images which can be processed to recover some of the characteristics of the 3D world which was imaged.

Steps

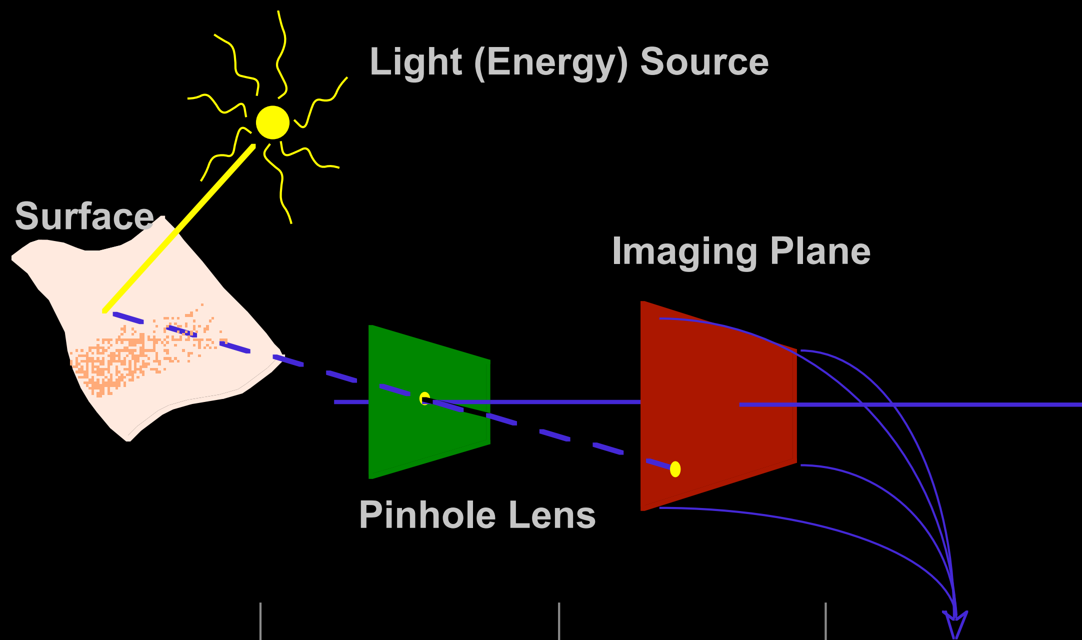
World → Optics → Sensor → Signal → Digitizer → Digital Representation

World	reality
Optics	focus {light} from world on sensor
Sensor	converts {light} to {electrical energy}
Signal	representation of incident light as continuous electrical energy
Digitizer	converts continuous signal to discrete signal
Digital Rep.	final representation of reality in computer memory

Factors in Image Formation

- Geometry
  - concerned with the relationship between points in the three-dimensional world and their images
- Radiometry
  - concerned with the relationship between the amount of light radiating from a surface and the amount incident at its image
- Photometry
  - concerned with ways of measuring the intensity of light
- Digitization
  - concerned with ways of converting continuous signals (in both space and time) to digital approximations

Image Formation



World	Optics	Sensor	Signal
		B&W Film	Silver Density
		Color Film	Silver density in three color layers
		TV Camera	Electrical

- Geometry describes the projection of:

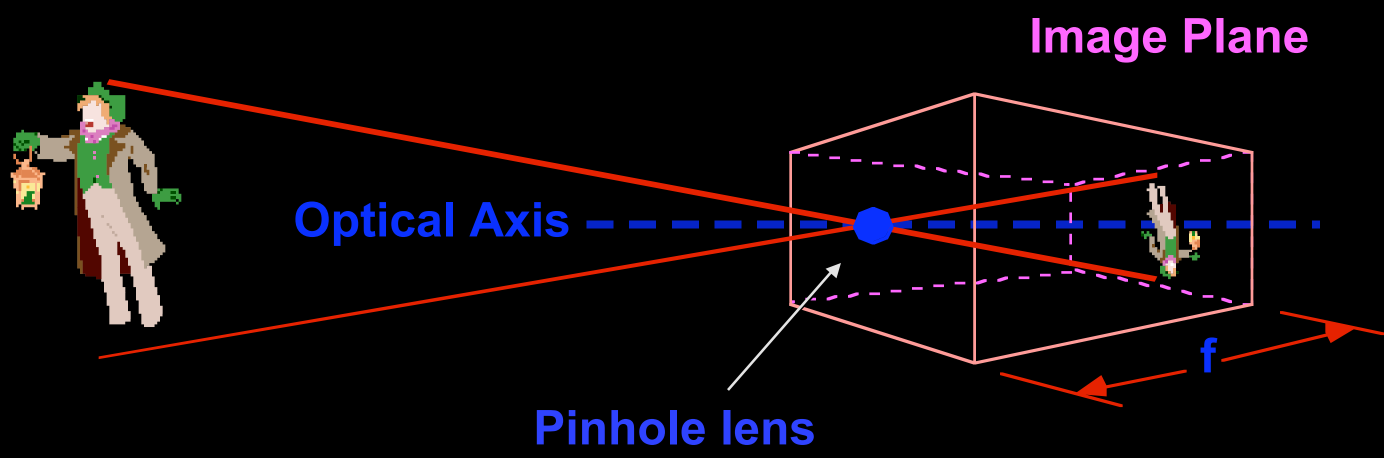
three-dimensional (3D) world ➡ two-dimensional (2D) image plane.

- Typical Assumptions
  - Light travels in a straight line
- Optical Axis: the perpendicular from the image plane through the pinhole (also called the central projection ray)
- Each point in the image corresponds to a particular direction defined by a ray from that point through the pinhole.
- Various kinds of projections:
  - \- perspective - oblique
  - \- orthographic - isometric
  - \- spherical

Basic Optics

- Two models are commonly used:
  - Pin-hole camera
  - Optical system composed of lenses
- Pin-hole is the basis for most graphics and vision
  - Derived from physical construction of early cameras
  - Mathematics is very straightforward
- Thin lens model is first of the lens models
  - Mathematical model for a physical lens
  - Lens gathers light over area and focuses on image plane.

Pinhole Camera Model



- World projected to 2D Image
  - Image inverted
  - Size reduced
  - Image is dim
  - No direct depth information
- f called the focal length of the lens
- Known as perspective projection

Pinhole camera image

Amsterdam

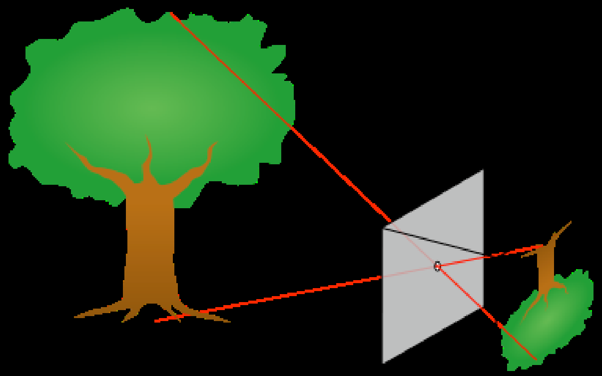
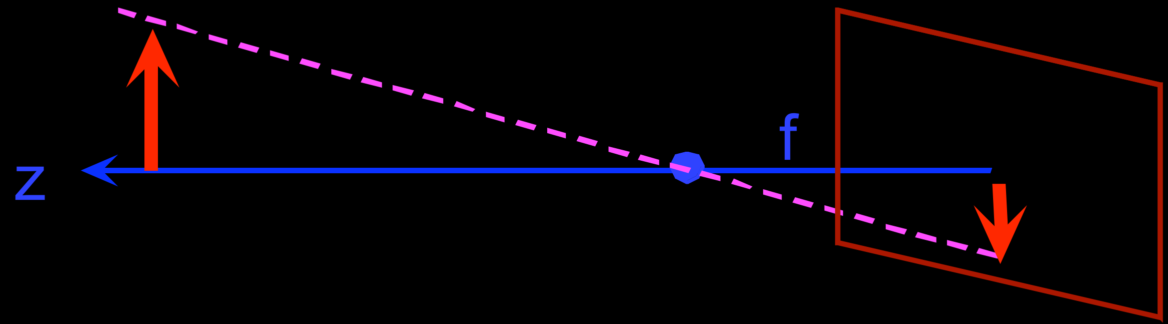


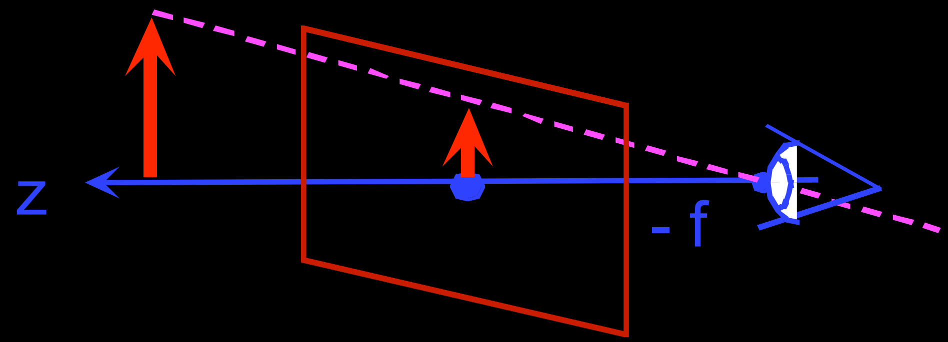
Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Equivalent Geometry

- Consider case with object on the optical axis:



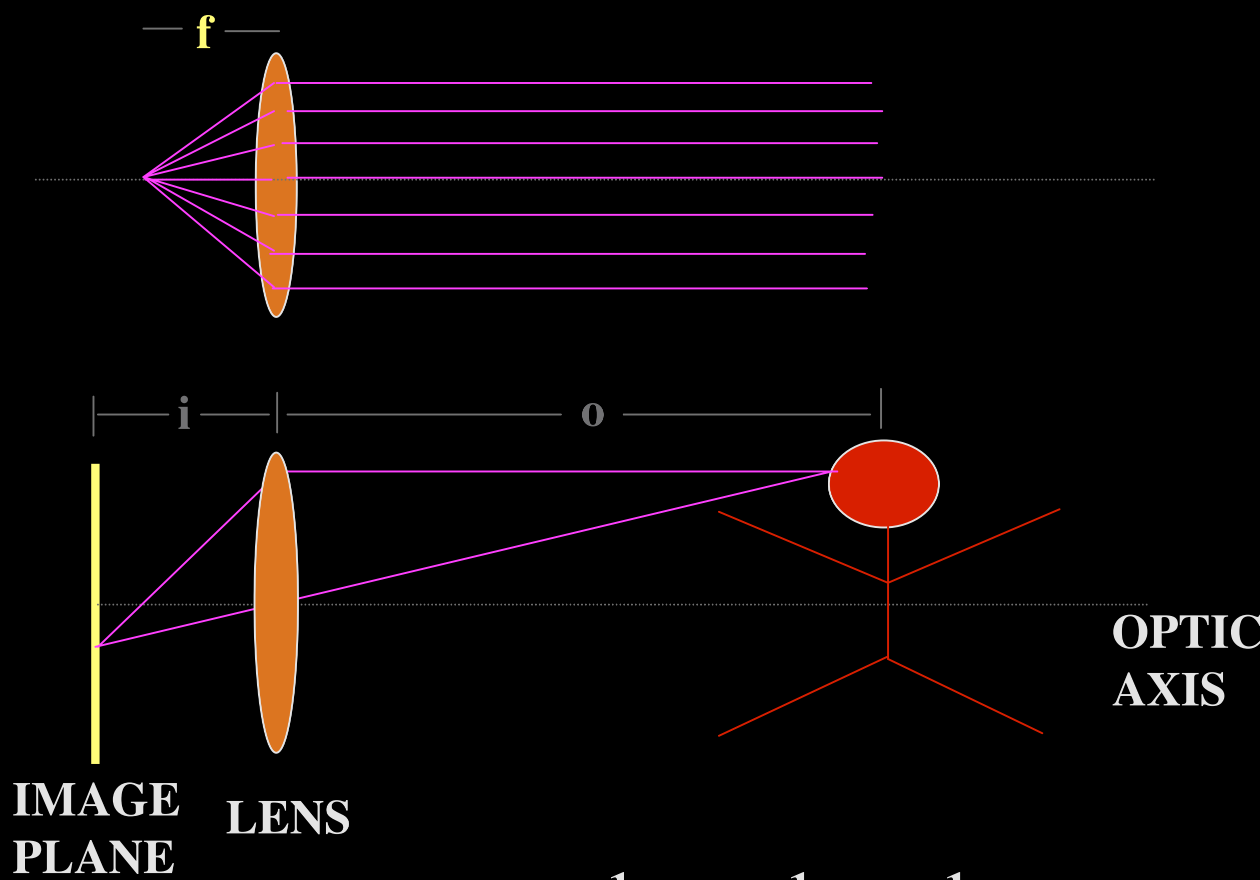
- More convenient with upright image:



- Equivalent mathematically

Thin Lens Model

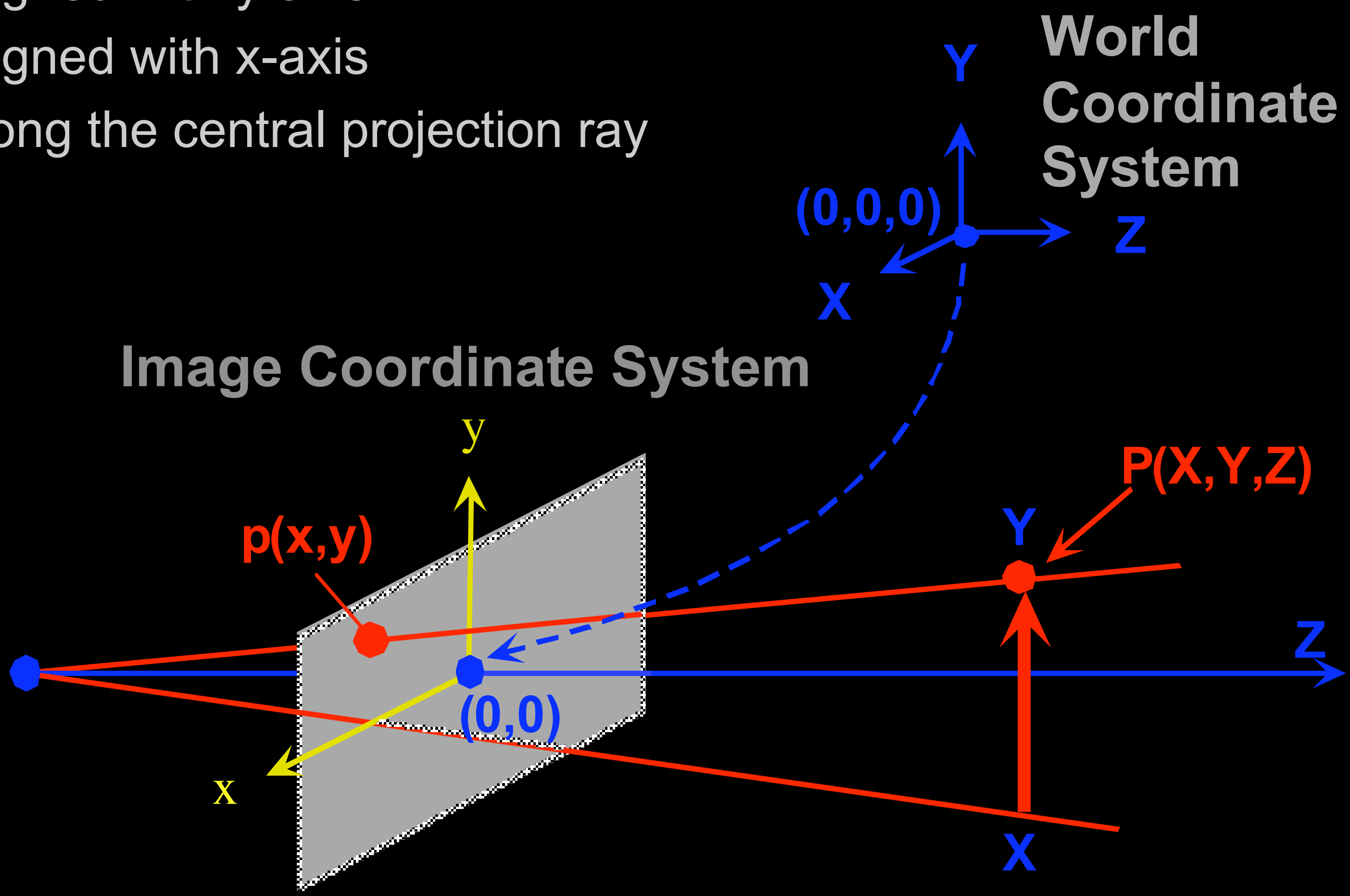
- Rays entering parallel on one side converge at focal point.
- Rays diverging from the focal point become parallel.



$$\frac{1}{f} = \frac{1}{i} + \frac{1}{o}$$ 'THIN LENS LAW'

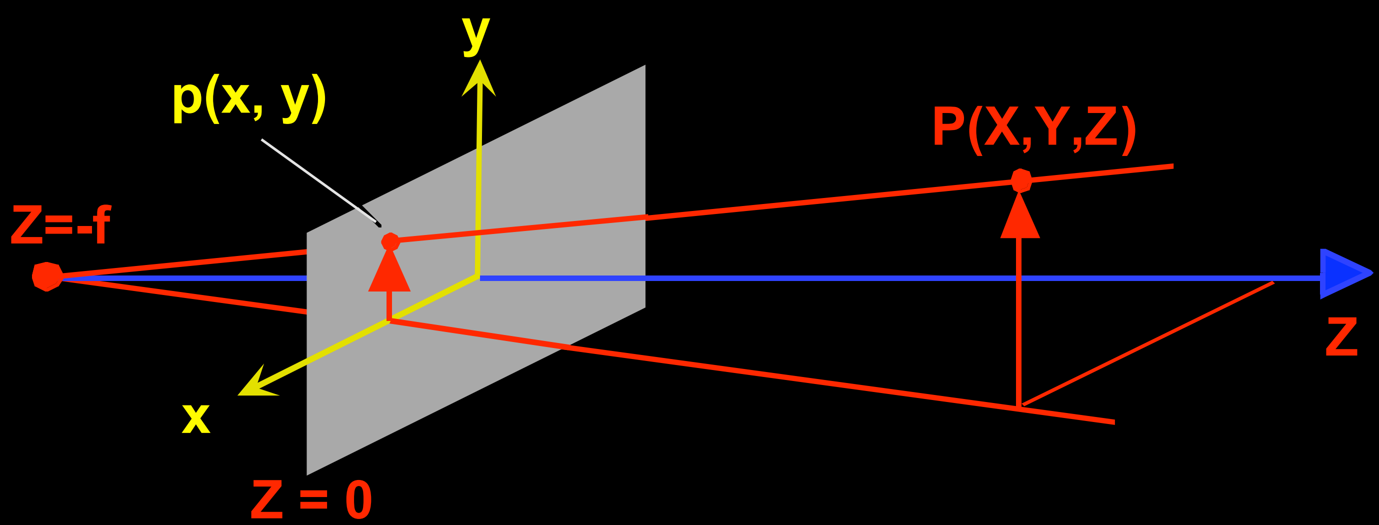
Coordinate System

- Simplified Case:
  - Origin of world and image coordinate systems coincide
  - Y-axis aligned with y-axis
  - X-axis aligned with x-axis
  - Z-axis along the central projection ray



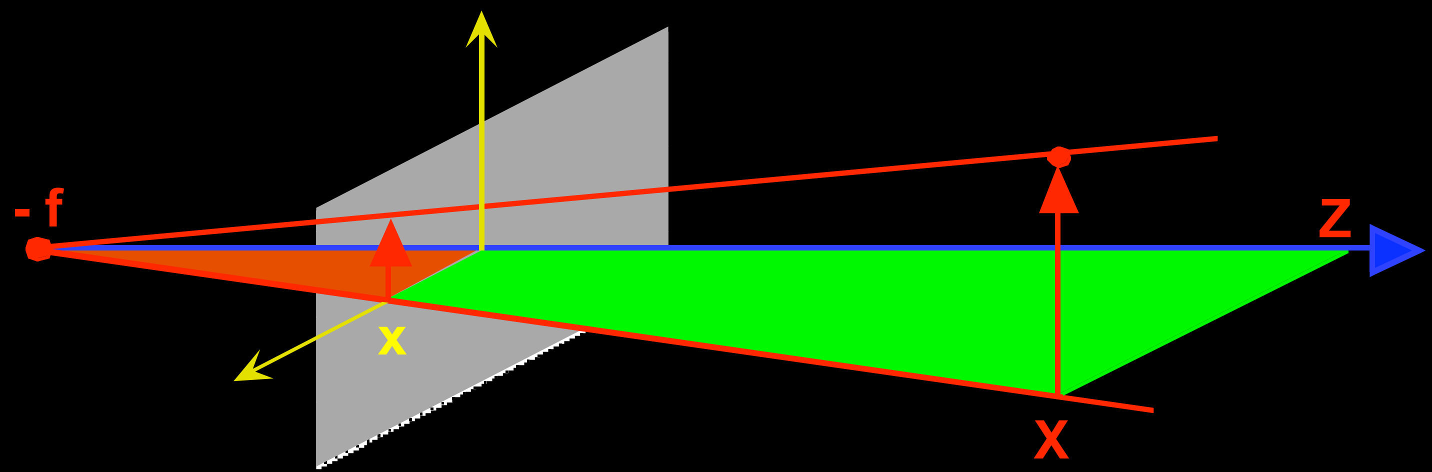
Perspective Projection

- Compute the image coordinates of p in terms of the world coordinates of P.



- Look at projections in x-z and y-z planes

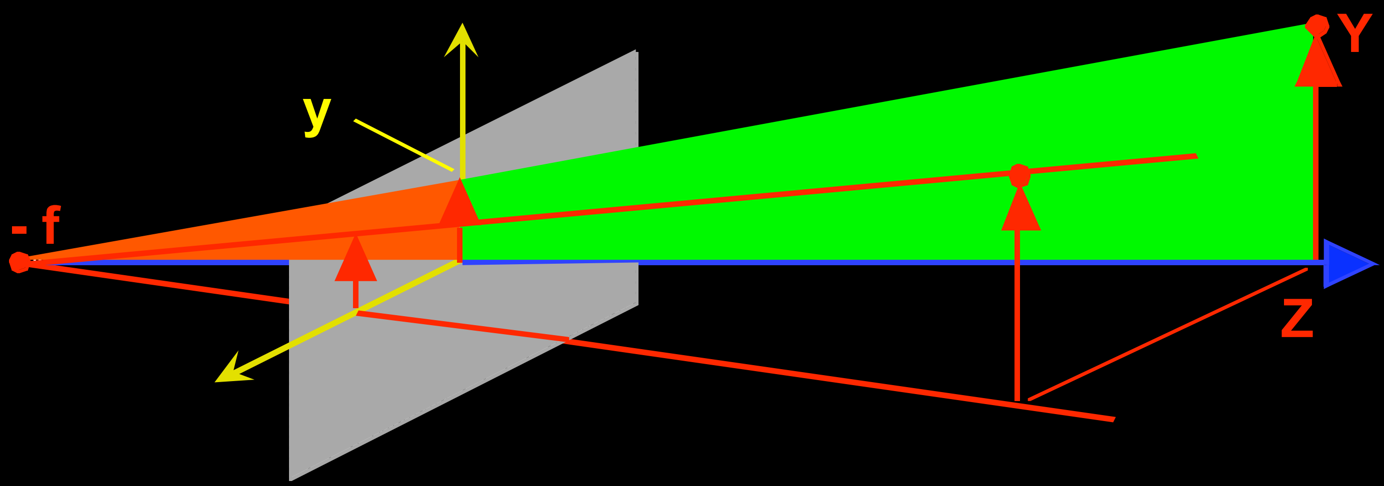
X-Z Projection



- By similar triangles: $\frac{x}{f} = \frac{X}{Z+f}$

$$x = \frac{fX}{Z+f}$$

Y-Z Projection



- By similar triangles: $\frac{y}{f} = \frac{Y}{Z+f}$

$$y = \frac{fY}{Z+f}$$

Perspective Equations

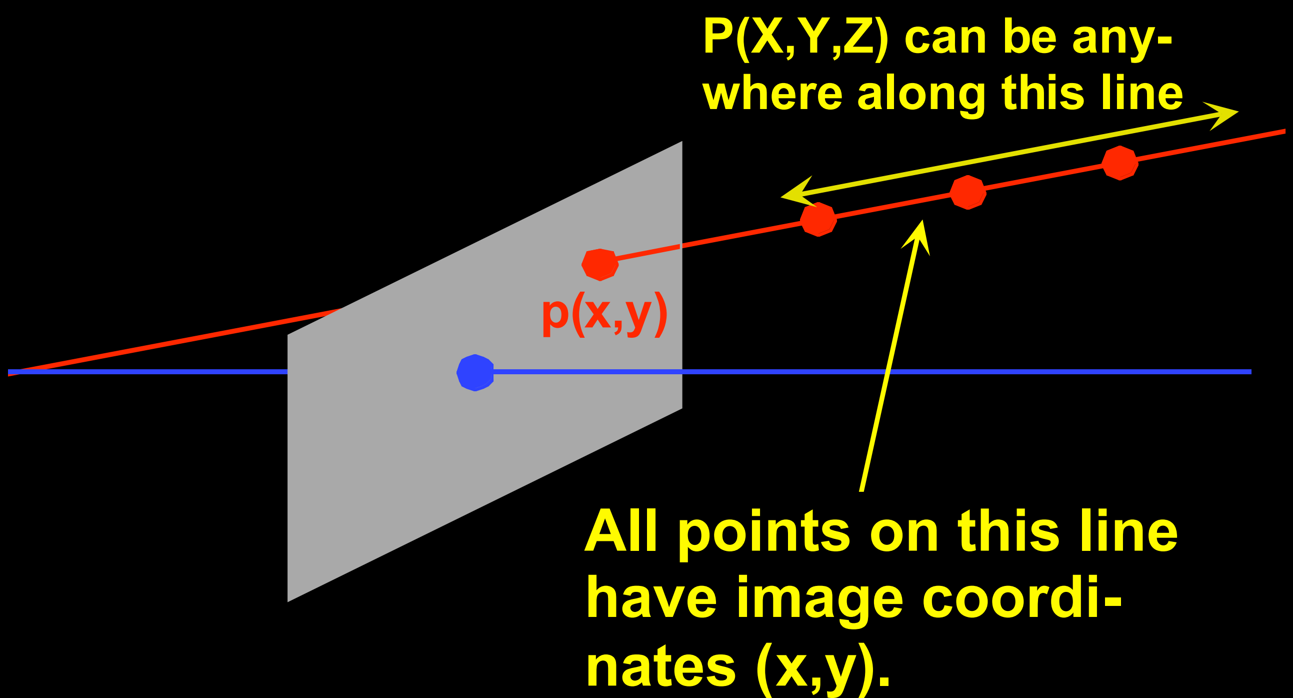
- Given point P(X,Y,Z) in the 3D world
- The two equations:

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

- transform world coordinates (X,Y,Z) into image coordinates (x,y)

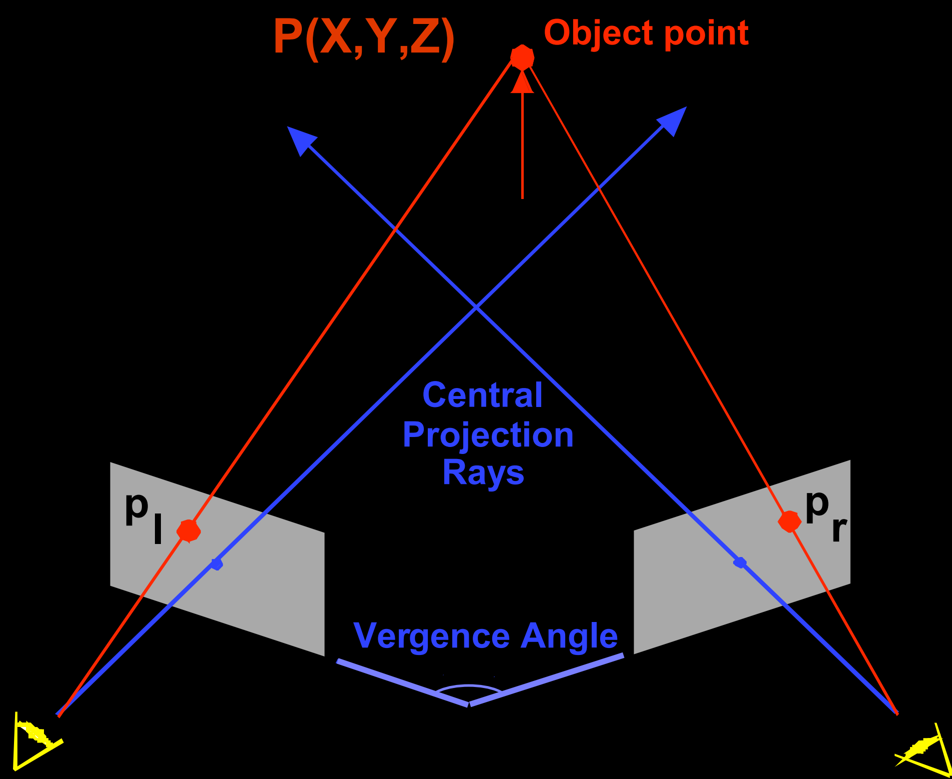
Reverse Projection

- Given a center of projection and image coordinates of a point, it is not possible to recover the 3D depth of the point from a single image.



In general, at least two images of the same point taken from two different locations are required to recover depth.

Stereo Geometry



- Depth obtained by triangulation
- Correspondence problem: p_l and p_r must correspond to the left and right projections of P, respectively.

- Consequences of image formation geometry for computer vision
  - What set of shapes can an object take on?
    - rigid
    - non-rigid
    - planar
    - non-planar
  - SIFT features
- Sensitivity to errors.

Radiometry

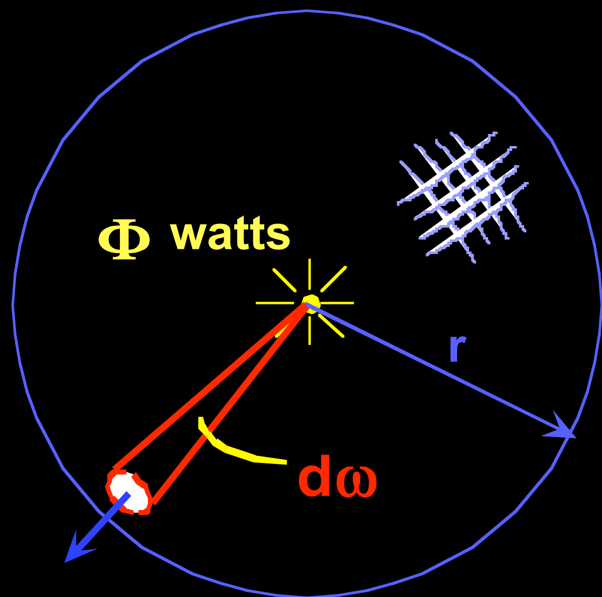
- Image: two-dimensional array of 'brightness' values.
- Geometry: where in an image a point will project.
- Radiometry: what the brightness of the point will be.
- Brightness: informal notion used to describe both scene and image brightness.
- Image brightness: related to energy flux incident on the image plane:

  IRRADIANCE
- Scene brightness: brightness related to energy flux emitted (radiated) from a surface.

  RADIANCE

Light

- Electromagnetic energy
- Wave model
- Light sources typically radiate over a frequency spectrum
- Φ watts radiated into 4π radians



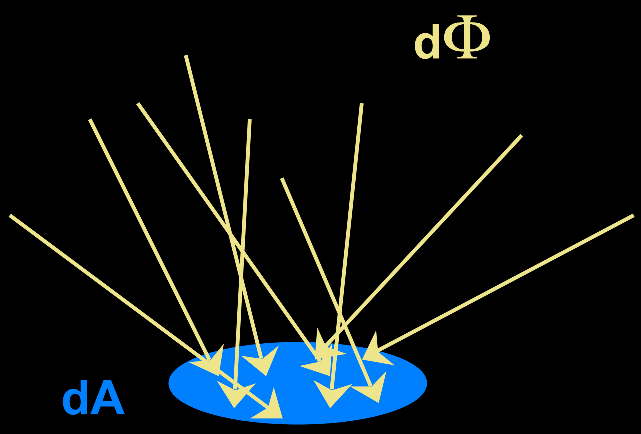
$$\Phi = \int_{\text{sphere}} d\Phi$$

R = Radiant Intensity = $\frac{d\Phi}{d\omega}$ Watts/unit solid angle (steradian)

(of source)

Irradiance

- Light falling on a surface from all directions.
- How much?



- Irradiance: power per unit area falling on a surface.

$$\text{Irradiance } E = \frac{d\Phi}{dA} \quad \text{watts/m}^2$$