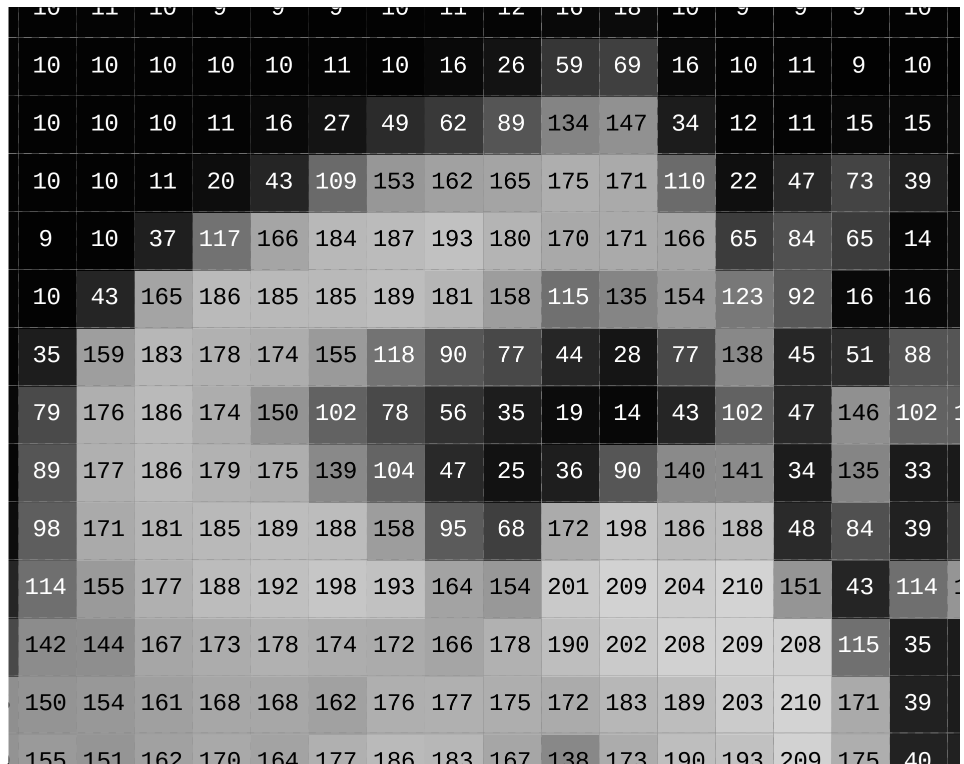
Sensors and Image Formation

- Different types of imaging sensors
- How images are formed (mathematical models)
- Coordinate systems

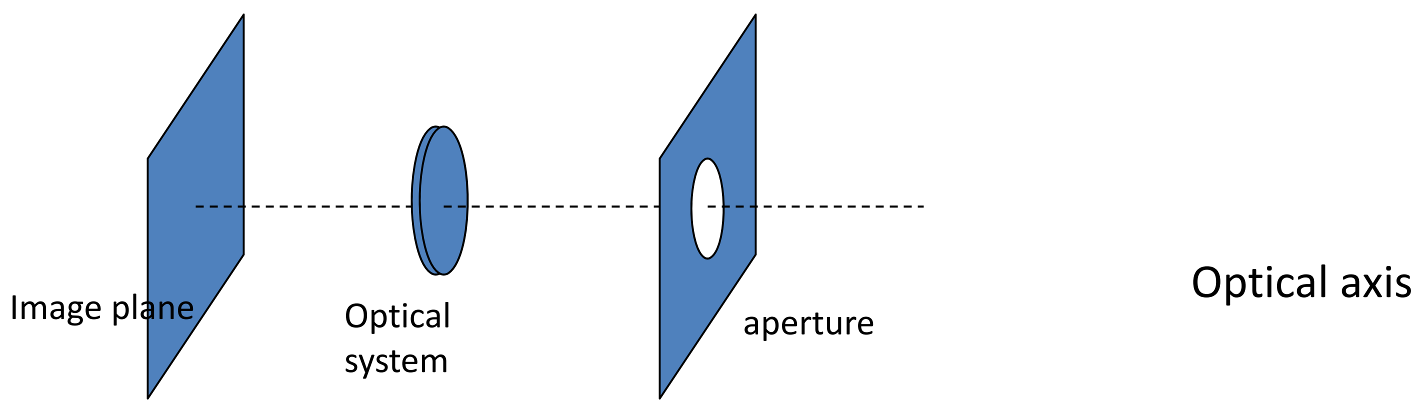
Digital Images

- Digital images are stored as arrays of numbers
- Numbers can represent
  - Intensity (gray level, or each color band)
  - Range
  - X-ray absorption coefficient
  - etc



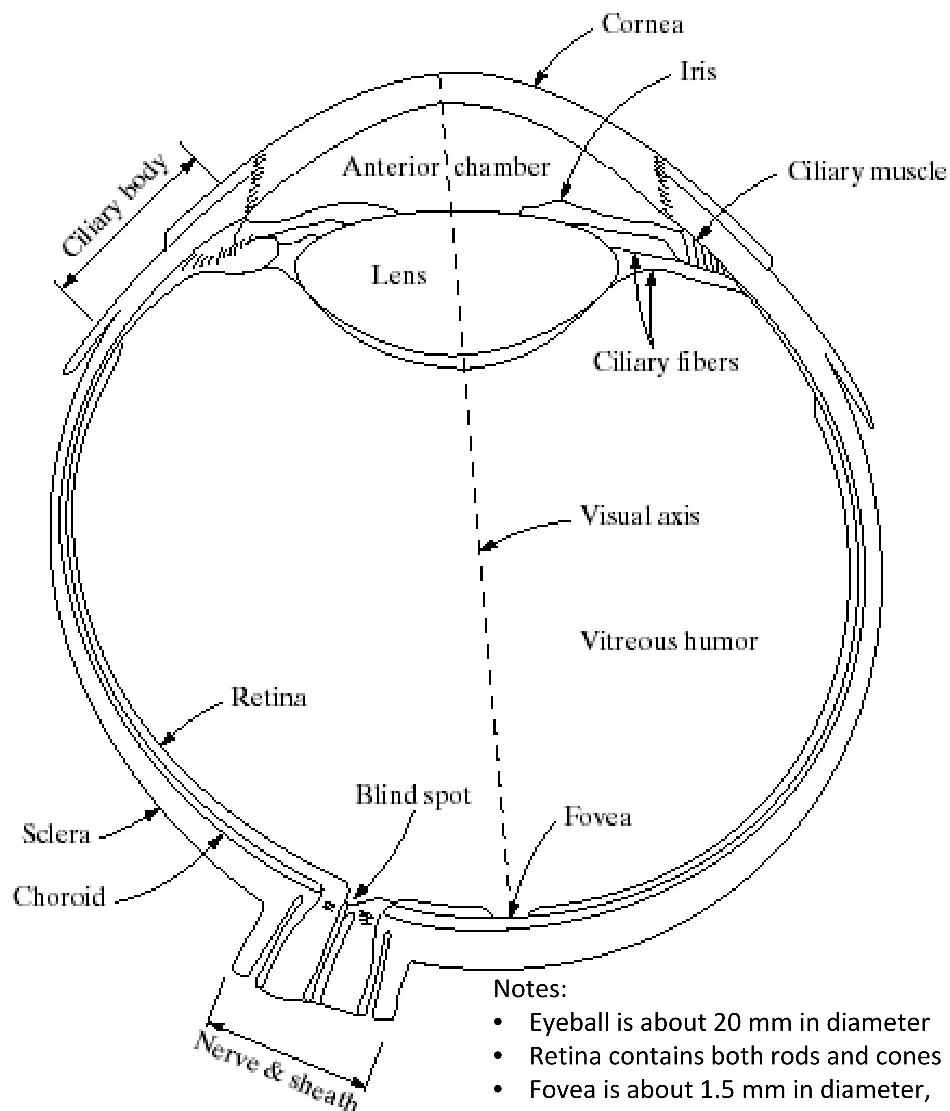
Intensity Image Sensors

- Basic elements of an imaging device
  - Aperture
    - An opening (or "pupil") to limit amount of light, and angle of incoming light rays
  - Optical system
    - Lenses - purpose is to focus light from a scene point to a single image point
  - Imaging photosensitive surface
    - Film or sensors, usually a plane



Biological Vision – A Guide to Developing Computer Vision?

- Unfortunately, it is pre-attentive
  - Difficult to measure and study detailed function
- Some information from experiments with
  - Animals (electrophysiology)
  - People's perception (psychophysics)
- A great deal of processing is done right in the retina
  - Data reduction of ~100 million receptors down to ~1 million optic nerve channels
- Further processing done in the visual cortex
  - Specialized detectors for motion, shape, color, binocular disparity
  - Evidence for maps in the cortex in correspondence to the image
- Still, it is an existence proof that it can be done, and well



Notes:
- Eyeball is about 20 mm in diameter
- Retina contains both rods and cones
- Fovea is about 1.5 mm in diameter, contains about 337,000 cones
- Focal length about 17 mm

Digital Camera

- Image plane is a 2D array of sensor elements
- CCD type (Charge coupled device)
  - Charge accumulates during exposure
  - Charges are transferred out to shift registers, digitized and read out sequentially
- CMOS type (complementary metal oxide on silicon)
  - Light affects the conductivity (or gain) of each photodetector
  - Digitized and read out using a multiplexing scheme
- Main design factors
  - Number and size of sensor elements
  - Chip size
  - ADC resolution

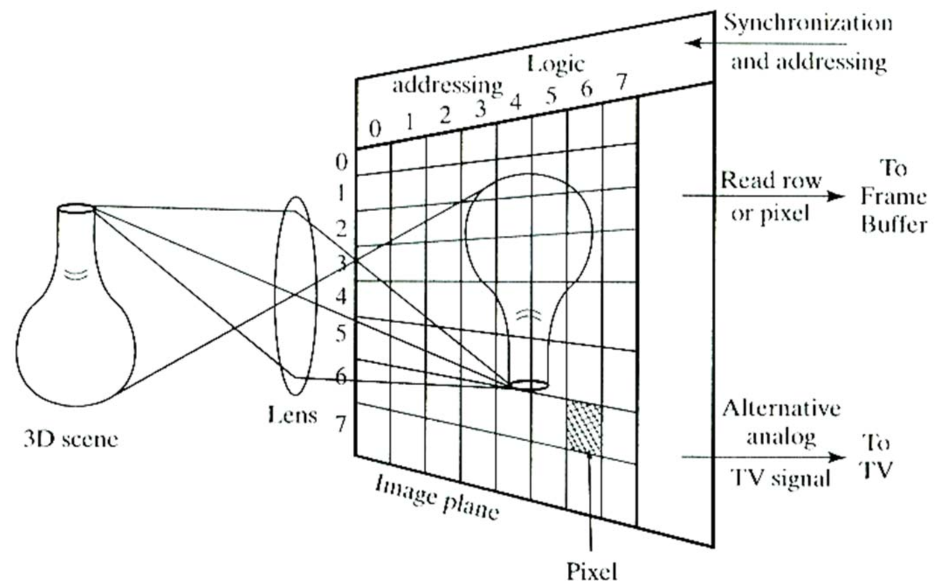
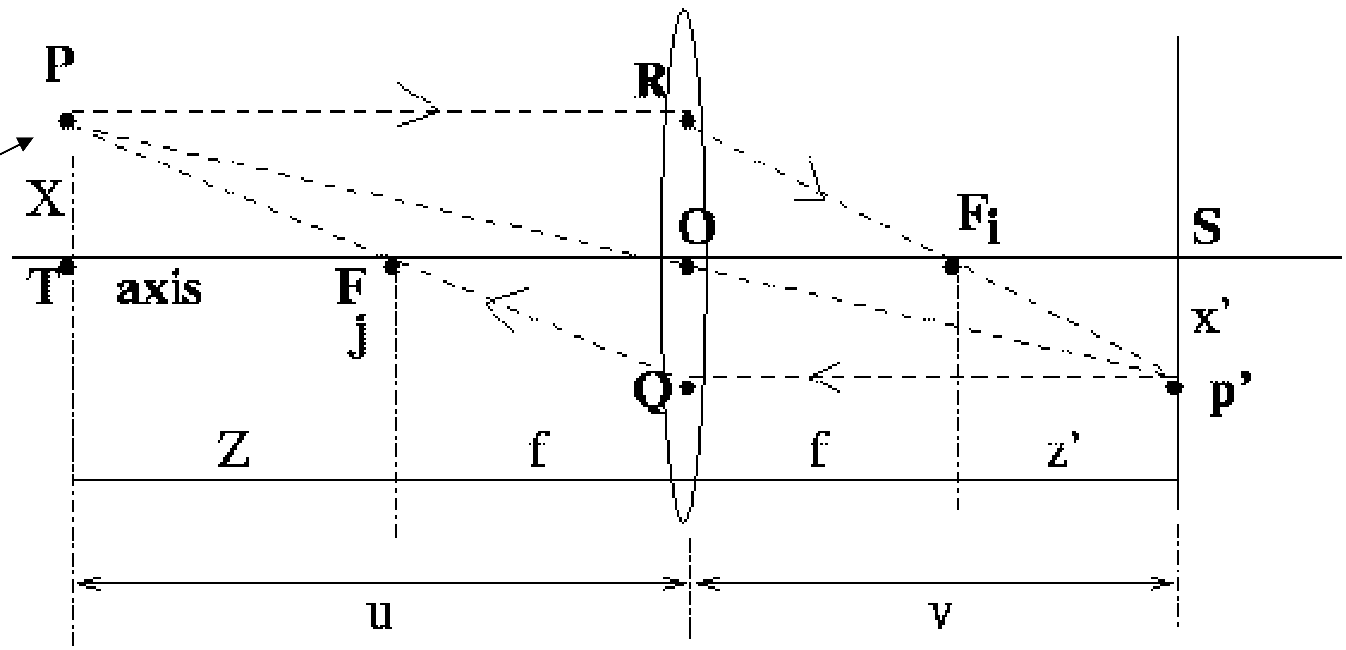


Figure 2.2 A CCD (charge-coupled device) camera imaging a vase; discrete cells convert light energy into electrical charges, which are represented as small numbers when input to a computer.

Thin Lens

- Rays parallel to the optical axis are deflected to go through the focus
- Rays passing through the center are undeflected

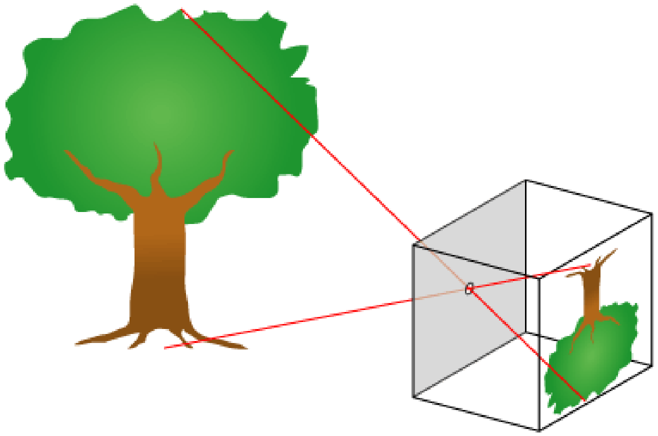
World point P is "in focus" at image point p'

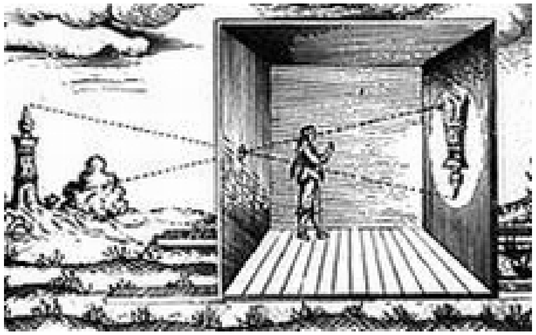


Equation of a thin lens: $$\frac{1}{Z}+\frac{1}{z'}=\frac{1}{f}$$

Pinhole Camera Model

- A good lens can be modeled by a pinhole camera; ie., each ray from the scene passes undeflected to the image plane
- Simple equations describe projection of a scene point onto the image plane ("perspective projection")
- We will use the pinhole camera model exclusively, except for a little later in the course where we model lens distortion in real cameras



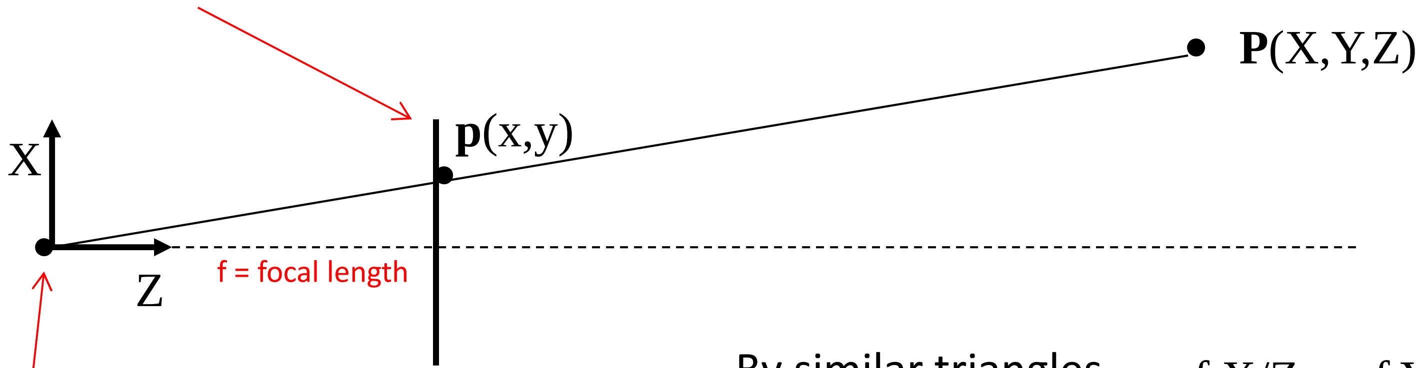


The pinhole camera ("camera obscura") was used by Renaissance painters to help them understand perspective projection

Perspective Projection Equations

For convenience (to avoid an inverted image) we treat the image plane as if it were in front of the pinhole

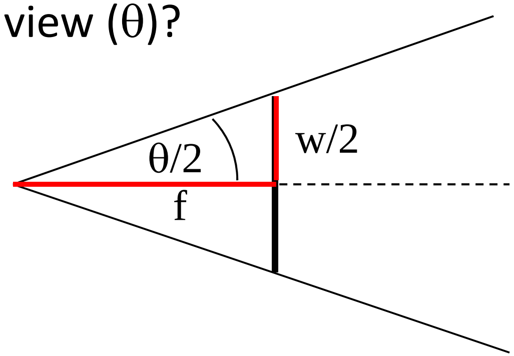
The XYZ coordinates of the point are with respect to the camera origin



We define the origin of the camera's coordinate system at the pinhole (note – this is a 3D XYZ coordinate frame

By similar triangles, $x = f\,X/Z,\ y = f\,Y/Z$

Field of view (θ)?

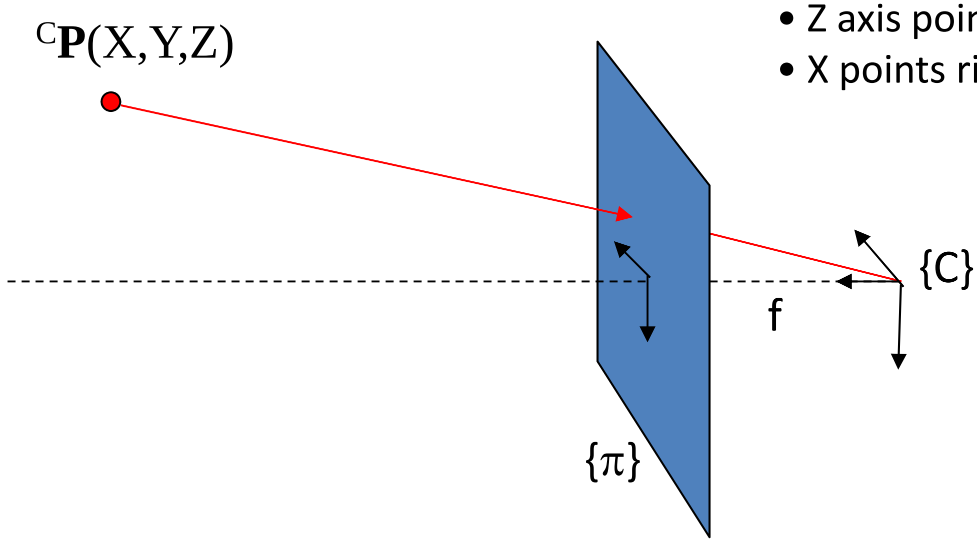


$\tan(\theta/2) = (w/2)/f$

Camera vs Image Plane Coords

Camera coordinate system {C}

- A 3D coordinate system (X,Y,Z) – units say, in meters
- Origin at the center of projection
- Z axis points outward along optical axis
- X points right, Y points down



$$x = f\,X/Z$$

$$y = f\,Y/Z$$

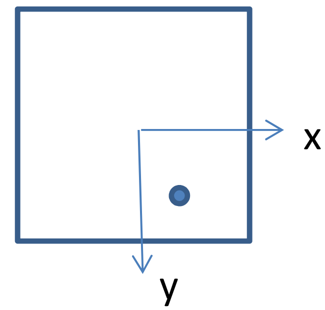
Image plane coordinate system {π}

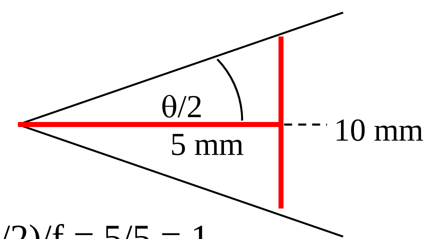
- A 2D coordinate system (x,y) – units in mm
- Origin at the intersection of the optical axis with the image plane
- In real systems, this is where the CCD or CMOS plane is

Examples

- Assume focal length = 5 mm
- A scene point is located at (X,Y,Z) = (1m, 2m, 5m)
  – What are the image plane coordinates (x,y) in mm?
  – If the image plane is 10mm x 10mm, what is the field of view?
  – A building is 100m wide. How far away do we have to be in order that it fills the field of view?

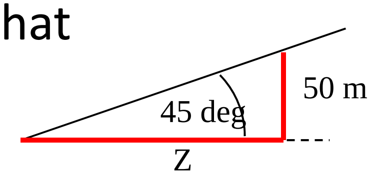
x = f X/Z = (5 mm) (1 m)/(5 m)
= 1 mm
y = f Y/Z = (5 mm) (2 m)/(5 m)
= 2 mm





$\tan(\theta/2) = (w/2)/f = 5/5 = 1$

so $\theta/2$ = 45 deg, fov is 90x90 deg

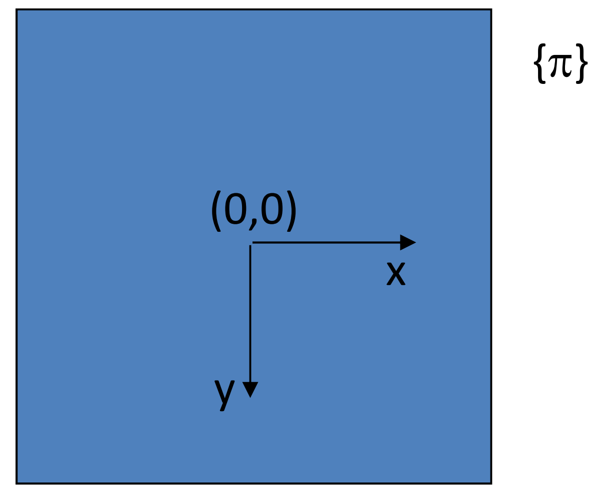


tan(45 deg) = (W/2)/Z= 50/Z

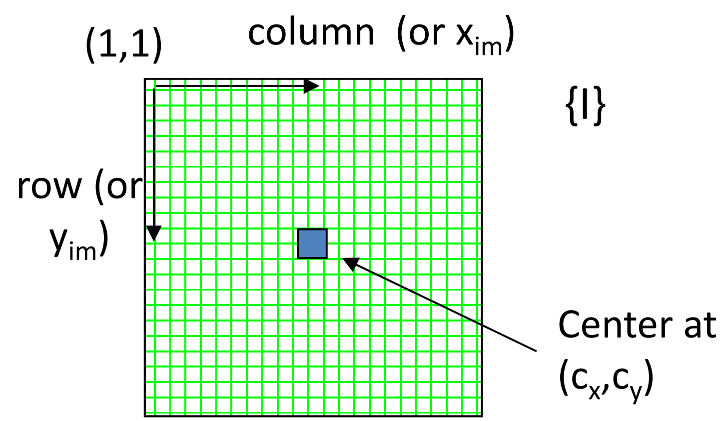
so Z = 50 m

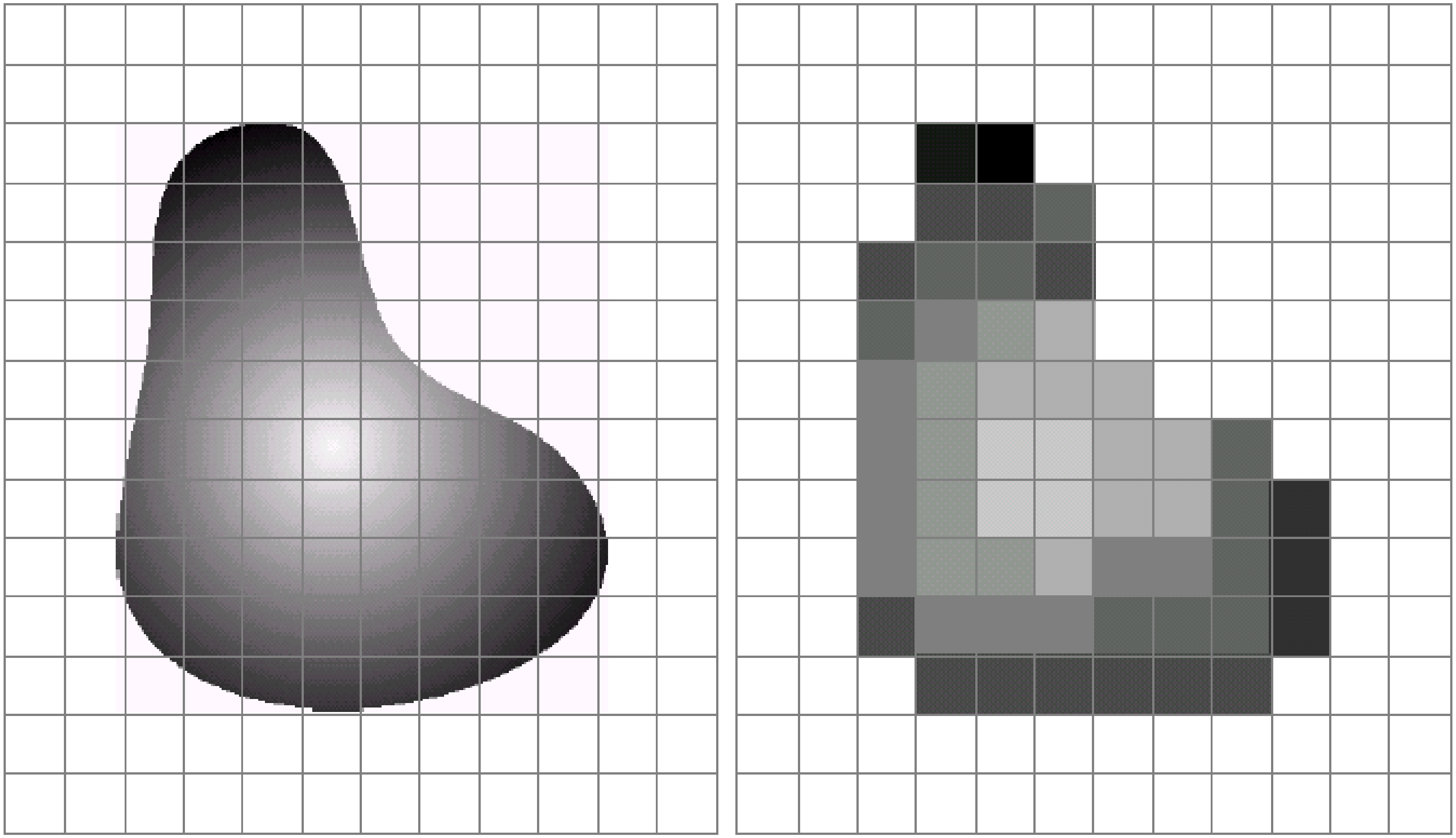
Image Buffer

- Image plane
  - The real image is formed on the CCD plane
  - (x,y) units in mm
  - Origin in center (principal point)



- Image buffer
  - Digital (or pixel) image
  - (row, col) indices
  - We can also use (x_{im}, y_{im})
  - Origin in upper left



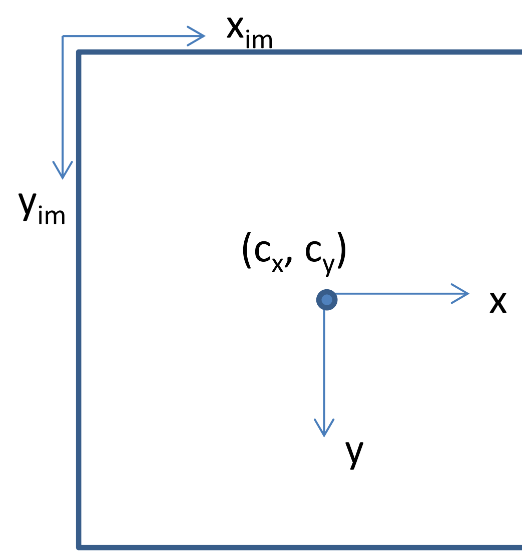


a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Conversion between real image and pixel image coordinates

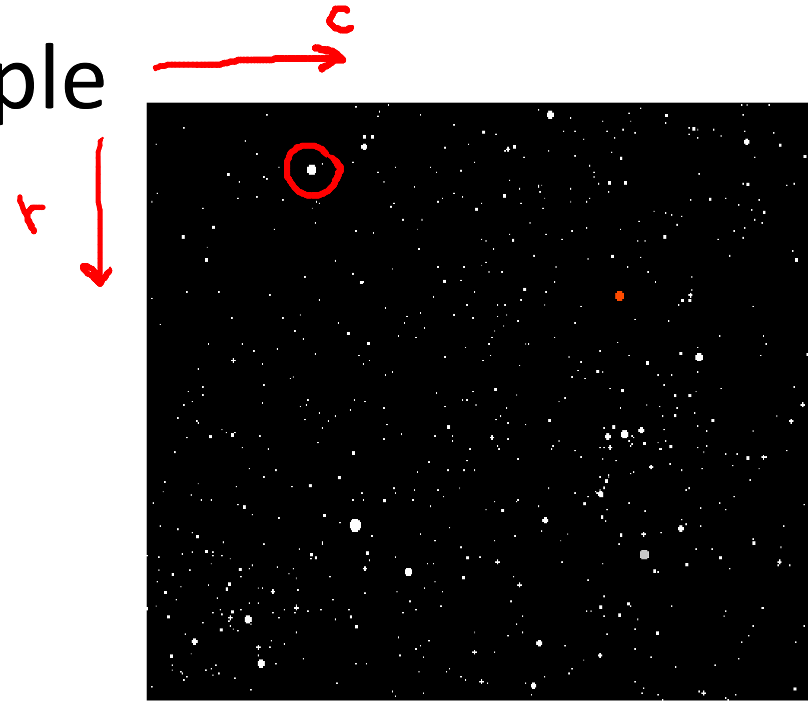
- Assume
  - The image center (principal point) is located at pixel (c_x,c_y) in the pixel image
  - The spacing of the pixels is (s_x,s_y) in millimeters



- Then

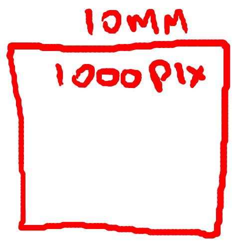
$$x = (x_{im} - c_x)\, s_x \qquad x_{im} = x/s_x + c_x$$
$$y = (y_{im} - c_y)\, s_y \qquad y_{im} = y/s_y + c_y$$

Example

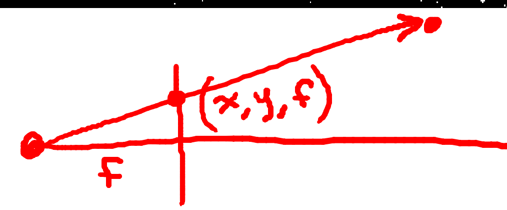


- A star is located at pixel (r,c)=(100,200) in a telescope image
- What is the 3D unit vector pointing at the star?

$$s_x = \frac{10\,mm}{1000\,pix} = .01\,mm$$



- Assume:
  - Image is 1000x1000 pixels
  - Optical center is in the center of the pixel image
  - CCD plane is 10 mm x 10 mm
  - Focal length is 1 m



$$x = (x_{im} - c_x) s_x$$
$$y = (y_{im} - c_y) s_y$$
$$c_x = c_y = 500$$

$$\vec{v} = \begin{pmatrix} x \\ y \\ f \end{pmatrix} = \begin{pmatrix} (200-500)(.01)\,mm \\ (100-500)(.01)\,mm \\ 1000\,mm \end{pmatrix}$$

$$\hat{r} = \vec{v} / |\vec{v}|$$

Note on focal length

- Recall

  $x = (x_{im} - c_x)\, s_x$

  $y = (y_{im} - c_y)\, s_y$
- or

  $x_{im} = x/s_x + c_x$

  $y_{im} = y/s_y + c_y$
- and

  $x = f\, X/Z$

  $y = f\, Y/Z$
- So

  $x_{im} = (f / s_x)\, X/Z + c_x$

  $y_{im} = (f / s_y)\, Y/Z + c_y$

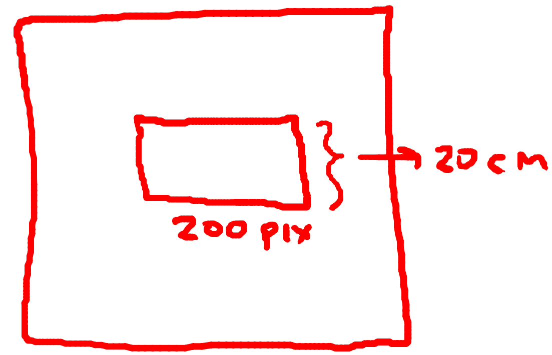
- All we really need is

  $f_x = (f / s_x)$

  $f_y = (f / s_y)$
- We don't need to know the actual values of f and s_x,s_y; just their ratios
- We can alternatively express focal length in units of pixels

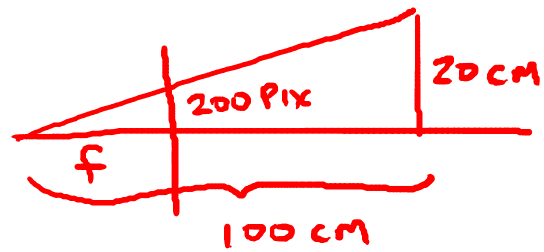
Example

- A camera observes a rectangle 1m away
  - The rectangle is known to be 20cm x 10cm
  - In the image, the rectangle measures 200 x 90 pixels



- Focal length in pixels?

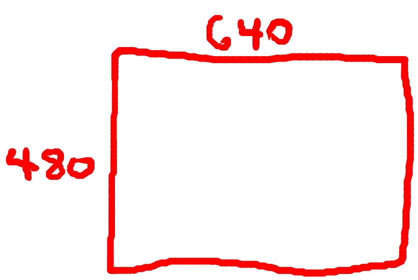
$$f_x = 1000 \text{ pix}$$

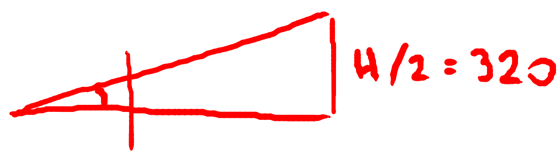


$$\frac{20\text{ cm}}{100\text{ cm}} = \frac{200\text{ pix}}{f} \rightarrow f_x = 1000 \text{ pix}$$

$$f_y: \quad \frac{10\text{ cm}}{100\text{ cm}} = \frac{90\text{ pix}}{f_y} \rightarrow f_y = 900 \text{ pix}$$

- If image size is 640x480 pixels, what is field of view?





$$\tan\frac{\theta_x}{2} = \frac{320\text{ pix}}{1000\text{ pix}} \rightarrow \theta_x = 35^\circ$$

$$\tan\frac{\theta_y}{2} = \frac{240\text{ pix}}{900\text{ pix}} \rightarrow \theta_y = 30^\circ$$

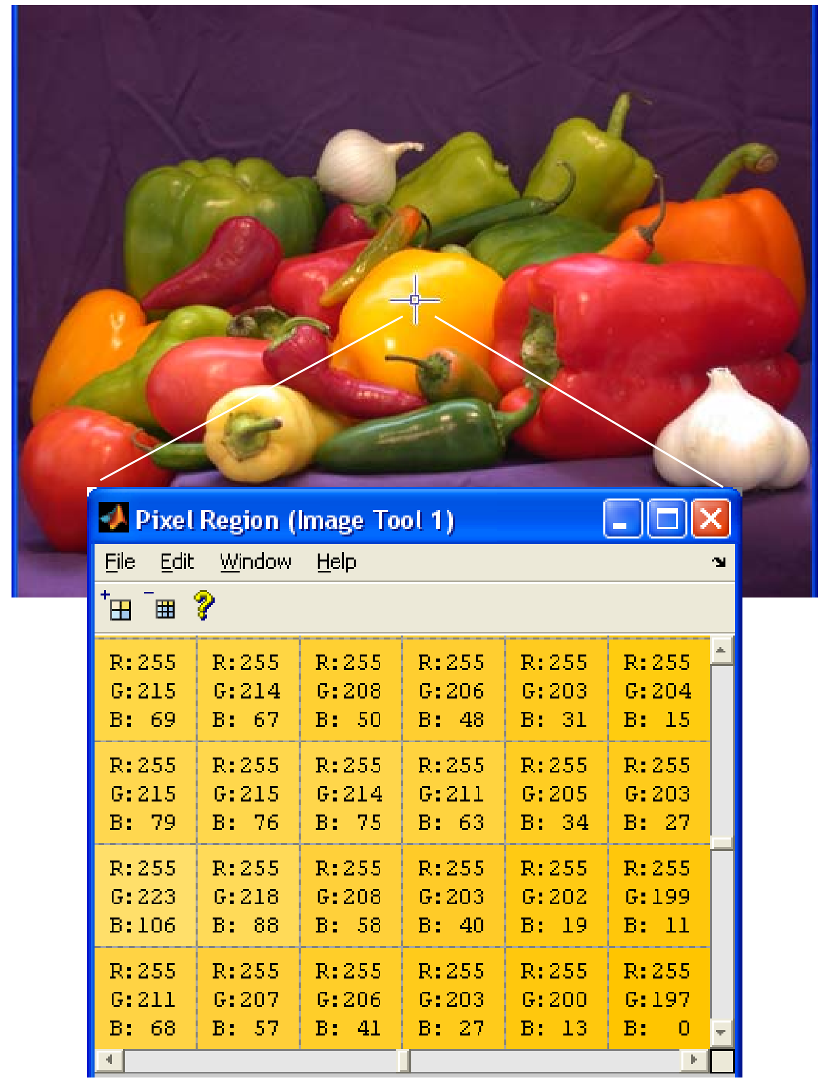
Camera Parameters

- Intrinsic parameters
  - Those parameters needed to relate an image point (in pixels) to a direction in the camera frame
  - f_x, f_y, c_x, c_y
  - Also lens distortion parameters (will discuss later)

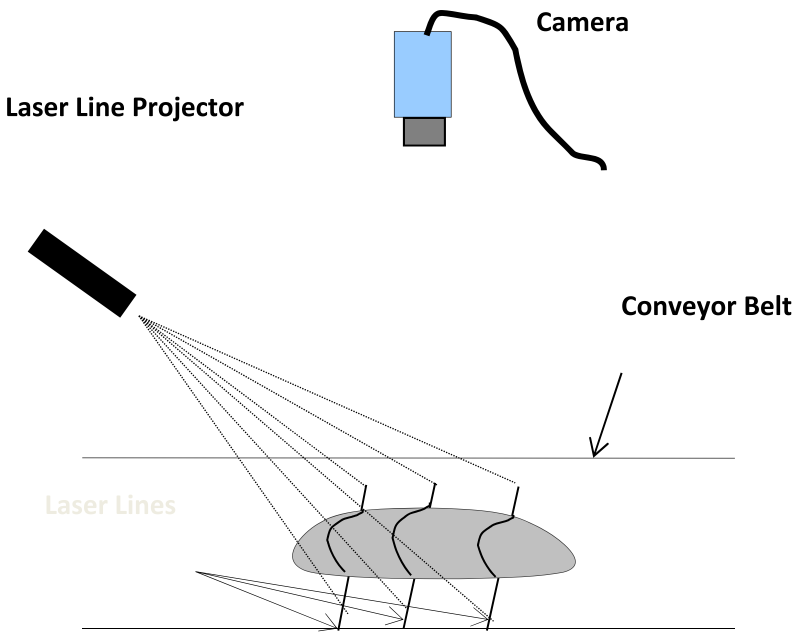
- Extrinsic parameters
  - Define the position and orientation (pose) of the camera in the world

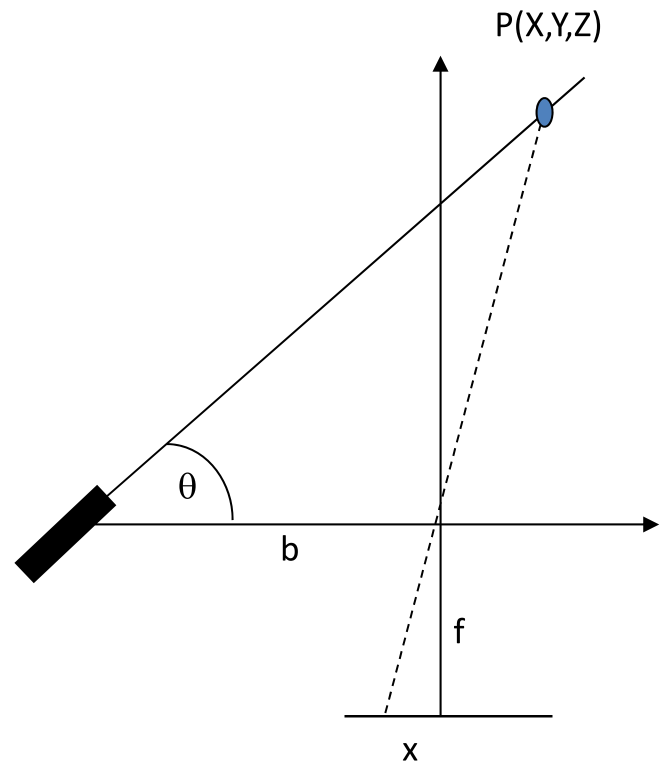
Other Common Types of Sensors

- Color
  - Each pixel is a triplet of red, green, blue values (RGB)
- Range
  - Each pixel value is the range to the nearest surface in the scene
  - Range images are also called depth maps, or 2.5D images
  - Types
    - Structured light systems
    - Time of flight - LIDAR (Light detection and ranging)



Structured Light (Active Triangulation)





Two equations, two unknowns (X,Z)

(1) $x/f = X/Z$

(2) $\tan\theta = Z/(b+X)$ or $\cot\theta = (b+X)/Z$

Solve for X,Z:

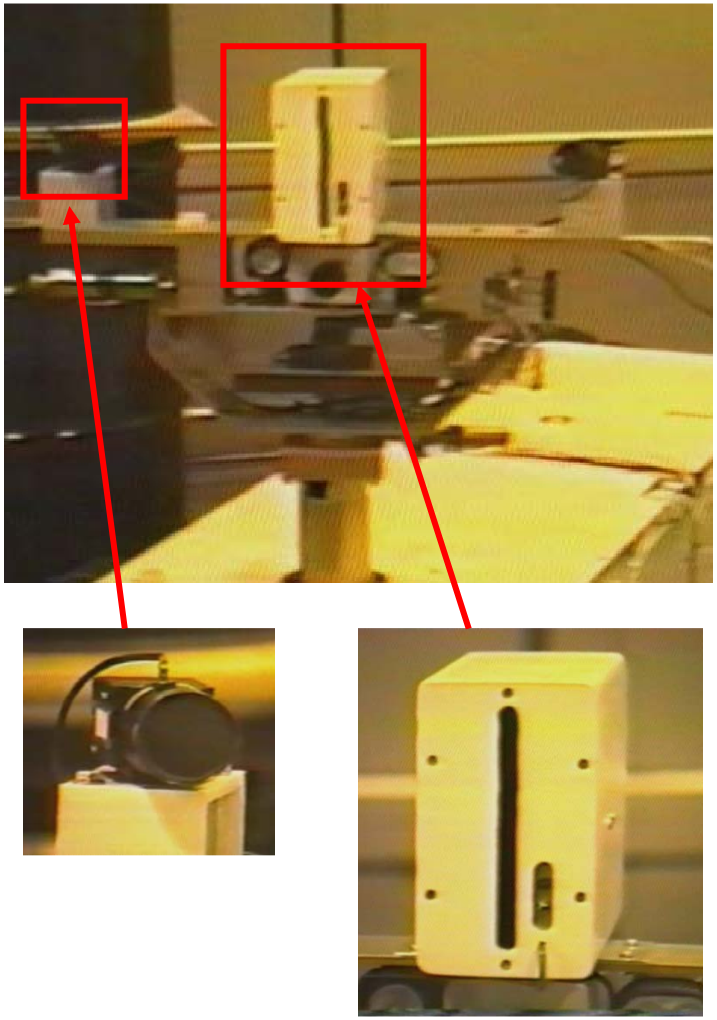
$X = Z\, x/f$

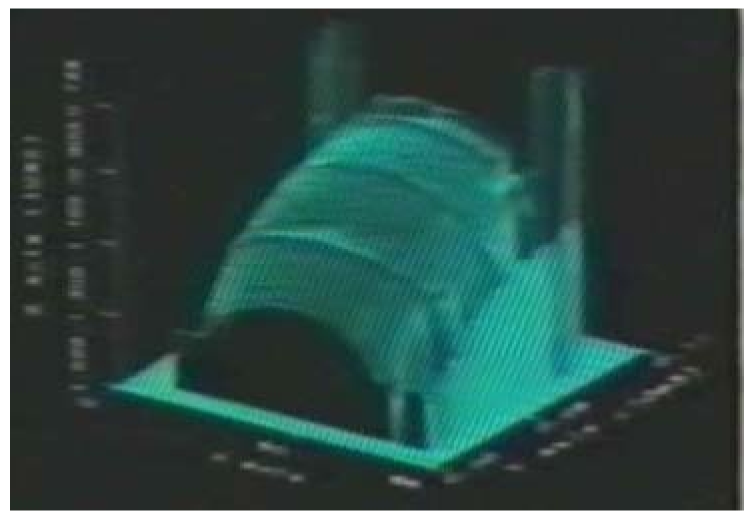
$\cot\theta = (b+Zx/f)/Z$, $\quad Z\cot\theta = b + Zx/f$

$Z(\cot\theta - x/f) = b$, $\quad Z = b/(\cot\theta - x/f)$

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{b}{f\cot\theta - x}\begin{pmatrix} x \\ y \\ f \end{pmatrix}$$

Structured light sensor to measure surfaces of drums

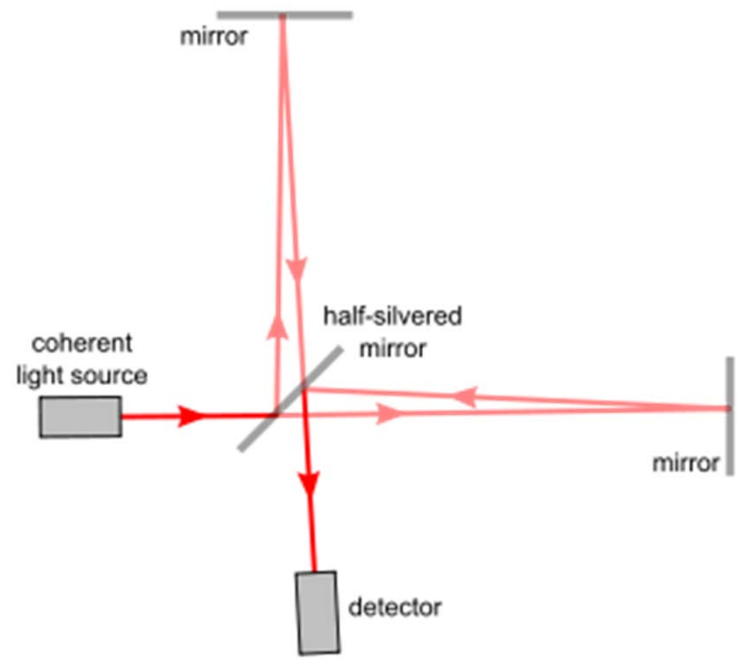




Intelligent Mobile Sensing System (IMSS) – Martin Marietta, 1991

Time of Flight Systems

- Pulse systems
  - Directly measure time of flight of a very short pulse
  - Many pulses are fired sequentially and the average response is used
  - Requires very accurate sub-nanosecond timing circuitry
- Interferometric systems
  - Measure the difference in phase between the sensed reflected beam and a reference beam
  - There is an "ambiguity interval"
- Sensors often produce an amplitude measurement as well as a range measurement



Frames of Reference

- Image frames are 2D; others are 3D
- The "pose" (position and orientation) of a 3D rigid body has 6 degrees of freedom