Image Formation

• The two parts of the image formation process

The geometry of image formation which determines where in the image plane the projection of a point in the scene will be located.

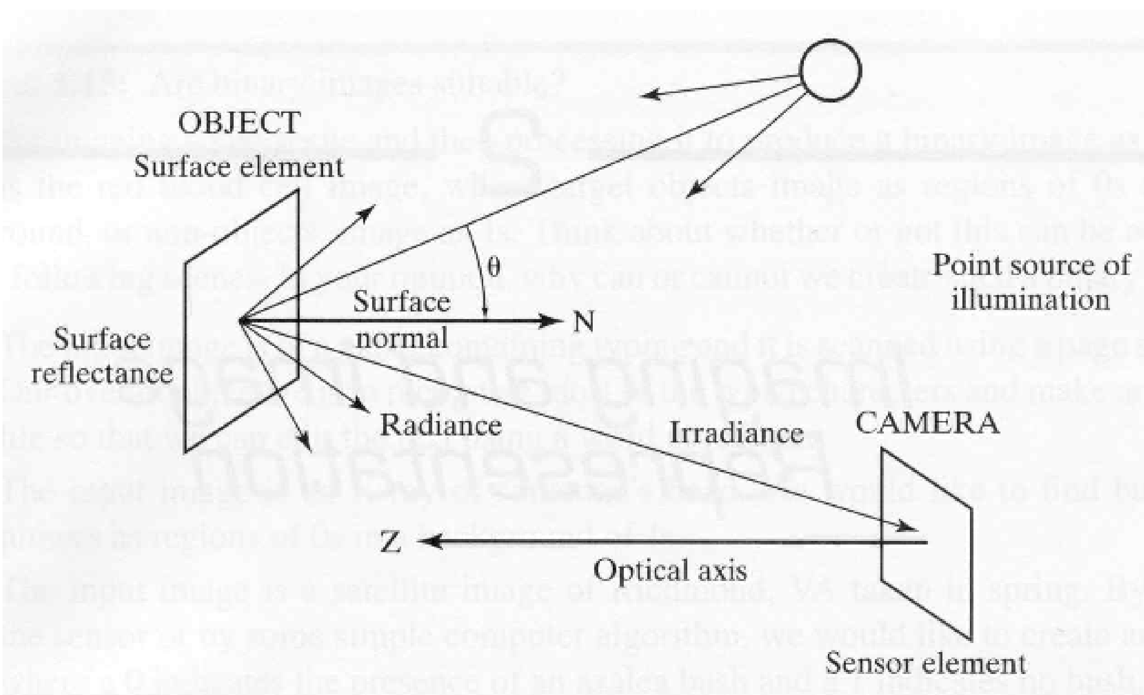
The physics of light which determines the brightness of a point in the image plane as a function of illumination and surface properties.

• A simple model

- The scene is illuminated by a single source.

- The scene reflects radiation towards the camera.

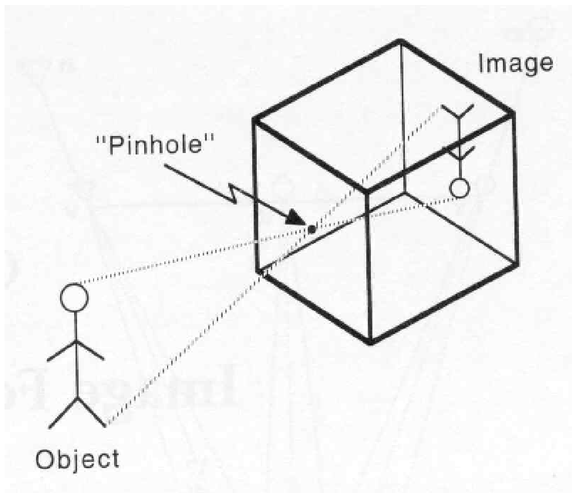
- The camera senses it via chemicals on film.



• Camera Geometry

- The simplest device to form an image of a 3D scene on a 2D surface is the "pinhole" camera.

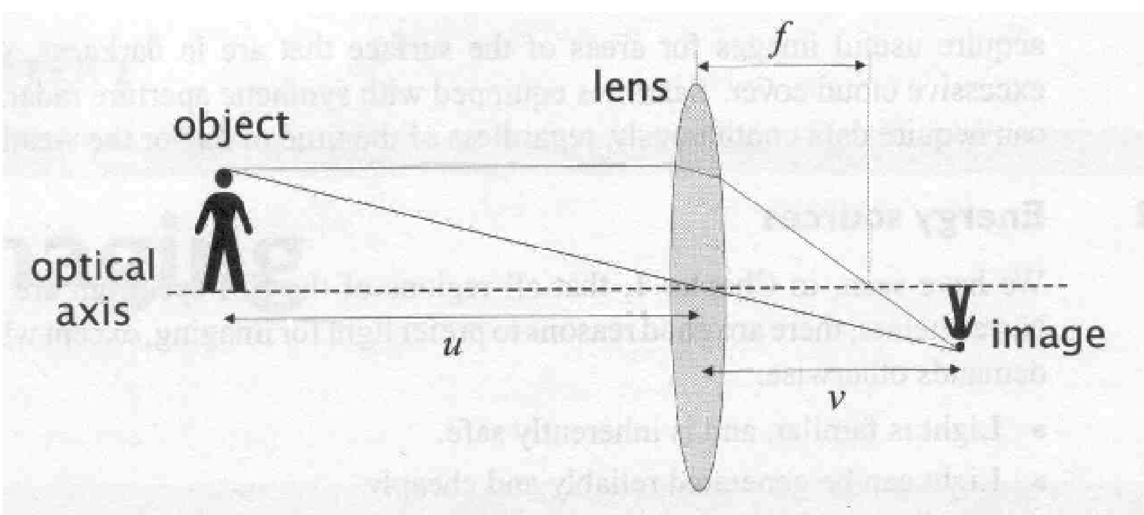
- Rays of light pass through a "pinhole" and form an inverted image of the object on the image plane.



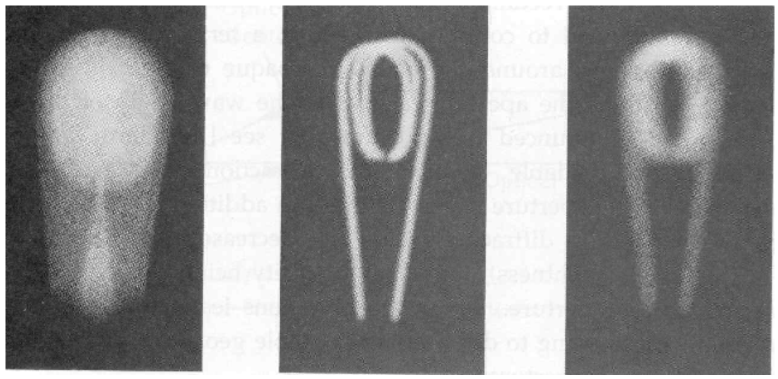
• Camera Optics

- In practice, the aperture must be larger to admit more light.

- Lens are placed in the aperture to <u>focus</u> the bundle of rays from each scene point onto the corresponding point in the image plane.



• Diffraction and Pinhole Optics



- If we use a wide pinhole, light from the source spreads across the image (i.e., not properly focused), making it blurry.

- If we narrow the pinhole, only a small amount of light is let in.

  * the image sharpness is limited by diffraction.

  * when light passes through a small aperture, it does not travel in a straight line.

  * it is scattered in many direction (this is a quantum effect).

- In general, the aim of using lens is to duplicate the pinhole geometry without resorting to undesirable small apertures.

• Human Vision

- At high light levels, pupil (aperture) is small and blurring is due to diffraction.

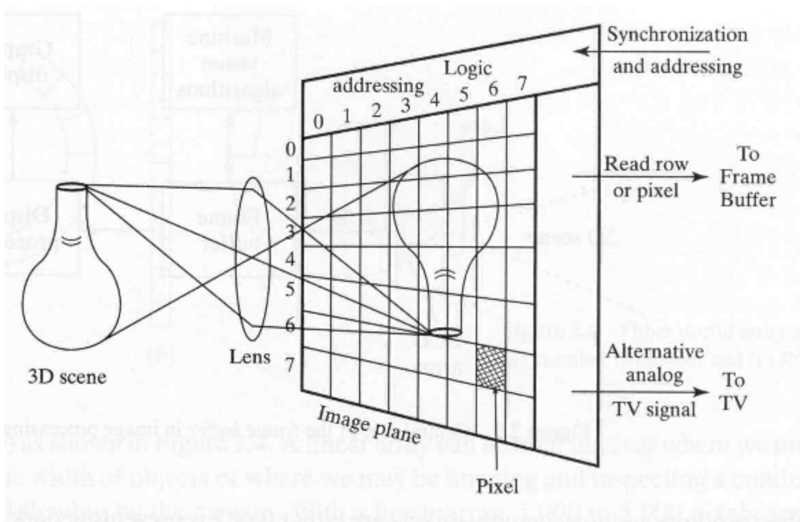
- At low light levels, pupil is open and blurring is due to lens imperfections.

• CCD Cameras

- An array of tiny solid state cells convert light energy into electrical charge.

- Manufactured on chips typically measuring about 1cm x 1cm (for a 512x512 array, each element has a real width of roughly 0.001 cm).

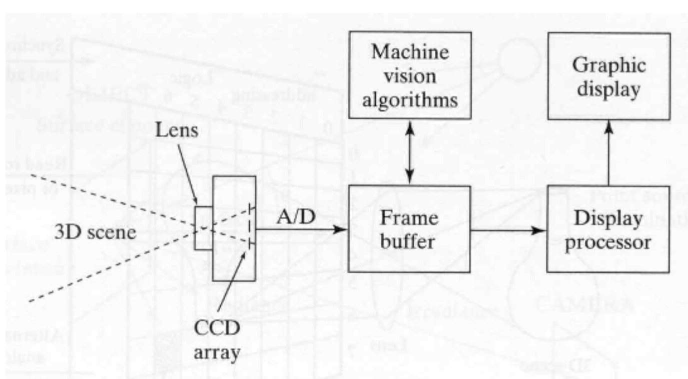
- The output of a CCD array is a continuous electric signal (*video signal*) which is generated by scanning the photo-sensors in a given order (e.g., line by line) and reading out their voltages.



• Frame grabber

- The video signal is sent to an electronic device called the frame grabber.

- The frame grabber digitizes the signal into a 2D, rectangular array N x M of integer values, stored in the *frame buffer*.



CCD array and frame buffer

- In a CCD camera, the *physical image plane* is the CCD array of nxm rectangular grid of photo-sensors.

- The *pixel image plane* (frame buffer) is an array of NxM integer values (pixels).

- The position of the same point on the image plane will be different if measured in CCD elements (x, y) or image pixels (x_{im}, y_{im}).

- In general, $n \neq N$ and $m \neq M$; assuming that the origin in both cases is the upper left corner we have:

$$x_{im} = \frac{N}{n} x \qquad\qquad y_{im} = \frac{M}{m} y$$

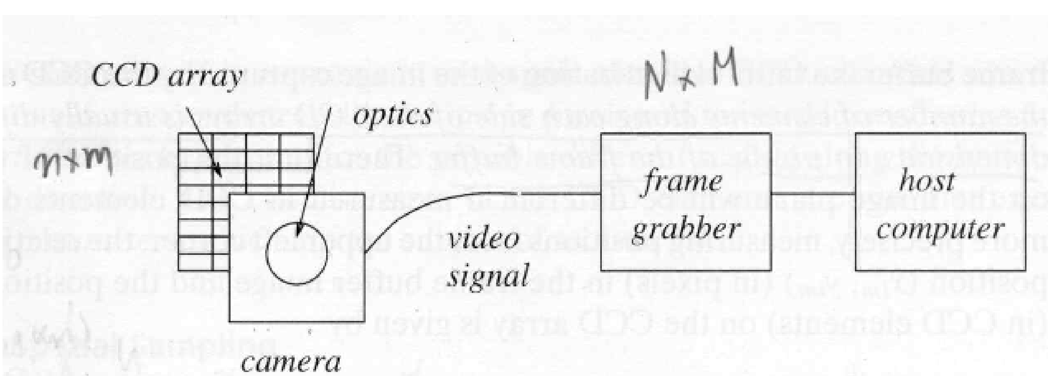
where (x_{im}, y_{im}) are the coordinates of the point in the pixel plane and (x, y) are the coordinates of the point in the CCD plane.

- In general, it is convenient to assume that the CCD elements are always in one-to-one correspondence with the image pixels.

- Units in each case:

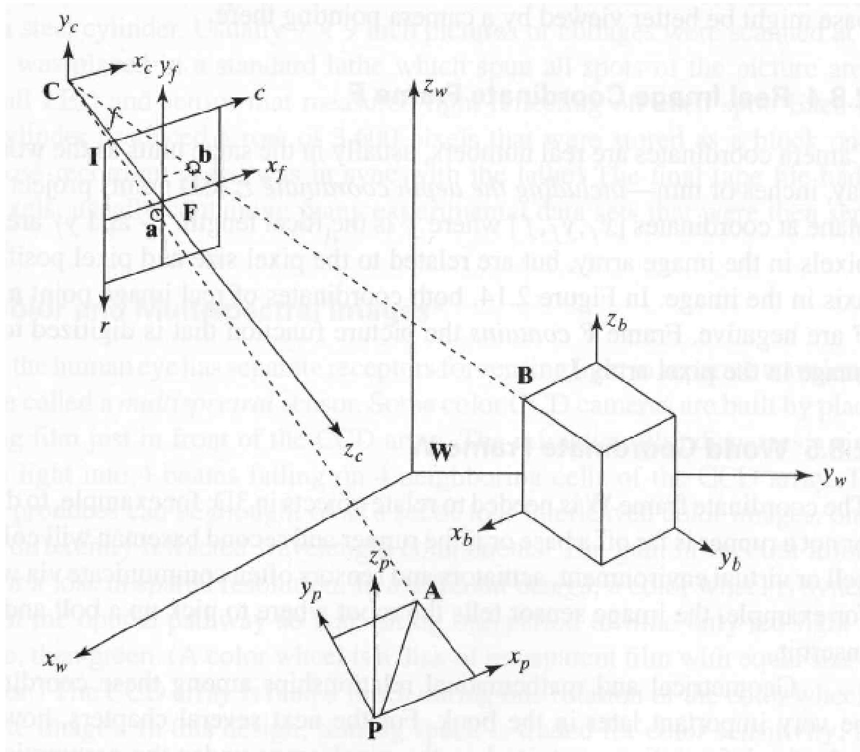
(x_{im}, y_{im}) is measured in pixels

(x, y) is measured, e.g., in millimeters.



Reference Frames

- Five reference frames are needed for general problems in 3D scene analysis.



• Object Coordinate Frame

- This is a 3D coordinate system: x_b, y_b, z_b

- It is used to model ideal objects in both computer graphics and computer vision.

- It is needed to inspect an object (e.g., to check if a particular hole is in proper position relative to other holes)

- The coordinates of 3D point B, e.g., relative to the object reference frame are $(x_b, 0, z_b)$

- Object coordinates do not change regardless how the object is placed in the scene.

$$\textit{Notation: } (X_o, Y_o, Z_o)^T$$

• World Coordinate Frame

- This is a 3D coordinate system: x_w, y_w, z_w

- The scene consists of object models that have been placed (rotated and translated) into the scene, yielding object coordinates in the world coordinate system.

- It is needed to relate objects in 3D (e.g., the image sensor tells the robot where to to pick up ta bolt and in which hole to insert it).

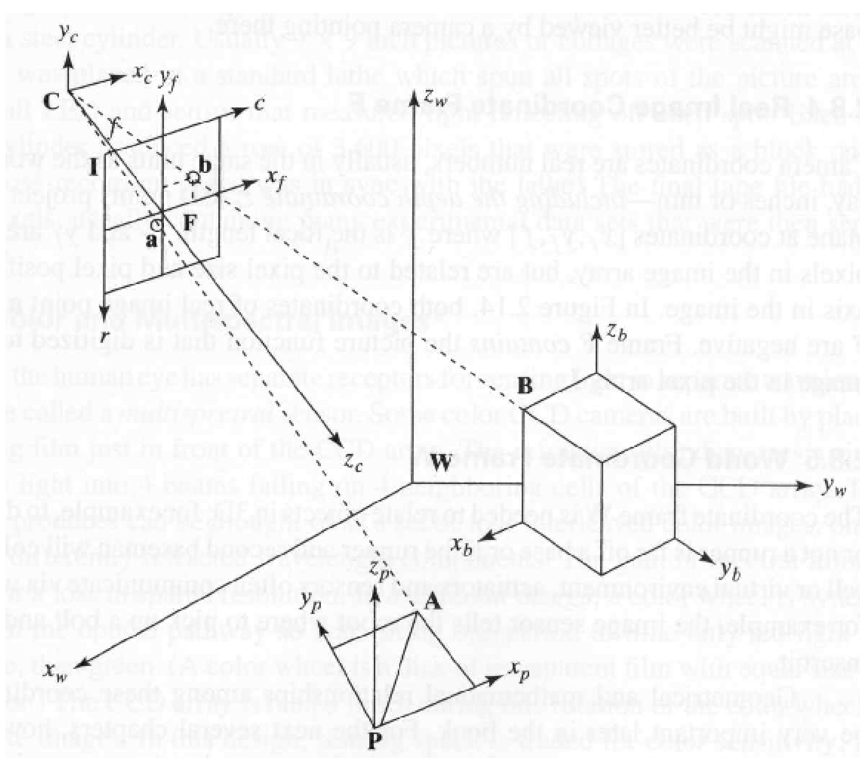
$$\textit{Notation: } (X_w, Y_w, Z_w)^T$$

• Camera Coordinate Frame

- This is a 3D coordinate system (x_c, y_c, z_c axes)

- Its purpose is to represent objects with respect to the location of the camera.

$$\textit{Notation: } (X_c, Y_c, Z_c)^T$$



• **Image Plane Coordinate Frame** (CCD plane)

- This is a 2D coordinate system (x_f, y_f axes)

- Describes the coordinates of 3D points projected on the image plane.

- The projection of A, e.g., is point a whose both coordinates are negative.

$$\textit{Notation: } (x, y)^T$$

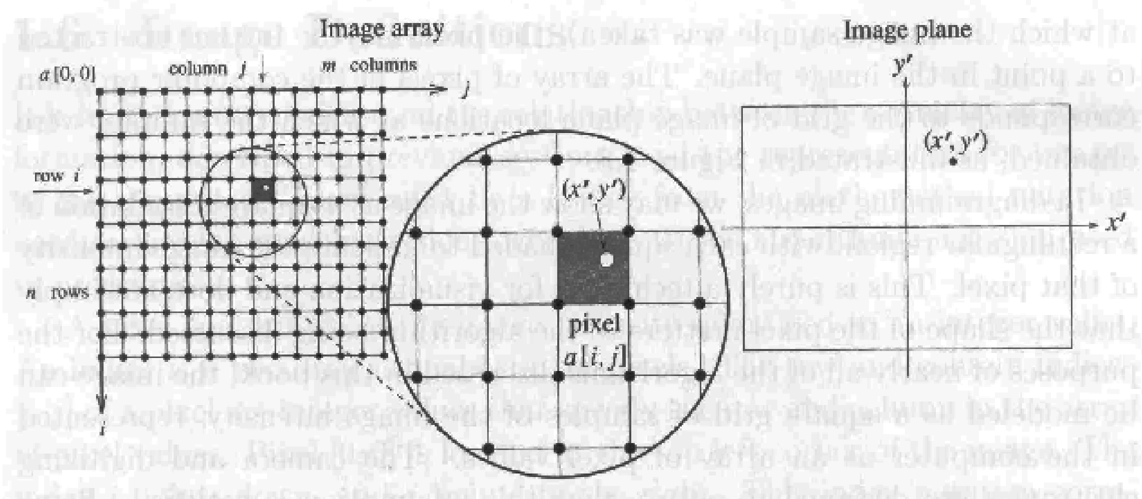
• **Pixel Coordinate Frame**

- This is a 2D coordinate system (r, c axes)

- Each pixel in this frame has an integer pixel coordinates.

- Point A, e.g., gets projected to image point (a_r, a_c) where a_r and a_c are integer row and column.

$$\textit{Notation: } (x_{im}, y_{im})^T$$



• **Transformations between frames**