Machine Vision

Lecture Set – 02

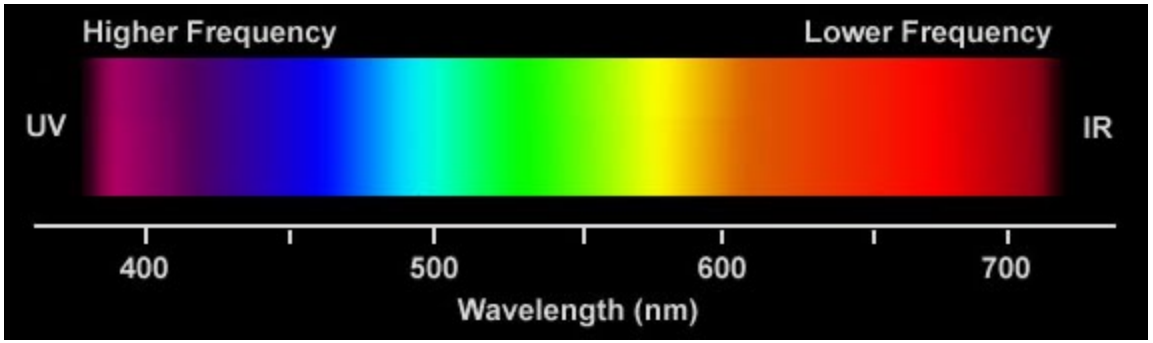
Digital Image Fundamentals

Huei-Yung Lin

Robot Vision Lab

Light

- A form of electromagnetic radiation
  - Speed = wavelength x frequency
- Speed of light in a vacuum = $3x10^8$ m/sec
- Visible spectrum
  - 400 nm - 700 nm
  - Longer wavelengths correspond to redder colors and shorter wavelengths to bluer color
- EM radiation comes in discrete packets called photons

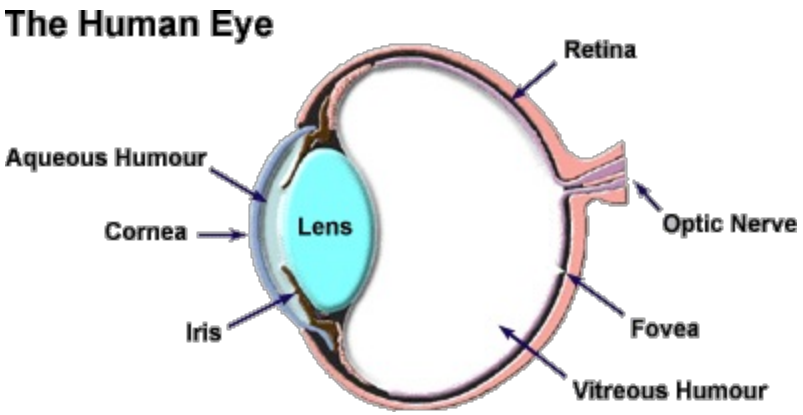


Light and Matter

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

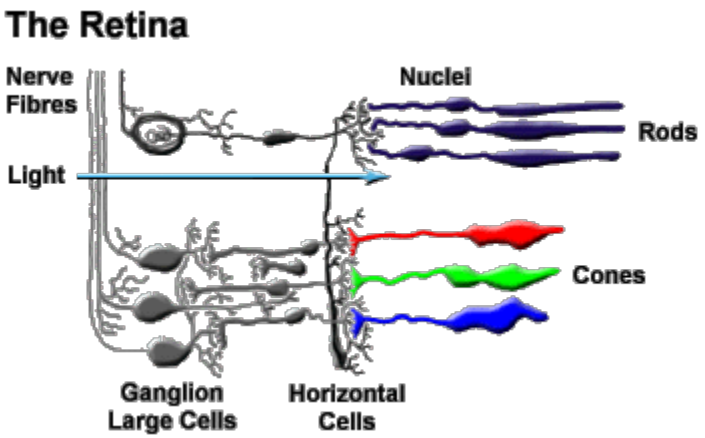
The Camera of the Mind

- Components of the human eye
  - Pupil
  - Lens
  - Retina
  - Fovea
  - Blind spot
  - Iris



The Retina

- There are two types of photosensitive cells in the retina, rods and cones
  - Cones come in three flavors which exhibit different sensitivities to different wavelengths of light, red, green and blue
  - Rods are not sensitive to variations in wavelength but they are more sensitive than cones and can pick up much dimmer light
- The fovea is populated entirely by cones



More Cells

- Ganglion Cells
  - The photosensitive cells transmit their information to ganglion cells which in turn transmit information to the brain via the optic nerve
- Numbers of cells
  - There are approximately 6 million cone cells, 120 million rods and 1 million optic nerve fibers

Image Formation

- Light
- Reflectance
- Image capture
  - Camera
  - Lens
  - Sensor
  - Projection models
  - Camera system parameters

The Anatomy of a Modern Camera

- Lens
- Shutter (exposure time)
- Diaphragm (aperture)
- Focusing Control

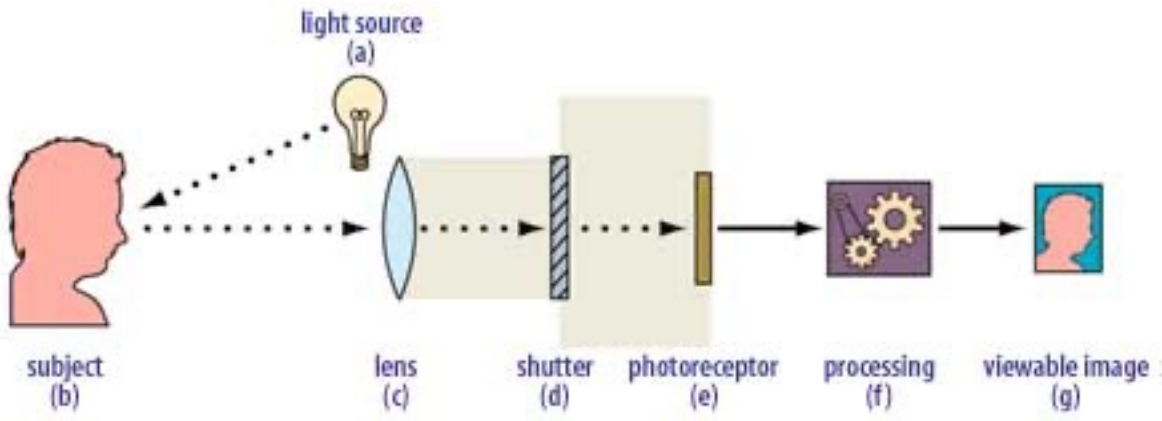


Image Geometry

- Model for the projection of scene point to image point
  - Center of projection (COP)
  - Focal length
  - Line of sight

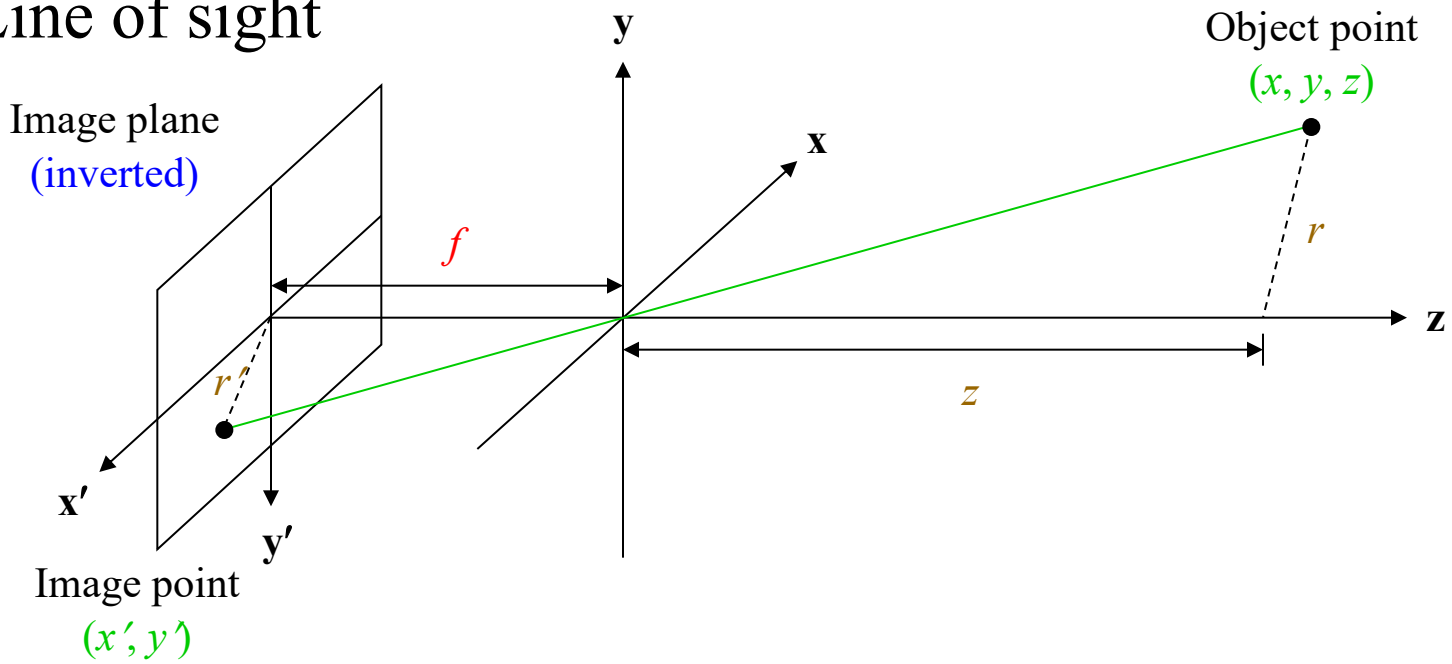
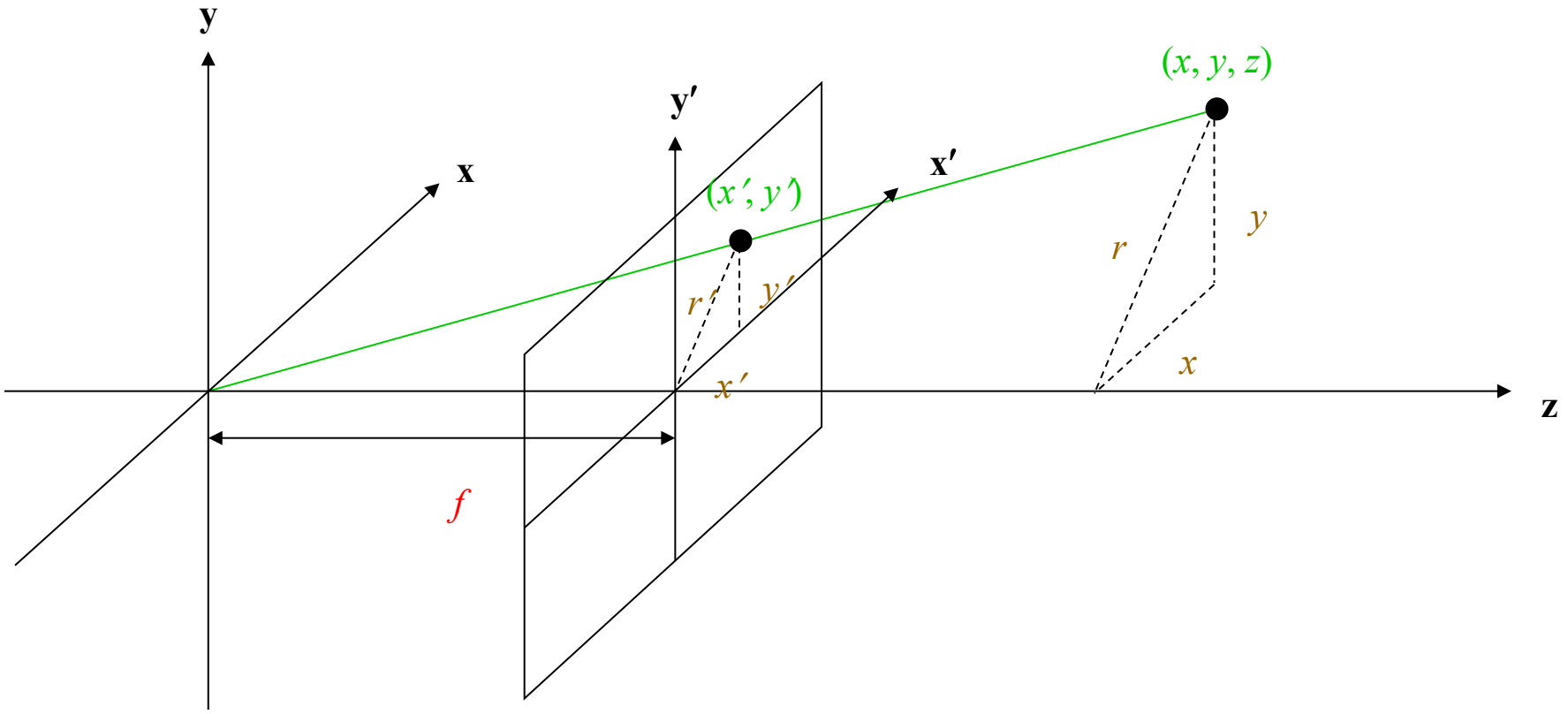


Image Geometry

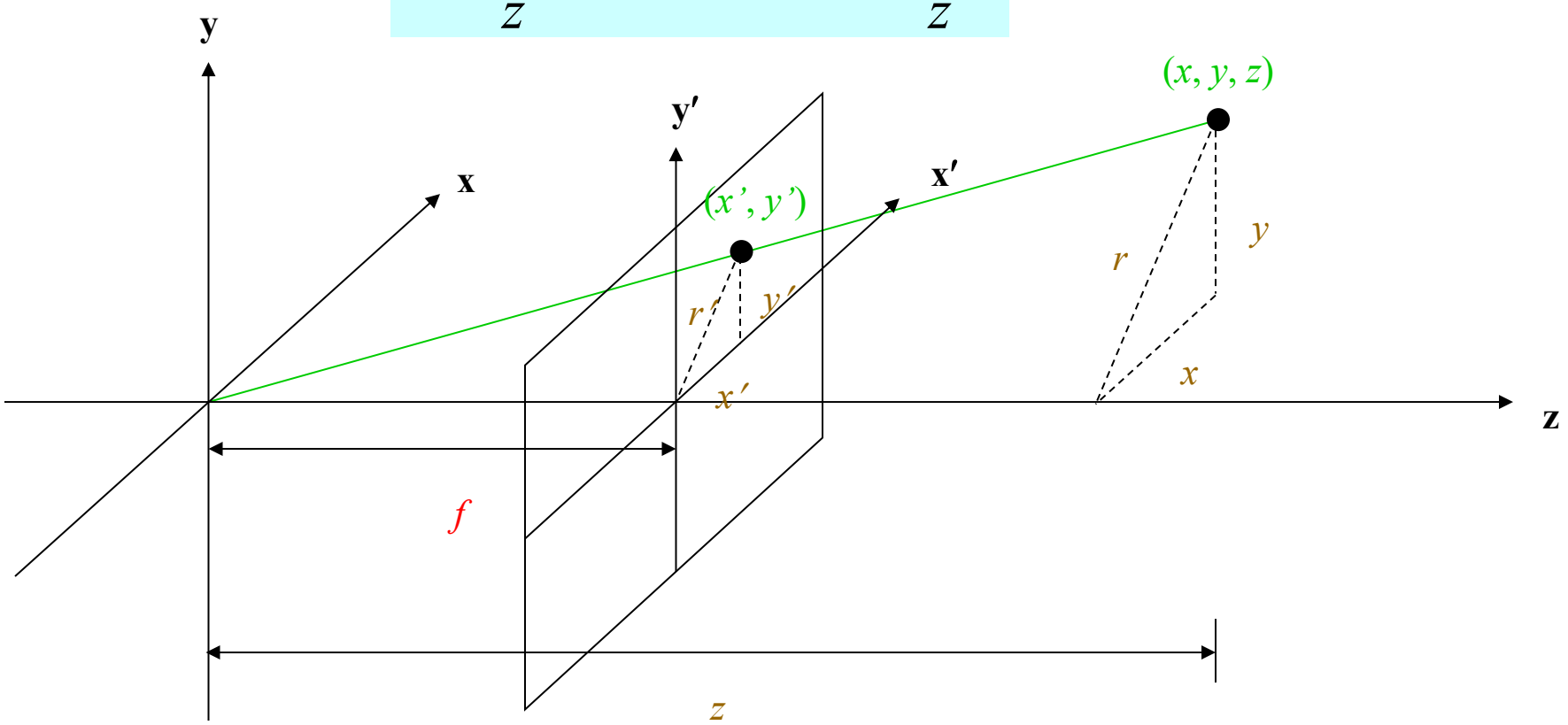
- Avoid inversion by placing image plane in front of center of projection
- Invert **x′** and **y′** also



Perspective Projection

- From similar triangles: $\frac{f}{z} = \frac{r'}{r}$ and $\frac{x'}{x} = \frac{y'}{y} = \frac{r'}{r}$
- We have

$$x' = \frac{f}{z}x \quad \text{and} \quad y' = \frac{f}{z}y$$



3/2/2023

- Teaching Assistants
  - 林鈺琴, idayulin93@gmail.com
  - 陳泳慈, a40123anny@gmail.com
- Office Hours
  - TBA
  - Lab 1421

Coordinate Systems

- There are four coordinate systems
  - World coordinate system
    - Also called absolute or object coordinates
    - (x_w, y_w, z_w) or (x_a, y_a, z_a)
  - Object coordinate system
    - Also called model coordinates
  - Camera coordinate system
    - (x_c, y_c, z_c)
  - Image coordinate system
    - (x', y')
- To use the image geometry, coordinates have to be transformed first!

Digital Images

- Digital image
  - 2-D array (matrix) of numbers
  - Numbers can be light intensity, distance (range), etc.
- Intensity image
  - Photograph-like images encoding light intensities, acquired by cameras
  - Measure the amount of light impinging on a photosensitive device
- Range image
  - Encoding shape and distance acquired by special sensors like sonar or laser scanners
  - Estimate directly the 3-D structure of the viewed scene through variety of techniques

Intensity Image Formation (1/2)

- Optical parameters
  - Characterize the sensor's optics
    - Lens type, focus length, field of view, angular aperture

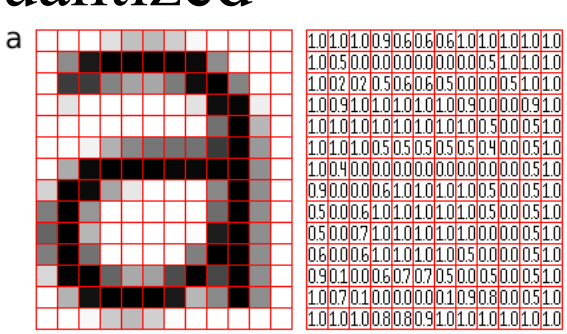
- Photometric parameters
  - Appear in models of the light energy reaching the sensor after being reflected from the objects in the scene
    - Type, intensity, direction of illumination
    - Reflectance properties of the viewed surface
    - Effects of the sensor's structure on the amount of light reaching the photoreceptors

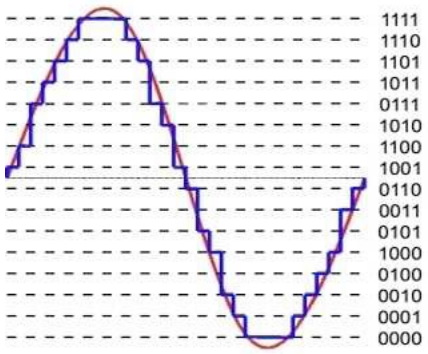
Intensity Image Formation (2/2)

- Geometric parameters
  - Determine the image position on which a 3-D point is projected
    - Type of projections
    - Position and orientation of camera in space
    - Perspective distortion introduced by the imaging process

Digital Image Representations

- A digital image is an array of numbers indicating the image irradiance at various points on the image plane
- Image intensities are spatially sampled
- Intensity values are quantized
  - 8-bits
  - 10-bit, 12-bits, etc.

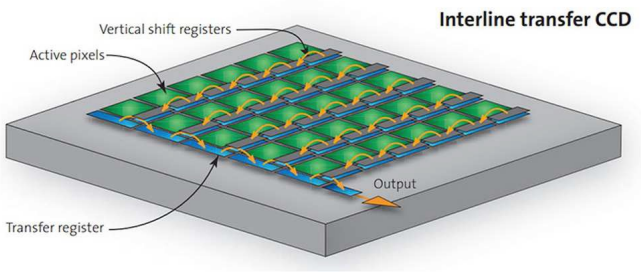


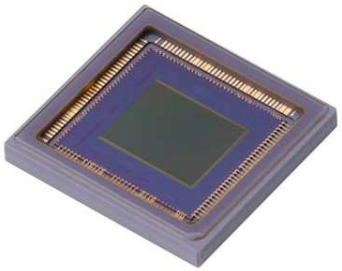


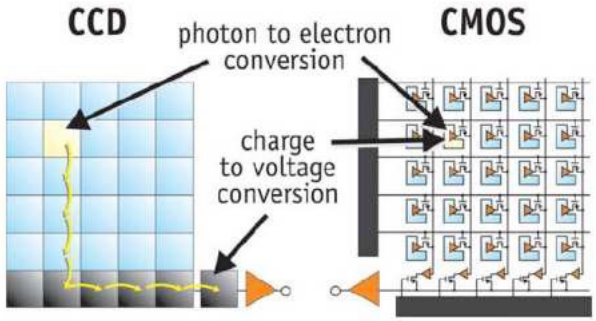
- Acquisition: CCD arrays
- Storage: Usually in computer memory
- Display: Computer hardware (video boards) and monitors

Image Sensors

- Images are formed by the interaction of the incident image irradiance with light sensitive elements on the image plane
- Light sensitive elements
  - Film
  - CCD (Charge Coupled Device)
  - CMOS Imaging element





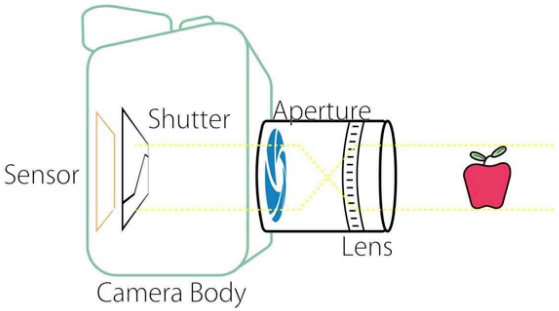


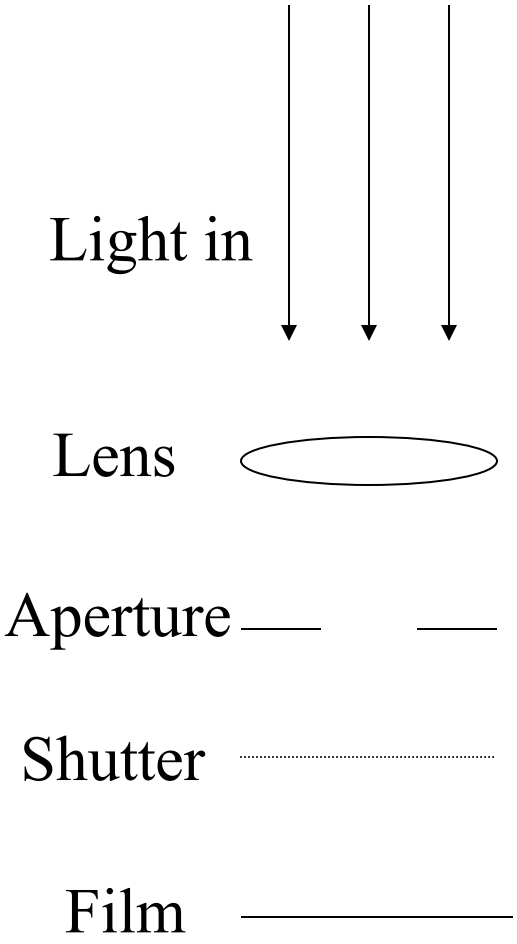
Film

- Film *samples* (records) the intensity of light that strikes each point
- What is intensity?
  - The scientific term for that roughly corresponds to brightness
  - It can be physically measured and there are many different units, such as lumens
- Real film is not perfect:
  - It has a finite dynamic range:
    - It cannot simultaneously record very dark and very bright regions - this is a big issue in photography
  - It has finite resolution:
    - If you blow it up large enough, you can see grains - this is rarely an issue in photography

Cameras

- A camera is a device for mediating the way light strikes film
- Lens lets light in while maintaining focus
- Aperture controls proportion of the light that gets to the film
- Shutter controls how long light is allowed to get to the film





Images as Samples

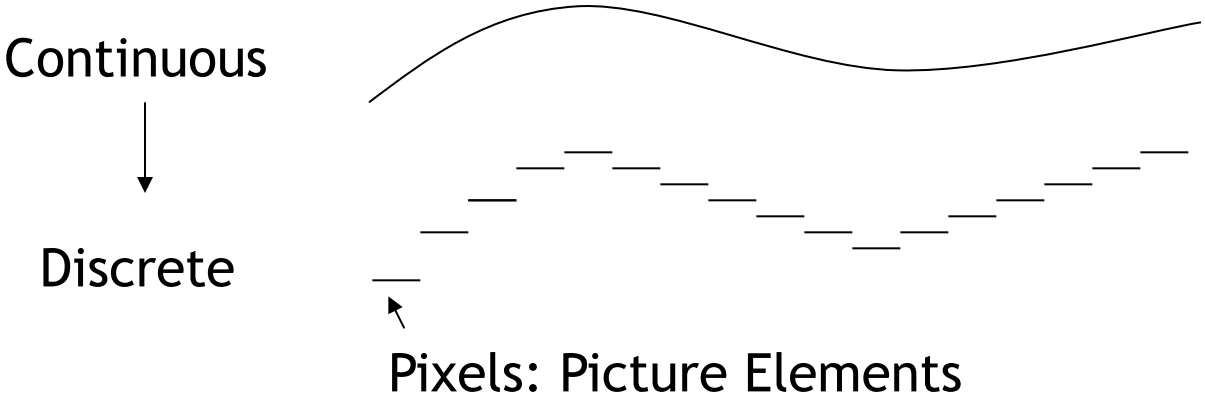
- A photograph is a sample of the light that fell onto the film
  - It's a very large set of samples, one for each point on the film
- The camera controls precisely what is sampled
  - Which period of time is sampled
  - Which region of space is sampled (which part of the light field)
  - Which region of the electromagnetic spectrum is sampled
  - Which range of intensity is sampled most accurately
- The idea of image as sample is central to many aspects of computer graphics

More on Film

- Spatial continuity:
  - In the real world, light tends to change smoothly over space
  - Film captures this smoothness quite well, with its high resolution
- Intensity continuity:
  - The real world contains a continuous range of intensities, from bright to dark
  - Film can capture a sub-range very well, but not outside the range
- Temporal continuity:
  - In the real world, light tends to vary smoothly over time
  - Movie film captures a discrete set of images over time

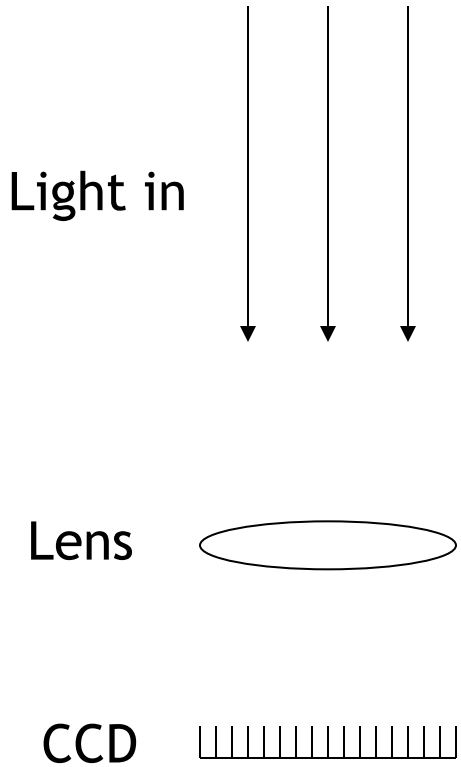
Digital Images

- Computers work with discrete pieces of information
- How do we digitize a continuous image?
  - Break the continuous space into small areas, pixels
  - Use a single value for each pixel - the pixel value no longer continuous in space or intensity



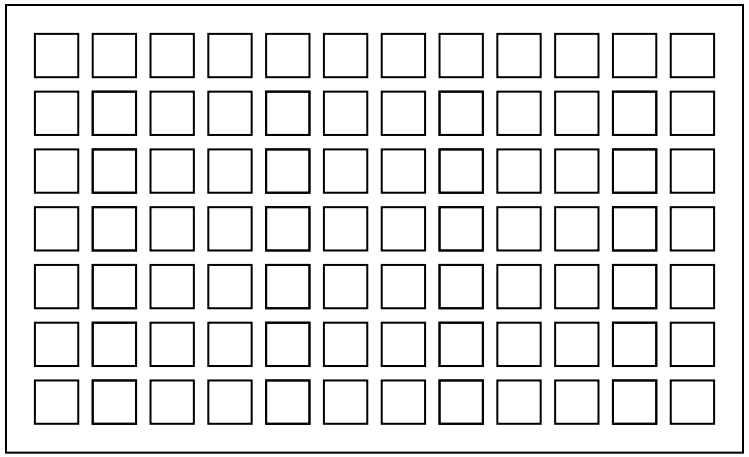
Digital Cameras

- CCD stores a charge each time a photon hits it
  - "Bins" have discrete area, one per pixel
  - Spatially discrete
- Camera "reads" the charges out of the bins at some frequency
- Convert charges to discrete value
  - Discrete in intensity
- Store values in memory - the image
- Still have issues of motion blur, depth of field, dynamic range, etc



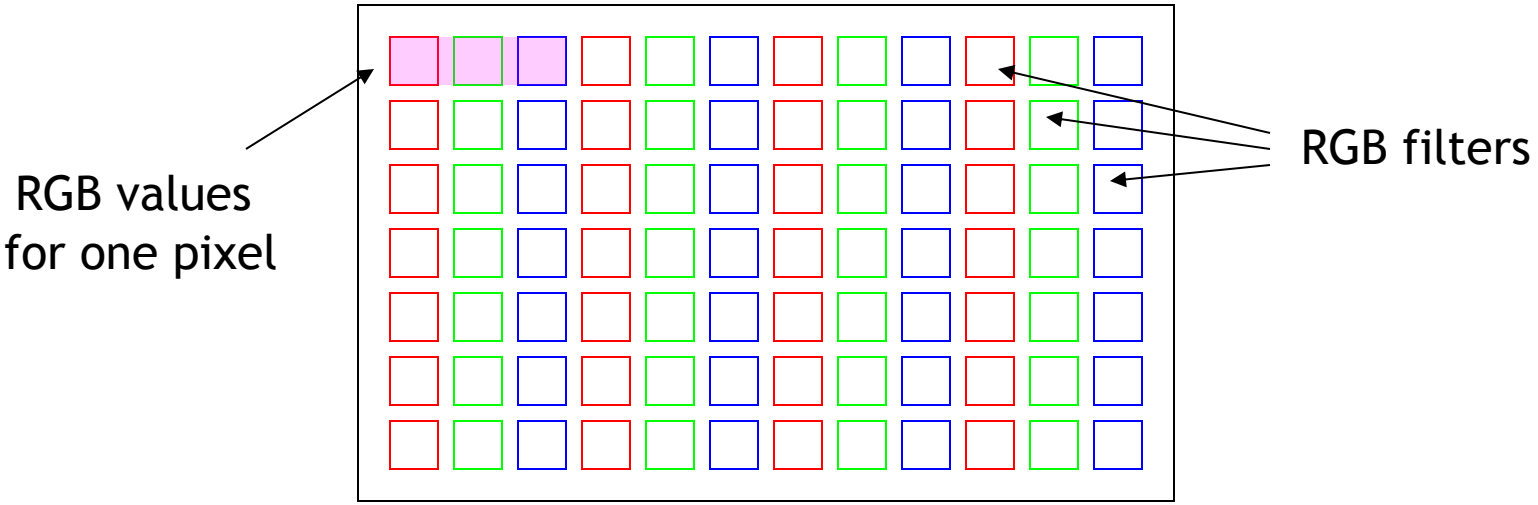
Digital Black & White Cameras

- CCDs consist of a (usually) 2-D array of photo-sensitive cells, each corresponding to a pixel
- Light falling onto a cell's surface causes the generation of a voltage roughly proportional to intensity of incident light
- Voltage reading of each cell is converted to a digital signal within a CCD-specific range (usually an 8 or 10-bit number)



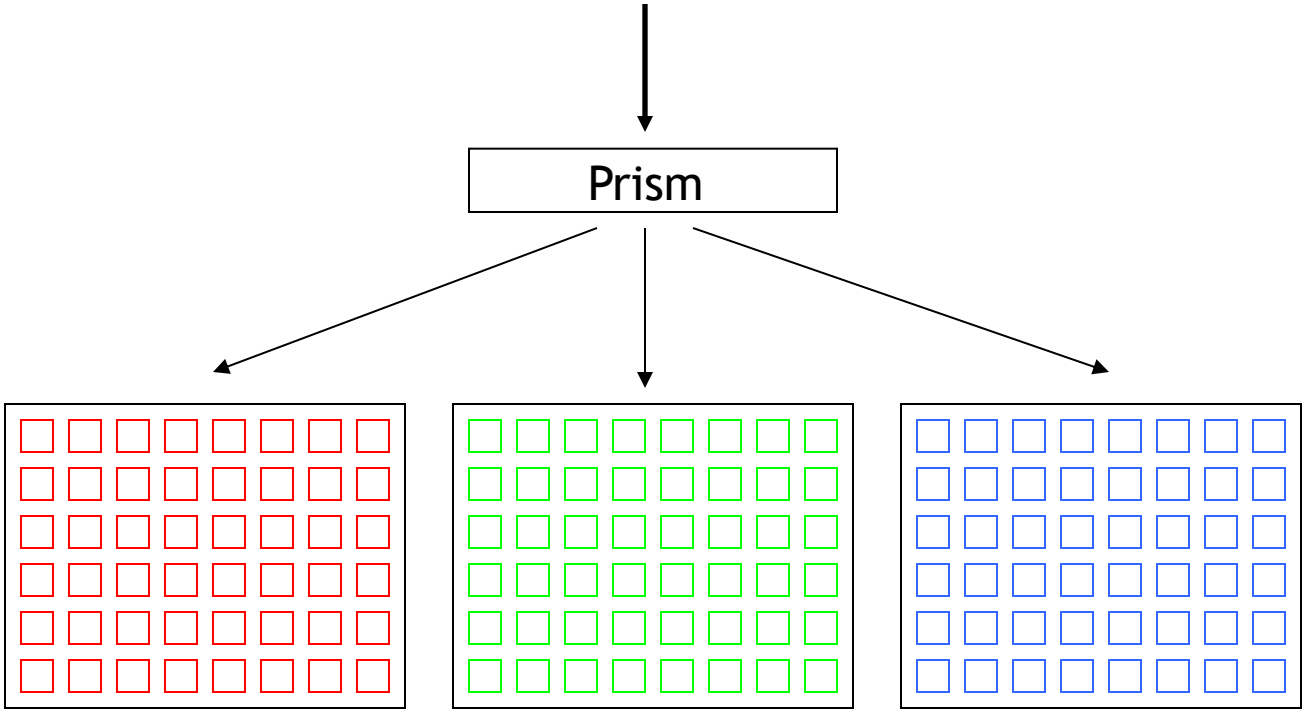
Digital Color Cameras

- 3-filter "mask" is placed on top of the CCD array, each filter permitting only red, green, or blue to go through
- Interpolation algorithms assign 3-band colors to every pixel

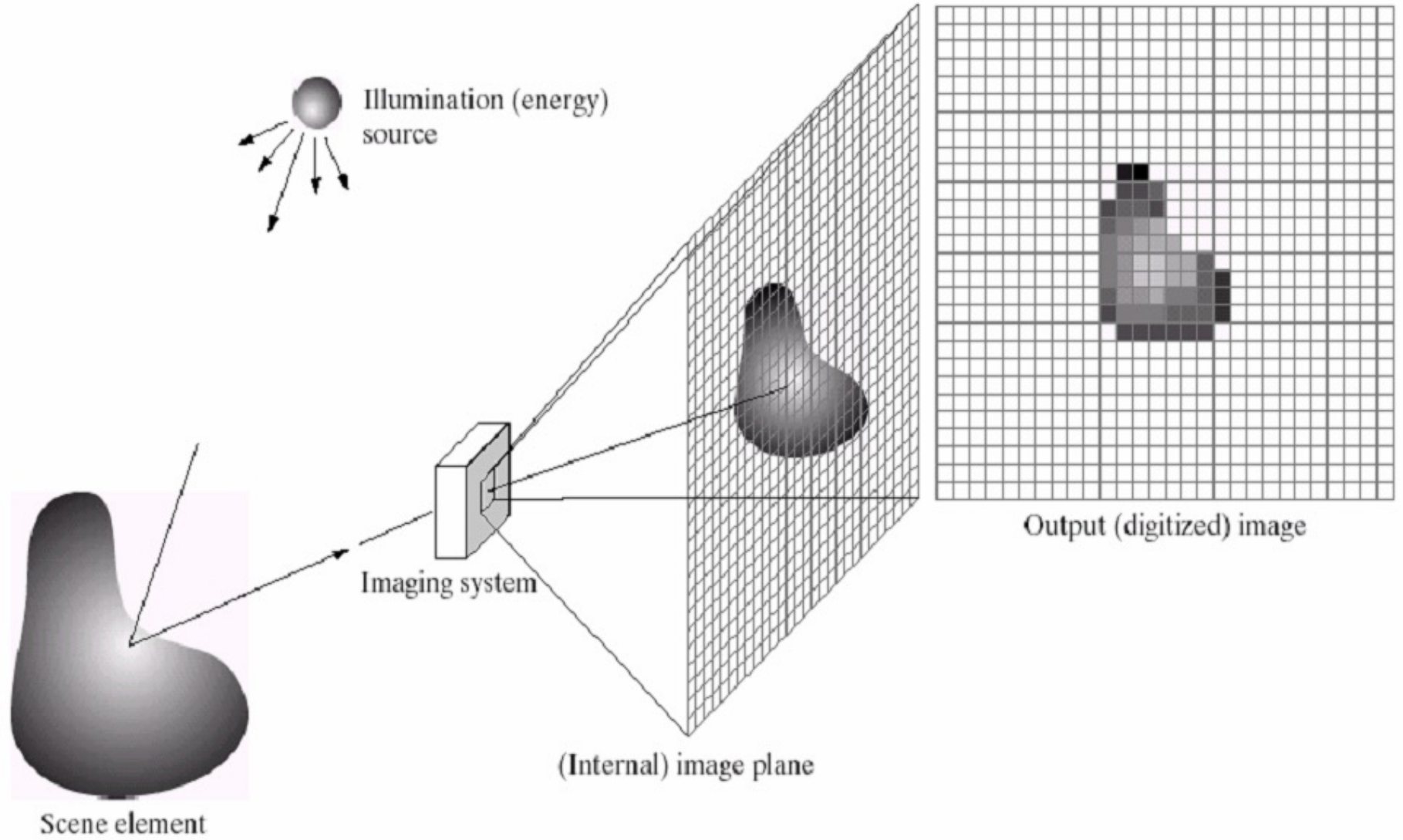


Digital Color Cameras

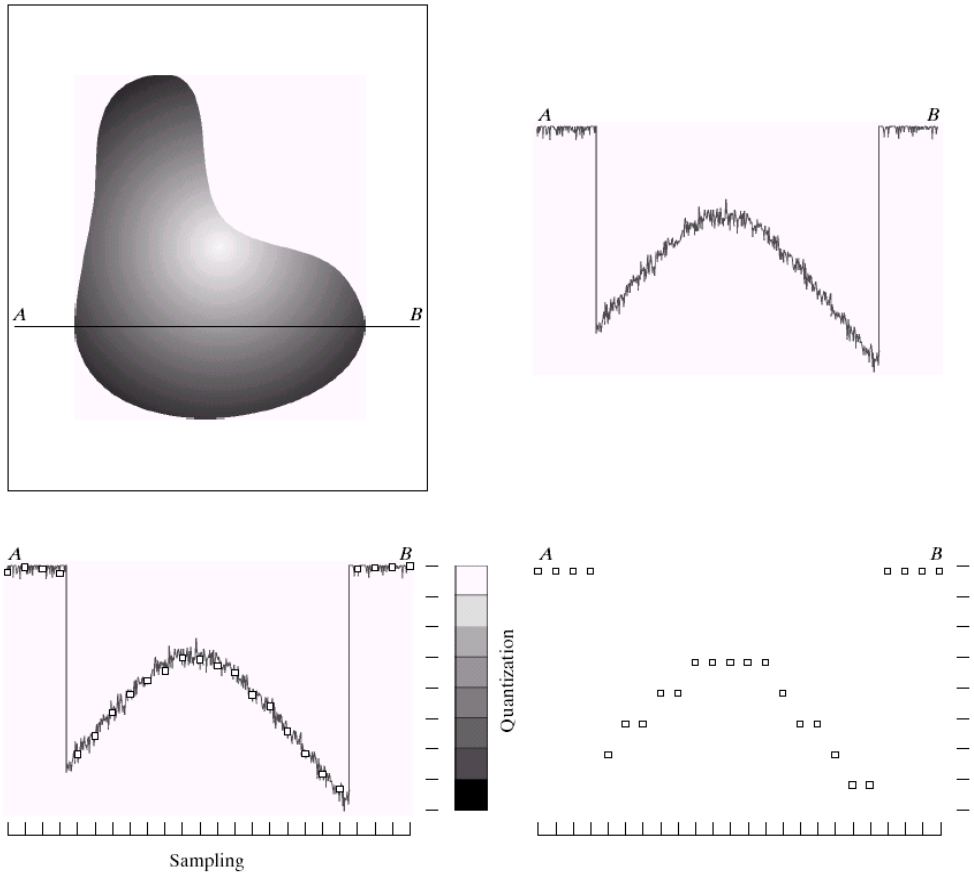
- Three-CCD color cameras precisely aligned
- Each array covered by single color filter (R, G, B)



Digital Image Acquisition Process



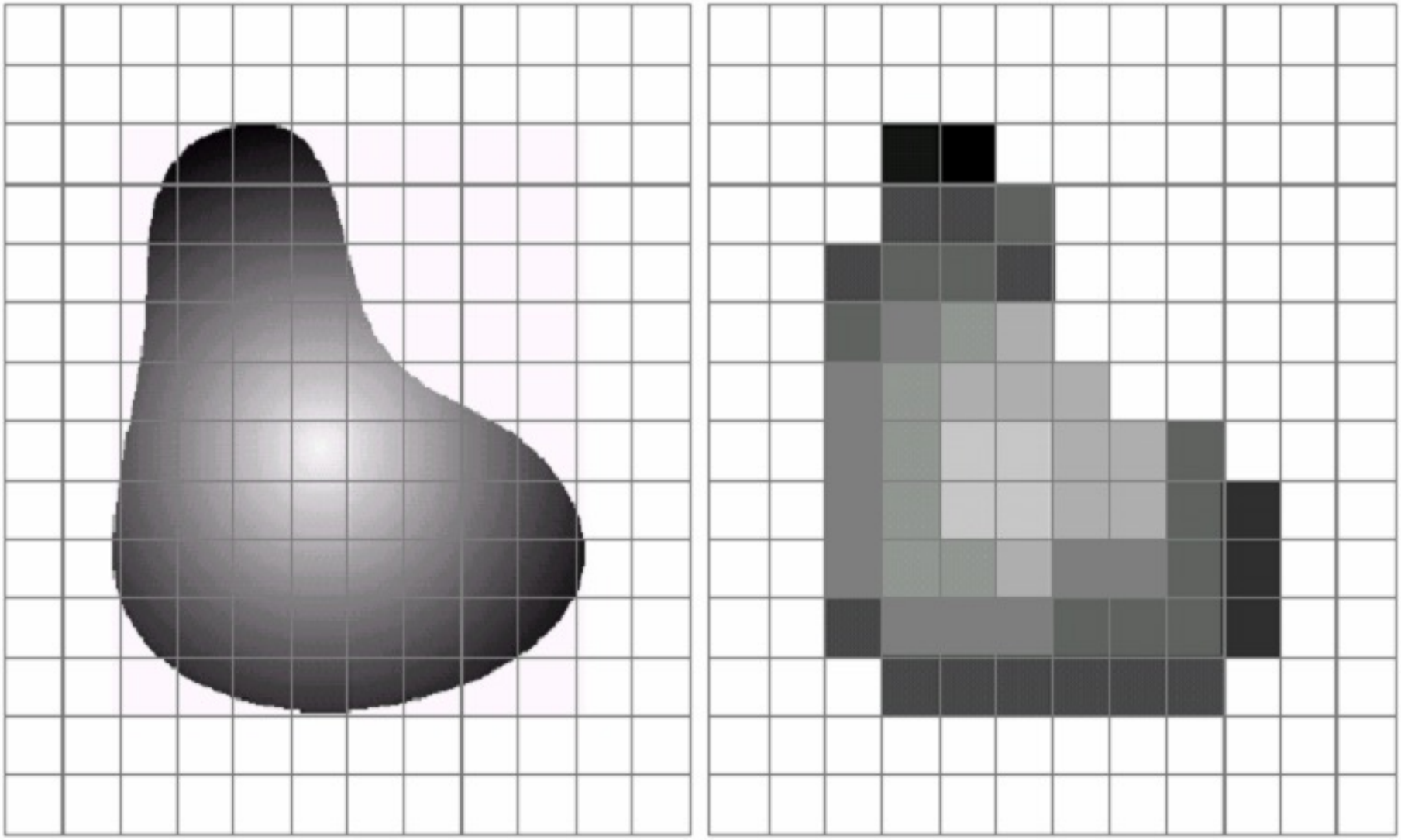
Generating a Digital Image



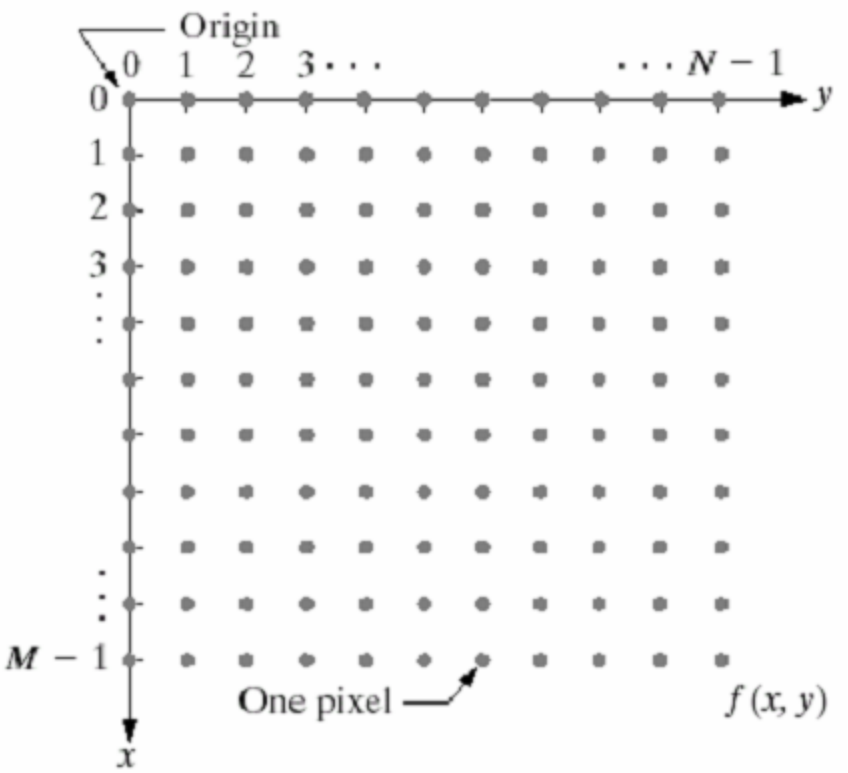
a b
c d

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Image Sampling and Quantization



Coordinate Conversion



Digital Image

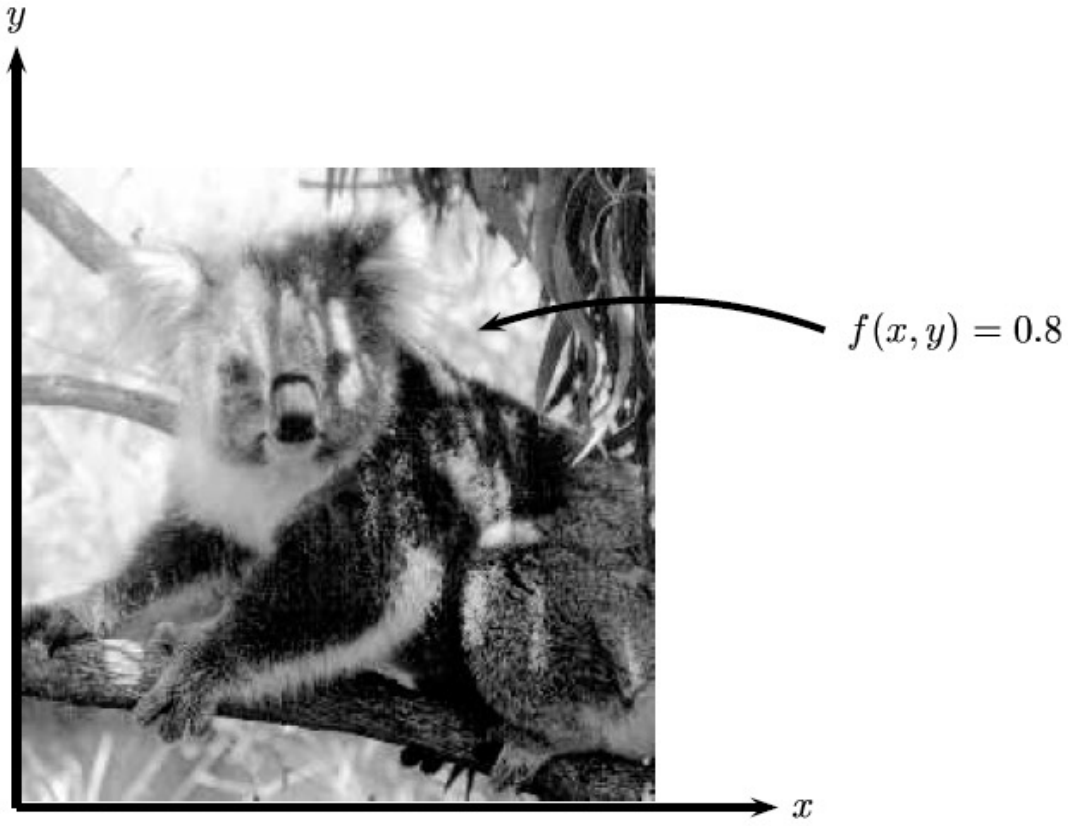


Figure 1.13: An image as a function

Image As Two Dimensional Array

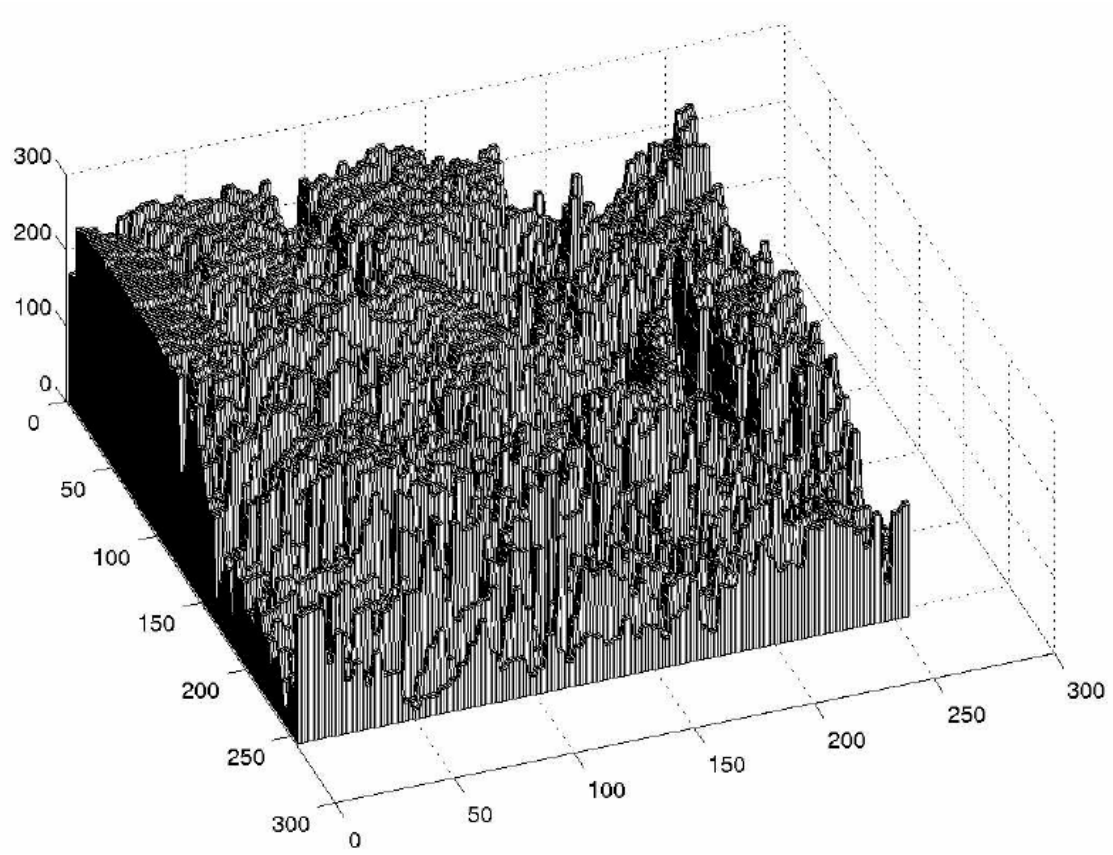


Figure 1.14: The image of figure 1.13 plotted as a function of two variables

Image As Two Dimensional Array

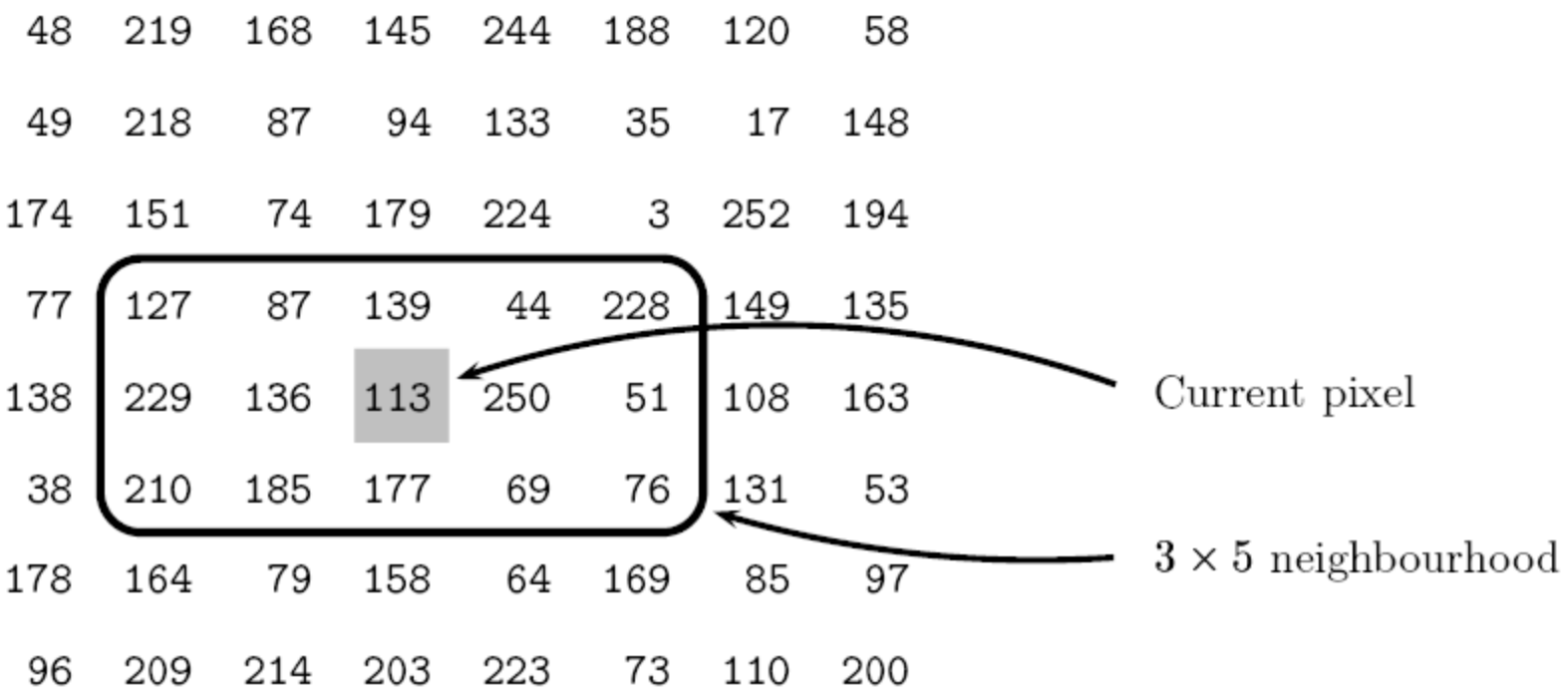
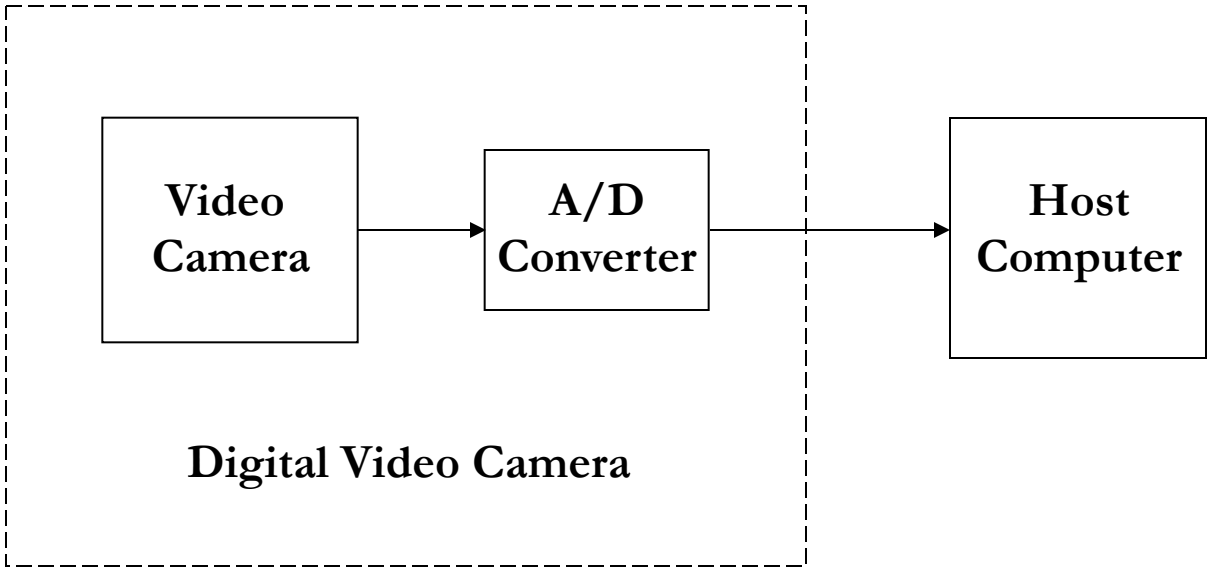


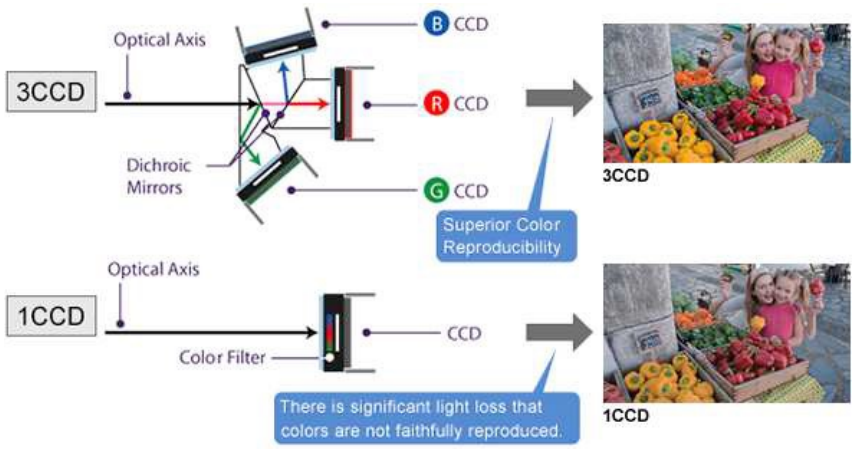
Figure 1.15: Pixels, with a neighbourhood

Digital Imaging Systems



Video Images

- Images are taken sequentially by opening and closing the shutter 30 times per second
  - Temporal aliasing
  - Motion blur



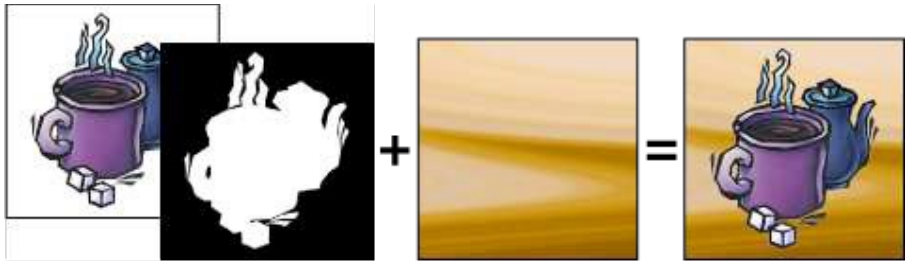
- Sensing Color
  - In a 3 CCD video camera the light path is split into three components which are passed through colored light filters and then imaged
  - As a result – a color image contains three channels of information: red, green and blue image intensities

3/7/2023

- Homework #1 description will be given this Thursday
- Due 3/23

Image Storage

- Images are stored as continuous string of bytes
  - As a two-dimensional array of pixels
  - Way of storage can affect performance of algorithms
- Color images can be stored using either 24 or 8 bits
  - When we have more colors, we can use a lookup table
- 4-bytes per pixel is more efficient than 3, because it aligns better with integer boundaries in memory
  - 4th pixel is often the alpha component between 0 and 1
  - Final color is [aR aG aB]

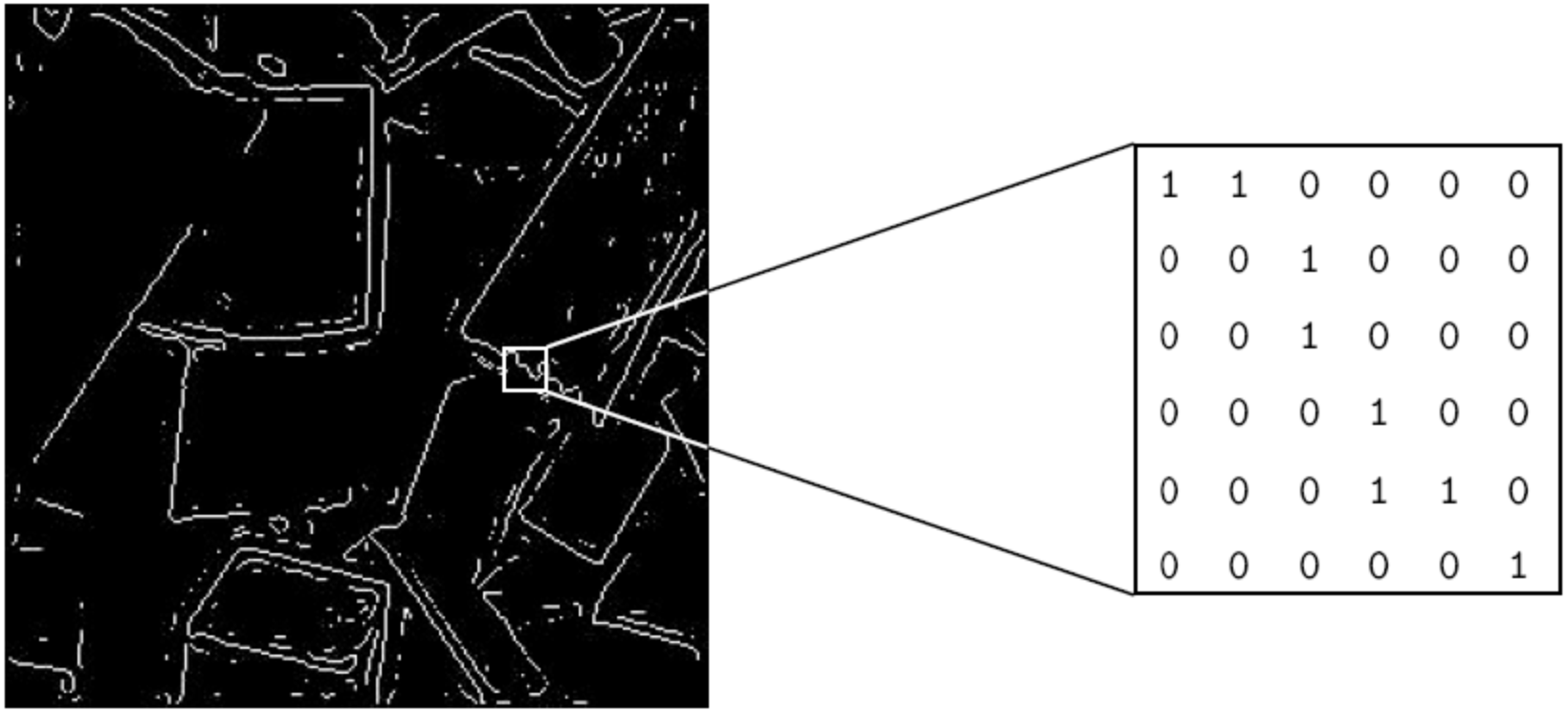


Types of Digital Images

- Binary image – 1 bit/pixel
- Grayscale image – 8 bits/pixel
- True color or RGB image – 24 bits/pixel
- Indexed image – 8 bits/pixel

Binary Image

- Efficient in terms of storage
- Document processing, handwriting, fingerprint



Grayscale Image

- The range is usually a power of 2
- 256 levels are sufficient for most applications

230	229	232	234	235	232	148
237	236	236	234	233	234	152
255	255	255	251	230	236	161
99	90	67	37	94	247	130
222	152	255	129	129	246	132
154	199	255	150	189	241	147
216	132	162	163	170	239	122

True Color Image

- RGB image (red-green-blue)
- 16,777,216 different possible colors

49	55	56	57	52	53
58	60	60	58	55	57
58	58	54	53	55	56
83	78	72	69	68	69
88	91	91	84	83	82
69	76	83	78	76	75
61	69	73	78	76	76

Red

64	76	82	79	78	78
93	93	91	91	86	86
88	82	88	90	88	89
125	119	113	108	111	110
137	136	132	128	126	120
105	108	114	114	118	113
96	103	112	108	111	107

Green

66	80	77	80	87	77
81	93	96	99	86	85
83	83	91	94	92	88
135	128	126	112	107	106
141	129	129	117	115	101
95	99	109	108	112	109
84	93	107	101	105	102

Blue

Indexed Image

- The image has an associated color map, which is simply a list of all the colors used in that image
- GIF, PNG formats, etc.

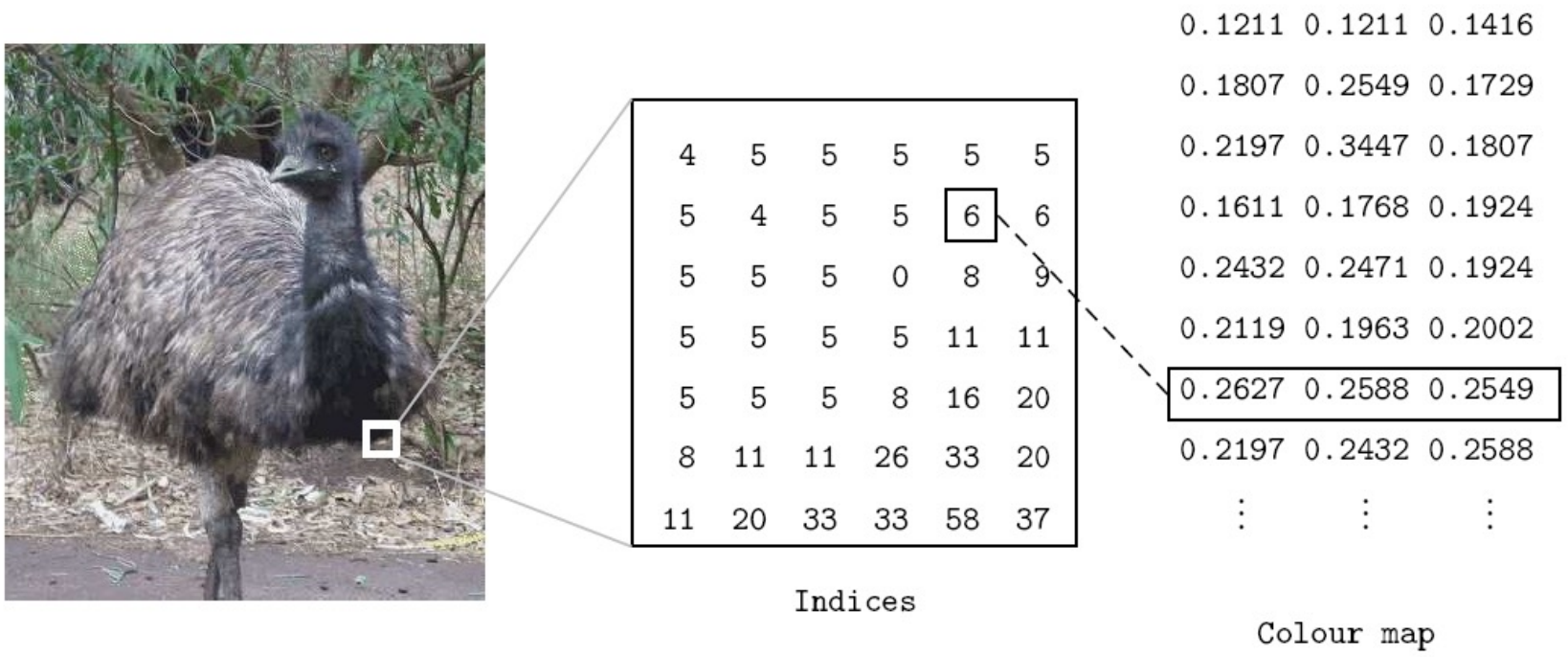


Image File Size

- The number of bits (or bytes) used in the image

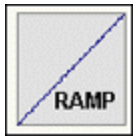
- 512×512 binary image
  - 512×512×1 = 262,144 bits = 32,768 bytes ≈ 0.033 Mb
- 512×512 grayscale image
  - 512×512×1 = 262,144 bytes ≈ 0.262 Mb
- 512×512 RGB image
  - 512×512×3 = 786,432 bytes ≈ 0.786 Mb

Gamma Correction

- The light output of a monitor's phosphors is generally not proportional to the voltage applied
- Brightness function $I=aV^g$
  - Macintosh - 1.8
  - SGI - 1.5
  - SUN and PC - 2.5
- Image colors look "distorted"
- Gamma correction color lookup table

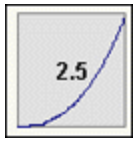
Gamma Correction

Sample Input to Monitor



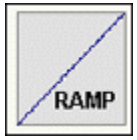
Graph of Input

Output from Monitor



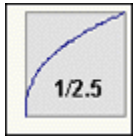
Graph of Output $I=V^{2.5}$

Sample Input



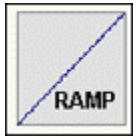
Graph of Input

Gamma Corrected Input



Graph of Correction $I=V^{1/2.5}$

Monitor Output



Graph of Output

Pixel Representations

- Resolution
- Quantization
- Efficiency in access

3/9/2023

- Homework #2 description will be given later at about 11:50
- Due in two weeks

Image Resolution

- An image with different spatial resolution
- (How to get this? Zoom-in? Take average?)

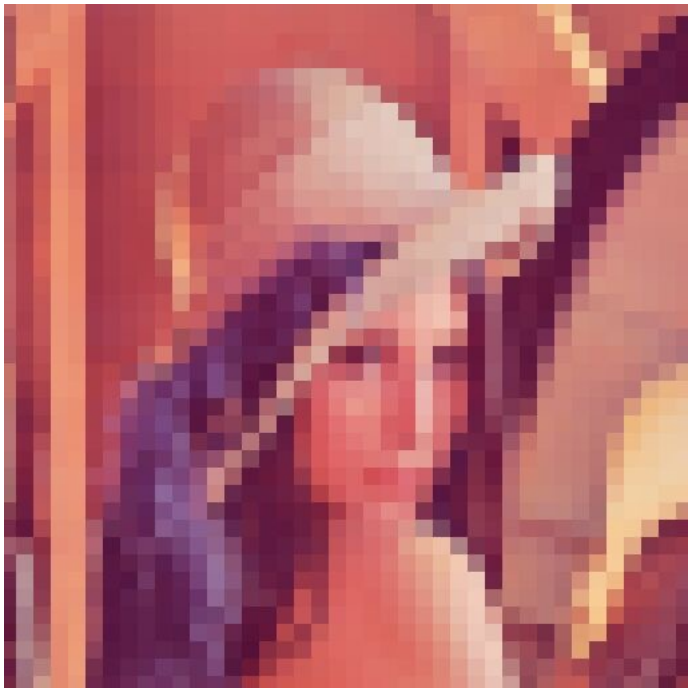


Image Quantization

- Number of bits per pixel (or color) determines number of possible values
- Dithering?

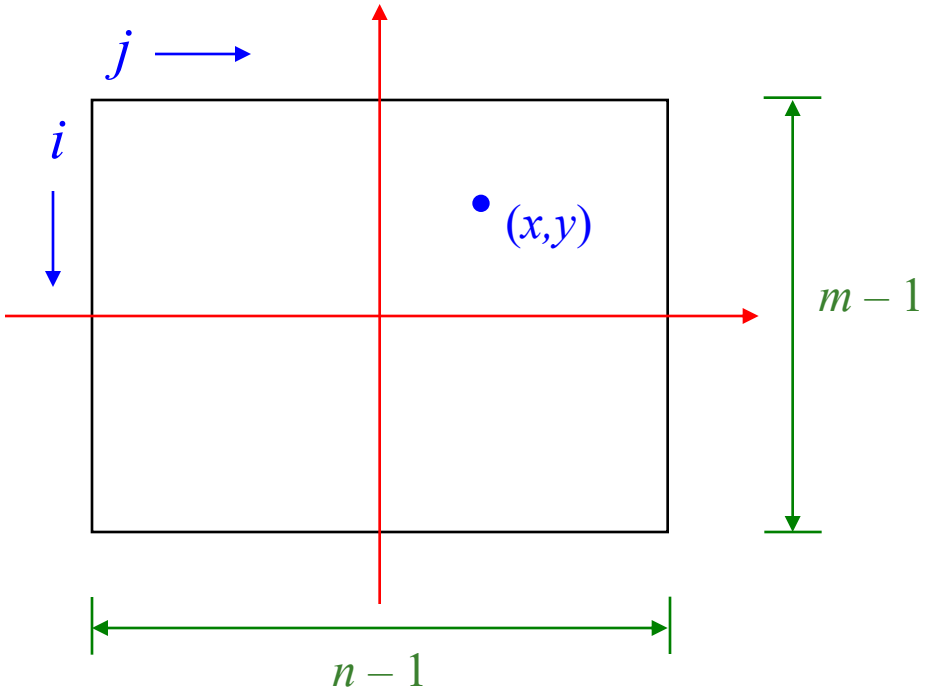
Image Definitions

- The relationship between geometry of image formation (x,y) and representation for digital image $[i,j]$
  - A pixel in an image is represented by a 2-D index $[i,j]$ starting from top left corner
  - A point in image plane is represented by (x,y) with origin at the center of the image
- Conversion formula:

$$x = j - \frac{n-1}{2}$$

$$y = -\left(i - \frac{m-1}{2}\right)$$

where m, n are the number of row and column of the digital image



Levels of Computation

- Point level : $f_B[i,j] = O_{\text{point}}\{f_A[i,j]\}$
  - Thresholding, grayscale, etc.
  - Can be speeded up by "lookup table"
- Local level : $f_B[i,j] = O_{\text{local}}\{f_A[i_k,j_l] : [i_k,j_l] \in N[i,j]\}$
  - Smoothing, edge detection, etc.
  - Can be speeded up by "parallel architecture"
- Global level : $P = O_{\text{global}}\{f[i,j]\}$
  - Intensity histogram, Fourier transform, etc.
  - Should be avoided if possible
- Object level – Size, shape, etc.

Reading

- Chapter 1 of Jain's book
  - Sections 1.1 – 1.6