ECG782: MULTIDIMENSIONAL DIGITAL SIGNAL PROCESSING
DIGITAL IMAGE FUNDAMENTALS
OUTLINE

- Image Formation and Models
- Pixels
- Pixel Processing
- Color
Use mathematical models to describe signals
- A function depending on some variable with a physical meaning

1D signal
- E.g. speech, audio, voltage, current
  - Dependent on “time”

2D signal
- E.g. image
  - Dependent on spatial coordinates in a plane

3D signal
- E.g. volume in space, video

M-D signal
- E.g. ???
Incoming light energy is focused and collected onto an image plane.
IMAGE FORMATION MODEL

- Imaging takes the 3D world and projects it onto a 2D image
- Simple model for the process is called the pinhole camera

\[ X = [x, y, z]^T \]
- represents point in world 3D space

\[ u = [u, v]^T \]
- represents a 2D point on image plane

\[ f \] - focal length of camera

World-image relationship
- \( u = \frac{xf}{z} \)
- \( v = \frac{yf}{z} \)
PERSPECTIVE PROJECTION

- Pinhole camera causes perspective distortion
  - Loss of information from perspective projection
  - The transform is not one-to-one
    - A line in space gets mapped to the same point
    - Need depth information to resolve ambiguity

- Orthographic (parallel) projection
  - Linear approximation with $f \to \infty$
  - This is how far away objects $z \to \infty$ are mapped onto image plane
IMAGE REPRESENTATION

- Multiple equivalent representations
- Image
- Surface
- Matrix
Image representation

- Image $f(x, y)$ is a 2D function
  - $f$ – amplitude, gray level, or brightness
  - $(x, y)$ – spatial coordinates
  - Conceptually, $(x, y)$ are continuous but are discrete in practice

- In general, the function can be vector-valued
  - E.g. color images represented by (red, green, blue)
  - $f(x, y) = [r, g, b]^T$

- The image function can be M-dimensional
  - E.g. computed tomography (CT) images are 3D
    - $f(x, y, z)$ represents x-ray absorption at point $(x, y, z)$
Think of an image as a function, $f$, that maps from $\mathbb{R}^2$ to $\mathbb{R}$
- $0 < f(x, y) < \infty$ is the intensity at a point $(x, y)$
- In reality, an image is defined over a rectangle with a finite range of values
  - $f: [a, b] \times [c, d] \to [0,1]$
- Computationally, $[0,1]$ range is convenient but usually we have an 8-bit quantized representation
  - $0 < f(x, y) < 255$
- Color image is just three separate functions pasted together
  - $f(x, y) = [r(x, y); g(x, y); b(x, y)]$
Images are usually represented by matrices
- $M \times N$ dimension

Be aware that images can have different origin definitions
- Bottom left - typical Cartesian coordinates
- Upper left – typical image definition (matrix or table notation)
- Matlab uses $(1,1)$ for origin not $(0,0)$

Index an element either by
- $(x, y)$
- $(row, col)$
MATRICE NOTATION

- Mathematical
- Notation starts with $f(0,0)$

- Matlab
- Notation starts with $I(1,1)$
  - No zero indexing
  - Swapped axis

$M - 1 \quad N - 1 \quad x \quad y$

$1 \quad N \quad x \quad y$

$(3,4) \rightarrow I(4,3)$
A continuous image is sampled and ordered into an image grid. Each grid element is known as a pixel. Voxel for volume element.

Consider the pixel as the smallest unit in an image. This is not quite a delta because it has a finite size on the CMOS sensor. It is possible to do sub-pixel processing (e.g., corner detection).
Quantization gives the number of output levels $L$

- Continuous image
- Scan line from A to B
- Sampling (horizontal bar) and quantization (vertical bar)
- Digital scan line – resulting effect of sampling and quantization
QUANTIZATION LEVELS

- $L = \text{number of output levels}$
- $k = \text{number of bits per pixel}$
- Output range of image
  - $[0, L - 1] = [0, 2^k - 1]$

- Image storage size
  - $b = M \times N \times k$
  - Number of bits to store image with dimensions $M \times N$

- 8-bits per channel is typical
  - Provide enough resolution to provide quality visual reproduction

*Figure 2.3: Brightness levels. (a) 64. (b) 16. (c) 4. (d) 2. © Cengage Learning 2015.*
RESOLUTION

- Spatial resolution → smallest discernible detail in an image
  - Controlled by the sampling factor (the size $M \times N$ of the CMOS sensor)

- Gray-level resolution → smallest discernible change in gray level
  - Based on number of bits for representation
The pixel neighborhood corresponds to nearby pixels

- **4-neighbors**
  - Horizontal and vertical neighbors

- **8-neighbors**
  - Include 4-neighbors and the diagonal pixels
CONNECTIVITY

- Path exists between pixels

- 4-connected

- 8-connected
**IMAGE PROCESSING**

- Usually the first stage of computer vision applications
  - Pre-process an image to ensure it is in a suitable form for further analysis
- Typical operations include:
  - Exposure correction, color balancing, reduction in image noise, increasing sharpness, rotation of an image to straighten
**2D SIGNAL PROCESSING**

- Image processing is an extension of signal processing to two independent variables
  - Input signal $\rightarrow$ output signal
- General system
  
  $x \rightarrow f \rightarrow y$

- Image processing
  
  $f(x, y) \rightarrow w \rightarrow g(x, y)$

- Linear operators
  
  - $H(af + bg) = aH(f) + bH(g)$
  - Input is an image, output is an image
- Important class of operators for image processing because of the wealth of theoretical and practical results
  
  - E.g. signal processing
- However, non-linear operations can provide better performance but not always in predictable ways.
POINT OPERATORS/PROCESSES

- Output pixel value only depends on the corresponding input pixel value

- Often times we will see operations like dividing one image by another
  - Matrix division is not defined
  - The operation is carried out between corresponding pixels in the two image
  - Element-by-element dot operation in Matlab
    - `>> I3 = I1./I2`
    - Where I1 and I2 are the same size
• Gain and bias (Multiplication and addition of constant)
  • \( g(x, y) = a(x, y)f(x, y) + b(x, y) \)
  • \( a \) (gain) controls contrast
  • \( b \) (bias) controls brightness
    • Notice parameters can vary spatially (think gradients)
• Linear blend
  • \( g(x) = (1 - \alpha)f_0(x) + \alpha f_1(x) \)
  • We will see this used later for motion detection in video processing
COLOR TRANSFORMS

- Usually we think of a color image as three images concatenated together
  - Have a red, green, blue slice corresponding to the notion of primary colors

- Manipulations of these color channels may not correspond directly with desired perceptual response
  - Adding bias to all channels may actually change the apparent color instead of increasing brightness

- Need other representations of color for mathematical manipulation
COLOR IMAGES

- Color comes from underlying physical properties

- However, humans do not perceive color in the same physical process
  - There is some subjectivity (e.g. color similarity)
HUMAN COLOR PERCEPTION

- Cones in human retina are sensitive to color
  - In the center of eye
  - 3 different types for different EM frequency sensitivity
- Rods are monochromatic
  - On outside of the eye and good for low lighting and motion sensing
COLORSPACES

- Uniform method for defining colors
- Can transform from one to another
  - Want to take advantage of properties and color gamut
- XYZ
  - International absolute color standard
  - No negative mixing
- RGB
  - Additive color mixing for red, green, and blue
  - Widely used in computers
- CMYK
  - Cyan, magenta, yellow, black
  - Used for printers and based off of reflectivity
- HSV
  - Hue, saturation, and value = color, amount, brightness
  - Closer to human perception