DIGITAL IMAGE PROCESSING

MODULE 1:
INTRODUCTION TO IMAGE PROCESSING
FUNDAMENTAL STEPS IN IMAGE PROCESSING

Image Acquisition:
Origin of Image Processing
Includes pre-processing
IMAGE ACQUISITION

For more study materials: WWW.KTUSTUDENTS.IN
FUNDAMENTAL STEPS IN IMAGE PROCESSING

Image Enhancement:
Manipulating an image suitable for a specific application
Different for different applications
FUNDAMENTAL STEPS IN IMAGE PROCESSING

Image Restoration:
Improving appearance of the image

Enhancement is subjective and restoration is objective

Based on mathematical or probabilistic models of the image degradation

Image Display
can take place at the output of any stage

Knowledge base

Outputs of these processes generally are image attributes

Morphological processing tools for extracting image components

Compression for reducing the storage required

Wavelets and multi-resolution processing

Color image processing

Color model

Color processing

Image enhancement

Restoration of degraded images

Image restoration

Problem domain

Image acquisition preprocessing involved

Outputs of these processes generally are images
FUNDAMENTAL STEPS IN IMAGE PROCESSING

Colour Image Processing:
Colour models and basic colour processing in an image

Use of internet increased the need of colour image processing

Outputs of these processes generally are images

- Color image processing
  - Color model
  - Color processing
- Wavelets and multi-resolution processing
- Compression
  - For reducing the storage required
- Morphological processing
  - Tools for extracting image components
- Image restoration
  - Restoration of degraded images
- Image enhancement
- Knowledge base
- Segmentation
  - Partition an image into objects
- Representation & description
- Object recognition

Image Display
- Can take place at the output of any stage

Problem domain
- Image acquisition
  - Preprocessing involved

For more study materials: WWW.KTUSTUDENTS.IN
Wavelets and multiresolution processing:

Wavelets are the foundation for representing images in various degree of resolution.

Images subdivided successively into smaller regions.
FUNDAMENTAL STEPS IN IMAGE PROCESSING

Image Compression:
Reducing the storage required to save an image or bandwidth required to send image over internet

JPEG is an example
**FUNDAMENTAL STEPS IN IMAGE PROCESSING**

Morphological Processing:

Deals with tools for extracting image components that are useful in the representation and description of shape.

For more study materials: [WWW.KTUSTUDENTS.IN](http://WWW.KTUSTUDENTS.IN)
FUNDAMENTAL STEPS IN IMAGE PROCESSING

Segmentation:
Partitions an image into its constituent parts or objects.
FUNDAMENTAL STEPS IN IMAGE PROCESSING

Representation & Description:
Follows the output of segmentation stage.
Usually is raw pixel data
E.g.: pixels of boundary of a region. Or all the points in the Region etc.
Object Recognition:
Is the process of assigning a label to the object based on the descriptors.
COMPONENTS OF AN IMAGE PROCESSING SYSTEM

Network

- Image displays
- Computer
- Mass storage
- Hardcopy
- Specialized image processing hardware
- Image processing software
- Image sensors
- Problem domain

Monitor
LCD
Display Card

Printers
Film cameras
Digital Units

Short-Term-Memory
On-Line HD, CD-ROMs
Archival Tapes, CD-ROMs
Sensing:

Contains Two elements.

First is Physical device that is sensitive to the energy radiated by the object we wish to image.

Second called a digitizer for converting output of the above device into digital form.
Sometimes consists of digitizer

H/w that performs primitive operations such as ALU (Arithmetic Logic Operations parallel in entire image.)
May be a mobile phone, pc to a super computer
Image Processing Software:

Consists of specialized modules that perform specific task.
A must in image processing

High resolution image will have large size

1) Short storage during processing (Buffer memory)
2) Online storage for relatively fast recalling
3) Archival storage for infrequent access
Flat screen TV, LED display, Mobile screen led display etc....
Mainly photo printers....

- Monitor
  - LCD
  - Display Card

- Printers
  - Film cameras
  - Digital Units

- Image displays
- Computer
- Mass storage
  - Short-Term-Memory
  - On-Line-HD, CD-ROMs
  - Archival-Tapes, CD-ROMs

- Hardcopy
- Specialized image processing hardware
- Image processing software

- Image sensors

- Problem domain

For more study materials: WWW.KTUSTUDENTS.IN
HUMAN EYE

Functions much like a camera:

Aperture (i.e., pupil), lens, mechanism for focusing (zoom in/out)

Surface for registering images (i.e., retina)
HUMAN EYE (CONT’D)

- In a camera, focusing at various distances is achieved by varying the distance between the lens and the imaging plane.
- In the human eye, the distance between the lens and the retina is fixed (i.e., 14mm to 17mm).
Focusing is achieved by varying the shape of the lens (i.e., flattening of thickening).
HUMAN EYE (CONT’D)

- Retina contains light sensitive cells that convert light energy into electrical impulses that travel through nerves to the brain.
- Brain interprets the electrical signals to form images.
HUMAN EYE (CONT’D)

- Two kinds of light-sensitive cells: rods and cone (unevenly distributed).

- Cones (6 – 7 million) are responsible for all color vision and are present throughout the retina, but are concentrated toward the center of the field of vision at the back of the retina.

- Fovea – special area
  - Mostly cones.
  - Detail, color sensitivity, and resolution are highest.

For more study materials: WWW.KTUSTUDENTS.IN
HUMAN EYE (CONT’D)

- Three different types of cones; each type has a special pigment that is sensitive to wavelengths of light in a certain range:
  - Short (S) corresponds to blue
  - Medium (M) corresponds to green
  - Long (L) corresponds to red
- Ratio of L to M to S cones:
  - approx. 10:5:1
- Almost no S cones in the center of the fovea
HUMAN EYE (CONT’D)

- Rods (120 million) more sensitive to light than cones but cannot discern color.
  - Primary receptors for night vision and detecting motion.
  - Large amount of light overwhelms them, and they take a long time to “reset” and adapt to the dark again.
  - Once fully adapted to darkness, the rods are 10,000 times more sensitive to light than the cones.
Images denoted by two-dimensional functions $f(x,y)$

Value of amplitude of $f$ at $(x,y)$: positive scalar quantity

Image generated by physical process: intensity values are proportional to the energy radiated by a physical source $\Rightarrow 0 < f(x,y) < \infty$

$f(x,y)$ may be characterized by 2 components:

1. The amount of source illumination \textit{incident} on the scene: \textit{illumination} $i(x,y)$

2. The amount of illumination \textit{reflected} by the objects of the scene: \textit{reflectance} $r(x,y)$

\[ f(x,y) = i(x,y) \cdot r(x,y), \text{ where } 0 < i(x,y) < \infty \text{ and } 0 < r(x,y) < 1 \]

\text{total absorption}

\text{total reflectance}
Example of typical ranges of illumination $i(x,y)$ for visible light (average values):

- Sun on a clear day: $\sim 90,000$ lm/m$^2$, down to $10,000$ lm/m$^2$ on a cloudy day
- Full moon on a clear evening: $\sim 0.1$ lm/m$^2$
- Typical illumination level in a commercial office: $\sim 1000$ lm/m$^2$

Typical values of reflectance $r(x,y)$:

- 0.01 for black velvet
- 0.65 for stainless steel
- 0.8 for flat white wall paint
- 0.9 for silver-plated metal
- 0.93 for snow
Monochrome image

Intensity \( l \): \( L_{\text{min}} \leq l \leq L_{\text{max}} \). In practice: \( L_{\text{min}} = i_{\text{min}} \cdot r_{\text{min}} \) and \( L_{\text{max}} = i_{\text{max}} \cdot r_{\text{max}} \)

Typical limits for indoor values in the absence of additional illumination: \( L_{\text{min}} \approx 10 \) and \( L_{\text{max}} \approx 1000 \)

\([L_{\text{min}}, L_{\text{max}}]\) is called the gray (or intensity) scale

Common practice: shift to \([0, L-1]\), where \( l=0 \) is considered black and \( l=L-1 \) is considered white.
Discretizing coordinate values is called Sampling
Discretizing the amplitude values is called Quantization

Method of sampling determined by the sensor arrangement:

- Single sensing element combined with motion:
  spatial sampling based on number of individual mechanical increments
- Sensing strip: the number of sensors in the strip establishes the sampling limitations in one image direction; in the other: same value taken in practice
- Sensing array: the number of sensors in the array establishes the limits of sampling in both directions
The quality of a digital image is determined to a large degree by the number of samples and discrete intensity levels used in sampling and quantization.

However image content is also an important consideration in choosing these parameters.
DIGITAL IMAGE DEFINITION:

- A digital image $f(m,n)$ described in a 2D discrete space is derived from an analog image
- $f(x,y)$ in a 2D continuous space through a sampling process that is frequently referred to as digitization.
- The 2D continuous image $f(x,y)$ is divided into $N$ rows and $M$ columns
- The intersection of a row and a column is termed a pixel
- The value assigned to the integer coordinates $(m,n)$ with $m=0,1,2..N-1$ and $n=0,1,2..N-1$ is $f(m,n)$. In fact, in most cases, is actually a function of many variables including depth, color and time ($t$).
The result of sampling and quantization is matrix of real numbers.

Assume that an image $f(x,y)$ is sampled so that the resulting digital image has $M$ rows and $N$ Columns.

The values of the coordinates $(x,y)$ now become discrete quantities thus the value of the coordinates at origin become $f(x,y) = f(0,0)$

The next Coordinates value along the first signify the image along the first row. It does not mean that these are the actual values of physical coordinates when the image was sampled.
Due to processing storage and hardware consideration, the number gray levels typically is an integer power of 2

\[ L = 2^k \]

Where \( L \) is the number of grey levels and \( k \) is the number of binary digits need to represent an image pixel.
SPATIAL AND INTENSITY RESOLUTION

- Spatial resolution is a measure of the smallest discernible detail in an image.
- Quantitatively spatial resolution can be stated in **Lines Pair Per Unit Distance and Dots/Pixel Per Unit Distance**.
- A widely used definition of image resolution is the largest number of discernible (visible) lines per unit distance (Eg: 100 lines pair per mm).
- Dots/Pixel Per Unit Distance is widely used (Eg: iphone 8 specification of screen 401 ppi (pixels per inch)).
- Image resolution 1024 X 1024 pixels is not a meaningful statement without stating the spatial dimensions. (eg: iphone8 spec 1920X1080-pixel resolution at 401 ppi)

For more study materials: WWW.KTUSTUDENTS.IN
Intensity resolution is the smallest discernible change in the intensity level, that is number of samples used to sample digital image.

It is common practice to refer Number of bits used to quantize the intensity level as Intensity resolution.

Eg: 256 intensity levels with 8 bits of intensity resolution.
EXAMPLE OF SPATIAL RESOLUTION

1250 dpi

300 dpi

150 dpi

72 dpi

For more study materials: WWW.KTUSTUDENTS.IN
INTENSITY RESOLUTION

**Figure 2.21**
(a) 452 × 374, 256-level image.
(b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.
INTENSITY RESOLUTION

FIGURE 2.21 (Continued)
(c)-(h) Image displayed in 16, 8, 4, and 2 gray levels (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)
FIGURE 2.22  (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

FIGURE 2.23  Representative isopreference curves for the three types of images in Fig. 2.22.
Image Interpolation

- Using known data to estimate values at unknown locations
- Used for zooming, shrinking, rotating, and geometric corrections

- Nearest Neighbor interpolation
  - Use closest pixel to estimate the intensity
  - Simple but has tendency to produce artifacts

- Bilinear interpolation
  - Use 4 nearest neighbor to estimate the intensity
  - Much better result
  - Equation used is \( v(x, y) = ax + by + cxy + d \)

- Bicubic interpolation
  - Use 16 nearest neighbors of a point
  - Equation used is
    \[
    v(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j
    \]
Zooming using Various Interpolation Techniques

**FIGURE 2.24** (a) Image reduced to 72 dpi and zoomed back to its original size (3692 × 2812 pixels) using nearest neighbor interpolation. This figure is the same as Fig. 2.20(d). (b) Image shrunk and zoomed using bilinear interpolation. (c) Same as (b) but using bicubic interpolation. (d)–(f) Same sequence, but shrinking down to 150 dpi instead of 72 dpi [Fig. 2.24(d) is the same as Fig. 2.20(c)]. Compare Figs. 2.24(c) and (f), especially the latter, with the original image in Fig. 2.20(a).
NEIGHBORS OF A PIXEL

- A pixel \( p \) at coordinates \((x,y)\) has four horizontal and vertical neighbors having coordinates \((x+1, y)\), \((x, y+1)\), \((x-1, y)\) and \((x, y-1)\).
- These are denoted by \( N_4(p) \)
- The four diagonal neighbors of \( p \) are \((x+1, y+1)\), \((x-1, y+1)\), \((x+1, y-1)\) and \((x-1, y-1)\)
- These are denoted by \( N_D(p) \)
- Together these two groups are called \( N_8(p) \)
- For Boundary pixels number of Neighboring pixels will be less
If $V$ is set of intensity values used to define adjacency, two pixels $p$ and $q$ are adjacent with values from $V$:

- 4-adjacency: $q$ is in the set of $N_4(p)$
- 8-adjacency: $q$ is in the set of $N_8(p)$
- M-adjacency: if
  - $q$ is in $N_4(p)$, or
  - $q$ is in $N_8(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixel whose values are from $V$
**FIND ADJACENCY ON GREY SCALE**

<table>
<thead>
<tr>
<th></th>
<th>250</th>
<th>100</th>
<th>268</th>
<th>100</th>
<th>110</th>
<th>140</th>
<th>240</th>
<th>200</th>
<th>90</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>103</td>
<td>101</td>
<td>254</td>
<td>110</td>
<td>110</td>
<td>230</td>
<td>240</td>
<td>150</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>180</td>
<td>255</td>
<td>200</td>
<td>130</td>
<td>120</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>103</td>
<td>100</td>
<td>150</td>
<td>50</td>
<td>200</td>
<td>220</td>
<td>250</td>
<td>230</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>205</td>
<td>100</td>
<td>230</td>
<td>200</td>
<td>100</td>
<td>200</td>
<td>210</td>
<td>250</td>
<td>150</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>201</td>
<td>220</td>
<td>150</td>
<td>50</td>
<td>103</td>
<td>250</td>
<td>180</td>
<td>160</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>210</td>
<td>270</td>
<td>110</td>
<td>230</td>
<td>150</td>
<td>200</td>
<td>150</td>
<td>140</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>200</td>
<td>130</td>
<td>100</td>
<td>110</td>
<td>210</td>
<td>120</td>
<td>130</td>
<td>130</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>130</td>
<td>120</td>
<td>200</td>
<td>210</td>
<td>110</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>140</td>
<td>100</td>
<td>130</td>
<td>140</td>
<td>130</td>
<td>120</td>
<td>110</td>
<td>100</td>
<td>90</td>
<td></td>
</tr>
</tbody>
</table>

$V = \{100, 101, 102, 103\}$
PATH

A path from \( p(x,y) \) to \( q(s,t) \) is a sequence of distinct points with coordinates \((x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\) Where \((x_i, y_i)\) and \((x_{i-1}, y_{i-1})\) are adjacent for \( 1 \leq i \leq n \).

Here \( n \) is the length of the path.

If \((x_0, y_0) = (x_n, y_n)\), the path is closed path.

We can define 4-, 8-, and m-paths based on the type of adjacency used.

Connected in \( S \), Let \( S \) represent a subset of pixels in an image. Two pixels \( p \) with coordinates \((x_0, y_0)\) and \( q \) with coordinates \((x_n, y_n)\) are said to be connected in \( S \) if there exists a path \((x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\) contains entirely pixels in \( S \).
Examples: Adjacency and Path

\[ V = \{1, 2\} \]

\[
\begin{array}{ccc}
0_{1,1} & 1_{1,2} & 1_{1,3} \\
0_{2,1} & 2_{2,2} & 0_{2,3} \\
0_{3,1} & 0_{3,2} & 1_{3,3}
\end{array}
\]

8-adjacent

\[
\begin{array}{ccc}
0 & 1 & 1 \\
0 & 2 & 0 \\
0 & 0 & 1
\end{array}
\]

m-adjacent

The 8-path from (1,3) to (3,3):
(i) (1,3), (1,2), (2,2), (3,3)
(ii) (1,3), (2,2), (3,3)
Let S represent a subset of pixels in an image.

Two pixels p and q are said to be connected in S if there exists a path between them consisting entirely of pixels in S. For every pixel p in S, the set of pixels in S that are connected to p is called a connected component of S.

If S has only one connected component, then S is called a Connected Set.

We call R a region of the image if R is a connected set. Two regions, R_i and R_j are said to be adjacent if their union forms a connected set.

Regions that are not to be adjacent are said to be disjoint.

Boundary (or border): The boundary of the region R is the set of pixels in the region that have one or more neighbors that are not in R. If R happens to be an entire image, then its boundary is defined as the set of pixels in the first and last rows and columns of the image.

Foreground and background: An image contains K disjoint regions, R_k, k = 1, 2, ..., K. Let R_u denote the union of all the K regions, and let (R_u)^c denote its complement. All the points in R_u is called foreground; All the points in (R_u)^c is called background.
DISTANCE MEASURES.

- Given pixels p, q and z with coordinates (x, y), (s, t), (u, v) respectively, the distance function D has following properties:
  - \(D(p, q) \geq 0\) \([D(p, q) = 0, \text{if } p = q]\)
  - \(D(p, q) = D(q, p)\)
  - \(D(p, z) \leq D(p, q) + D(q, z)\)

- The following are the different Distance measures:
  - Euclidean Distance: \(D_e(p, q) = [(x-s)^2 + (y-t)^2]^{1/2}\)
  - City Block Distance: \(D_4(p, q) = |x-s| + |y-t|\)
  - Chess Board Distance: \(D_8(p, q) = \max(|x-s|, |y-t|)\)
Array versus Matrix Operations

- Consider two $2 \times 2$ images
  \[
  \begin{bmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
  \end{bmatrix}
  \text{ and }
  \begin{bmatrix}
  b_{11} & b_{12} \\
  b_{21} & b_{22}
  \end{bmatrix}
  \]

- Array Product is:
  \[
  \begin{bmatrix}
  a_{11}b_{11} & a_{12}b_{12} \\
  a_{21}b_{21} & a_{22}b_{22}
  \end{bmatrix}
  \]

- Matrix Product is:
  \[
  \begin{bmatrix}
  a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\
  a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22}
  \end{bmatrix}
  \]

For more study materials: WWW.KTUstudents.IN
MATHEMATICAL TOOLS USED FOR DIP

Linear vs. Nonlinear Operation

\[ H[f(x, y)] = g(x, y) \]
\[ H[a_i f_i(x, y) + a_j f_j(x, y)] \]
\[ = H[a_i f_i(x, y)] + H[a_j f_j(x, y)] \]
\[ = a_i H[f_i(x, y)] + a_j H[f_j(x, y)] \]
\[ = a_i g_i(x, y) + a_j g_j(x, y) \]

H is said to be a **linear operator**;
ARITHMETIC OPERATIONS

- Arithmetic operations between images are array operations. The four arithmetic operations are denoted as

\[ s(x,y) = f(x,y) + g(x,y) \]
\[ d(x,y) = f(x,y) - g(x,y) \]
\[ p(x,y) = f(x,y) \times g(x,y) \]
\[ v(x,y) = f(x,y) \div g(x,y) \]