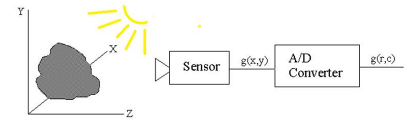
Computer Vision and Image understanding

(Digital image fundamentals)

Bhabatosh Chanda
bchanda57@gmail.com

Image formation



- Two transformations that simultaneously take place in the camera
  - Geometric transformation (perspective projection)
  - Photometric transformation (light transportation)

Photometric model

- Let light intensity $\mathscr{I}$ is transported ideally to image plane *(ξ, ζ)*.
- Scene reflectance $\mathfrak{R}$ is transferred to the image plane.
- Thus image intensity mapped is ideal one and is given by

$$f(\xi, \zeta) = \mathfrak{R}(\xi, \zeta)\, \mathscr{I}(\xi, \zeta)$$
$$(\xi, \zeta) \equiv (x, y)$$

- However, light energy transport to real image plane undergoes some transformation. This may be modeled by photometric transformation.

$$g(x, y) = \mathrm{T}[f(x, y)]$$

that satisfies the condition: $f(x, y) \geq 0$ and $g(x, y) \geq 0$.

Digitization

The process of converting infinite number of possible values to finite ones is called ***digitization***.

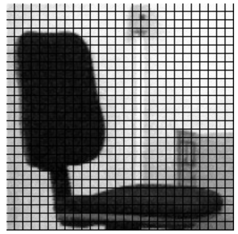
Image digitization process has two steps:

- **Sampling**
  - deals with two-dimensional spatial coordinates
- **Quantization**
  - deals with magnitude at each position

Sampling

Sampling process is

- Dividing the spatial domain into a finite number of cells, each of which can be represented by an integer coordinate in discrete domain.
- Each cell is called a ***pixel***.
- A pixel is the smallest accessible unit of an image.

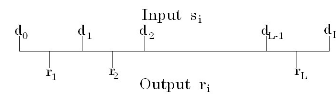


Quantization

- Average intensity over a cell may have any value between the lowest (black) to the highest (white).

The process of ***quantization*** is

- Dividing the entire intensity range into a finite number of bins and representing an intensity value by the integer index of the bin in which it falls.

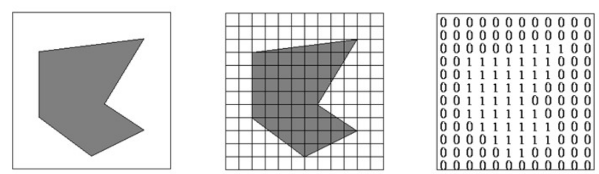


Digitization: Two terms

- **Resolution** – denotes area of the scene represented by a pixel
  - Higher resolution → smaller cell size, so finer details are better preserved.
- **Levels** (or graylevels) – number of bins that divides the intensity range
  - More levels → narrower bins, shades are better preserved.

In both the cases, image is of better quality, but needs more space and processing time.

Example of binary image



Pixel is the smallest rectangular area under the grid. It is smallest unit of a digital image that can be accessed or processed.

Example of graylevel image

Digital graylevel image $g(r,c)$

Discrete domain $D \subset Z^2$, where

$$D = \{(r,c) \mid r = 0,1,2,........, M-1; c = 0,1,2,........, N-1\}$$

$$g(r,c) = \begin{bmatrix} g(0,0) & g(0,1) & ..... & ..... & g(0,N-1) \\ g(1,0) & g(1,1) & ...... & ...... & g(1,N-1) \\ ...... & ...... & ...... & ...... & ...... \\ ...... & ...... & ...... & ...... & ...... \\ g(M-1,0) & g(M-1,1) & ...... & ...... & g(M-1,N-1) \end{bmatrix}$$

where $g(r,c) = \{0,1,2,..........., L-1\}$

Colour models

- Colour is a complex perceptual phenomenon.
- Sensation of colour arises due to response of three neurochemical sensors or receptors in the retina to the visible light.

$$R = \int C(\nu)\, h_R(\nu)\, d\nu$$

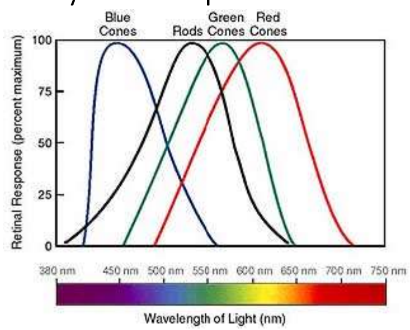
$$G = \int C(\nu)\, h_G(\nu)\, d\nu$$

$$B = \int C(\nu)\, h_B(\nu)\, d\nu$$

Colour models

- Wavelength of visible spectrum ranges from 0.38 µm to 0.76 µm (approx.)
- Response characteristics $h_B(\nu)$ for blue attains maximum at about 0.44 µm or 6.8 x 10^{14} Hz.
- Response characteristics $h_G(\nu)$ for green attains maximum at about 0.52 µm or 6.8 x 10^{14} Hz.
- Response characteristics $h_R(\nu)$ for red attains maximum at about 0.70 µm or 4.3 x 10^{14} Hz.

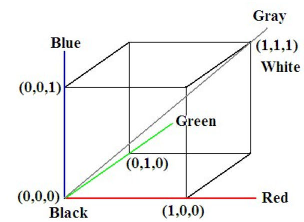
Color sensitivity of receptors



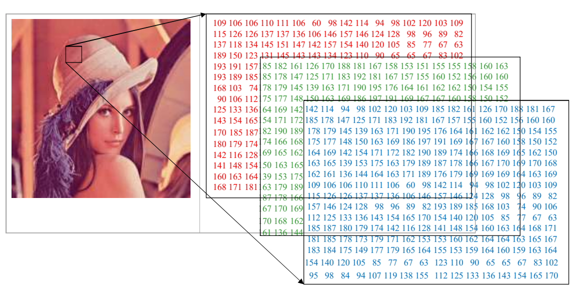
RGB colour model

- A unit cube in Cartesian coordinate system.
- ***Additive primaries***, i.e.,

  $$R + G + B = W$$
- Used in devices like camera, monitor.
- (0,0,0) → Black
- (1,1,1) → White



Example of colour image



<u>Digital colour image</u> $g_C(r,c)$

Discrete domain $D \subset Z^2$, where

$$D = \{(r,c) \mid r = 0,1,2,\ldots\ldots\ldots, M-1; c = 0,1,2,\ldots\ldots\ldots, N-1\}$$

$$[g_C(r,c)]_{MxN} = \begin{bmatrix} g_C(0,0) & g_C(0,1) & \ldots\ldots & \ldots\ldots & g_C(0,N-1) \\ g_C(1,0) & g_C(1,1) & \ldots\ldots & \ldots\ldots & g_C(1,N-1) \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ g_C(M-1,0) & g_C(M-1,1) & \ldots\ldots & \ldots\ldots & g_C(M-1,N-1) \end{bmatrix}$$

where $g_C(r,c) = [g_R(r,c) \;\; g_G(r,c) \;\; g_B(r,c)]^T$

and $g_R(r,c)$, $g_G(r,c)$ and $g_B(r,c) \in \{0,1,2,\ldots\ldots\ldots\ldots, L-1\}$

CMY colour model

- Complementary to RGB colour model, so uses same system for representation.
- Subtractive primaries, i.e.,

  Cyan = 1 – red

  Magenta = 1 – green

  Yellow = 1 – blue
- Used in printing devices
- (0,0,0) → white and (1,1,1) → black

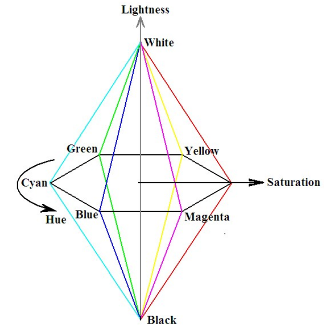
HLS colour model

- Model is represented by a double-hex cone.
- User oriented model (for interaction) based on intuitive colour notion.

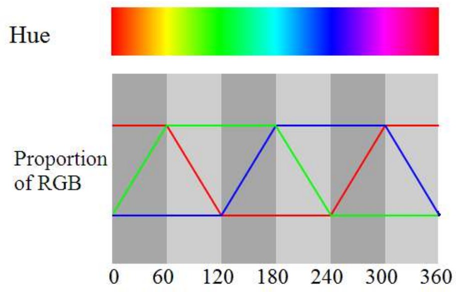
$$0^o \leq S \leq 100$$

$$0 \leq L \leq 1$$

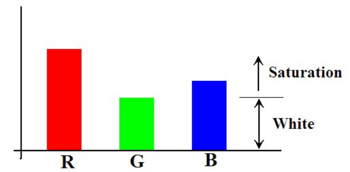
$$0^o \leq H \leq 360^o \quad (circular)$$



Hue: mixing RGB

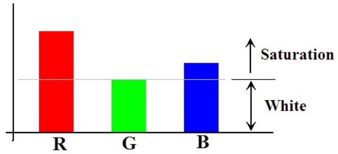


Saturation: Purity



Saturation: Purity

- Hue shows pure colours are formed by combining two of the additive primaries
- Minimum of R, G and B combines with others to form whiteness in colour, e.g., pink = red + white
- Saturation suggests the purity of colour



RGB to YC_bC_r

- Present days all digital videos and static images are compression standards are based on

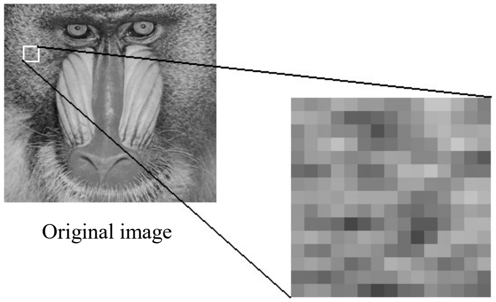
$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.81312 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}$$

- *R′, G′, B′* are gamma-corrected *R, G, B*.
- *Y* is intensity component. Others represent colors.

Digital geometry

- Digital image is defined over a 2D discrete domain.
- Here a point has integer coordinate values representing a small rectangular region of the image domain.
- Relationship between the points and trans-formations in discrete domain may be handled by the concept of digital geometry.

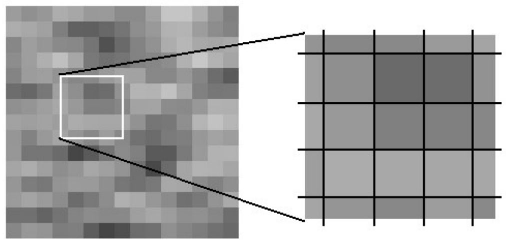
Digital geometry (contd.)



Original image

Enlarged block

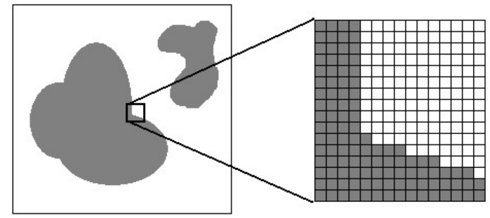
Digital geometry (contd.)



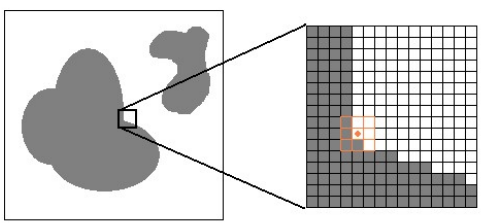
Enlarged block

Pixels (boundary overlayed)

Digital geometry (contd.)



Digital geometry (contd.)



Neighborhood

p_4	p_3	p_2
p_5	p_0	p_1
p_6	p_7	p_8

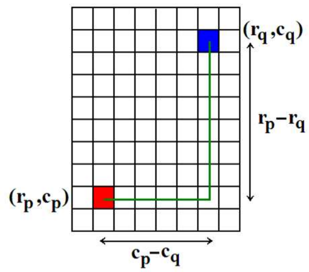
p_0 is the candidate pixel

p_1, p_3, p_5, p_7 are 4-neighbors of p_0

$p_1, p_2, \ldots\ldots, p_8$ are 8-neighbors of p_0

Distance

- Coordinates of two points are (r_p, c_p) and (r_q, c_q)
- Horizontal distance between them is $r_p - r_q$
- Vertical distance is $c_p - c_q$



Distance

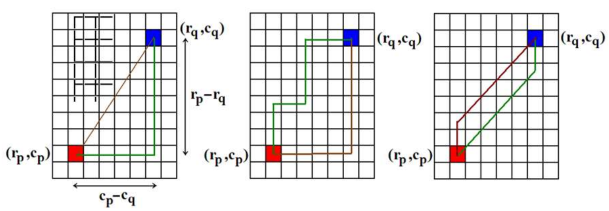
Distance between two pixels (r_p, c_p) and (r_q, c_q) may be defined as

Euclidean: $de(p,q) = \sqrt{(r_p - r_q)^2 + (c_p - c_q)^2}$

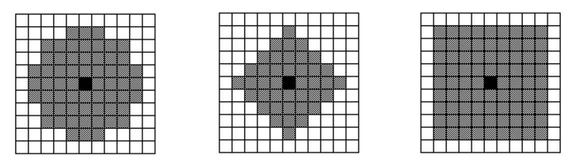
City-block: $d_4(p,q) = | r_p - r_q | + | c_p - c_q |$

Chess-board: $d_8(p,q) = \max\{| r_p - r_q |, | c_p - c_q |\}$

Distance (contd.)



Eucildean City-block Chess-board



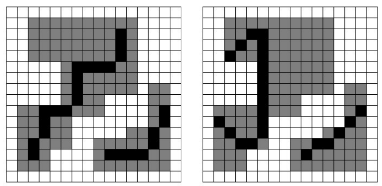
(a) $d_e(p,q) \le r$ (b) $d_4(p,q) \le r$ (c) $d_8(p,q) \le r$

Disk of radius 4 using (a) Euclidean distance, (b) city-block distance and (c) chessboard distance.

Path

Path is a sequence of pixels $p_0, p_1, p_2, \ldots\ldots\ldots, p_n$ where all pixels have same value and every pair of p_i and p_{i+1} $(i = 0, 1, 2, \ldots\ldots\ldots, n-1)$ are neighbors.

Example:



A path may be 4-connected or 8-connected depending on whether p_i and p_{i+1} are 4-neighbors or 8-neighbors.

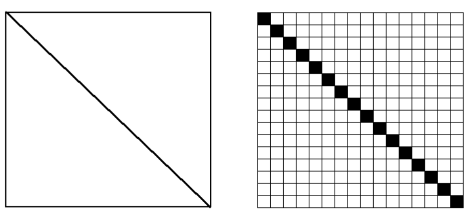
Connected component

Connected component is a set of pixel where every pair of pixels are connected by a 4- or 8-connected path

- 4-connected region
- 8-connected region

If objects are 8-connected then background should be 4-connected and *vice versa*.

4- and 8-neighbour conflict!



Thank you!

Any question?