Basic Tools in Image Processing

Digitalization: Sampling and Quantization

Sampling is the process of converting a signal (e.g., a continuous function of time or space) into a numeric sequence (e.g., a function of discrete time or space).
Quantization is the discretization of the intensity value. Typically, 256 levels (for each color) suffice to represent the intensity.

Histogram

A histogram of a gray scale image $\{x\}$ is a graph representing the number of occurrences n_i of each gray i in the image: $p_x(i) = p(x = i) = \frac{n_i}{n}$ where i is the gray level $in\,\{0 \dots L-1\}$ (L is usually 256) and n is the total number of pixels.
Histogram equalization is technique used to adjust image intensities to enhance contrast:

- Histogram: $p_x(i)$
- Cumulative Histogram: $ch(i) = \sum_{j=0}^{i} p_x(j)$
- $i \leftarrow floor\left(ch(i) \times \frac{L-1}{n}\right)$

Thresholding is a method of segmentation to create a binary image: $g(x, y) = f(x, y) > T\ ?\ 1{:}\,0$ where T is the threshold value ($T = 128$ for average grey-level).

Fourier Transform for Image Processing

The Fourier Transform is the series expansion of an image function in term of "cosine" image basis functions.

- Continuous: $F(u, v) = \int\int f(x, y) \exp\big(-2i\pi(ux + vy)\big)\, dx\, dy$
- Discrete: $F(u, v) = \sum\sum f(x, y) \exp\left(-2i\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right)$

We usually display $\log(1 + |F(u, v)|)$ instead of $|F(u, v)|$
What do frequencies represent in FT?

- Low: uniform areas
- High: edges and noise

Image Filtering

- Via FFT: $f(x, y) \rightarrow FT \rightarrow F(u, v) \rightarrow F(u, v) \times H(u, v) \rightarrow FT^{-1} \rightarrow f'(x, y)$
- Via convolution: $f(x, y) \rightarrow f(x, y) * h(x, y) \rightarrow f'(x, y)$

Low pass filtering to reduce noise

- Average filtering: $\frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $\frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$
- Median filtering (nonlinear filter):
  - Classify $S = \{f(x_j, y_j), (x_j, y_j) \in W\}$
  - $f'(x_i, y_i) = med(S)$

High pass filtering to detect edges

Gradient and Laplacian to sharpen the image and detect edges:

- $\nabla f(x, y) = (\partial_x f, \partial_y f) \in \mathbb{R}^2$

- $|\nabla f(x,y)| = \sqrt{(\partial_x f)^2 + (\partial_y f)^2} \in \mathbb{R}$
- $\nabla^2 f(x,y) = \partial^2_{x^2} f(x,y) + \partial^2_{y^2} f(x,y) \in \mathbb{R}$

Gradient pattern:

- Gradient: $G_x = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ and $G_y = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$
- Prewitt: $\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}$
- Sobel: $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}$

And check if $|\nabla f(x,y)| > threshold$ at (x_0, y_0)

Note: $\frac{1}{16}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \sim \frac{1}{4}(1 \quad 2 \quad 1) * \frac{1}{4}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

Local Binary pattern: $LBP = \sum \mathrm{H}(g_p - g_c)\, 2^p$ where $H(x) = 1?\, x \geq 0 : 0$

Laplacian pattern: $\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

Hough Transform

Designed for line detection and can be used for any analytical curve (circle, ellipse, etc.).

A curve defined by pixels and described by a parametric equation $f(x, y, a_1, \ldots, a_n) = 0$.

- Line $\{y = \boldsymbol{a}x + \boldsymbol{b} \mid \boldsymbol{\rho} = x\cos(\boldsymbol{\theta}) + y\sin(\boldsymbol{\theta})\}$: 2 parameters
- Circle $\{(x - \boldsymbol{a})^2 + (y - \boldsymbol{b})^2 = \boldsymbol{c}^2\}$: 3 parameters

A line in I ⇌ A point in H

A point in I ⇌ A sinusoid in H

From 2 points $M(x_i, y_i)$ and $M(x_j, y_j)$ in I to one point (ρ, θ) in H: $\rho = \frac{|x_i y_j - x_j y_i|}{\sqrt{(y_j - y_i)^2 + (x_j - x_i)^2}}$ and $\theta = -\arctan \frac{x_j - x_i}{y_j - y_i}$

Quad Tree

- Split: split the picture until reaching uniform areas
- Merge: merge neighbor areas according to similarity criterion (mean, variance, etc.)

Mathematical Morphology

Given an object A and a structuring element B_p:

- Erosion: $er(A, B_p) = \{p \mid B_p \subset A\}$
- Dilatation: $dil(A, B_p) = \{p \mid A \cap B_p \neq \emptyset\}$
- Opening: $dil(er(A, B_p), B_p) \Rightarrow$ cut narrows, remove small islands
- Closing: $er(dil(A, B_p), B_p) \Rightarrow$ fill narrow canals

Properties:

- Translation invariant
- err and dil are not inverse of each other
- Duality: $dil(A, B_p) = [er(A^C, B_p)]^C$

- Increasing: $X \subset X' \Rightarrow T(X) \subset T(X')$

Optical Flow

Apparent motion: optical flow between two consecutive image frames taken at t and $t' = t + \delta t$.

- STEP 1: local estimation. Calculate displacement $d_i\left(d_{x_i}, d_{y_i}\right)$ for each pixel $p_i(x_i, y_i)$ where $d_{x_i} = x_i' - x_i$ and $d_{y_i} = y_i' - y_i$.
- STEP 2: global interpretation. $(x_i' \quad y_i') = z_f.(x_i \quad y_i) + (pan_x \quad pan_y)$
  - z_f: zoom factor (0: no zoom, $\gg 1$: backward, $\ll 1$: forward)
  - pan_x and pan_y: vertical and horizontal pan (mean displacement of the image)

To estimate the optical flow, we assume that the luminance of each pixel is constant over time:

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \Rightarrow \frac{\partial I}{\partial x}\delta x + \frac{dI}{dy}\delta y + \frac{dI}{dt}\delta t = 0 \Rightarrow \frac{\partial I}{\partial x}V_x + \frac{dI}{dy}V_y = -\frac{\partial I}{\partial t}$$

Now we can compute V_x and V_y via an iterative process:

$$V^{(i)} = V^{(i-1)}(p, t) - \epsilon.DFD\left(p, t, V^{(i-1)}\right)\nabla I\left(x - V_x^{(i-1)}, y - V_y^{(i-1)}, t - 1\right)$$

Introduction to Colorimetry

Color perception depends on:

1. Illumination source
2. Characteristics of the object
3. Characteristics of the human eye
4. Human brain

Light is electromagnetic radiation.

Visible spectrum: radiations with wavelength from $380\,nm$ to $780\,nm$.

Means to produce light:

- Incandescence (sun, candle)
- Gas discharge (sodium, mercury)
- Photoluminescence (florescent light tube)
- Chemical reactions (light without heat)

Addictive color mixing

Primary colors: red, green and blue

Secondary colors: yellow, cyan and magenta

Grassman's laws:

- Law 1: any color is a linear combination of RGB. $C = r_c(R) + g_c(G) + b_c(B)$
- Law 2: $C = C_1 + C_2 = [r_1 + r_2](R) + [g_1 + g_2](G) + [b_1 + b_2](B)$
- Law 3: proportionality. $k.C = k.r_c(R) + k.g_c(G) + k.b_c(B)$

RGB color model

Not easy to determine RGB values for a given color

Color making attributes:

- Hue: common definition of a color (red, orange, etc.)
- Lightness/value: total quantity of light reaching the eye

- Chroma/saturation: intensity of a hue

Munsell color system: $C(r\colon saturation\,, \theta\colon hue\,, z\colon value)$ in a cylindrical coordinate system

Hue Saturation Value / Hue Lightness Saturation color models.

CIE XYZ color model: $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.41 & 0.36 & 0.18 \\ 0.21 & 0.72 & 0.07 \\ 0.02 & 0.12 & 0.95 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$

CIE chromaticity diagram: $\left(x = \frac{X}{X+Y+Z}, y = \frac{Y}{X+Y+Z}, z = \frac{Z}{X+Y+Z}\right)$

The CIE XYZ color space to determine dominant and complementary wavelengths.

Color mixing: $P = \alpha P_1 + \beta P_2 + \gamma P_3$ where $\alpha = \frac{a}{a+b+c}$, etc.

A MacAdam ellipse is a region which contains indistinguishable colors to the human eye (scales are not uniform).

CIE USC diagram: $\left(u = \frac{4X}{X+15Y+3Z}, v = \frac{9Y}{X+15Y+3Z}\right)$ (more uniform than (x, y) diagram)

CIE L*a*b* color model (uniform), given $white = (X_n, Y_n, Z_n)$, $f(t) = t > \delta^3\ ?\ \sqrt[3]{t} : \frac{t}{3t^2} + \frac{4}{29}$ and $\delta = \frac{6}{29}$:

- Luminosity: $L^* = 166 f\left(\frac{Y}{Y_n}\right) - 16$
- Red/Green axis: $a^* = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right)$
- Blue/Yellow axis: $b^* = 200\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right)$

YIQ color model: Y for luminance and I,Q for color (chromaticity):

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.273 & -0.322 \\ 0.212 & -0.522 & 0.315 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

YUV and YCrCb color models: used for digital video:

$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.144 \\ -0.169 & -0.331 & -0.500 \\ 0.500 & -0.419 & -0.018 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, Cr = R - Y$ and $Cb = B - Y$