

XỬ LÝ ẢNH

Trần Quang Đức

Thông tin giảng viên

- Thông tin liên hệ
 - Phòng B1-801 (2 pm - 3 pm – Thứ Ba Hàng Tuần)
 - Bộ môn Truyền Thông và Mạng Máy Tính
 - Viện Công Nghệ Thông Tin và Truyền Thông
 - Đại học Bách Khoa Hà Nội
 - E-mail: ductq@soict.hust.edu.vn
 - Điện thoại: (+84) (4) 38682596

Nội dung

- **Chương 1:** Giới thiệu chung
- **Chương 2:** Hình thành và biểu diễn ảnh
- **Chương 3:** Thu nhận và số hóa ảnh
- **Chương 4:** Cơ sở lý thuyết xử lý ảnh
- **Chương 5:** Cải thiện và nâng cấp ảnh
- **Chương 6:** Phát hiện và tách biên ảnh
- **Chương 7:** Phân vùng ảnh
- **Chương 8:** Xử lý ảnh nhị phân
- **Chương 9:** Mã hóa và nén ảnh

Tài liệu tham khảo

- R.C. Gonzalez, and R. E. Woods, “Digital Image Processing,” vol. 2, Prentice Hall, 2002.
- A. K. Jain, “Fundamentals of Digital Image Processing,” vol. 3, Englewood Cliffs, Prentice Hall, 1989.
- L. M. Bá, N. T. Thủy, “Nhập môn xử lý ảnh,” Trường Đại học Bách Khoa Hà Nội, 2008.
- R.C. Gonzalez, R. E. Woods, “Digital Image Processing using Matlab,”

GIỚI THIỆU CHUNG

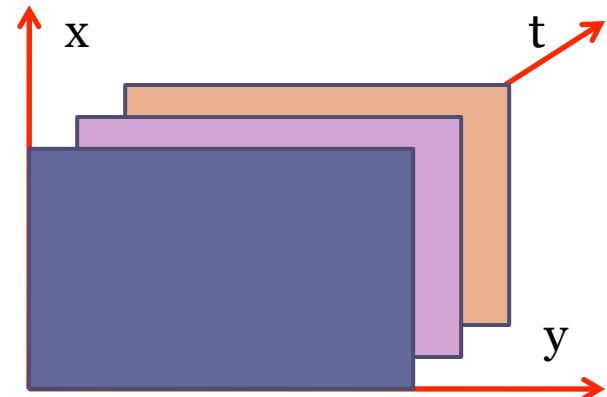
Trần Quang Đức

Các khái niệm

- **Ảnh:** Thông tin về vật thể hay quang cảnh được chiếu sáng mà con người quan sát và cảm nhận bằng mắt và hệ thần kinh thị giác.
- Đối tượng của xử lý ảnh là xử lý các ảnh tự nhiên, ảnh chụp, dữ liệu ảnh có nguồn gốc từ tín hiệu ảnh đặc trưng bởi biên độ và dải tần số. Có sự phân biệt giữa xử lý ảnh với đồ họa.
- Hệ thống xử lý ảnh thu nhận khung cảnh hoặc ảnh ở đầu vào, thực hiện các phép xử lý để tạo ra một ảnh ở đầu ra thỏa mãn các yêu cầu về cảm thụ hoặc trích rút các đặc trưng của ảnh.

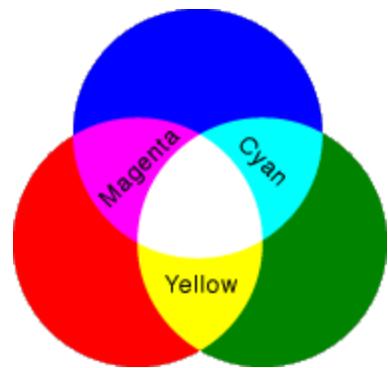
Ảnh tĩnh và chuỗi ảnh

- **Ảnh tĩnh (Still Image):** Biểu diễn bởi hàm độ chói của các biến toàn bộ trong mặt phẳng ảnh $I(x,y)$.
- **Chuỗi ảnh (Sequence of Images):** Hàm độ chói của các biến tọa độ mặt phẳng và biến thời gian $I(x,y,t)$.

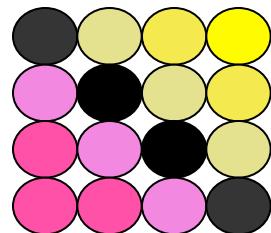


Ảnh tĩnh

- Ảnh nhị phân: 1 bit/pixel
- Ảnh xám: 8 bits/pixel
- Ảnh màu: 16-24 bit/pixel
- RGB, YUV, HSL, YCbCr



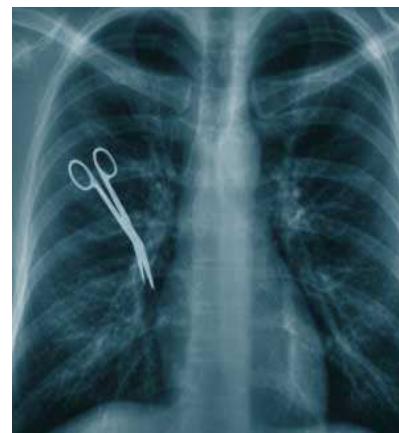
Pixel (RGB)



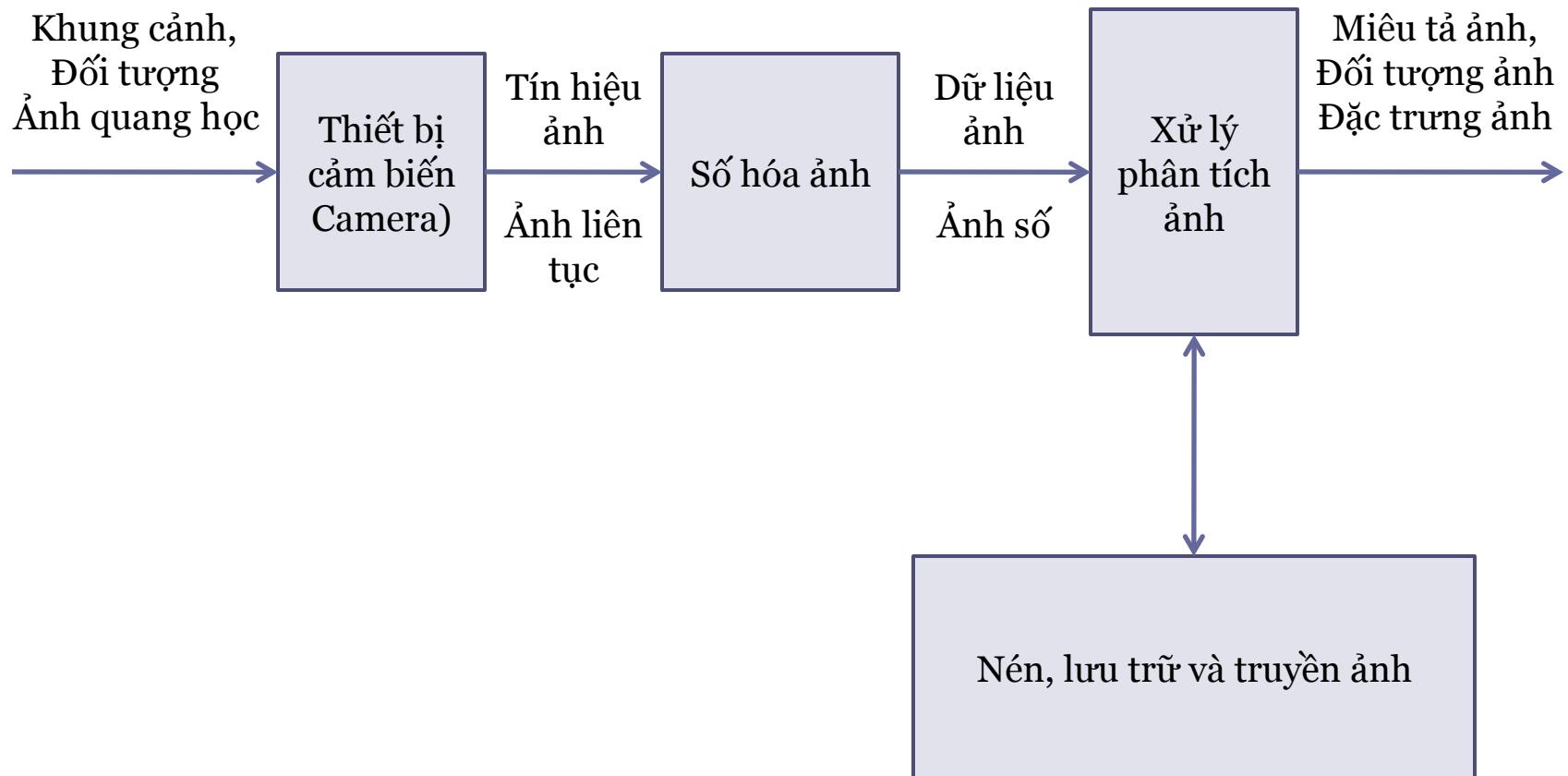
Chuỗi ảnh

- **Video:** Chuỗi các ảnh (khung hình), quan hệ thời gian giữa các khung hình biểu diễn ảnh động.
- Tần số và Độ phân giải
 - NTSC (525 dòng, 30 khung hình/giây)
 - PAL (625 dòng, 25 khung hình/giây)
 - SECAM (625 dòng, 25 khung hình/giây)
 - Phim (24 khung hình/giây)
 - HDTV (16:9, 720 dòng, 60 khung hình/giây)
 - SVGA: 72 khung hình/giây, 1024x720 pixel

Các ví dụ về ảnh số



Hệ thống xử lý ảnh



Các vấn đề của xử lý ảnh

- **Thu nhận ảnh, chụp ảnh và số hóa ảnh**
 - Hệ thống chụp ảnh và tín hiệu ảnh
 - Hệ thống số hóa ảnh: Lấy mẫu,Lượng tử hóa
- **Phân tích ảnh và thị giác máy tính**
 - Cải thiện nâng cấp ảnh, sửa lỗi, khôi phục ảnh
 - Phân tách đặc trưng: tách biên, phân vùng ảnh
 - Biểu diễn và xử lý đặc trưng hình dạng đối tượng ảnh
 - Nhận dạng đối tượng ảnh, phân tích cảnh và hiểu cảnh
- **Mã hóa, nén ảnh**
 - Các phương pháp nén và các chuẩn nén

Ứng dụng của xử lý ảnh

- Thông tin ảnh, truyền thông ảnh
- Xử lý ảnh vệ tinh, viễn thám
- Thiên văn, nghiên cứu không gian, vũ trụ
- Người máy, tự động hóa
- Máy thông minh, thị giác máy nhân tạo
- Sinh học, y học
- Giám sát kiểm soát, Quân sự

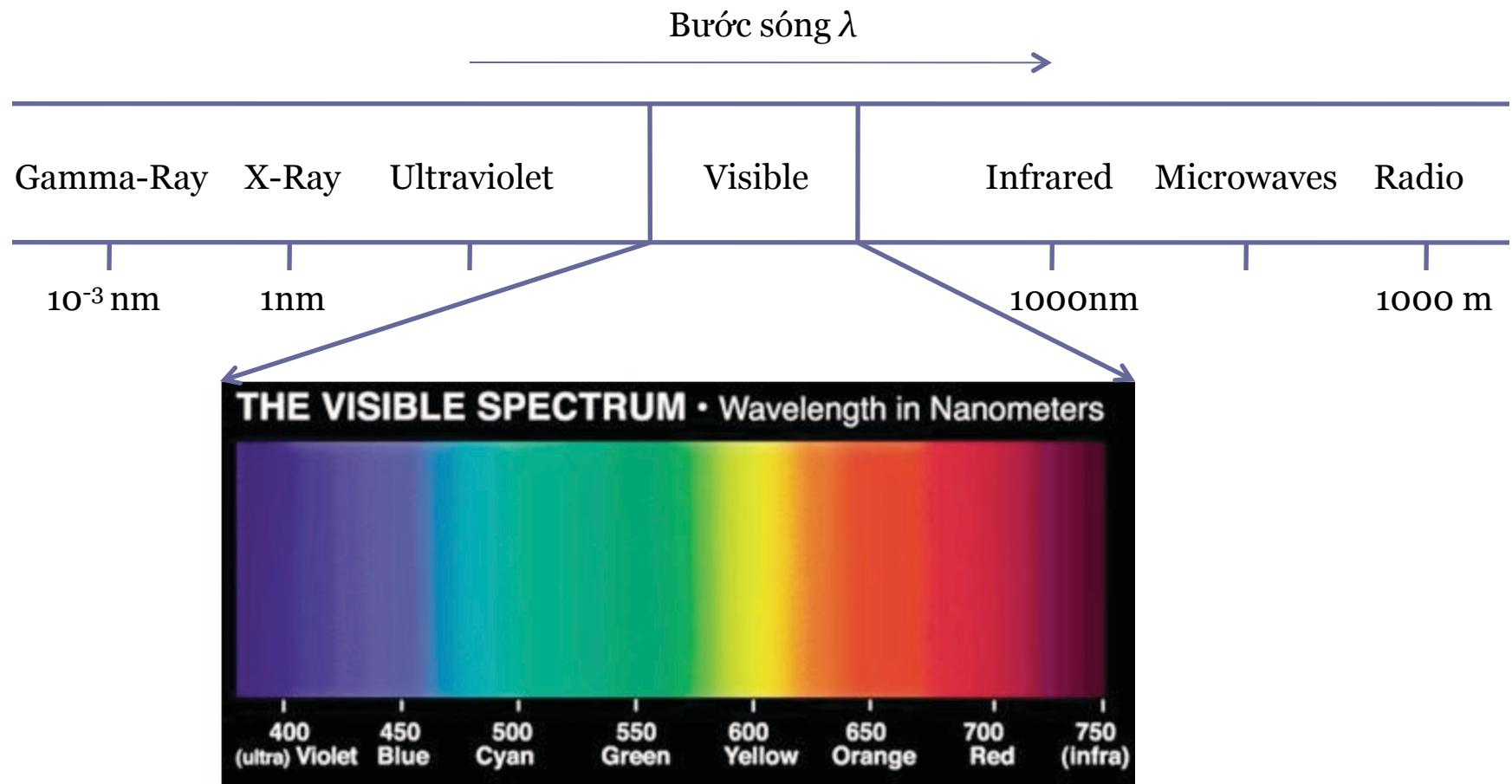
Ứng dụng của xử lý chuỗi ảnh

- Nén video, truyền thông video, truyền hình số
- Giám sát theo dõi phát hiện chuyển động
- Điều khiển lưu lượng chuyển động
- Người máy chuyển động
- Quốc phòng an ninh
- Y học, hóa học

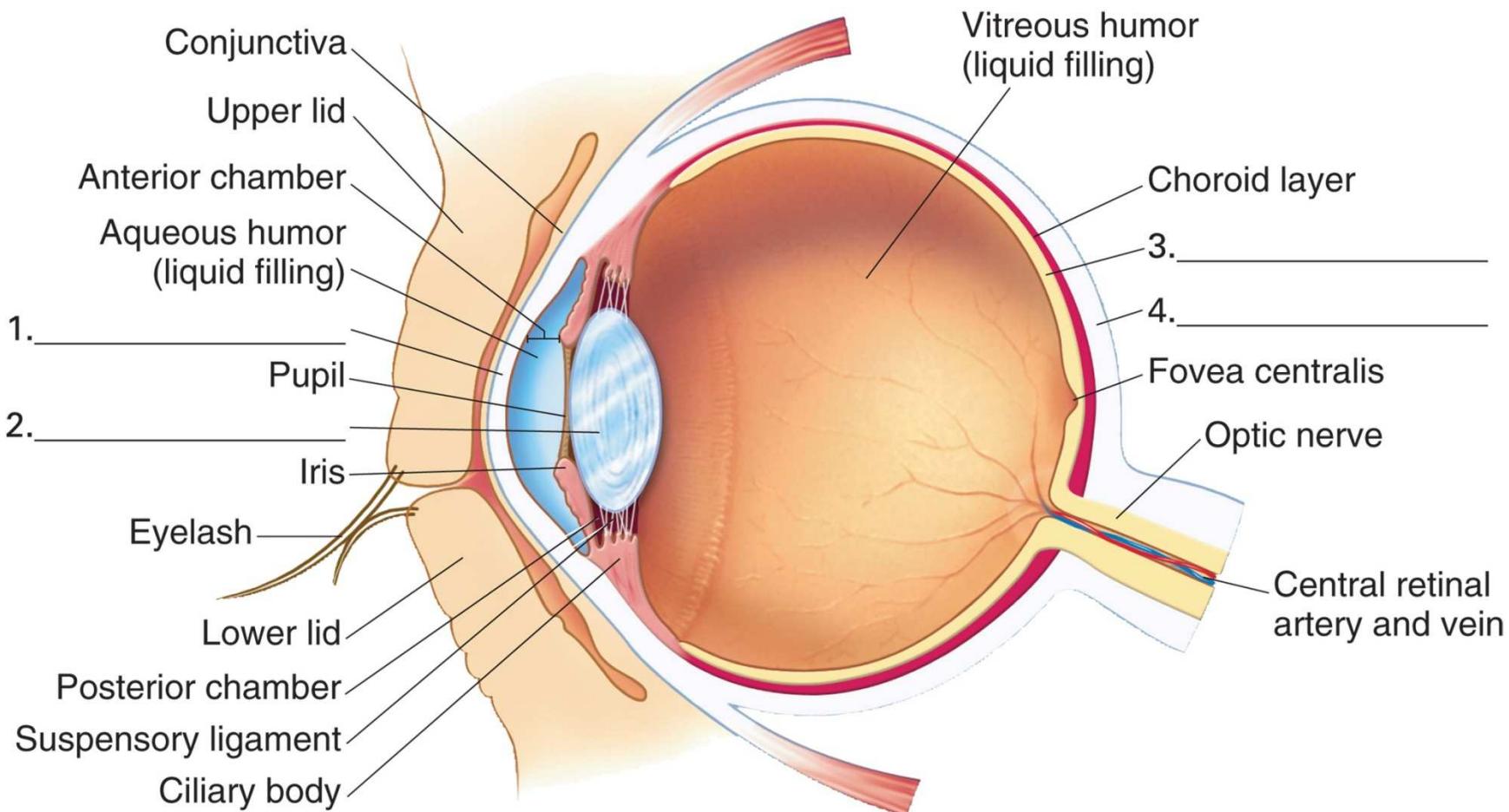
HÌNH THÀNH VÀ BIỂU DIỄN ẢNH

Trần Quang Đức

Ánh sáng và sóng điện từ



Hệ thống thị giác



Hệ thống thị giác

- **Tế bào que**

- Có từ 75-150 triệu
- Rất nhạy cảm với ánh sáng
- Cảm nhận trên dải rộng
- Cung cấp khả năng nhìn đêm
- Cảm nhận độ chói
- Độ phân giải cao

- **Tế bào nón**

- Có từ 6-7 triệu
- Tập trung chủ yếu tại điểm vàng tại trung tâm võng mạc
- Cảm nhận trên dài hẹp
- Ba loại tế bào cảm nhận màu xanh lam, xanh lục và đỏ

Cảm nhận và biểu diễn màu

- **Độ chói (Radiance)**

- Tổng năng lượng của chùm tia từ nguồn

- **Độ rọi (Luminance)**

- Độ đo năng lượng ánh sáng thu thập được từ nguồn sáng
 - Biến thiên theo khoảng cách từ nguồn sáng, bước sóng
 - Không phụ thuộc vào môi trường

$$L(x, y) = \int_0^{\infty} f(x, y, \lambda) V(\lambda) d\lambda$$

- $f(x, y, \lambda)$ Phân bố ánh sáng trong không gian
 - $V(\lambda)$ Hàm hiệu suất độ rọi tương đối của hệ thống thị giác

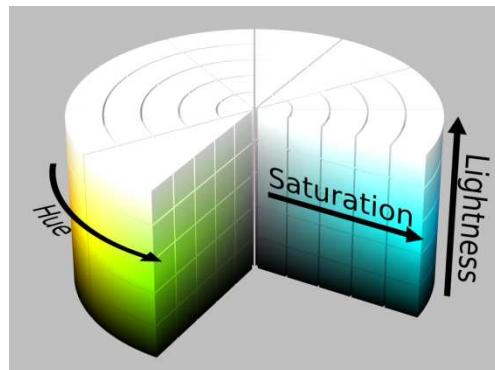
Màu sắc

- Cảm nhận màu sắc phụ thuộc vào phổ của ánh sáng. Ánh sáng nhìn thấy với dải phổ rất hẹp. Ánh sáng với tất cả các thành phần phổ nhìn thấy có năng lượng bằng nhau sẽ được cảm nhận là ánh sáng trắng.
- Một màu có thể tạo nên bằng cách trộn 3 màu cơ bản tương đương với 3 dạng tế bào cảm nhận màu sắc (Đỏ, Lục, Lam). Phân biệt màu sắc dựa trên độ sáng, sắc độ và độ bão hòa.

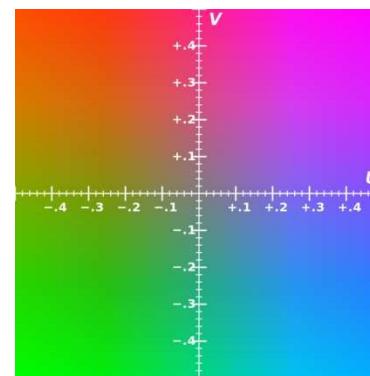
Độ sáng, sắc độ và độ bão hòa

- **Độ sáng (Brightness)**
 - Thuộc tính chủ quan, đặc trưng cho khả năng cảm nhận độ rọi
 - Phụ thuộc vào độ rọi của môi trường xung quanh
- **Đặc trưng màu (Chrominance)**
 - Sắc độ (Hue) là thuộc tính liên quan đến bước sóng chủ yếu trong hỗn hợp của các bước sóng ánh sáng. Sắc độ đặc trưng cho màu sắc chủ đạo được cảm nhận.
 - Độ bão hòa (Saturation) đặc trưng cho độ thuần khiết tương đối. Độ bão hòa phụ thuộc vào độ rộng của phổ ánh sáng và thể hiện lượng màu trắng được trộn với sắc độ.
 - Sắc độ và độ bão hòa gọi là đặc trưng màu.

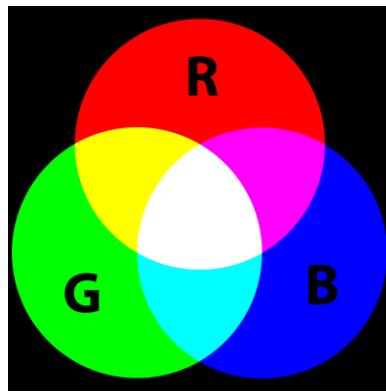
Các mô hình màu



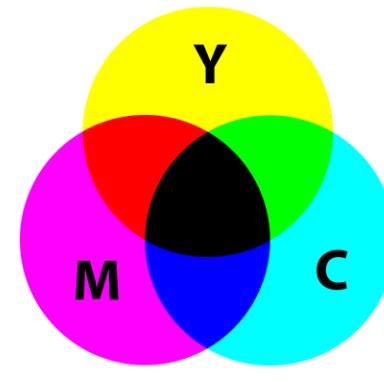
HSV



YCbCr ($Y=0.5$)



RGB



CMYK

Các mô hình màu

- Red, Green, Blue tại 3 trục nhận giá trị [0 255]. Mô hình RGB có thể biểu diễn hơn 16 triệu màu. Trong đó, R (255, 0, 0), G (0, 255, 0) và B (0, 0, 255).
- CMYK (Cyan, Magenta, Yellow, Black) với C = 255-R, M = 255-G, và Y = 255-B.
- YCbCr

$$Y = \frac{16 + 219(0.299R + 0.587G + 0.114B)}{255}$$

$$Cb = \frac{128 + 224(-0.169R - 0.331G + 0.5B)}{255}$$

$$Cr = \frac{128 + 224(0.5R - 0.419G - 0.081B)}{255}$$

Biểu diễn ảnh số

9	7	1	1	1	2	2	1
8	9	9	7	1	1	1	1
7	5	195	7	2	3	1	1
8	6	5	6	3	1	2	3
9	7	6	9	1	2	2	1
5	8	7	9	1	1	2	1
8	9	8	9	1	2	1	2

Ảnh là một mảng hai chiều gồm M hàng và N cột với $f(x,y)$ là giá trị mức xám tại tọa độ (x,y) .

$N_4(p), N_D(p), N_8(p)$

Quan hệ 4-lân cận (4-adjacency)

Quan hệ 8-lân cận (8-adjacency)

Quan hệ m -lân cận (m -adjacency)

0	1	1
0	1	0
0	0	1

0	1	1
0	1	0
0	0	1

0	1	1
0	1	0
0	0	1

Độ sai khác

- **L₂-norm error**

$$MSE = \frac{1}{N} \sum_{i=0}^{N-1} E(|X_i - \hat{X}_i|^2)$$

- **L₁-norm error**

$$MAD = \frac{1}{N} \sum_{i=0}^{N-1} E(|X_i - \hat{X}_i|)$$

- **L_∞-norm error**

$$MaxError = \max_i E(|X_i - \hat{X}_i|)$$

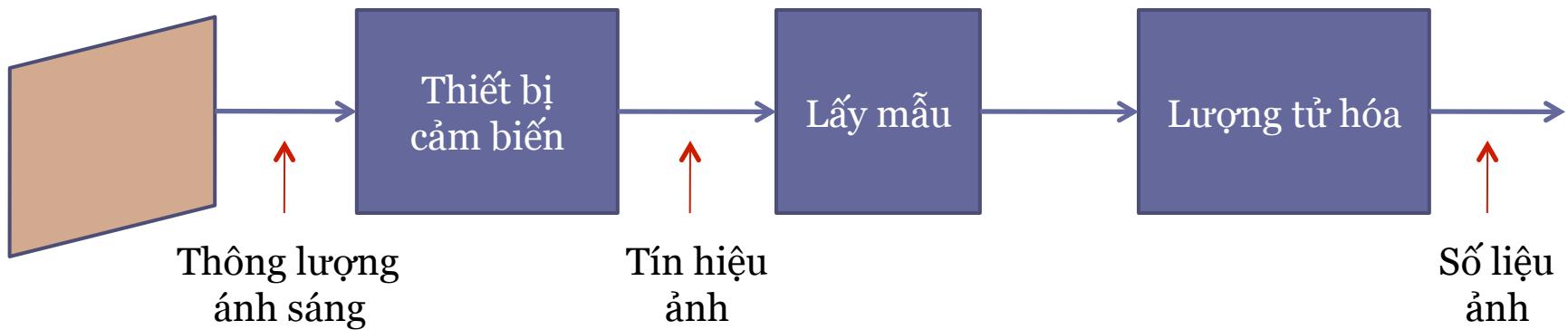
- **PSNR**

$$PSNR = 10 \log_{10} \frac{M^2}{MSE}$$

THU NHẬN VÀ SỐ HÓA ẢNH

Trần Quang Đức

Hệ thống số hóa ảnh



Thiết bị cảm biến: Charge Coupled Device >4000x4000 phần tử.

Lấy mẫu được xây dựng dựa trên tỷ số Nyquist

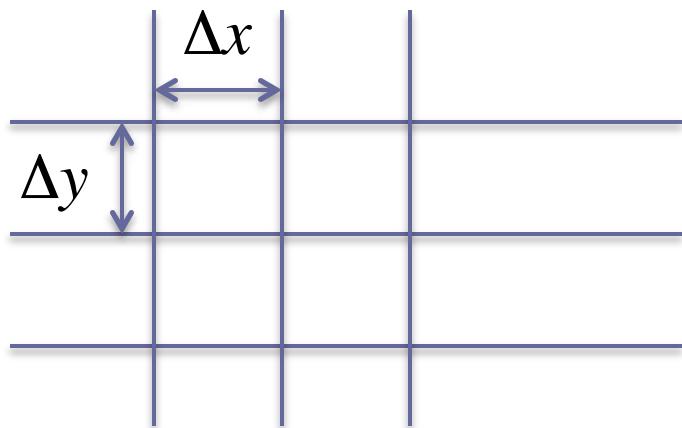
Lượng tử hóa: Lượng tử hóa Lloyd-Max

Lấy mẫu

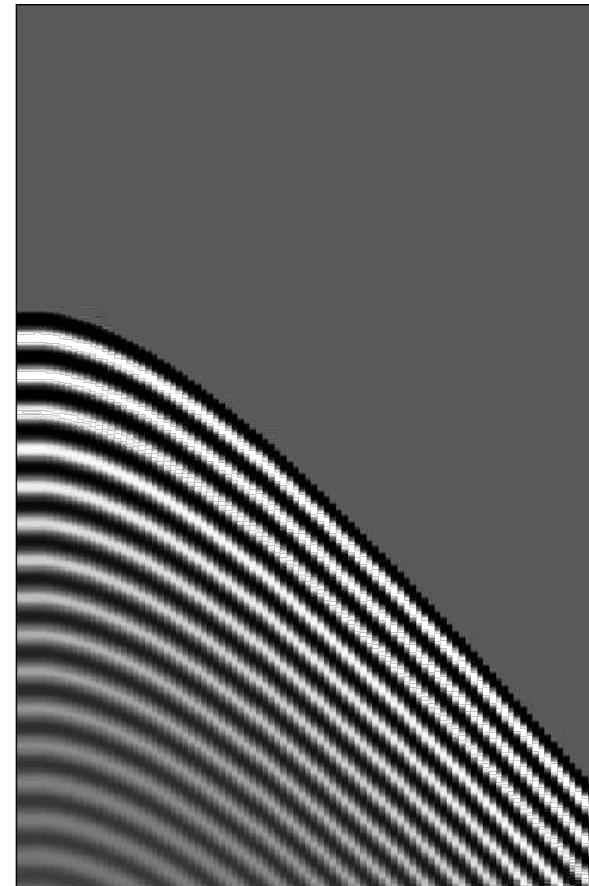
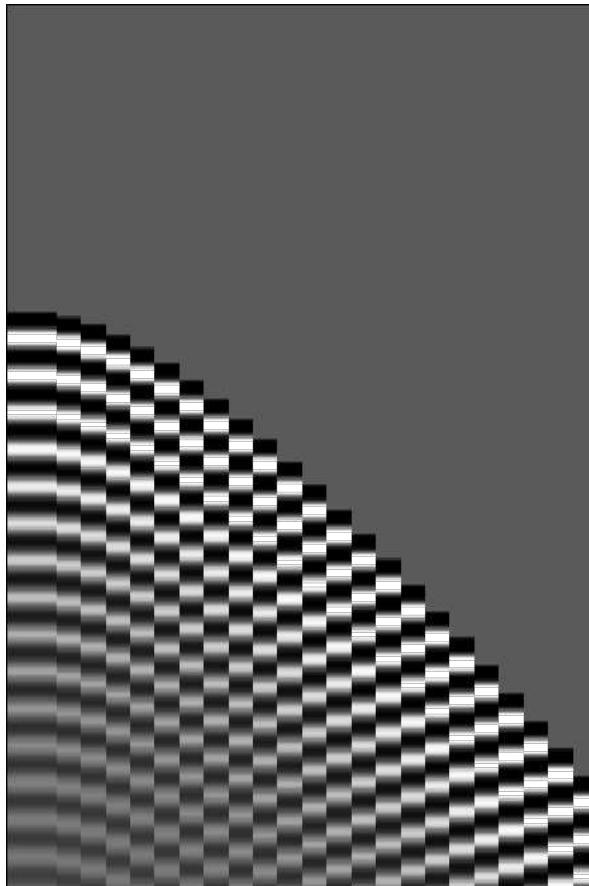
- Ảnh với dải giới hạn (Band limited images): Một hàm $f(x,y)$ gọi là giải giới hạn nếu khai triển Fourier $F(u,v)$ của nó là 0 bên ngoài miền bao, cụ thể $F(u,v) = 0$ nếu $|u| > u_o$ và $|v| > v_o$.
- Tỷ số Nyquist: Tần số lấy mẫu theo x và y tương đương hoặc lớn hơn 2 lần dải giới hạn u_o và v_o

$$\frac{1}{\Delta x} \geq 2u_o$$

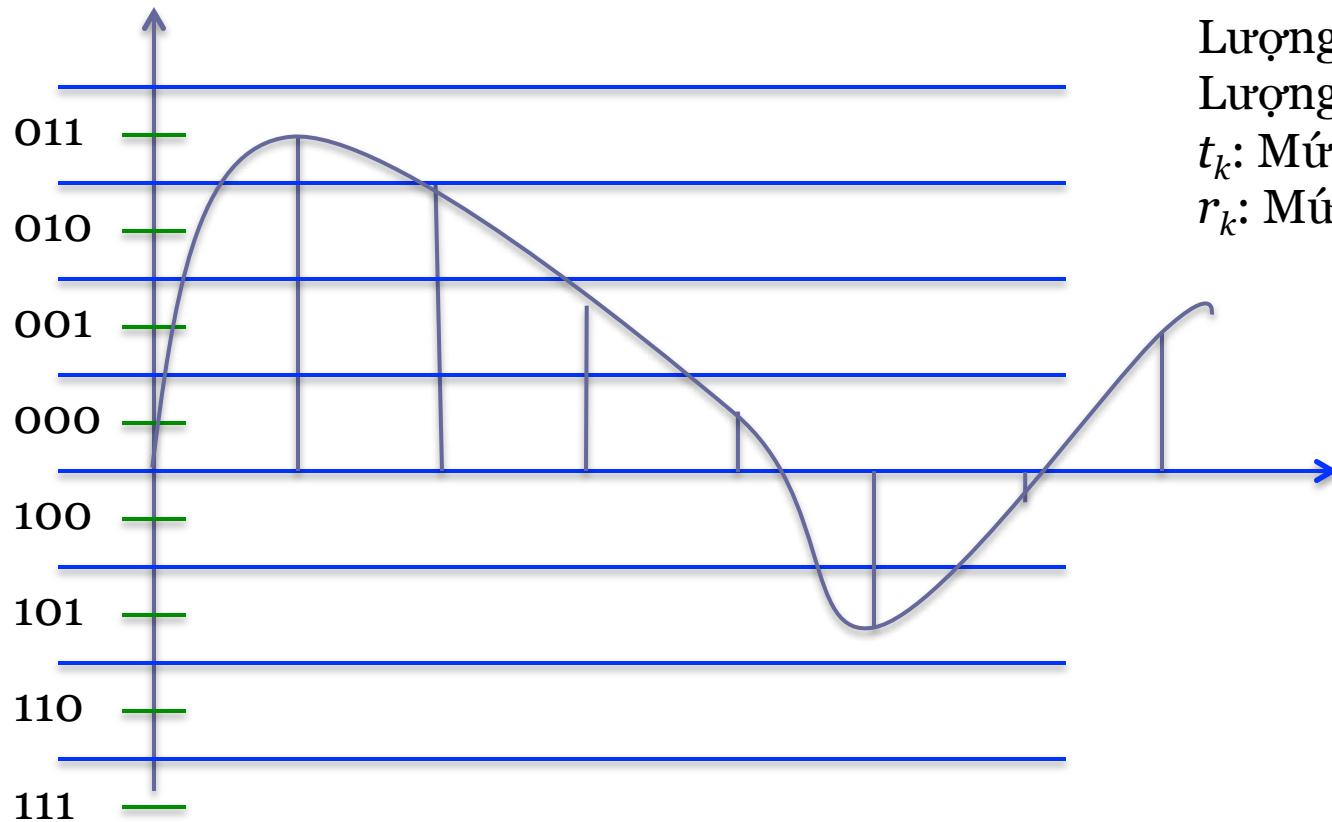
$$\frac{1}{\Delta y} \geq 2v_o$$



Hiện tượng răng cưa



Lượng tử hóa



Lượng tử hóa đều
Lượng tử hóa Lloyd-Max
 t_k : Mức quyết định
 r_k : Mức dựng lại

Lượng tử hóa Lloyd-Max

- u là một biến thực ngẫu nhiên với hàm mật độ liên tục $p(u)$. Giá trị t_k và r_k của bộ lượng tử hóa L mức được xác định sao cho sai số trung bình bình phương là nhỏ nhất.

$$\varepsilon = \text{E}[(u - r_i)^2] = \int_{t_1}^{t_L} [u - r_i]^2 p(u) du$$

$$\varepsilon = \sum_{i=1}^{L-1} \int_{t_i}^{t_{i+1}} [u - r_i]^2 p(u) du$$

Lượng tử hóa Lloyd-Max

- Để tính r_k , ta thực hiện đạo hàm cấp một theo t_k và r_k , và cho giá trị đạo hàm bằng 0.

$$(t_k - r_{k-1})^2 p(t_k) - (t_k - r_k)^2 p(t_k) = 0$$

$$2 \int_{t_k}^{t_{k+1}} [u - r_k] p(u) du = 0$$

- Từ đó, ta có

$$t_k = \frac{r_k + r_{k-1}}{2}$$

$$r_k = \frac{\int_{t_k}^{t_{k+1}} u p(u) du}{\int_{t_k}^{t_{k+1}} p(u) du}$$

Lượng tử hóa đều

- Lượng tử hóa đều là lý tưởng với hàm mật độ liên tục đều. Mức quyết định và mức khôi phục được tính như sau

$$p(u) = \begin{cases} \frac{1}{b-a} & a \leq u \leq b \\ 0 & otherwise \end{cases}$$
$$t_k = \frac{r_k + r_{k-1}}{2}$$
$$r_k = \frac{t_{k+1} + t_k}{2}$$

- Sai số lượng tử

$$\frac{q}{12}$$

$$q = r_k - r_{k-1} = t_{k+1} - t_k$$

Thay đổi kích thước ảnh

- **Giảm kích thước ảnh**

- Giảm số lượng mẫu có thể dẫn đến hiện tượng răng cưa nếu tần số lấy mẫu không thỏa mãn điều kiện của tỷ số Nyquist.
- Sử dụng bộ lọc thông thấp để loại bỏ những thành phần tần số cao (thành phần có tần số không thỏa mãn điều kiện về tỷ số Nyquist). Tuy nhiên sử dụng bộ lọc thông thấp có thể dẫn đến mờ ảnh.

- **Tăng kích thước ảnh**

- Tăng kích thước ảnh bằng các phương pháp nội suy như Nearest Neighbor, Bilinear hoặc Bicubic.

Tăng kích thước ảnh

- **Nearest Neighbor**
- **Bilinear**

$$F[m, n'] = (1 - a)I[m, n] + aI[m, n + 1]$$

$$O[m', n'] = (1 - b)F[m, n'] + bF[m + 1, n']$$

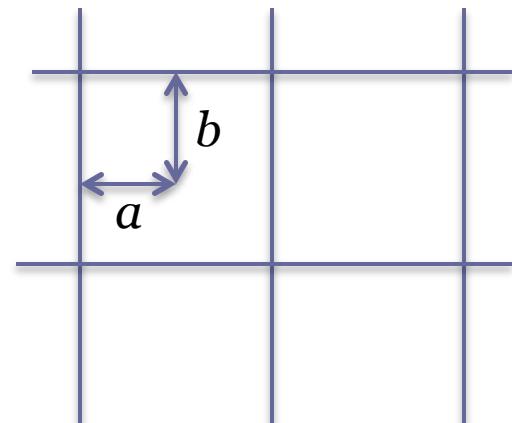
- **Bicubic**

$$F[m, n'] = -a(1 - a)^2 I[m, n - 1] + (1 - 2a^2 + a^3)I[m, n]$$

$$+a(1 + a - a^2)I(m, n + 1) - a^2(1 - a)I[m, n + 2]$$

$$O[m', n'] = -b(1 - b)^2 F[m - 1, n'] + (1 - 2b^2 + b^3)F[m, n']$$

$$+b(1 + b - b^2)F(m + 1, n') - b^2(1 - b)F[m + 2, n']$$



CƠ SỞ LÝ THUYẾT XỬ LÝ ẢNH

Trần Quang Đức

Tổng quan

- Toán tử tuyến tính
- Tích chập (2D Convolution)
- Biến đổi Fourier (Fourier Transform)
- Biến đổi KL (Karhunen-Loeve Transform)
- Biến đổi Wavelet (Wavelet Transform)

Toán tử tuyến tính

- Xử lý điểm ảnh bằng ánh xạ biến đổi

$$g(m,n) = T(f(m,n)), \quad m \in [1, M] \text{ và } n \in [1, N]$$

- Hệ thống được gọi là tuyến tính khi và chỉ khi tổ hợp tuyến tính của hai tín hiệu đầu vào $f_1(m,n)$ và $f_2(m,n)$ tạo nên tổ hợp tuyến tính của tín hiệu đầu ra $g_1(m,n)$ và $g_2(m,n)$.

$$\begin{aligned} T(\alpha_1 f_1(m,n) + \alpha_2 f_2(m,n)) &= \alpha_1 T(f_1(m,n)) + \alpha_2 T(f_2(m,n)) \\ &= \alpha_1 g_1(m,n) + \alpha_2 g_2(m,n) \end{aligned}$$

Tích chập

- Tích chập - Trường hợp liên tục

$$\begin{aligned} g(x, y) &= f(x, y) * h(x, y) \\ &= \int_{\tau_1=-\infty}^{+\infty} \int_{\tau_2=-\infty}^{+\infty} f(\tau_1, \tau_2) h(x - \tau_1, y - \tau_2) d\tau_1 d\tau_2 \end{aligned}$$

- Tích chập - Trường hợp rời rạc

$$\begin{aligned} g[x, y] &= f[x, y] * h[x, y] \\ &= \sum_{n_1=-\infty}^{+\infty} \sum_{n_2=-\infty}^{+\infty} f[n_1, n_2] h[x - n_1, y - n_2] \end{aligned}$$

Tích chập

- Tính giao hoán

$$f(x, y) * g(x, y) = g(x, y) * f(x, y)$$

- Tính kết hợp

$$f(x, y) * [g(x, y) * h(x, y)] = [f(x, y) * g(x, y)] * h(x, y)$$

- Tính phân phối

$$f(x, y) * [g(x, y) + h(x, y)] = [f(x, y) * g(x, y)] + [f(x, y) * h(x, y)]$$

Biến đổi Fourier

- Biến đổi Fourier rời rạc cho tín hiệu một chiều

$$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-2\pi j ux / M}$$
$$f(x) = \sum_{u=0}^{M-1} F(u) e^{2\pi j ux / M}$$

$$e^{j\theta} = \cos\theta + j\sin\theta$$

- Biến đổi Fourier rời rạc cho tín hiệu hai chiều

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi j (ux/M + vy/N)}$$

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi j (ux/M + vy/N)}$$

Biến đổi Fourier

- Nếu $f(x,y)$ chỉ chứa giá trị thực, thì phổ của $f(x,y)$ sẽ đối xứng qua gốc tọa độ

$$F(u,v) = F^*(-u,-v) \quad |F(u,v)| = |F(-u,-v)|$$

- $F(u,v)$ là hàm tuần hoàn với chu kỳ M và N

$$F(u,v) = F(u+M,v) = F(u,v+N) = F(u+M,v+N)$$

- Liên hệ với tích chập

$$f(x,y)^* g(x,y) \Leftrightarrow F(u,v)G(u,v)$$

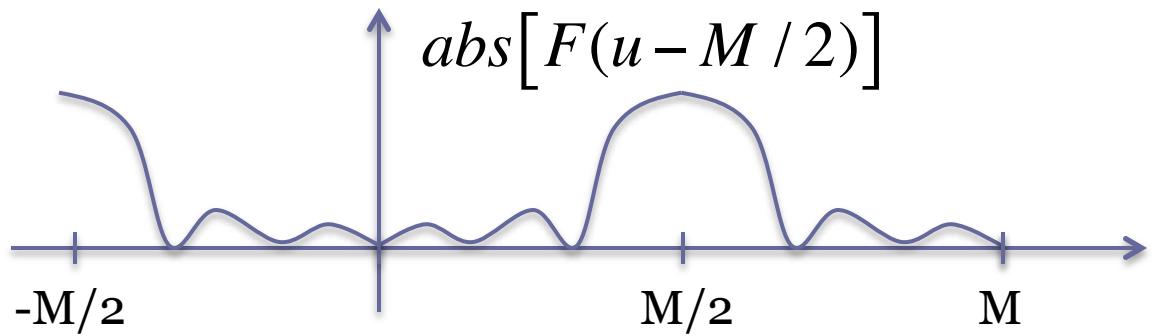
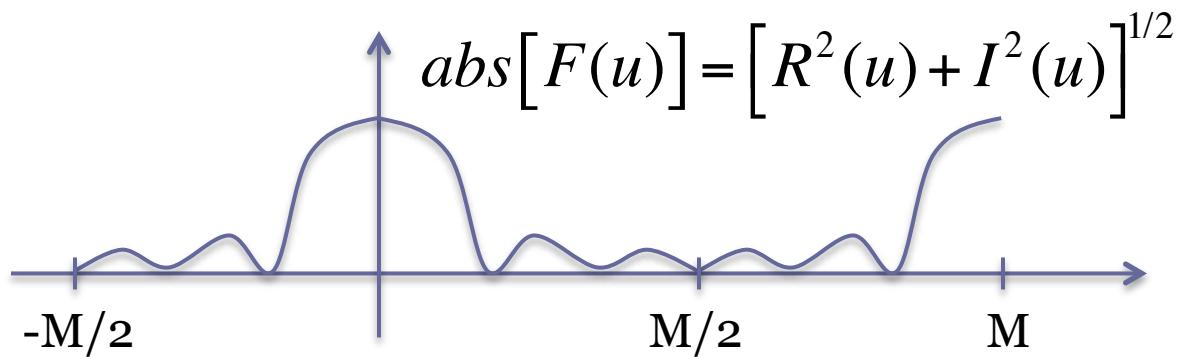
$$f(x,y)g(x,y) \Leftrightarrow F(u,v)^* G(u,v)$$

Biến đổi Fourier

$$P(u) = [F(u)]^2$$

$$\phi(u) = \tan^{-1} \left[\frac{I(u)}{R(u)} \right]$$

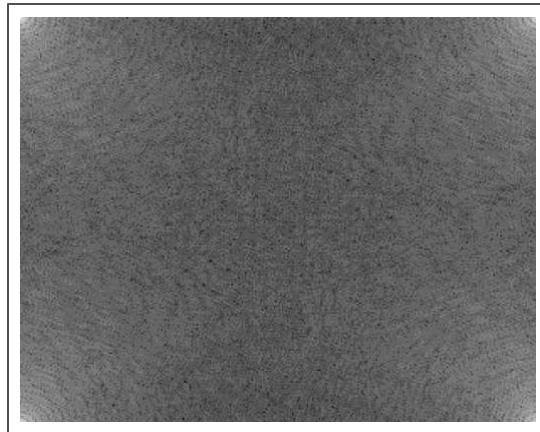
$$\Im[f(x)(-1)^x]$$



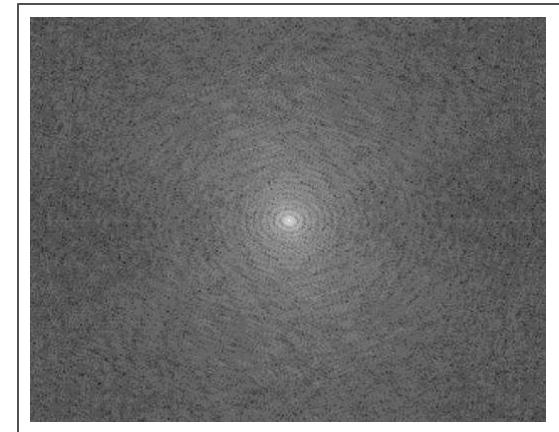
Biến đổi Fourier



(a)



(b)



(c)

(a) Ảnh gốc $f(x,y)$

(b) Biến đổi Fourier của $f(x,y)$, $F(u,v)$

(c) Biến đổi Fourier của $(-1)^{x+y}f(x,y)$, $F(u-M/2, v-N/2)$

$$F(u - M/2, v - N/2) = \Im[f(x, y)(-1)^{x+y}]$$

Biến đổi KL

- Biến đổi KL được sử dụng trong các hệ thống nhận dạng hoặc nén ảnh. Nó cho phép giảm lượng thông tin dư thừa trong ảnh thông qua quá trình giảm tính tương quan giữa các điểm ảnh nằm lân cận nhau. Cho một khối ảnh gồm N điểm ảnh. Gọi \mathbf{x} là véc tơ chứa giá trị của các điểm ảnh trên

$$\mathbf{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{Bmatrix} \quad m_x = E[\mathbf{x}]$$

Biến đổi KL

- Từ đó, ta có ma trận hiệp phương sai

$$C_x = E[(\mathbf{x} - \mathbf{m}_x)(\mathbf{x} - \mathbf{m}_x)^T] = E[\mathbf{xx}^T] - E[\mathbf{m}_x \mathbf{m}_x^T]$$

- Giá trị $c_{i,j}$ của C_x thể hiện mối tương quan giữa hai điểm ảnh x_i và x_j . Gọi \mathbf{U} là ma trận của biến đổi tuyến tính $\mathbf{y} = \mathbf{U}^T(\mathbf{x} - \mathbf{m}_x)$ sao cho ma trận hiệp phương sai C_y là ma trận chéo.
- Giả sử u_i là véc tơ cơ sở của không gian biến đổi. Các phần tử của \mathbf{y} có thể xác định như sau

$$y_i = u_i^T (\mathbf{x} - \mathbf{m}_x)$$

Biến đổi KL

- Để giảm hoàn toàn tính tương quan giữa các điểm ảnh trong không gian biến đổi, \mathbf{U} phải là ma trận trực giao

$$\mathbf{U}^T \mathbf{U} = \mathbf{I} = \mathbf{U} \mathbf{U}^T$$

- Từ đó, có thể chứng minh được năng lượng của khối ảnh được bảo toàn trước và sau khi thực hiện biến đổi

$$\begin{aligned}\|\mathbf{y}\|^2 &= \mathbf{y}^T \mathbf{y} \\ &= (\mathbf{x} - \mathbf{m}_x)^T \mathbf{U} \mathbf{U}^T (\mathbf{x} - \mathbf{m}_x) \\ &= \|\mathbf{x} - \mathbf{m}_x\|^2\end{aligned}$$

Biến đổi KL

- Do ma trận hiệp phương sai C_y là ma trận chéo

$$C_y = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}$$
$$\begin{aligned} C_y &= E[(\mathbf{y} - \mathbf{m}_y)(\mathbf{y} - \mathbf{m}_y)^T] = E[\mathbf{y}\mathbf{y}^T] \\ &= E[\mathbf{U}^T(\mathbf{x} - \mathbf{m}_x)(\mathbf{x} - \mathbf{m}_x)^T\mathbf{U}] \\ &= \mathbf{U}^T C_x \mathbf{U} \end{aligned}$$

- Có thể thấy rằng

$$C_x \mathbf{U} = \mathbf{U} C_y \Rightarrow C_x u_i = \lambda_i u_i$$

Biến đổi KL

- Trong biểu thức trên, u_i là véc tơ riêng của C_x và λ_i là giá trị riêng tương ứng của u_i . Véc tơ \mathbf{x} có thể khôi phục dựa trên phép biến đổi

$$\mathbf{y} = \mathbf{U}^T (\mathbf{x} - m_x) \Rightarrow \mathbf{x} = \mathbf{U}\mathbf{y} + m_x$$

- Quá trình nén ảnh chỉ lưu K ($K < N$) phần tử đầu tiên của \mathbf{y} . Véc tơ \mathbf{x} ban đầu có thể được xấp xỉ như sau

$$\tilde{\mathbf{x}} = \mathbf{U}\tilde{\mathbf{y}} + m_x$$

- Với

$$\tilde{\mathbf{y}} = [y_1, y_2, \dots, y_K, 0, \dots, 0]^T$$

Biến đổi KL

- Trung bình bình phương của sai số được xác định bởi

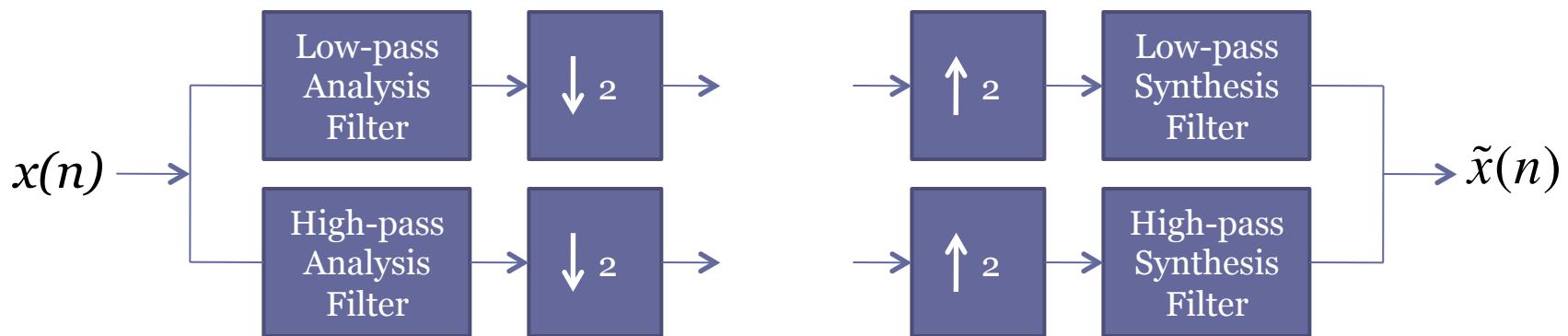
$$\begin{aligned} E\left[\frac{1}{N} \sum_{i=1}^N (x_i - \tilde{x}_i)^2\right] &= \frac{1}{N} E[|\mathbf{x} - \tilde{\mathbf{x}}|^2] \\ &= \frac{1}{N} E[|\mathbf{Uy} - \mathbf{U}\tilde{\mathbf{y}}|^2] \\ &= \frac{1}{N} E[|\mathbf{y} - \tilde{\mathbf{y}}|^2] = \frac{1}{N} \sum_{i=K+1}^N \lambda_i^2 \end{aligned}$$

Biến đổi KL

- Năng lượng của khối ảnh trước và sau khi biến đổi được bảo toàn.
- Ma trận hiệp phương sai C_y là ma trận chéo. Do vậy các điểm ảnh trong không gian biến đổi không tương quan với nhau. Lượng thông tin dư thừa trong khối ảnh ban đầu bị giảm một cách tối đa.
- Nén ảnh dựa trên biến đổi KL có tỷ lệ nén cao. Tuy nhiên việc tính toán véc tơ riêng và giá trị riêng tương ứng rất phức tạp với kích thước khối ảnh lớn. Biến đổi KL cũng phụ thuộc vào đặc trưng riêng của từng ảnh.

Biến đổi Wavelet

- Biến đổi Wavelet (Wavelet Transform) được sử dụng trong xử lý đa phân giải và nén ảnh JPEG2000. Biến đổi Wavelet dựa trên quá trình mã hóa dải tần con (sub-band coding).



$h_o(n)$: Low-pass Analysis Filter
 $h_i(n)$: High-pass Analysis Filter

$g_o(n)$: Low-pass Synthesis Filter
 $g_i(n)$: High-pass Synthesis Filter

Biến đổi Wavelet

- Định nghĩa biến đổi Z:

$$X(z) = \sum_{-\infty}^{+\infty} x(n)z^{-n}$$

- Giảm mẫu 2 lần trên miền thời gian

$$x_{down}(n) = x(2n) \Leftrightarrow X_{down}(z) = \frac{1}{2} [X(z^{1/2}) + X(-z^{1/2})]$$

- Tăng mẫu 2 lần trên miền thời gian

$$x_{up}(n) = \begin{cases} x(n/2) & n = 0, 2, 4, \dots \\ 0 & Otherwise \end{cases} \Leftrightarrow X_{up}(z) = X(z^2)$$

Biến đổi Wavelet

- Giảm mẫu và tăng mẫu trên miền thời gian:

$$\hat{X}(z) = \frac{1}{2} [X(z) + X(-z)]$$

- Quá trình mã hóa bằng tần con có thể được minh họa như sau

$$\hat{X}(z) = \frac{1}{2} G_0(z) [H_0(z)X(z) + H_0(-z)X(-z)]$$

$$+ \frac{1}{2} G_1(z) [H_1(z)X(z) + H_1(-z)X(-z)]$$

$$\hat{X}(z) = \frac{1}{2} [H_0(z)G_0(z) + H_1(z)G_1(z)]X(z)$$

$$+ \frac{1}{2} [H_0(-z)G_0(z) + H_1(-z)G_1(z)]X(-z)$$

Biến đổi Wavelet

- Từ đó, ta có

$$H_0(-z)G_0(z) + H_1(-z)G_1(z) = 0$$

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 2$$

- Biểu diễn dưới dạng nhận ma trận

$$\begin{bmatrix} G_0(z) & G_1(z) \end{bmatrix} \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} = \begin{bmatrix} 2 & 0 \end{bmatrix}$$

- Quadrature mirror filters (QMF)

$$H_1(z) = H_0(-z) \quad G_0(z) = H_0(z) \quad G_1(z) = -H_0(-z)$$

CẢI THIỆN VÀ NÂNG CẤP ẢNH

Trần Quang Đức

Tổng quan

Cải thiện và nâng cấp ảnh

Toán tử trên điểm ảnh

Ảnh âm bản
Biến đổi sử dụng hàm log
Biến đổi sử dụng hàm mũ
Dãn và co độ tương phản
Cắt theo mức
Trích chọn bit
Cân bằng Histogram
Histogram đặc trưng
Trù ảnh
Trung bình ảnh

Toán tử trên miền không gian

Lọc trung bình
Lọc trung bình trọng số
Lọc trung vị
Lọc cực đại
Lọc cực tiểu
Lọc sắc nét
Lọc High-boost

Toán tử trên miền tần số

Lọc thông thấp
Lọc thông cao
Lọc High-boost
Laplace trên miền tần số

TOÁN TỬ TRÊN ĐIỂM ẢNH

Trần Quang Đức

Phác họa nội dung

Cải thiện và nâng cấp ảnh

Toán tử trên điểm ảnh

Ảnh âm bản
Biến đổi sử dụng hàm log
Biến đổi sử dụng hàm mũ
Dãn và co độ tương phản
Cắt theo mức
Trích chọn bit
Cân bằng Histogram
Histogram đặc trưng
Trù ảnh
Trung bình ảnh

Toán tử trên miền không gian

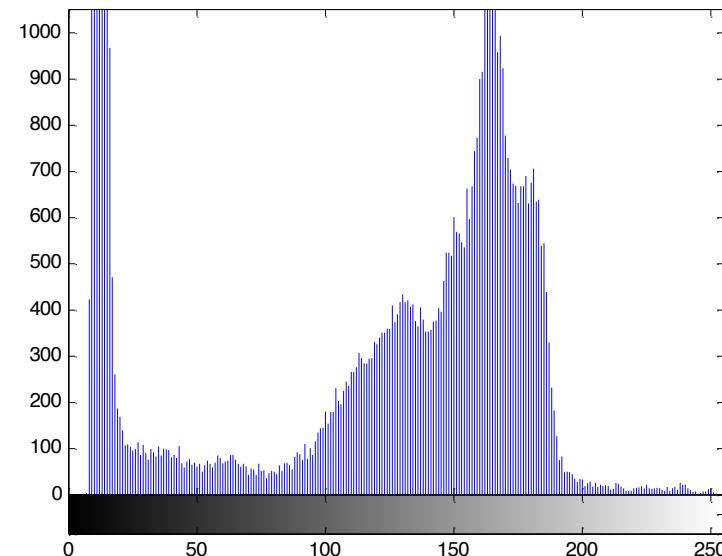
Lọc trung bình
Lọc trung bình trọng số
Lọc trung vị
Lọc cực đại
Lọc cực tiểu
Lọc sắc nét
Lọc High-boost

Toán tử trên miền tần số

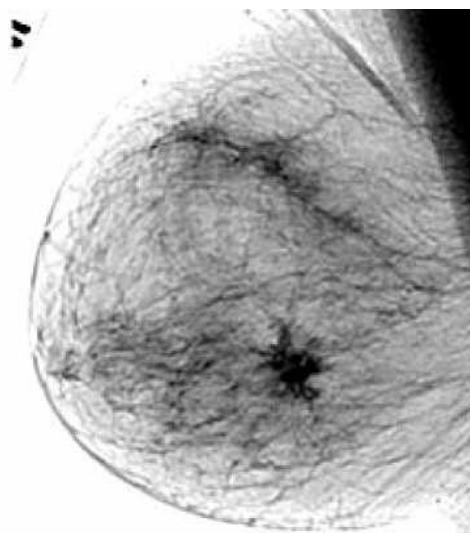
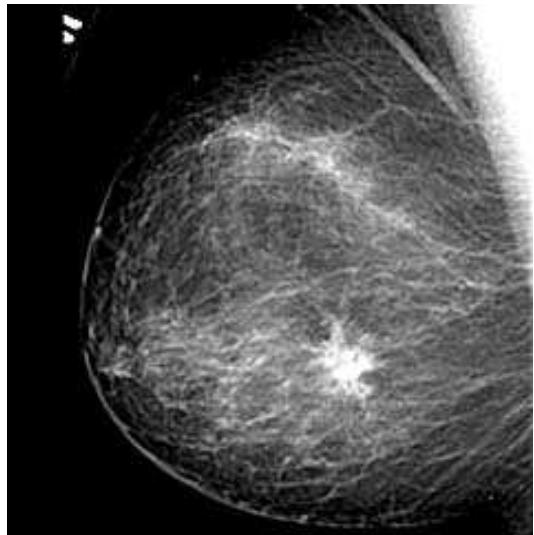
Lọc thông thấp
Lọc thông cao
Lọc High-boost
Laplace trên miền tần số

Định nghĩa Histogram

- Histogram (Lược đồ mức xám) biểu diễn bằng một hàm rời rạc $h(r_k)=n_k$, trong đó r_k là mức xám thứ k và n_k là số lượng điểm ảnh có mức xám r_k . Thông thường histogram được tiêu chuẩn hóa bằng cách chia $h(rk)$ cho n , với $n=\sum n_k$.



Ảnh âm bản



$$g(m, n) = L - 1 - f(m, n)$$

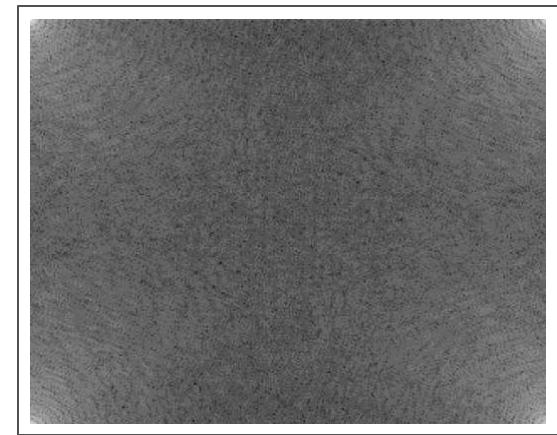
Biến đổi sử dụng hàm log



(a)



(b)



(c)

(a) Ảnh gốc $f(x,y)$

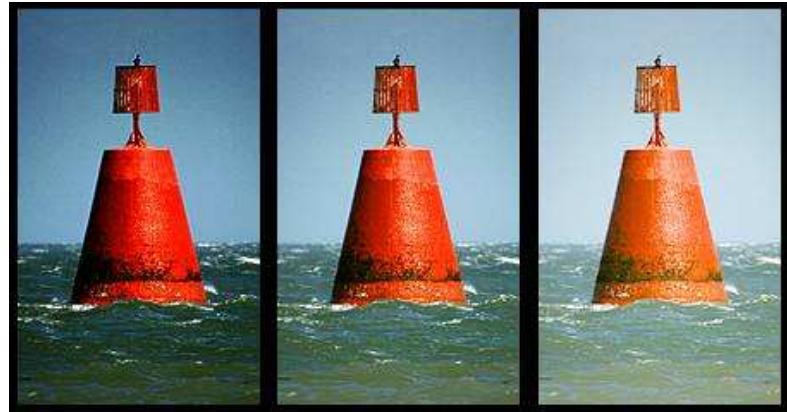
(b) Biến đổi Fourier của $f(x,y)$, $F(u,v)$

(c) Biến đổi sử dụng hàm log ($c=1$)

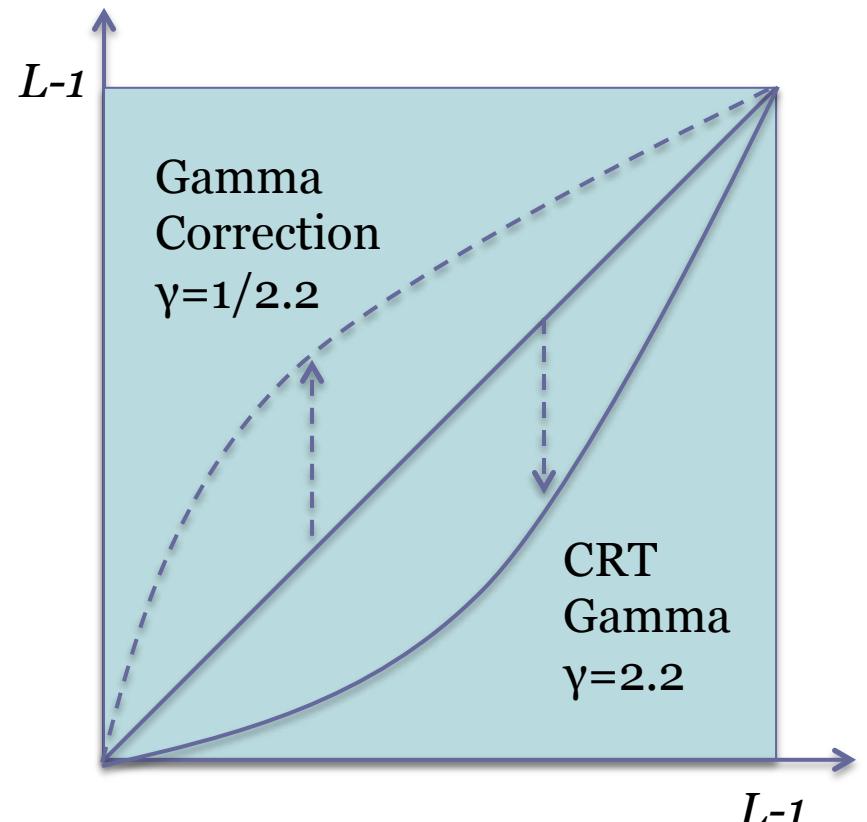
$$g(m,n) = c \log[1 + f(m,n)]$$

$$G(u,v) = c \log[1 + |F(u,v)|]$$

Biến đổi sử dụng hàm mũ



$$g(m, n) = cf(m, n)^{\gamma}$$

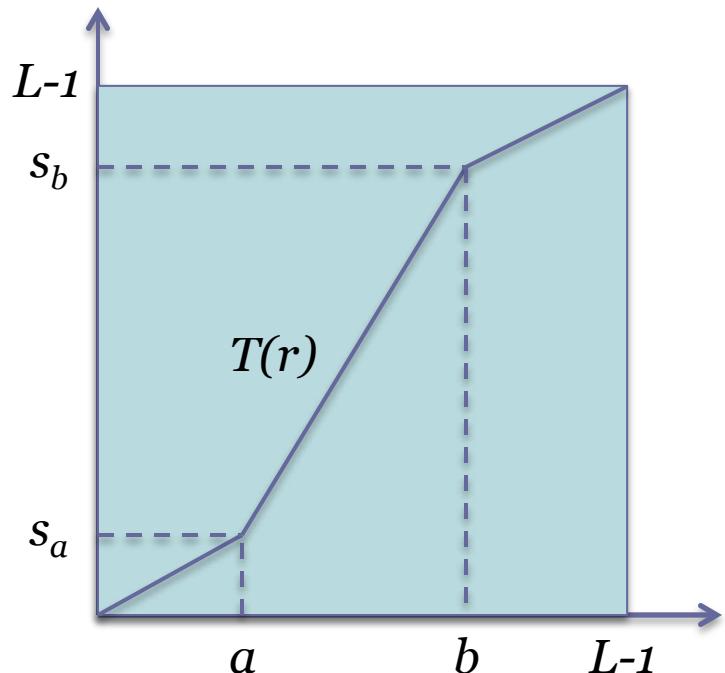


$$g(m, n) = cf(m, n)^{1/\gamma}$$

Dãn và co độ tương phản

$$T(r) = \begin{cases} ar & r < a \\ \beta(r - a) + s_a & a \leq r < b \\ \gamma(r - b) + s_b & b \leq r \end{cases}$$

Dải tần nhạy sáng là giới hạn dải mức xám xuất hiện trong ảnh. Dải tần nhạy sáng được đo bằng độ tương phản.



Ảnh có độ tương phản thấp, dải tần nhạy sáng hẹp, việc chuyển tiếp giữa các sắc mức xám không được sống động và linh hoạt. Hàm $T(r)$ được dùng để tăng độ tương phản. Ví dụ, với $a, \beta, \gamma > 1$, $T(r)$ thực hiện dãn độ tương phản.

Dãn và co độ tương phản



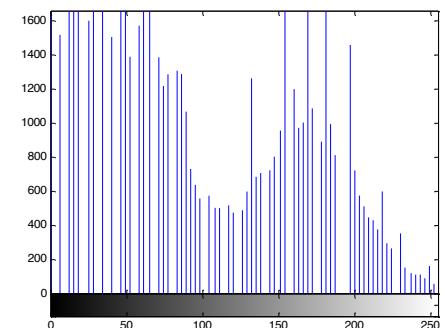
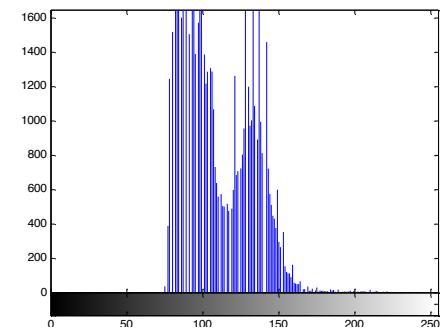
a



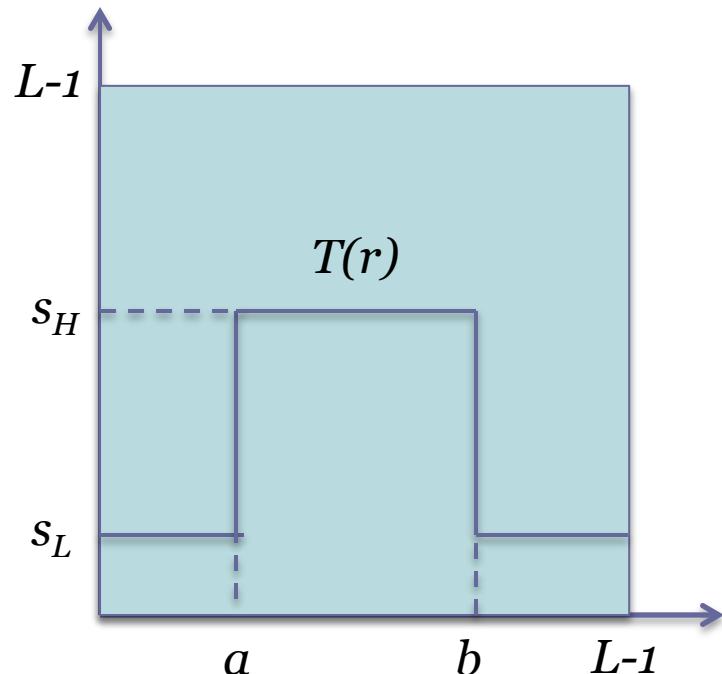
b

Ảnh (b) thu được sau khi thực hiện tăng cường độ tương phản ảnh (a). Histogram của ảnh (a) và (b).

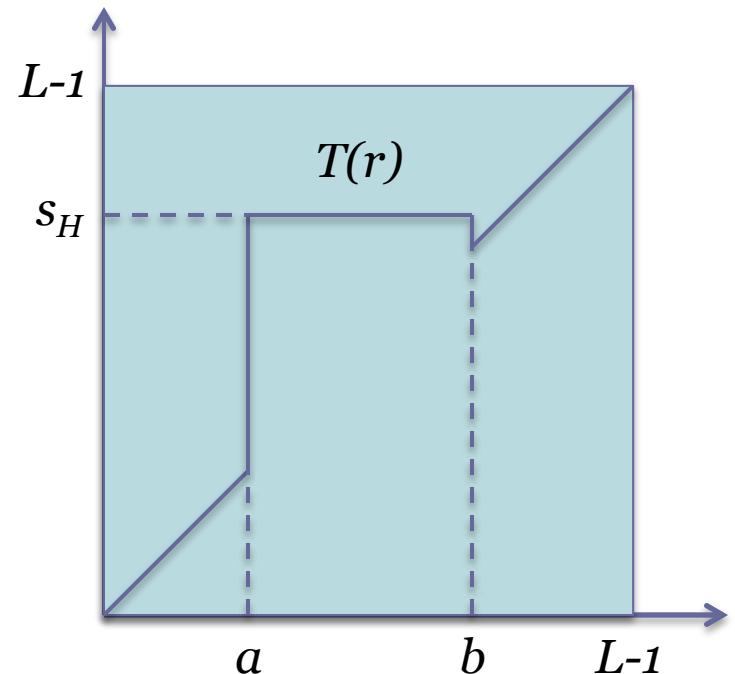
Ảnh (a) có mức xám trong dải [74,224], ảnh (b) có mức xám trong dải [0,255].



Cắt theo mức



$$T(r) = \begin{cases} s_H & a \leq r \leq b \\ s_L & Otherwise \end{cases}$$



$$T(r) = \begin{cases} s_H & a \leq r \leq b \\ r & Otherwise \end{cases}$$

Cắt theo mức



a



b



c

- a) Ảnh gốc
- b) Cắt theo mức không nền với $a=100$ và $b=200$
- c) Cắt theo mức có nền với $a=100$ và $b=200$

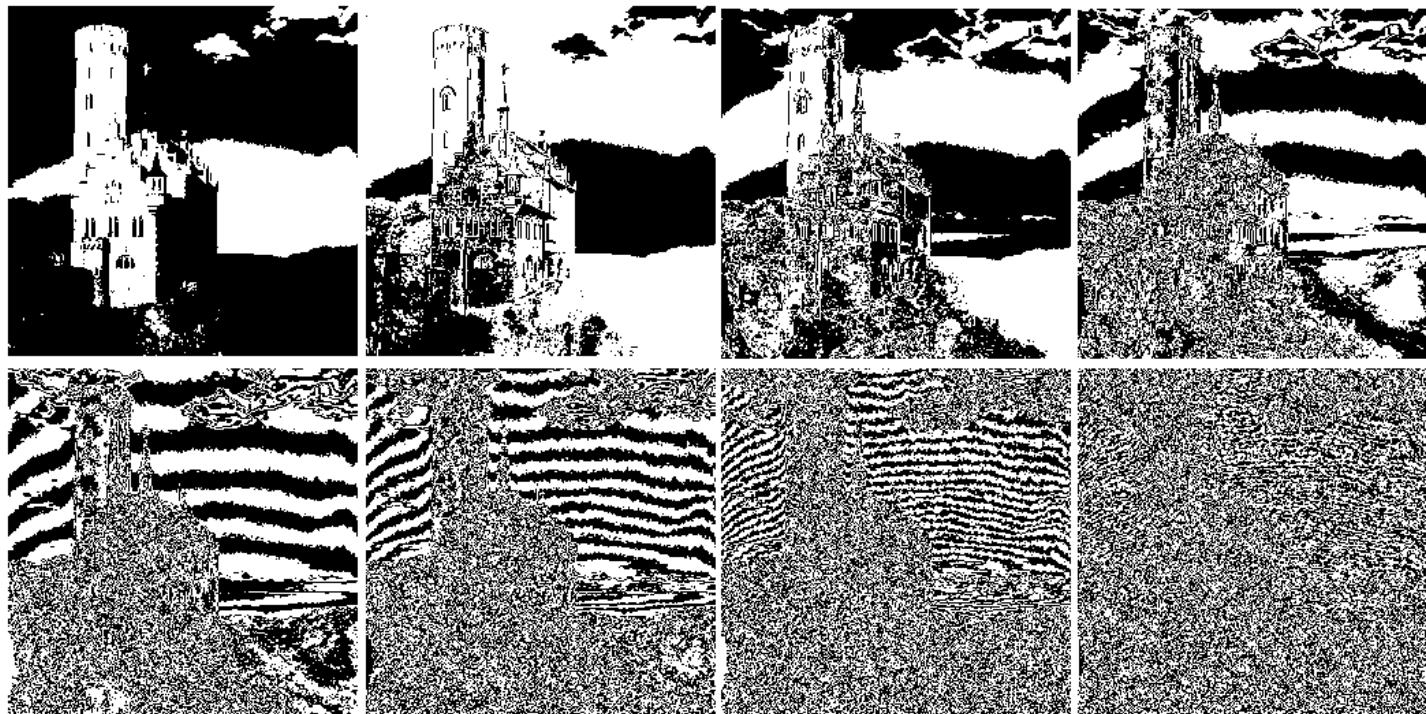
Trích chọn bit

Mỗi điểm ảnh được biểu diễn bằng 8 mặt phẳng bit từ 0 đến 7, xếp chồng lên nhau theo thứ tự từ cao xuống thấp.



Những mặt phẳng bit phía trên tập trung phần lớn dữ liệu ý nghĩa của ảnh. Dữ liệu chi tiết ảnh tập trung ở những mặt phẳng bit phía dưới. Phân tích sử dụng phương pháp trích chọn bit cho phép xác định những mặt phẳng bit quan trọng và số lượng bit yêu cầu của quá trình lượng tử hóa.

Trích chọn bit



Minh họa 8 mặt phẳng bit. (a), (b), (c), (d), (e), (f),
(g), (h) tương đương mặt phẳng bit từ 7 đến 0

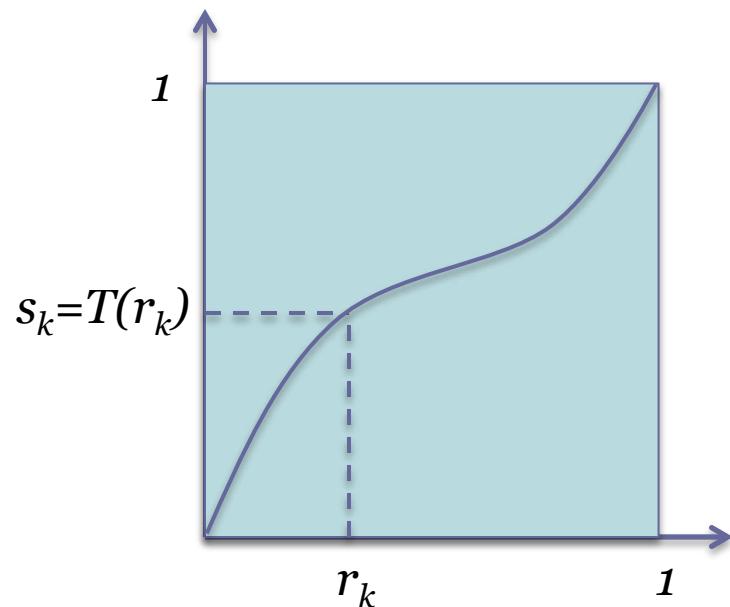
a b c d
e f g h

Cân bằng Histogram

- Giả sử r được chuẩn hóa $0 \leq r \leq 1$, $s = T(r)$ với $T(r)$ là hàm một biến đơn điệu tăng và $0 \leq T(r) \leq 1$. Điều kiện trên đảm bảo sự tồn tại của ánh xạ ngược của $r = T^{-1}(s)$ và $0 \leq s \leq 1$.

Nếu p_s và p_r là hàm mật độ xác suất của hai biến s và r . Do số lượng điểm ảnh là không đổi, nên $p_s(s)ds = p_r(r)dr$. Giả sử

$$s = T(r) = \int_0^r p_r(w)dw$$



Cân bằng Histogram

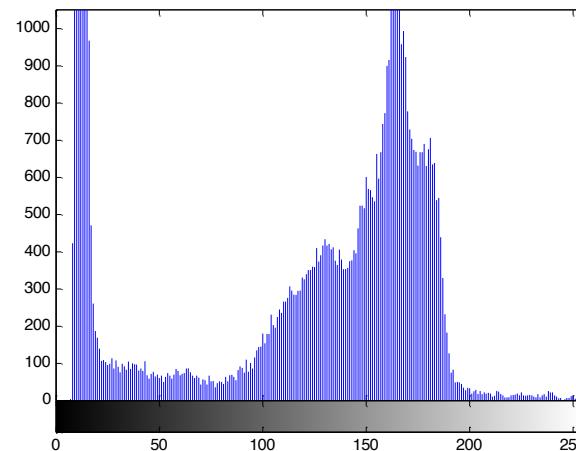
- Như vậy $ds = p_r(r)dr$. Từ đó, $p_s(s) = 1$ với $0 \leq s \leq 1$ và là hàm phân bố đều. Cân bằng histogram cho phép tạo ra histogram mới có phân bố đều.

$$p_r(r_k) = \frac{n_k}{n} \quad s_k = T(r_k) = \sum_{j=1}^k p_r(r_j) = \sum_{j=1}^k \frac{n_j}{n}$$

- Trên thực tế, giá trị mức xám mới được xác định như sau

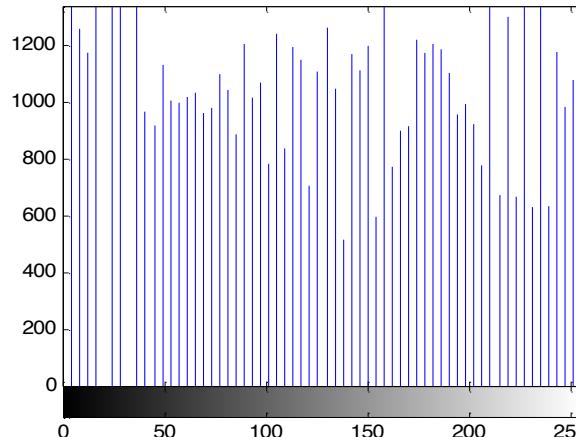
$$s_k = round\left(\frac{s_k - \min\{s_k\}}{\max\{s_k\} - \min\{s_k\}}(L-1)\right)$$

Cân bằng Histogram



a b
c d

- (a) Ảnh gốc
- (b) Histogram của ảnh gốc
- (c) Ảnh với cân bằng Histogram
- (d) Histogram của ảnh (c)



Histogram đặc trưng

- Phương pháp histogram đặc trưng được áp dụng để tạo ảnh đầu ra với histogram có hình dạng cho trước. Giả sử p_r và p_z là hàm mật độ xác suất của ảnh đầu vào và ảnh đầu ra. Gọi $T(r)$ và $G(z)$ là phép biến đổi sao cho

$$s = T(r) = \int_0^r p_r(w)dw \quad G(z) = \int_0^z p_z(t)dt = s$$

- $T(r)$ thực hiện cân bằng histogram của ảnh đầu vào, trong khi $G(z)$ có thể xác định dựa trên p_z . Từ đó, ta có

$$s = T(r) = G(z) \Rightarrow z = G^{-1}(T(r))$$

Histogram đặc trưng

- 1) Xây dựng histogram của ảnh đầu vào
- 2) Thực hiện cân bằng histogram của ảnh đầu vào
- 3) Xác định $G(z)$ dựa trên p_z
- 4) Tìm giá trị z_k nhỏ nhất ứng với s_k sao cho
$$[G(z_k) - s_k] \geq 0$$
- 5) Mỗi điểm ảnh có mức xám r_k , tìm s_k tương ứng (bước (2)). Từ s_k có thể tìm z_k dựa trên kết quả của bước (4).

Trừ ảnh



$$g(x,y) = f(x,y) - h(x,y)$$

a b

c d

(a) $f(x,y)$

(b) $h(x,y)$

(c,d) $g(x,y) = f(x,y) - h(x,y)$

Trung bình ảnh

$$f(x, y) = g(x, y) + \eta(x, y)$$

$$\bar{g}(x, y) = \frac{1}{K} \sum_{i=1}^K g_i(x, y)$$



a



b

Trung bình ảnh được sử dụng giảm nhiễu bằng cách tăng tỷ lệ SNR. Ảnh (b) thu được dựa trên trung bình ảnh. Lưu ý: nhiễu là một quá trình ngẫu nhiên.

$$\mathbb{E}[\bar{g}(x, y)] = f(x, y)$$

$$\sigma_{\bar{g}(x, y)} = \frac{1}{\sqrt{K}} \sigma_{\eta(x, y)}$$

TOÁN TỬ TRÊN MIỀN KHÔNG GIAN

Trần Quang Đức

Phác họa nội dung

Cải thiện và nâng cấp ảnh

Toán tử trên điểm ảnh

Ảnh âm bản
Biến đổi sử dụng hàm log
Biến đổi sử dụng hàm mũ
Dãn và co độ tương phản
Cắt theo mức
Trích chọn bit
Cân bằng Histogram
Histogram đặc trưng
Trù ảnh
Trung bình ảnh

Toán tử trên miền không gian

Lọc trung bình
Lọc trung bình trọng số
Lọc trung vị
Lọc cực đại
Lọc cực tiểu
Lọc sắc nét
Lọc High-boost

Toán tử trên miền tần số

Lọc thông thấp
Lọc thông cao
Lọc High-boost
Laplace trên miền tần số

Nhiễu

- Ảnh thường bị biến dạng do nhiễu ngẫu nhiên. Nhiễu xuất hiện trong quá trình thu nhận ảnh hoặc truyền tin. Các yếu tố môi trường, ví dụ điều kiện ánh sáng yếu, nhiệt độ của thiết bị cảm biến cũng ảnh hưởng đến sự xuất hiện của nhiễu.
- Nhiễu có thể phụ thuộc hoặc độc lập với nội dung ảnh và thường được biểu diễn bằng các thuộc tính thống kê. Xử lý nhiễu phụ thuộc vào nội dung ảnh thường có độ phức tạp cao (**ngoài chương trình**).
- Nhiễu có thể là nhiễu trắng, nhiễu Gauss, nhiễu đầm, nhiễu xung hoặc nhiễu muối tiêu.

Nhiễu lượng tử hóa

- Nhiễu lượng tử hóa được tạo ra trong quá trình lượng tử hóa (sử dụng chuỗi nhị phân để xấp xỉ giá trị biên độ của tín hiệu thu nhận). Nhiễu lượng tử có thể xấp xỉ bởi phân bố đều nên còn được gọi là nhiễu đều.

$$p(z) = \begin{cases} \frac{1}{b-a} & a \leq z \leq b \\ 0 & Otherwise \end{cases}$$

- Nhiễu lượng tử có đặc trưng sau

$$\mu = \frac{a+b}{2}$$

$$\sigma^2 = \frac{(b-a)^2}{12}$$

Nhiễu trắng

- Nhiễu trắng là nhiễu có phổ năng lượng không đổi. Nếu nhiễu có phổ năng lượng nhiều hơn ở một vài tần số, nhiễu được gọi là nhiễu màu. Hàm tương quan của nhiễu trắng là Dirac's delta. Do đó, nhiễu trắng không tương quan ở hai mảnh bất kỳ.

`imnoise(I,'localvar', V)`



Nhiễu Gauss

- Nhiễu Gauss là mô hình phổ biến để xấp xỉ nhiễu trong nhiều ứng dụng khác nhau. Mật độ phân bố xác suất của nhiễu là hàm Gauss, được đặc trưng bởi giá trị trung bình μ và phương sai σ^2 .

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/2\sigma^2}$$

$$[\mu - \sigma, \mu + \sigma]$$

$$[\mu - 2\sigma, \mu + 2\sigma]$$

imnoise(I,'gaussian',M,V)



Nhiễu Poisson

- Trong quá trình thu nhận, nếu số lượng lớn hạt photon tập trung vào một điểm, chúng sẽ tạo ra nhiễu tại điểm đó. Nhieu được đặc trưng bởi hàm mật độ phân bố xác suất Poisson, nên được gọi là nhiễu Poisson.

`imnoise(I,'poisson')`



Nhiễu muối tiêu

- **Nhiễu xung** đặc trưng bởi một điểm ảnh có giá trị mức xám khác biệt lớn so với những điểm lân cận. Xung của nhiễu có thể âm hoặc dương. Xung nhiễu âm đại diện cho điểm ảnh đen (pepper), trong khi xung nhiễu dương đại diện cho điểm ảnh trắng (salt). **Nhiễu muối tiêu** xuất hiện khi ảnh bị bão hòa bởi nhiễu xung.

imnoise(I, 'salt & pepper', 0.02)



Lọc trung bình

- Lọc trung bình

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

- Lọc thông thấp

$$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\frac{1}{(b+2)^2} \begin{bmatrix} 1 & b & 1 \\ b & b^2 & b \\ 1 & b & 1 \end{bmatrix}$$

Lọc trung bình



a



b



c

- (a) Ảnh bị nhiễu muối tiêu
- (b) Lọc trung bình
- (c) Lọc thông thấp (Lọc trung bình trọng số)

Lọc trung vị, cực tiểu và cực đại

8	89	8	8
8	7	8	8
7	7	8	8
8	8	8	8

W: 3x3, 5x5, 7x7

$$\hat{g}(x, y) = \underset{m, n \in W}{median}\{f(m, n)\}$$

$$\hat{g}(x, y) = \underset{m, n \in W}{\min}\{f(m, n)\}$$

$$\hat{g}(x, y) = \underset{m, n \in W}{\max}\{f(m, n)\}$$



min

max

Minh họa lọc trung vị



a



b



c

- (a) Ảnh bị nhiễu muối tiêu
- (b) Lọc trung vị cửa số 3x3
- (c) Lọc trung vị cửa số 7x7

Lọc sắc nét

Đạo hàm bậc nhất	Đạo hàm bậc hai
$\frac{\partial f}{\partial x} = f(x+1) - f(x)$	$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x)$

5	5	4	3	2	1	0	0	0	6	0	0	0	1	3	1	0	0	0	0	7	7	7	7
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Đạo hàm bậc nhất

0	-1	-1	-1	-1	-1	0	0	6	-6	0	0	0	1	2	-2	-1	0	0	0	7	0	0	0
---	----	----	----	----	----	---	---	---	----	---	---	---	---	---	----	----	---	---	---	---	---	---	---

Đạo hàm bậc hai

-1	0	0	0	0	0	1	0	0	6	-12	6	0	0	1	1	-4	1	1	0	0	7	-7	0	0
----	---	---	---	---	---	---	---	---	---	-----	---	---	---	---	---	----	---	---	---	---	---	----	---	---

Lọc sắc nét

$$\begin{aligned}\nabla^2 f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \\ &= f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)\end{aligned}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Lọc sắc nét

$$g(x,y) = \begin{cases} f(x,y) - \nabla^2 f(x,y) & \text{If center coefficient is negative} \\ f(x,y) + \nabla^2 f(x,y) & \text{If center coefficient is positive} \end{cases}$$

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$



$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$



Lọc High-boost

- Nguyên tắc cơ bản

$$f_{hb}(x, y) = Af(x, y) - \bar{f}(x, y) = (A - 1)f(x, y) - f_s(x, y)$$

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & A + 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$



$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & A + 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$



TOÁN TỬ TRÊN MIỀN TẦN SỐ

Trần Quang Đức

Phác họa nội dung

Cải thiện và nâng cấp ảnh

Toán tử trên điểm ảnh

Ảnh âm bản
Biến đổi sử dụng hàm log
Biến đổi sử dụng hàm mũ
Dãn và co độ tương phản
Cắt theo mức
Trích chọn bit
Cân bằng Histogram
Histogram đặc trưng
Trù ảnh
Trung bình ảnh

Toán tử trên miền không gian

Lọc trung bình
Lọc trung bình trọng số
Lọc trung vị
Lọc cực đại
Lọc cực tiểu
Lọc sắc nét
Lọc High-boost

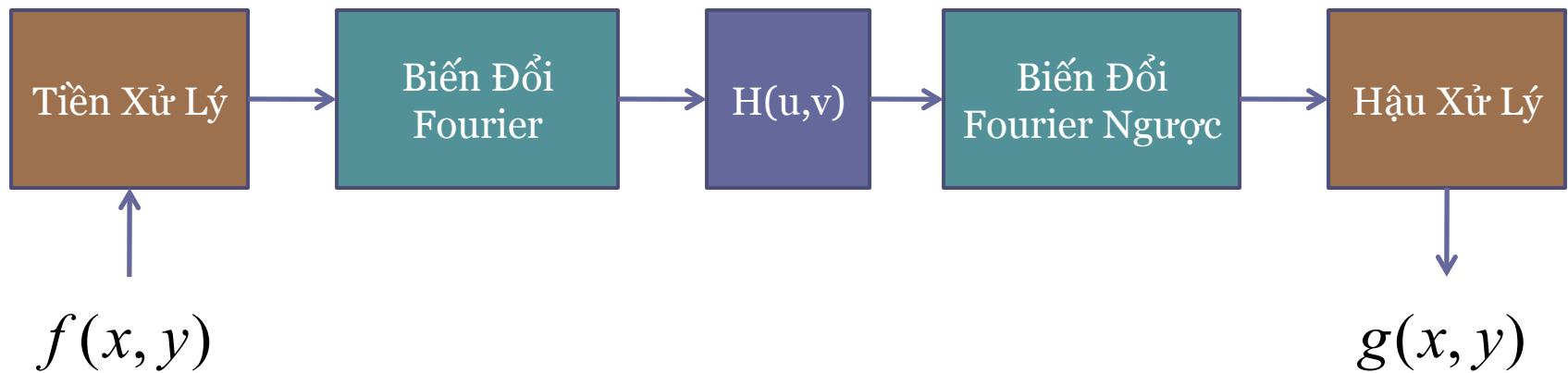
Toán tử trên miền tần số

Lọc thông thấp
Lọc thông cao
Lọc High-boost
Laplace trên miền tần số

Lọc trên miền tần số

- Những thành phần tần số thấp đại diện cho dữ liệu ảnh tại những vùng trơn mịn, trong khi những thành phần tần số cao đại diện cho dữ liệu chi tiết ảnh như biên ảnh và nhiễu.
- Lọc thông thấp chỉ cho những thành phần tần số thấp đi qua và loại bỏ những thành phần tần số cao. Vì vậy, ảnh sau khi lọc sẽ trơn mịn nhưng không được sắc nét bằng ảnh đầu vào.
- Lọc thông cao chỉ cho những thành phần tần số cao đi qua và loại bỏ những thành phần tần số thấp. Do đó, ảnh sau khi lọc sẽ có dải mức xám không đa dạng tại những vùng trơn mịn, nhưng sắc nét hơn so với ảnh đầu vào.

Hệ thống lọc trên miền tần số



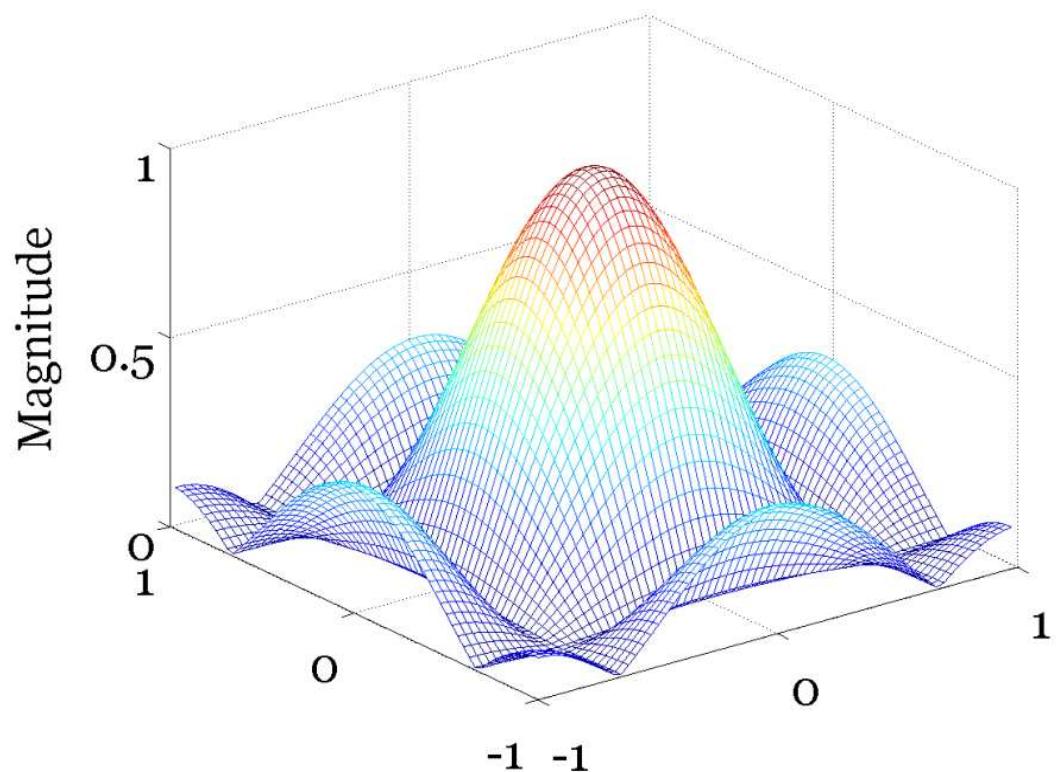
- Liên hệ với tích chập

$$f(x, y) * g(x, y) \Leftrightarrow F(u, v)G(u, v)$$

$$f(x, y)g(x, y) \Leftrightarrow F(u, v) * G(u, v)$$

Liên hệ với tích chập

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$



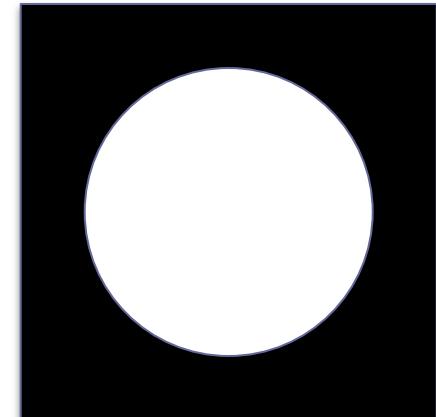
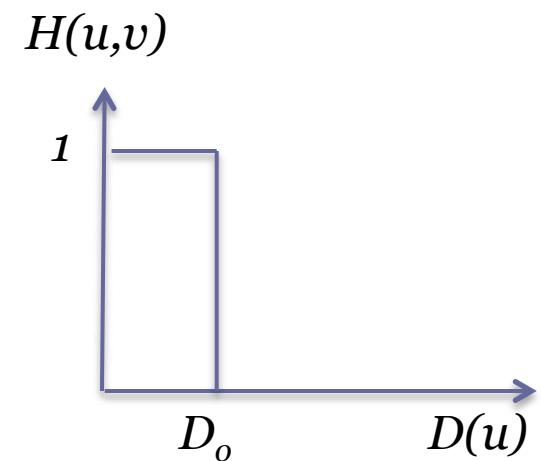
Lọc thông thấp lý tưởng

$$H(u, v) = \begin{cases} 1 & \text{if } D(u, v) \leq D_0 \\ 0 & \text{if } D(u, v) > D_0 \end{cases}$$

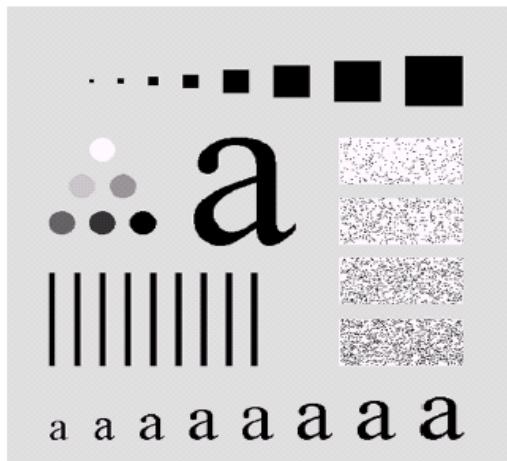
$$D(u, v) = \left[(u - M / 2)^2 + (v - N / 2)^2 \right]^{1/2}$$

$$P_T = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} P(u, v)$$

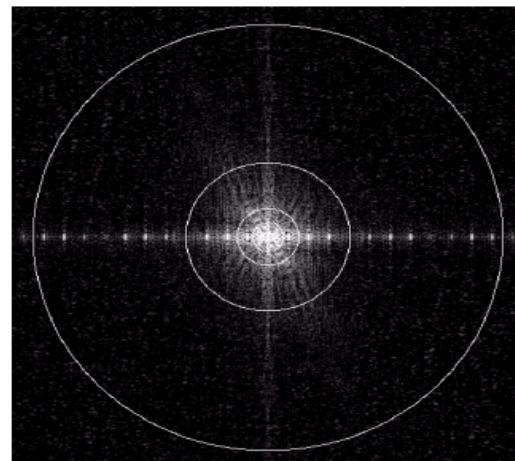
$$\alpha = 100 \left[\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} P(u, v) / P_T \right]$$



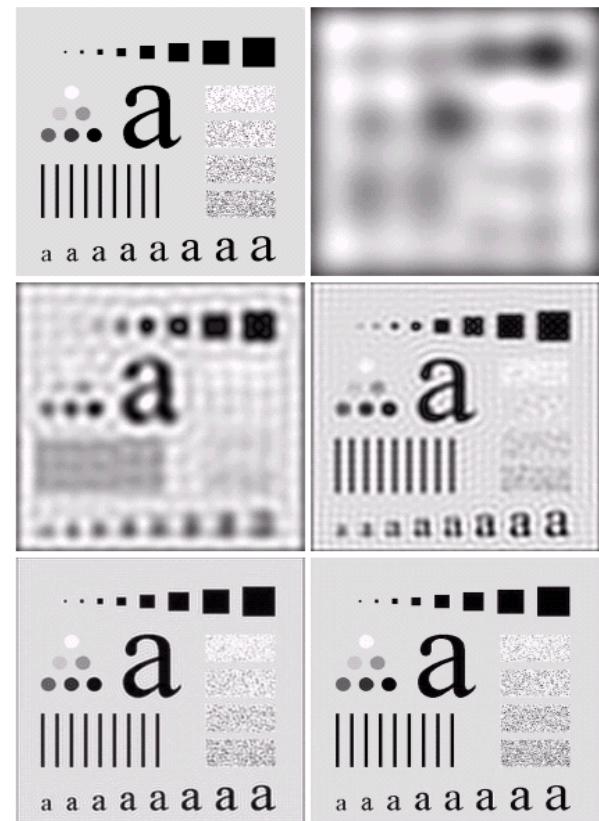
Lọc thông thấp lý tưởng



a



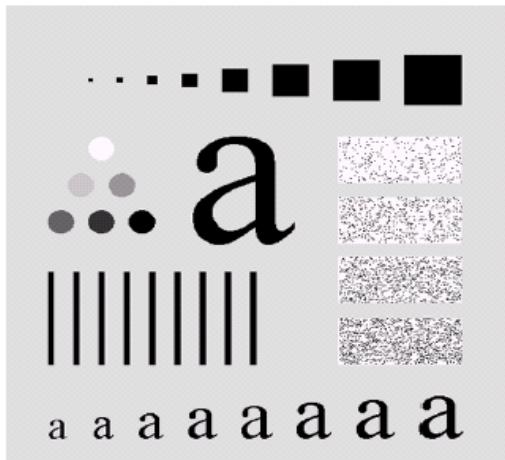
b



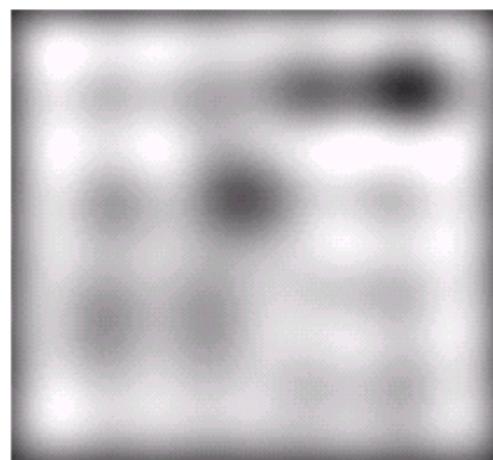
c

- (a) Ảnh gốc
- (b) Lọc thông thấp lý tưởng
- (c) Lọc thông thấp lý tưởng với $D_o=5,\dots,230$

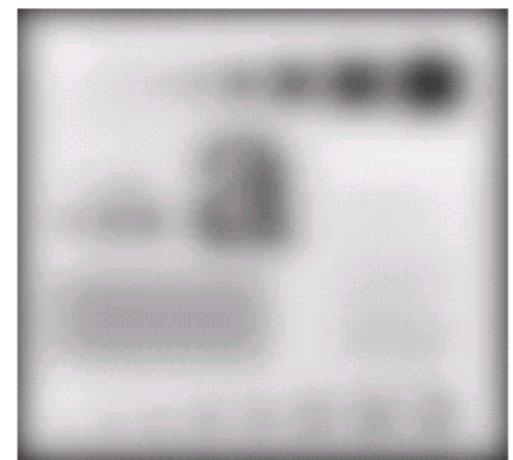
Lọc thông thấp Butterworth



a



b

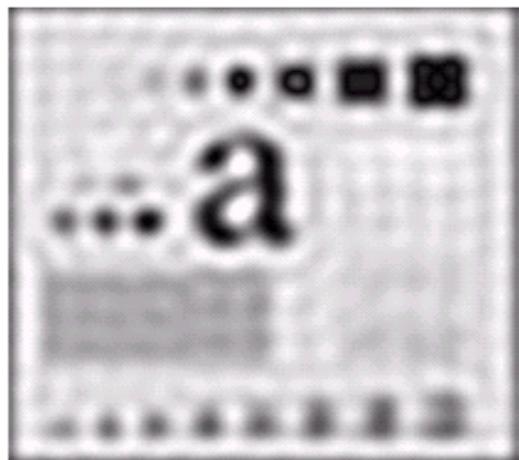


c

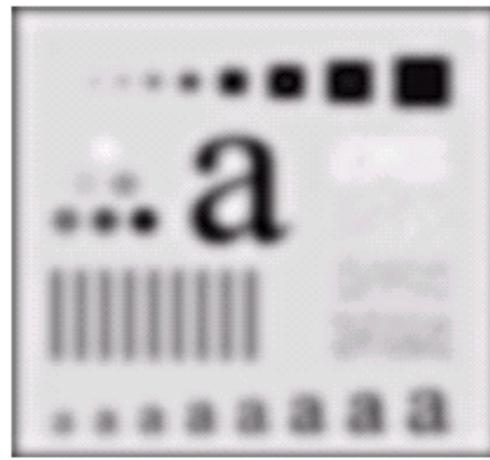
- (a) Ảnh gốc
- (b) Lọc thông thấp lý tưởng với $D_o=5$
- (c) Lọc thông thấp Butterworth với $D_o=5$

$$H(u, v) = \frac{1}{1 + (D(u, v)/D_0)^{2n}}$$

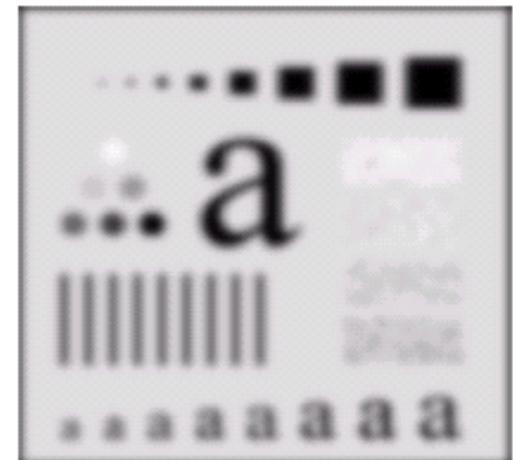
Lọc thông thấp Gauss



a



b



c

(a) Lọc thông thấp lý tưởng với $D_o=15$

(b) Lọc thông thấp Butterworth với $D_o=15$

(c) Lọc thông thấp Gauss với $D_o=15$

$$H(u, v) = e^{-D^2(u, v)/2D_0}$$

Lọc thông cao

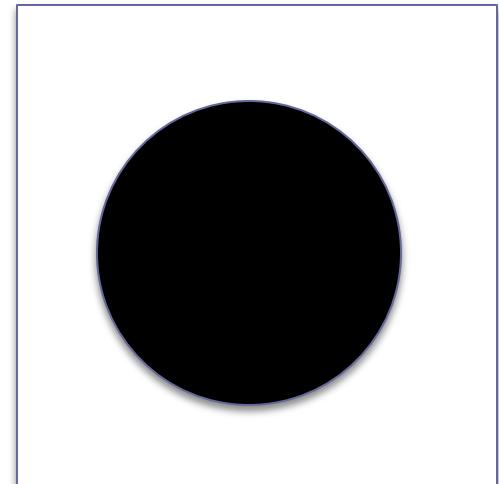
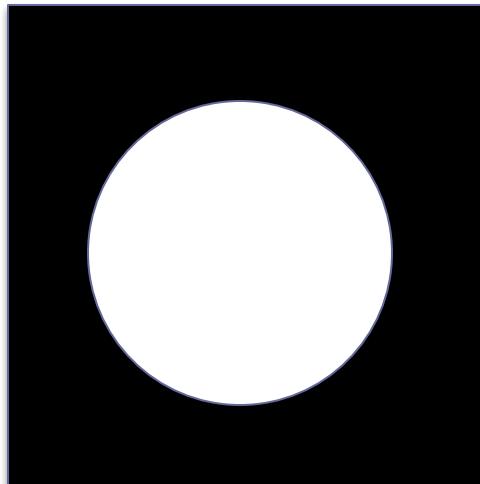
$$H_{hp}(u, v) = 1 - H_{lp}(u, v)$$

$$F(u, v)H_{hp}(u, v) = F(u, v)[1 - H_{lp}(u, v)]$$

$$F(u, v)H_{hp}(u, v) = F(u, v) - F(u, v)H_{lp}(u, v)$$

$$f_{hp}(x, y) = f(x, y) - f_{lp}(x, y)$$

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$



Lọc thông cao

- Lọc thông cao lý tưởng

$$H(u, v) = \begin{cases} 0 & \text{if } D(u, v) \leq D_0 \\ 1 & \text{if } D(u, v) > D_0 \end{cases}$$

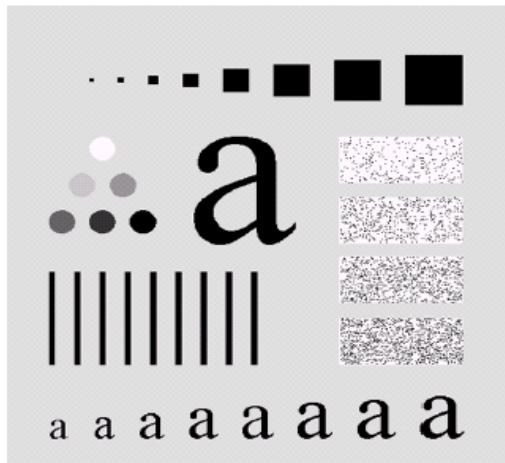
- Lọc thông cao Butterworth

$$H(u, v) = \frac{1}{1 + (D_0 / D(u, v))^{2n}}$$

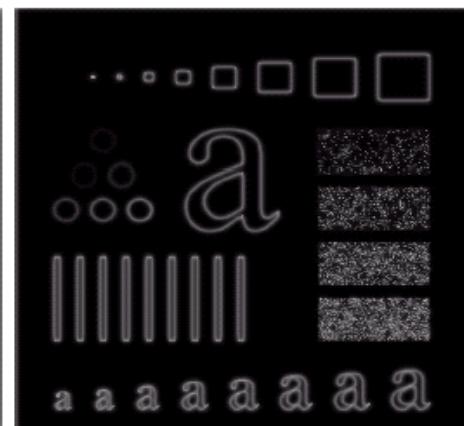
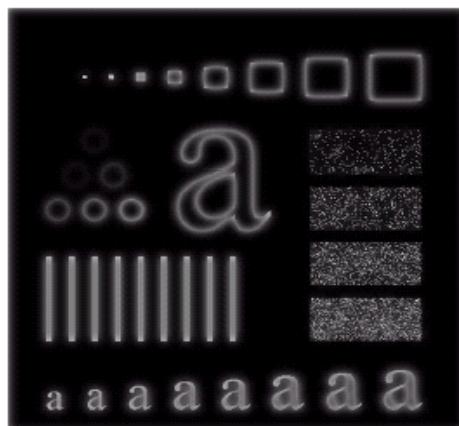
- Lọc thông cao Gauss

$$H(u, v) = 1 - e^{-D^2(u, v)/2D_0}$$

Lọc thông cao



- (a) Lọc thông cao Gauss với $D_o=15$
- (b) Lọc thông cao Gauss với $D_o=30$
- (c) Lọc thông cao Gauss với $D_o=80$



Lọc High-boost

$$f_{hb}(x, y) = Af(x, y) - f_{lp}(x, y)$$

$$f_{hb}(x, y) = (A - 1)f(x, y) + f(x, y) - f_{lp}(x, y)$$

$$f_{hb}(x, y) = (A - 1)f(x, y) + f_{hp}(x, y)$$

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & A+4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Với $A=1$, ảnh đầu ra tương đương với ảnh sau lọc sắc nét. Nếu giá trị của A tăng, vai trò của f_{hp} giảm dần. Giá trị A đủ lớn, f_{hb} sẽ tương đương với ảnh đầu vào nhân với một hằng số. Từ đó ta có

$$f_{hb}(x, y) = (A - 1)f(x, y) + h_{hp}(x, y) * f(x, y)$$

$$H_{hb}(u, v) = (A - 1) + H_{hp}(u, v)$$

Laplace trên miền tần số

- Biến đổi Fourier của đạo hàm bậc hai

$$\Im \left[\frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \right] = (ju)^2 F(u,v) + (jv)^2 F(u,v)$$

- Toán tử Laplace trong miền tần số

$$\nabla^2 f(x,y) \Leftrightarrow -[(u - M/2)^2 + (v - N/2)^2] F(u,v)$$

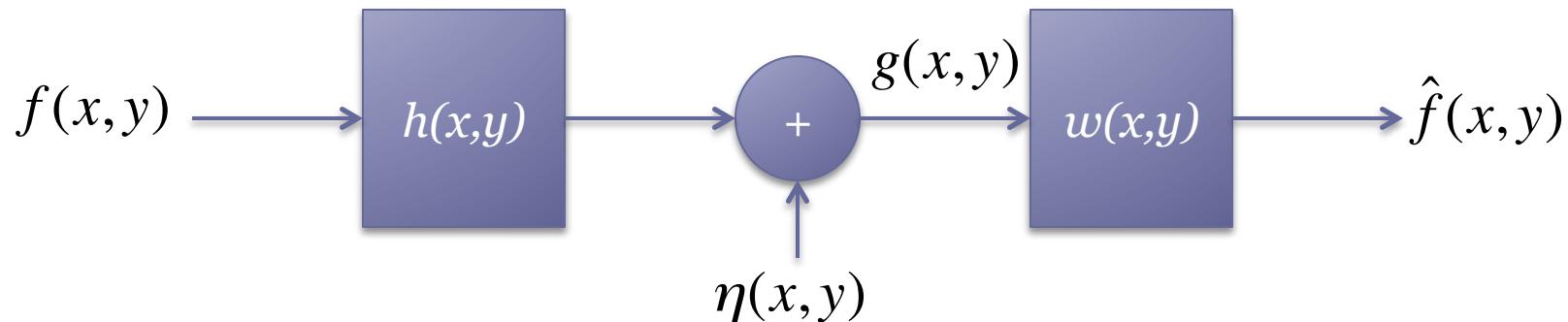
- Ảnh sau lọc sắc nét

$$g(x,y) = \Im^{-1} \left\{ [1 + (u - M/2)^2 + (v - N/2)^2] F(u,v) \right\}$$

KHÔI PHỤC ẢNH

Trần Quang Đức

Hệ thống khôi phục ảnh



$$G(u,v) = F(u,v)H(u,v) + N(u,v)$$

$$\hat{F}(u,v) = G(u,v)W(u,v)$$

Trong hệ thống trên, ta có $f(x,y)$ là ảnh đầu vào, $\hat{f}(x,y)$ là ảnh khôi phục tương ứng, $h(x,y)$ là hàm suy hao và $\eta(x,y)$ là nhiễu, xuất hiện trong quá trình thu nhận hoặc truyền ảnh. Khôi phục ảnh thông qua ước đoán hàm suy hao $h(x,y)$ còn có tên khác là **giải chập mù**.

Ước đoán hàm suy hao

- **Ước đoán dựa trên quan sát ảnh** ước đoán hàm suy hao bằng thông tin thu nhận trực tiếp từ ảnh. Ví dụ, để khôi phục ảnh bị nhòe và nhiễu, người ta tiến hành quan sát một vùng ảnh có cấu trúc đơn giản với giả thiết vùng ảnh sau khi khôi phục có nhiễu không đáng kể.
- **Ước đoán dựa trên thực nghiệm** sử dụng thiết bị tương tự như thiết bị thu nhận ảnh. Thực nghiệm được lặp lại nhiều lần với các thông số khác nhau cho đến khi ảnh khôi phục gần tương tự với ảnh đầu vào.
- **Ước đoán dựa trên mô hình hóa** sử dụng mô hình toán học để ước đoán hàm suy hao.

Minh họa về mô hình toán học

- Ảnh bị nhở do chuyển động tương đối của vật thể so với thiết bị cảm ứng trong quá trình thu nhận. Giả thiết T là thời gian mở và đóng cửa trap, $x_o(t)$ và $y_o(t)$ là các thành phần chuyển động theo phương x và y . Với $f(x,y)$ là ảnh đầu vào, ta có

$$g(x, y) = \int_0^T f(x - x_0(t), y - y_0(t)) dt$$

- Thực hiện biến đổi Fourier biểu thức trên,

$$G(u, v) = \int_0^T F(u, v) e^{-j2\pi[u x_0(t) + v y_0(t)]} dt$$

Minh họa về mô hình toán học

- Từ đó, hàm suy hao có thể viết như sau

$$H(u, v) = \int_0^T e^{-j2\pi[ux_0(t)+vy_0(t)]} dt$$

- Do chuyển động đều $y_o(t)=bt/T$ và $x_o(t)=at/T$, nên

$$H(u, v) = \frac{T}{\pi(ua + vb)} \sin[\pi(ua + vb)] e^{j\pi(ua + vb)}$$

- Ta có phương pháp lọc ngược với ảnh khôi phục có dạng

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)}$$

Lọc Weiner

- Lọc Weiner tìm $W(u,v)$ sao cho ảnh khôi phục có sai số trung bình bình phương nhỏ nhất so với ảnh đầu vào

$$\begin{aligned}\varepsilon &= E\left\{\left[F(u,v) - \hat{F}(u,v)\right]^2\right\} \\ &= E\left\{\left[F(u,v) - W(u,v)(F(u,v)H(u,v) + N(u,v))\right]^2\right\}\end{aligned}$$

- Giả thiết nhiễu độc lập với tín hiệu, ta có

$$\begin{aligned}\varepsilon &= [1 - W(u,v)H(u,v)][1 - W(u,v)H(u,v)]^* E\left\{\left[F(u,v)\right]^2\right\} \\ &\quad + W(u,v)W^*(u,v)E\left\{\left[N(u,v)\right]^2\right\}\end{aligned}$$

Lọc Weiner

- Thực hiện đạo hàm và cho giá trị đạo hàm bằng 0, ta có

$$W(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \left| \frac{N(u, v)}{F(u, v)} \right|^2} = \frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + \left| \frac{N(u, v)}{F(u, v)} \right|^2}$$

- Chú ý rằng khi không có nhiễu,

$$\left| \frac{N(u, v)}{F(u, v)} \right|^2 = \frac{1}{SNR} = 0 \Rightarrow W(u, v) = \frac{1}{H(u, v)}$$

Minh họa lọc Weiner



a



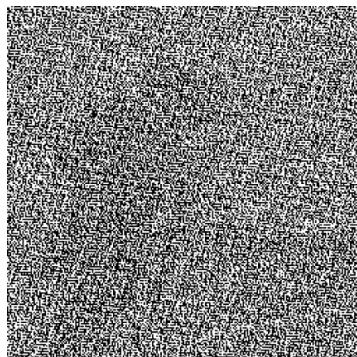
b



c



d



e



f

- (a) Ảnh gốc
- (b) Ảnh bị nhòe
- (c) Khôi phục ảnh (b)
- (d) Ảnh bị nhòe và nhiễu
- (e) Khôi phục ảnh (d) (không nhiễu)
- (f) Khôi phục ảnh (d) (có nhiễu)

PHÁT HIỆN VÀ TÁCH BIÊN ẢNH

Trần Quang Đức

Tổng quan

- Biên ảnh là tập hợp của những điểm tại đó giá trị mức xám thay đổi cục bộ đột ngột
- Sự biến thiên của một hàm liên tục được xác định thông qua đạo hàm.
- Ảnh số là một hàm rời rạc với biến là tọa độ trong mặt phẳng ảnh. Sự biến thiên vì vậy phải xác định thông qua các đạo hàm rời rạc, gồm đạo hàm bậc nhất (Gradient) và đạo hàm bậc hai (Laplace).

Đạo hàm bậc nhất

- Gradient theo hai hướng

$$G_x = \frac{\partial f}{\partial x}$$

$$G_y = \frac{\partial f}{\partial y}$$

- Biên độ của gradient

$$mag(\nabla f) = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{1/2}$$

$$mag(\nabla f) = \left[\left| \frac{\partial f}{\partial x} \right| + \left| \frac{\partial f}{\partial y} \right| \right]$$

- Tính gradient rời rạc

$$G_x = f(x+1, y) - f(x, y)$$

$$G_y = f(x, y+1) - f(x, y)$$

Đạo hàm bậc nhất

Toán tử Robert	$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$G_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
Toán tử Prewitt	$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
Toán tử Sobel	$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

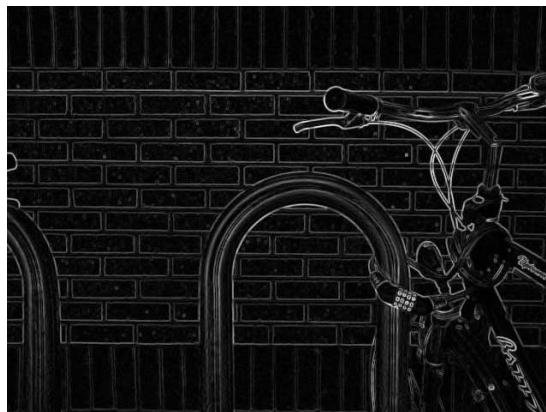
Đạo hàm bậc nhất



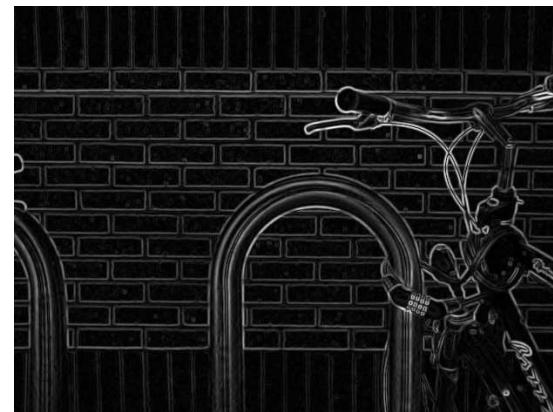
(a) Ảnh gốc



(b) Toán tử Robert



(c) Toán tử Prewitt



(d) Toán tử Sobel

Đạo hàm bậc nhất

- Toán tử Robert nhạy cảm với nhiều và những biến thiên nhỏ ở biên vì sử dụng ít điểm lân cận để xấp xỉ đạo hàm bậc nhất. Toán tử Robert đáp ứng cực đại với biên tạo góc 45° .
- Toán tử Prewitt và Sobel đáp ứng cực đại với biên ngang hoặc thẳng đứng. Bên cạnh tách biên toán tử Prewitt và Sobel còn có thể làm tròn nhiều.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

Đạo hàm bậc nhất

- Toán tử la bàn (Kirsch)
 - Ước lượng gradient của 8 hướng theo mặt nạ tổng chập
 - Kết quả của tổng chập cực đại chỉ hướng của gradient

$$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

$$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

Đạo hàm bậc nhất



(a) Ảnh gốc



(b) Toán tử Kirsch

$$\alpha = \tan^{-1} \left(\frac{G_y}{G_x} \right)$$

Toán tử Kirsch có thể tìm trực tiếp hướng của Gradient

Đạo hàm bậc hai

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

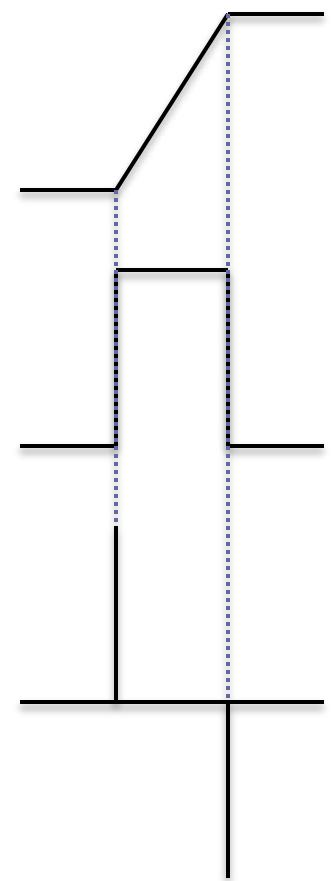
$$\frac{\partial^2 f}{\partial x^2} = f(x-1, y) + f(x+1, y) - 2f(x, y)$$

$$\frac{\partial^2 f}{\partial y^2} = f(x, y-1) + f(x, y+1) - 2f(x, y)$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Đạo hàm bậc hai

Toán tử Laplace nhạy cảm với nhiễu và không hỗ trợ tìm được hướng của đường biên. Toán tử Laplace được sử dụng để xác định vị trí của một điểm ảnh thuộc vùng tối hoặc vùng sáng của biên.



Laplace of Gaussian

$$h(r) = -e^{-\frac{r^2}{2\sigma^2}}$$

với $r^2 = x^2 + y^2$ và σ được gọi là độ lệch chuẩn

$$\nabla^2 h(r) = - \left[\frac{r^2 - \sigma^2}{\sigma^4} \right] e^{-\frac{r^2}{2\sigma^2}}$$

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$



Phương pháp Canny

- 1) Lọc làm tròn nhiễu với mặt nạ Gauss (5x5)

$$h(r) = -e^{-\frac{r^2}{2\sigma^2}}$$
$$H = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix}$$

- 2) Tính biên độ và hướng của gradient

$$G = |G_x| + |G_y|$$
$$\alpha = \tan^{-1} \left(\frac{G_y}{G_x} \right)$$

Phương pháp Canny

- 3) Xóa điểm biên không phải cực đại theo hướng gradient
- 4) Phân loại điểm biên (rõ, mờ, rất mờ) dựa theo ngưỡng
- 5) Kết nối các điểm biên mờ thuộc lân cận điểm biên rõ

2 ↑	3 ↑	5 ↑	4 ↑	6 ↑
6 ↑	5 ↑	18 ↑	18 ↑	16 ↑
5 ↑	11 ↑	4 ↑	3 ↑	2 ↑
16 ↑	4 ↑	4 ↑	4 ↑	3 ↗
5 ↑	4 ↑	3 ↑	2 ↗	12 ↑

2 ↑	3 ↑	5 ↑	4 ↑	6 ↑
6 ↑	5 ↑	18 ↑	18 ↑	16 ↑
5 ↑	11 ↑	4 ↑	3 ↑	2 ↑
16 ↑	4 ↑	4 ↑	4 ↑	3 ↗
5 ↑	4 ↑	3 ↑	2 ↗	12 ↑

2 ↑	3 ↑	5 ↑	4 ↑	6 ↑
6 ↑	5 ↑	18 ↑	18 ↑	16 ↑
5 ↑	11 ↑	4 ↑	3 ↑	2 ↑
16 ↑	4 ↑	4 ↑	4 ↑	3 ↗
5 ↑	4 ↑	3 ↑	2 ↗	12 ↑

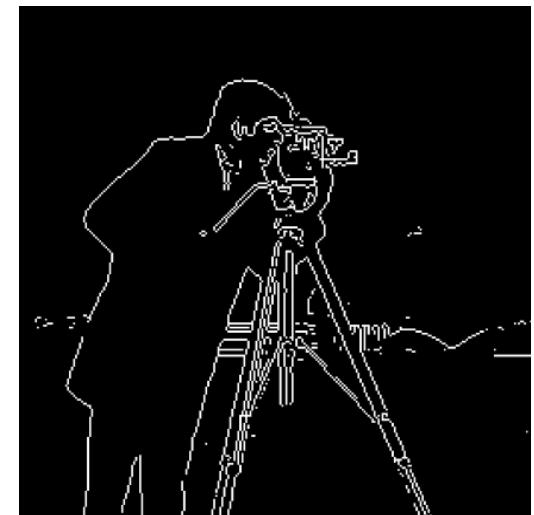
Phương pháp Canny



a



b



c

- (a) Phương pháp Canny
- (b) Laplace of Gaussian
- (c) Toán tử Sobel

PHÂN VÙNG ẢNH

Trần Quang Đức

Tổng quan

- Một vùng ảnh là tập hợp của các điểm ảnh có cùng chung một hoặc nhiều thuộc tính về đối tượng trong ảnh.
- Phân vùng ảnh là một trong những nhiệm vụ khó khăn nhất của xử lý ảnh. Trong một số trường hợp, người ta sử dụng các thiết bị đặc biệt, ví dụ máy chụp ảnh hồng ngoại để tăng hiệu quả của quá trình phân vùng.
- Quá trình phân vùng dựa trên hai tính không liên tục và tính đồng nhất của điểm ảnh. Các phương pháp gồm phân vùng dựa trên biên ảnh, phân vùng dựa trên ngưỡng và phân vùng dựa trên miền ảnh.

Phân vùng dựa trên biên ảnh

- Biên ảnh là tập hợp của những điểm ảnh nằm trên ranh giới giữa hai vùng ảnh. Biên ảnh được xác định dựa trên các phương pháp đạo hàm bậc nhất (gradient) và đạo hàm bậc hai (laplace).
- Tìm biên thông qua đạo hàm thường tạo ra biên không khép. Nhiều cũng tạo ra các điểm biên giả, gây khó khăn cho quá trình phân vùng ảnh.
- Kết nối các điểm biên có thể dựa trên các ***phương pháp xử lý cục bộ*** hoặc ***xử lý toàn cục sử dụng biến đổi Hough***.

Phương pháp xử lý cục bộ

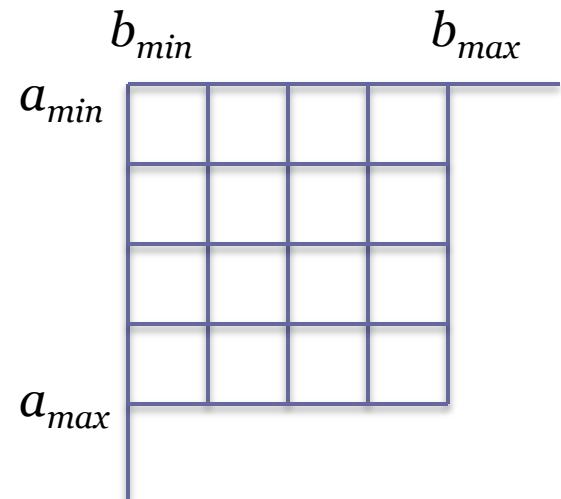
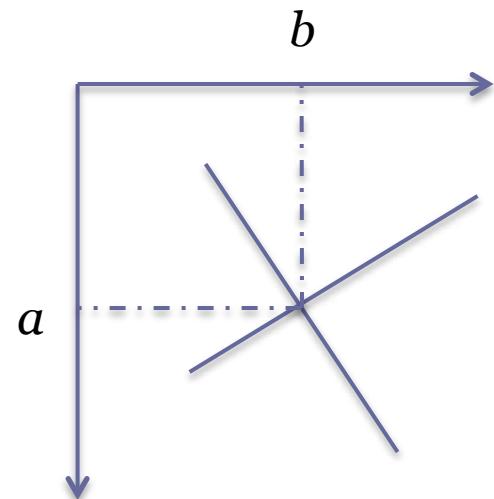
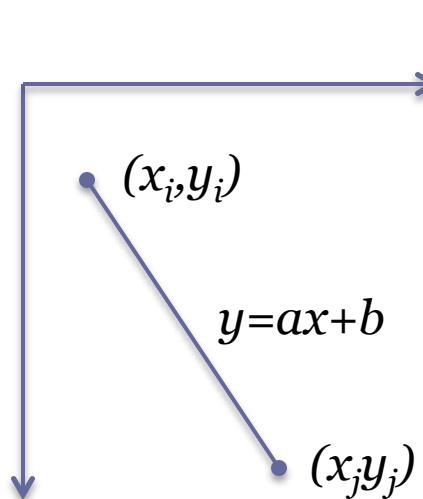
- Cho điểm biên (x_o, y_o) , điểm ảnh (x, y) thuộc lân cận của (x_o, y_o) được kết nối với (x_o, y_o) nếu

$$|\nabla f(x, y) - \nabla f(x_0, y_0)| \leq E \quad |\alpha(x, y) - \alpha(x_0, y_0)| \leq A$$

- Với E và A là các giá trị ngưỡng. Quá trình trên được lặp lại với tất cả các điểm ảnh. Các lân cận được xem xét thuộc mặt nạ kích thước 3×3 hoặc 5×5 .
- Kỹ thuật xử lý cục bộ không tính đến lịch sử tìm kiếm nên có thể tạo ra các điểm biên giả.

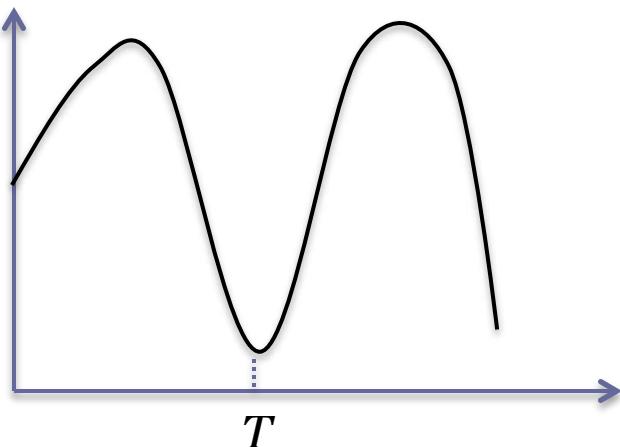
Phương pháp xử lý toàn cục

- Kỹ thuật xử lý toàn cục tìm đường thẳng xấp xỉ đi qua n điểm ảnh. Giả sử đường thẳng đi qua điểm ảnh (x_i, y_i) có dạng $y_i = ax_i + b$, ta có thể viết $b = y_i - ax_i$. Từ đó, đường thẳng được tìm kiếm trên mặt phẳng tham số ab .



Phân vùng dựa trên ngưỡng

- 1) Chọn giá trị khởi tạo T_o
- 2) Chia vùng ảnh dựa theo T_o (G_1 và G_2)
- 3) Tính μ_1 và μ_2 tương ứng của G_1 và G_2
- 4) Tính ngưỡng mới $T_1 = (\mu_1 + \mu_2)/2$
- 5) Lặp lại các bước 2-4 đến khi $T_i - T_{i-1} < Th$



$$g(x, y) = \begin{cases} 1 & \text{If } f(x, y) > T \\ 0 & \text{If } f(x, y) \leq T \end{cases}$$

Phân vùng dựa trên hàm xác suất

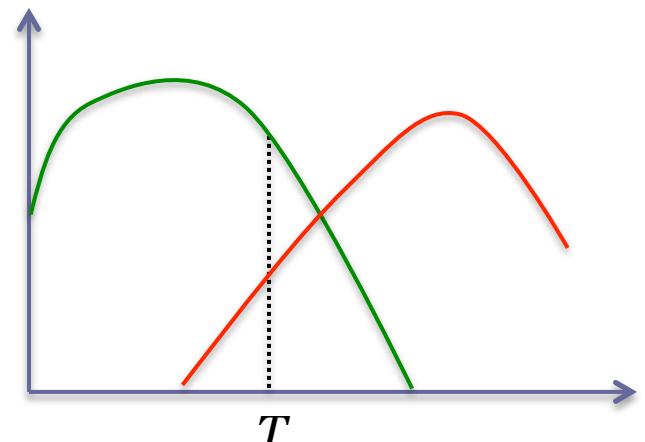
$$p(z) = P_1 p_1(z) + P_2 p_2(z)$$

$$P_1 + P_2 = 1$$

$$E_1(T) = \int_{-\infty}^T p_1(z) dz \quad E_2(T) = \int_{-\infty}^T p_2(z) dz$$

$$E(T) = P_1 E_1(T) + P_2 E_2(T)$$

$$\frac{\partial E}{\partial T} = 0 \Rightarrow P_2 p_1(T) = P_1 p_2(T)$$



$$p_i(z) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(z-\mu_i)^2}{2\sigma_i^2}}$$

$$T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2} \ln\left(\frac{P_2}{P_1}\right)$$

Phương pháp cải thiện Histogram

- **Nhận xét:** Ngưỡng thường được chọn dựa trên lược đồ mức xám. Do đó, cải thiện lược đồ mức xám có thể tăng hiệu quả của quá trình phân vùng.
- Cải thiện lược đồ mức xám bằng cách chỉ xem xét những điểm ảnh nằm lân cận biên để giảm sự chênh lệch số lượng điểm ảnh giữa các vùng.
- **Nhắc lại:** Đạo hàm bậc nhất (gradient) cho phép xác định biên ảnh trong khi đạo hàm bậc hai (laplace) xem xét một điểm ảnh thuộc vùng tối hay vùng sáng của biên.

Phương pháp cải thiện Histogram

$$s(x, y) = \begin{cases} 0 & \text{If } \nabla f < T \\ + & \text{If } \nabla f \geq T \text{ and } \nabla^2 f \geq 0 \\ - & \text{If } \nabla f \geq T \text{ and } \nabla^2 f < 0 \end{cases}$$

- Điểm ảnh không thuộc biên ảnh hoặc lân cận của biên ảnh được ký hiệu 0. Điểm ảnh lân cận của biên ảnh thuộc vùng tối được ký hiệu +. Điểm ảnh lân cận của biên ảnh thuộc vùng sáng được ký hiệu -.
- Biên ảnh tại những điểm có bước chuyển (-,+) hoặc (+,-).

Phân vùng dựa trên miền ảnh

- Vùng ảnh là tập hợp của các điểm ảnh có cùng chung một hoặc nhiều thuộc tính về đối tượng trong ảnh
- Phân vùng ảnh là quá trình chia tập các điểm ảnh thành những tập con R_i thỏa mãn các điều kiện sau

$$X = \bigcup R_i \quad R_i \cap R_j = \emptyset \quad \forall i, j, i \neq j$$

$$P(R_i) = TRUE \quad P(R_i \cup R_j) = FALSE$$

$$R_i \neq \emptyset$$

Phương pháp gia tăng vùng

- Phương pháp gia tăng vùng xét lân cận của những điểm ảnh hạt giống. Nếu chúng thỏa mãn điều kiện về độ tương đồng thì hợp lại thành một vùng.

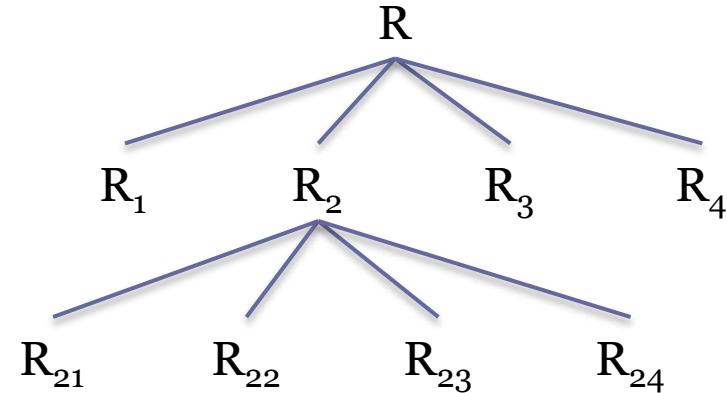
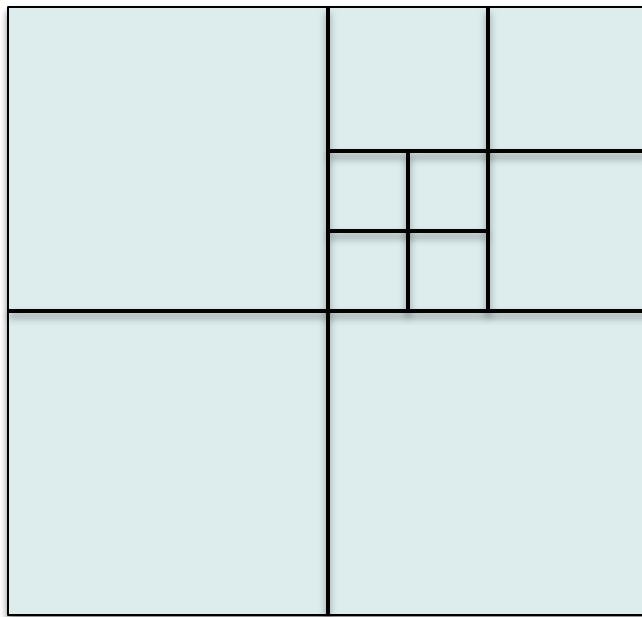
9	9	1	1	1	2	2	1
8	9	9	8	1	1	1	1
7	9	9	8	1	2	1	1
8	9	9	6	1	1	2	1
9	8	8	9	1	2	2	1
8	8	8	9	1	1	2	1
8	9	8	9	1	2	1	2

$$|f(x, y) - f(x_0, y_0)| \leq 2$$

Sử dụng quan hệ 4-lân cận

Phương pháp chia và hợp vùng

- Phương pháp chia và hợp vùng sử dụng phương pháp phân tách cây tứ phân để chia ảnh thành vùng nhỏ. Những vùng nhỏ lân cận được hợp lại nếu chúng thỏa mãn điều kiện về độ tương đồng.



XỬ LÝ ĐA PHÂN GIẢI

Trần Quang Đức

Biểu diễn đa phân giải

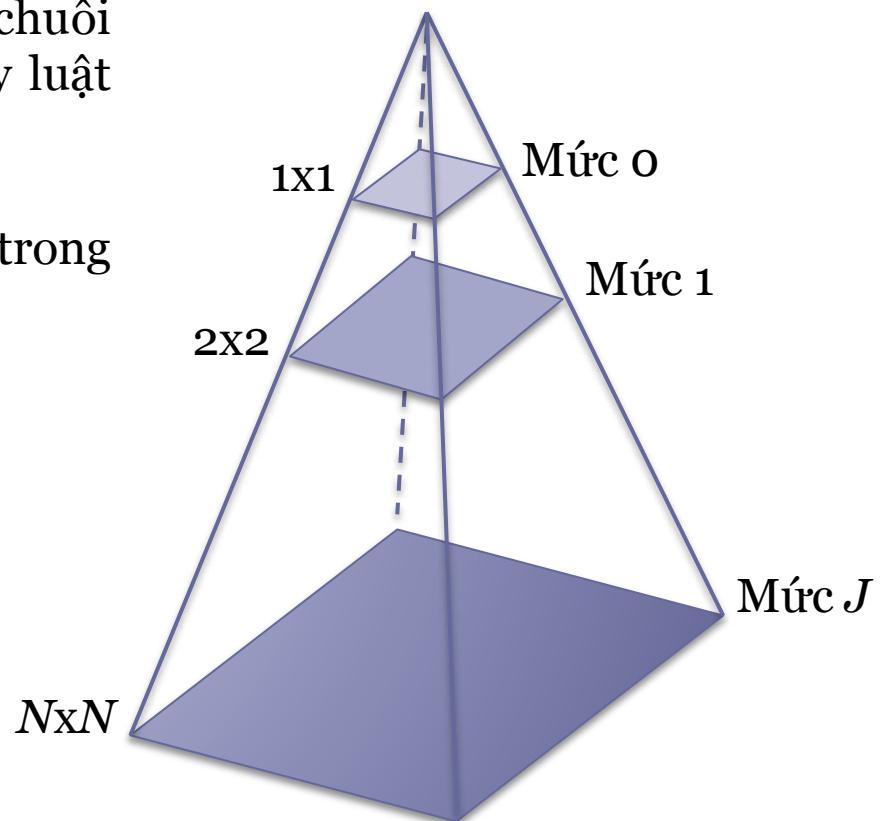
Biểu diễn đa phân giải biểu diễn chuỗi ảnh có độ phân giải giảm dần theo quy luật dựa trên cấu trúc tháp.

Biểu diễn đa phân giải được ứng dụng trong phân vùng ảnh và nén ảnh.

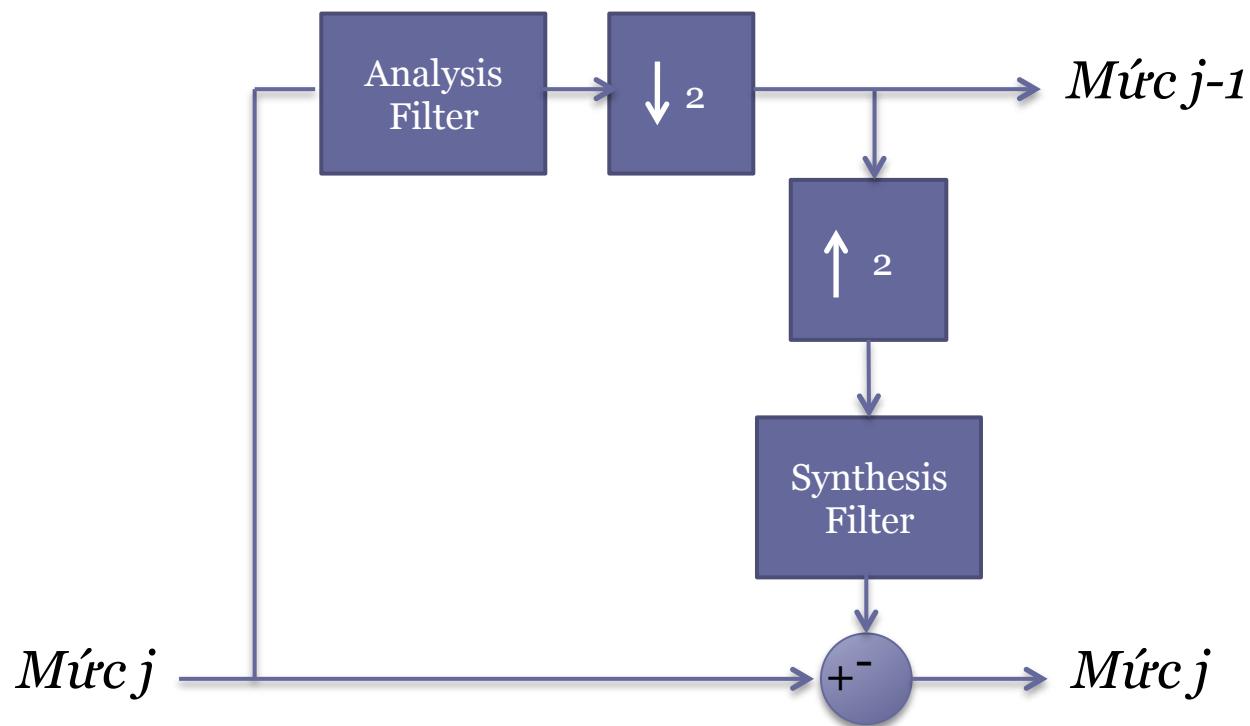
Ảnh tại mức j kích thước $2^j \times 2^j$, $0 \leq j \leq J$

Dung lượng lưu trữ $J+1$ mức

$$N^2 \left(1 + \frac{1}{4} + \dots + \frac{1}{4^J}\right) \leq \frac{4}{3} N^2$$



Biểu diễn đa phân giải



Biến đổi Haar

- Biến đổi Haar có tính đối xứng và phân tách $\mathbf{T}=\mathbf{H}\mathbf{F}\mathbf{H}$. Để tạo ra \mathbf{H} , ta định nghĩa số nguyên k , $k=2^p+q-1$, với $0 \leq p \leq n-1$, $q=0$ hoặc 1 nếu $p=0$ và $1 \leq q \leq 2^p$ nếu $p \neq 0$.

$$h_0(z) = h_{00}(z) = \frac{1}{\sqrt{N}}, \quad z \in [0,1]$$

$$h_{pq}(z) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & (q-1)/2^p \leq z < (q-0,5)/2^p \\ -2^{p/2} & (q-0,5)/2^p \leq z < q/2^p \\ 0 & Otherwise, z \in [0,1] \end{cases}$$

Biến đổi Haar

Nếu $N=4$, giá trị k, p, q như sau

k	p	q
0	0	0
1	0	1
2	1	1
3	1	2

$$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$H_2 \sim$ Bộ lọc phân tích QMF

$$G_0(z) = H_0(z) \quad G_1(z) = -H_0(-z)$$

$$H_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

Biến đổi Wavelet rời rạc

Giá trị của một hàm rời rạc có thể biểu diễn thành chuỗi các hệ số

$$W_\varphi(j_0, k) = \frac{1}{\sqrt{M}} \sum_n f[n] \varphi_{j_0, k}[n] \quad W_\psi(j, k) = \frac{1}{\sqrt{M}} \sum_n f[n] \psi_{j, k}[n]$$

$$f[n] = \frac{1}{\sqrt{M}} \sum_k W_\varphi(j_0, k) \varphi_{j_0, k}[n] + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j, k) \psi_{j, k}[n]$$

$$M=2^J, x=0,1,2,\dots, M-1; j=0,1,2,\dots, J-1 \text{ và } k=0,1,2,\dots, 2^j-1.$$

Ví dụ: $f(0)=1, f(1)=4, f(2)=-3, f(3)=0$

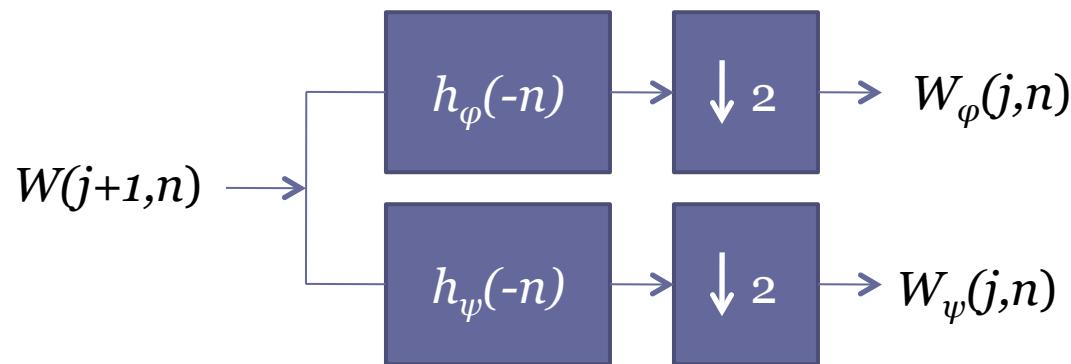
Biến đổi Wavelet nhanh

Trong biến đổi Wavelet nhanh, các hệ số có thể xác định như sau

$$W_\varphi(j,k) = h_\varphi(-n) * W_\varphi(j+1,n)$$

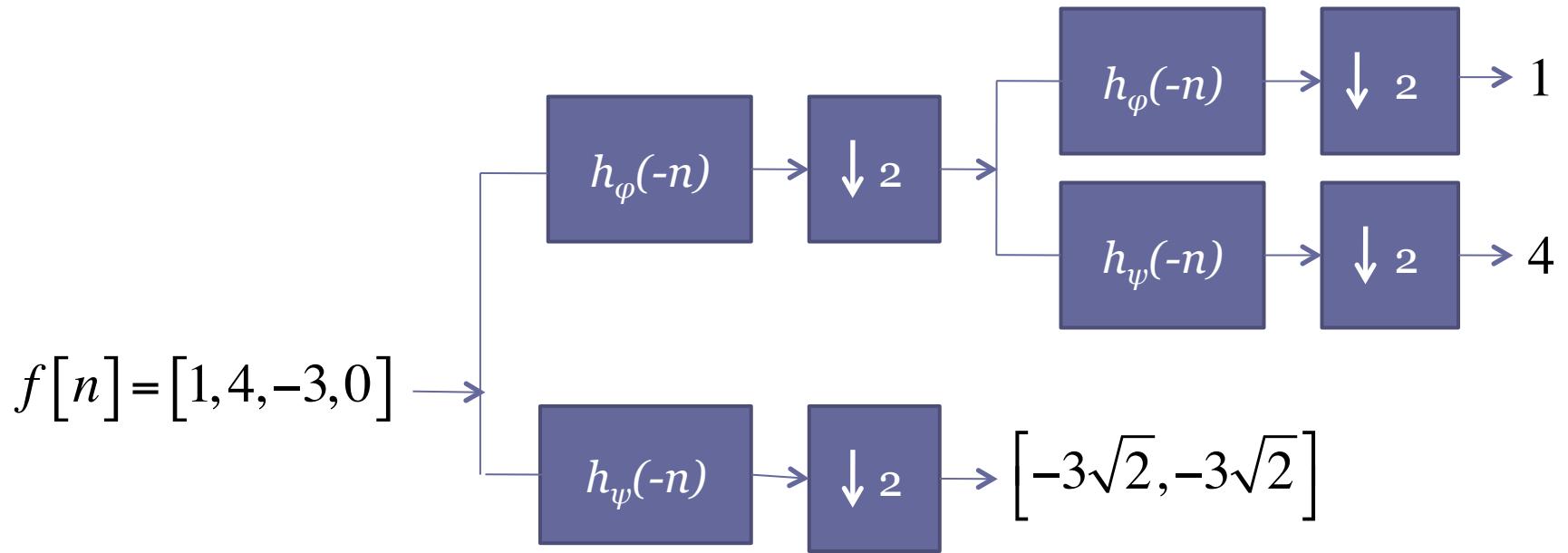
$$W_\psi(j,k) = h_\psi(-n) * W_\psi(j+1,n)$$

với $n=2k$ ($k \geq 0$)

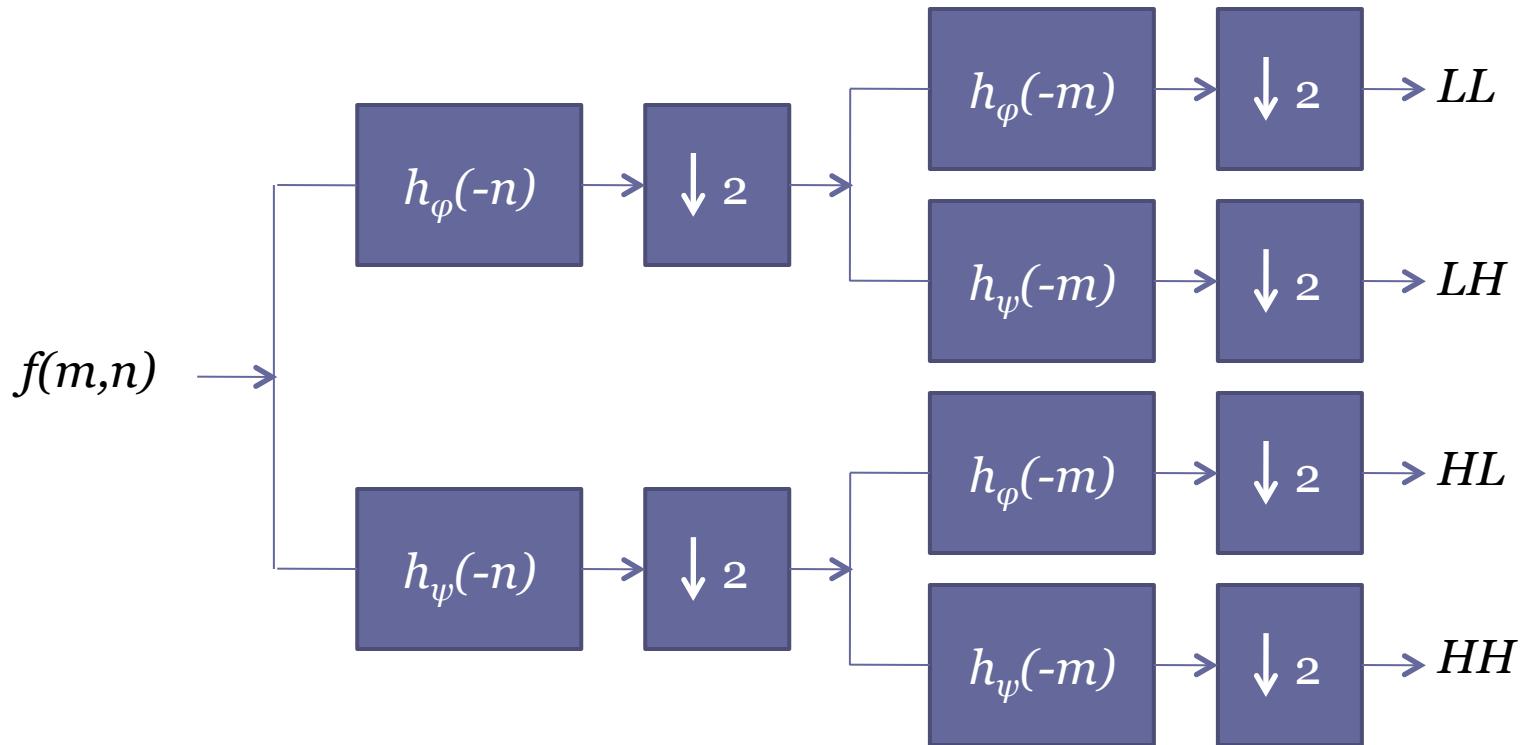


Minh họa

Ví dụ: $f(n) = [1, 4, -3, 0]$



Biến đổi Wavelet nhanh 2 chiều



XỬ LÝ ẢNH NHỊ PHÂN

Trần Quang Đức

Khái niệm

- Ảnh nhị phân gồm những điểm ảnh có giá trị 0 và 1. Giá trị 0 biểu diễn điểm ảnh thuộc đối tượng ảnh (điểm ảnh đen). Giá trị 1 biểu diễn điểm ảnh thuộc nền (điểm ảnh trắng).

$$g(x, y) = \begin{cases} 1 & \text{If } f(x, y) \geq T \\ 0 & \text{If } f(x, y) < T \end{cases}$$

Lý thuyết tập hợp

- A là một tập hợp thuộc không gian \mathbb{Z}^2 . Một phần tử $a=(a_1, a_2)$ thuộc A được ký hiệu

$$a \in A$$

- Tương tự, nếu phần tử a không thuộc A , ta ký hiệu

$$a \notin A$$

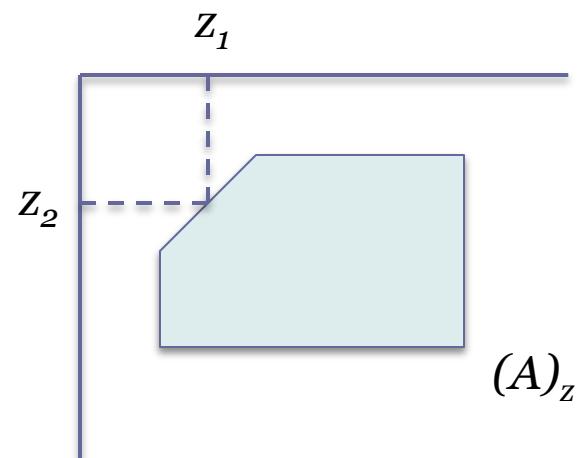
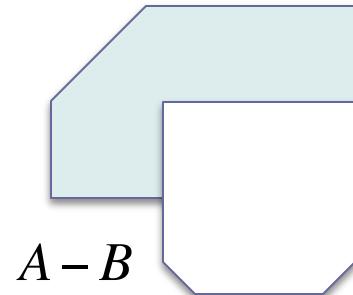
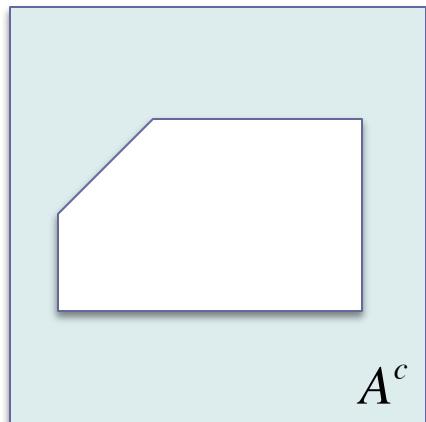
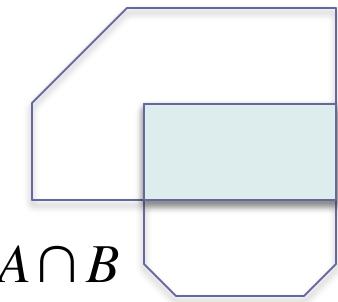
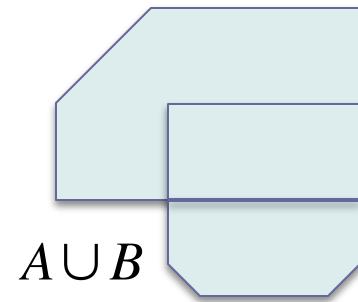
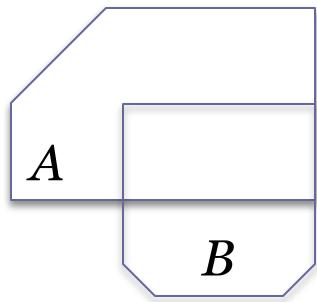
- Nếu A không chứa một phần tử nào, A là tập rỗng

$$A = \emptyset$$

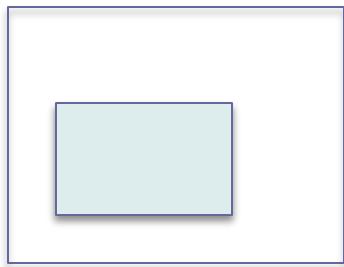
Lý thuyết tập hợp

A là tập con của B	$A \subseteq B$
Hợp của hai tập hợp A và B	$C = A \cup B$
Giao của hai tập hợp A và B	$C = A \cap B$
Phần bù của tập hợp A	$A^c = \{w \mid w \notin A\}$
Hiệu của hai tập hợp A và B	$A - B = \{w \mid w \in A, w \notin B\}$
Phản chiếu của tập hợp A	$\hat{A} = \{w \mid w = -a, \forall a \in A\}$
Dịch tập hợp A với $z=(z_1, z_2)$	$(A)_z = \{c \mid c = a + z, \forall a \in A\}$

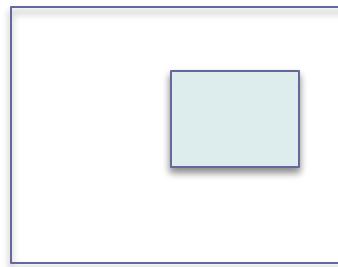
Lý thuyết tập hợp



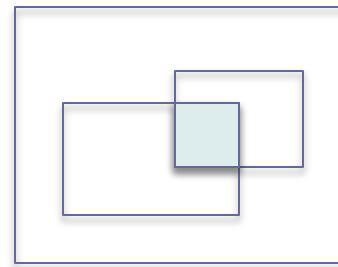
Toán tử Logic



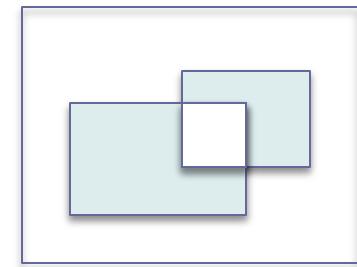
A



B



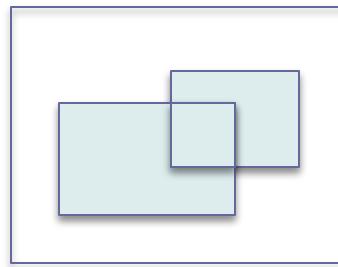
$A \text{ AND } B$



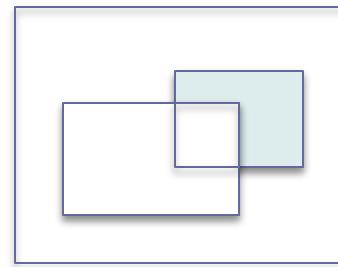
$A \text{ XOR } B$



$\text{NOT } A$



$A \text{ OR } B$



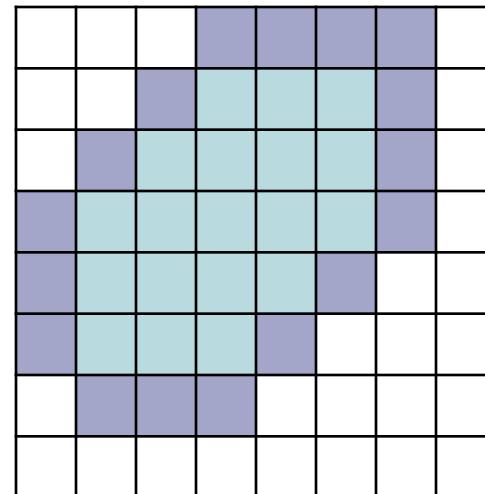
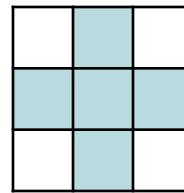
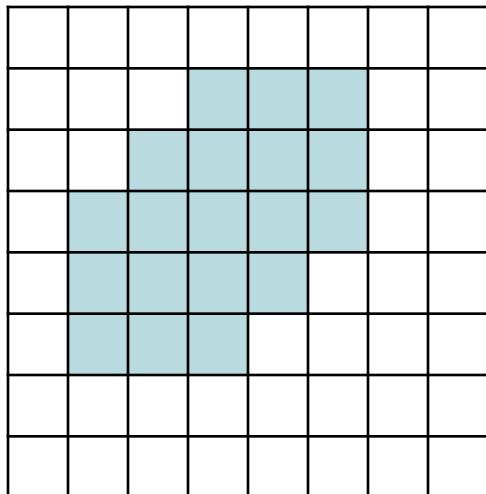
$[\text{NOT } A] \text{ AND } B$

Toán tử hình thái

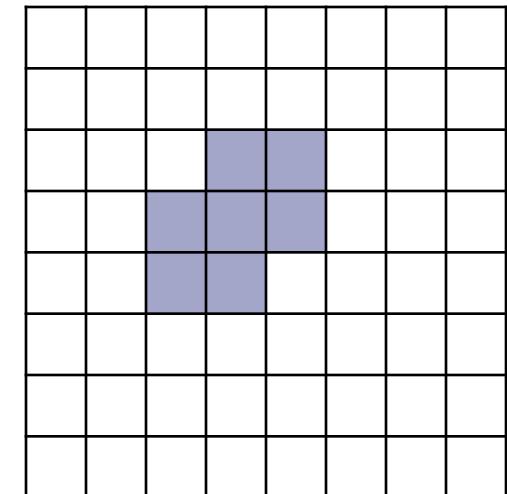
- Toán tử hình thái là công cụ toán học để xử lý hình dạng trong ảnh. Toán tử hình thái sử dụng hướng tiếp cận lý thuyết tập hợp.
- Toán tử hình thái bao gồm phép co (erosion), phép dãn (dilation), phép mở (opening), phép đóng (closing).
- Toán tử hình thái được ứng dụng trong tách biên ảnh, lấp đầy vùng ảnh, tạo kết nối giữa các vùng ảnh hoặc làm xương ảnh.

Phép dān và phép co

Phép dān	Phép co
$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\}$	$A \ominus B = \{z \mid (B)_z \subseteq A\}$



$A \oplus B$

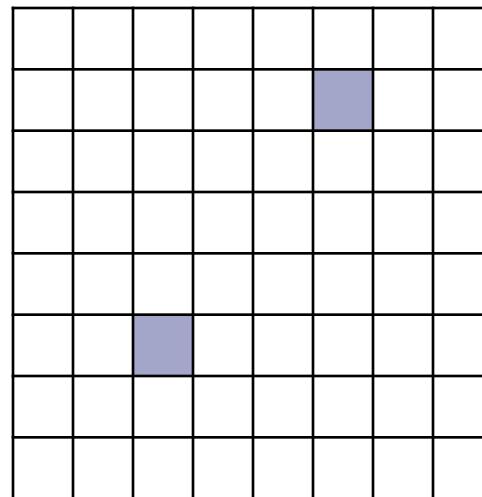
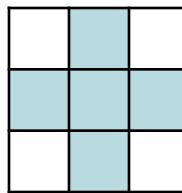
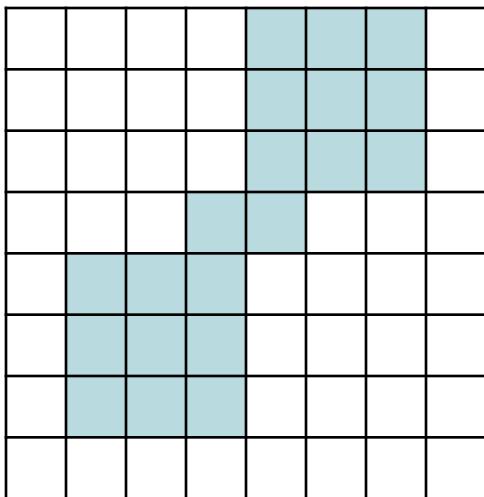


$A \ominus B$

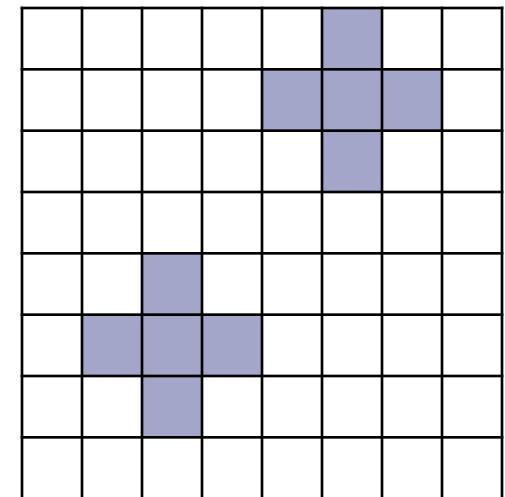
Phép mở

Phép mở xóa bỏ những đoạn mảnh, loại bỏ nhiễu nhưng làm tăng số đoạn đứt gãy

$$A \circ B = (A \ominus B) \oplus B$$



$$A \ominus B$$

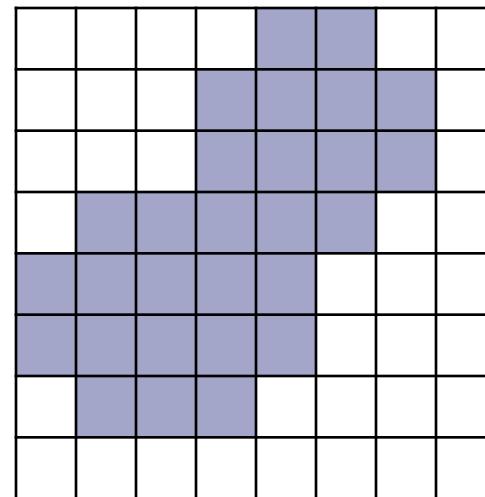
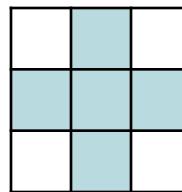
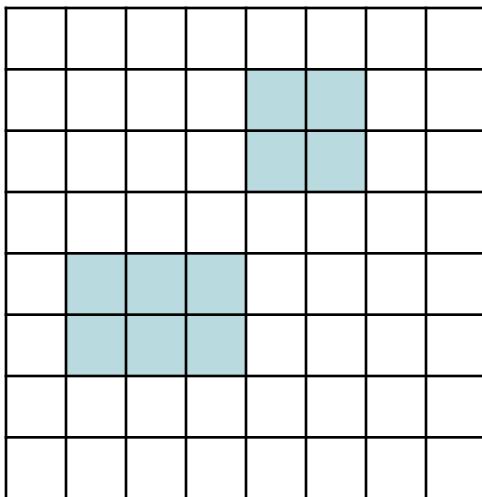


$$A \circ B$$

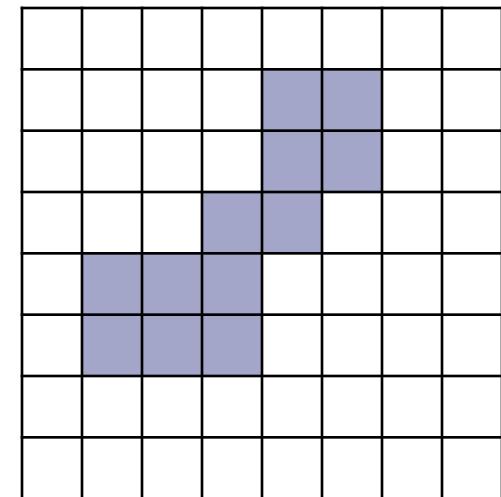
Phép đóng

Phép đóng có thể làm tròn biên ảnh và kết nối các vùng của cùng một đối tượng.

$$A \bullet B = (A \oplus B) \ominus B$$



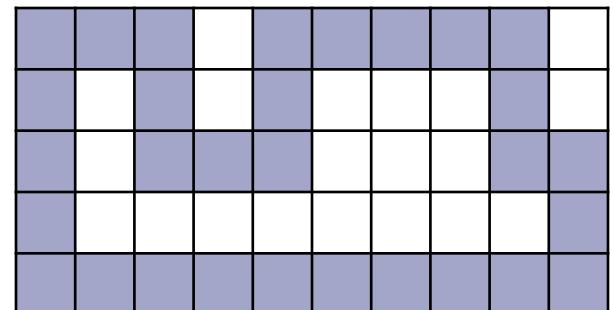
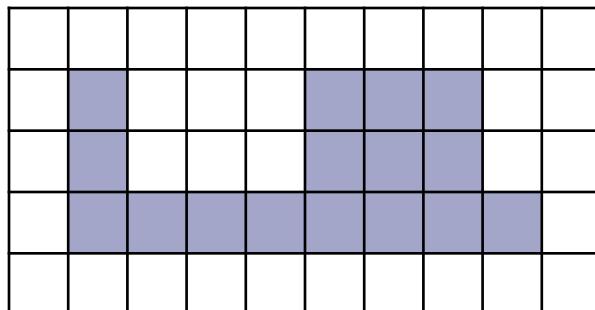
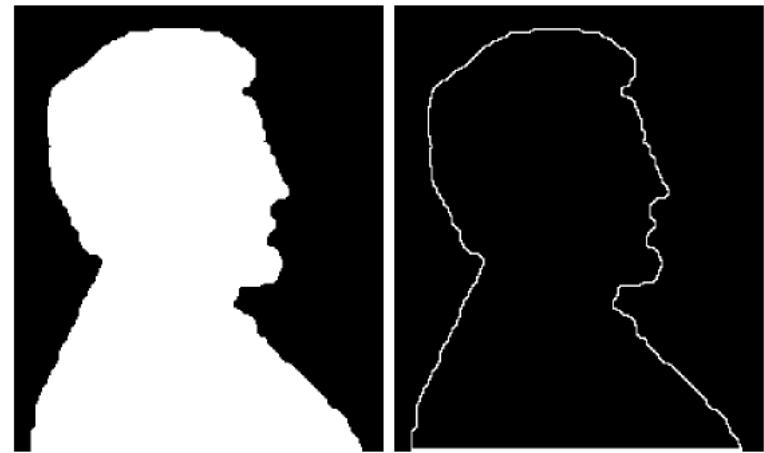
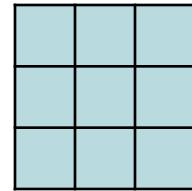
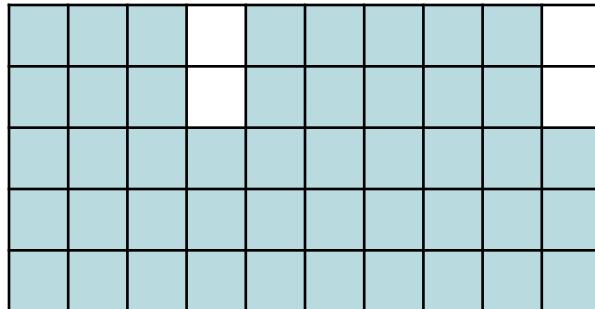
$$A \ominus B$$



$$A \circ B$$

Tách biên ảnh

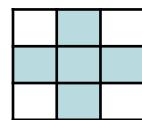
$$\beta(A) = A - (A \ominus B)$$



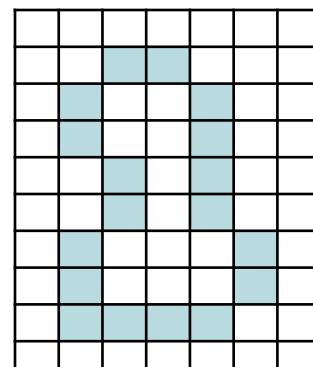
Lấp đầy vùng ảnh

$$X_k = (X_{k-1} \oplus B) \cap A^c, \quad k = 1, 2, 3 \dots$$

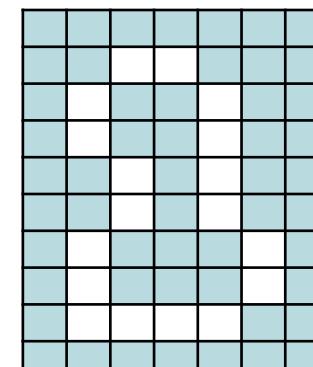
Lặp quá trình trên đến khi $X_{k-1} = X_k$



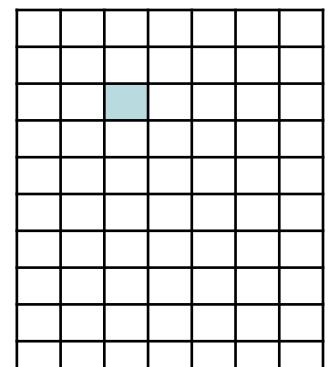
B



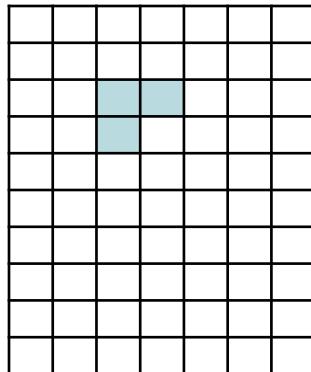
A



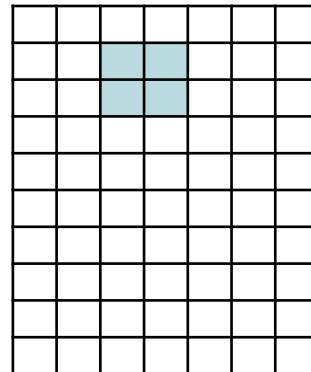
A^c



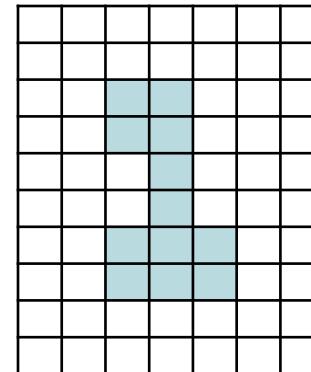
X_o



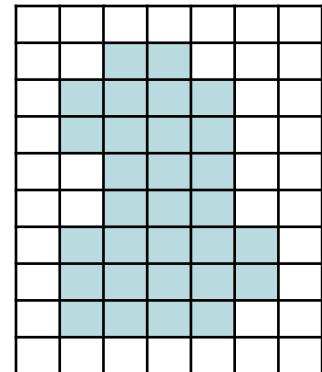
X_1



X_2

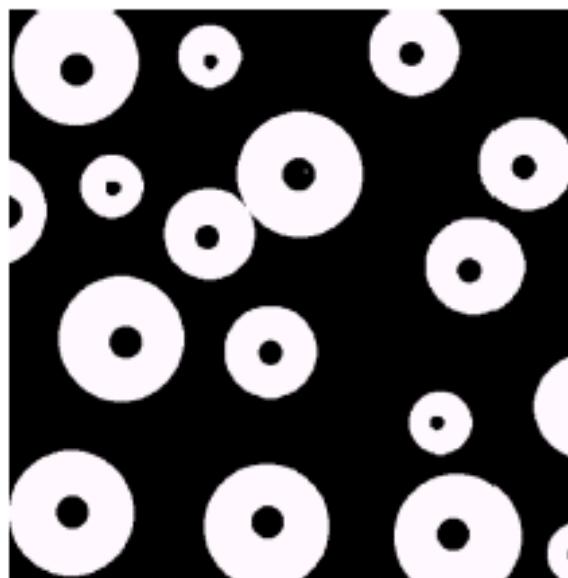


X_7

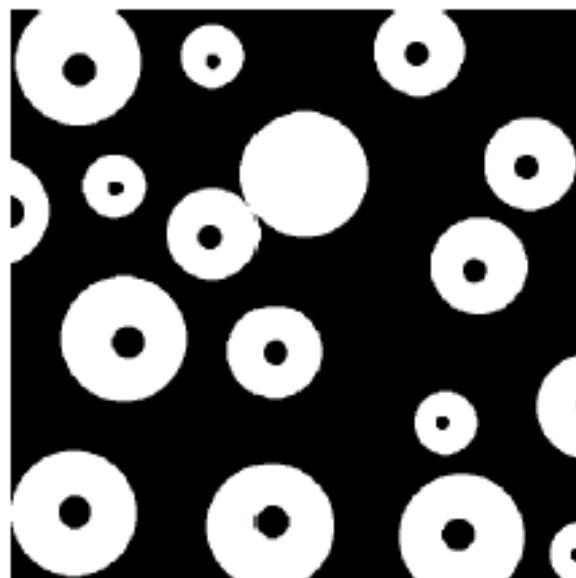


$X_7 \cup A$

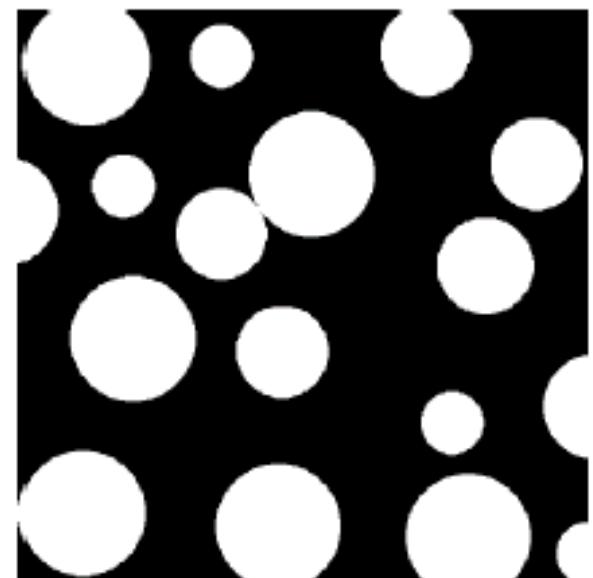
Lấp đầy vùng ảnh



Ảnh gốc



Lấp đầy một vùng ảnh



Lấp đầy vùng ảnh

MÃ HÓA VÀ NÉN ẢNH

Trần Quang Đức

Dung lượng thông tin

- Một trang văn bản: 2 KB.
- Một ảnh màu (800x600x24): 1,4 MB
- 30 phút âm thanh thoại số (8 kHz, 8 bits): 14 MB
- 30 phút audio CD (44,1 kHz, 16 bits, stereo): 316 MB
- 30 phút video (800x600x24, 25 khung hình/s): 64,8 GB

Khái niệm

- Nén dữ liệu ảnh: Biến đổi dòng thông tin ảnh thành từ mã nhằm giảm độ dư thừa thông tin. Các phương pháp nén khác nhau do định nghĩa các kiểu dư thừa thông tin khác nhau.
- Các kiểu dư thừa thông tin gồm: sự phân bố mức xám (Mã Huffman), sự lặp lại của các mức xám (RLC), những mẫu sử dụng tần xuất cao (LZW) hoặc độ dư thừa vị trí (mã hóa dự đoán)
- *Tỷ lệ nén=1/r (%)* trong đó r là kích thước dữ liệu gốc chia cho kích thước dữ liệu sau nén

Phân loại phương pháp nén

- Cách 1 (Dựa trên nguyên lý nén)
 - Nén không mất mát thông tin
 - Nén mất mát thông tin
- Cách 2 (Dựa trên cách thức thực hiện nén)
 - Phương pháp không gian
 - Phương pháp dựa trên biến đổi
- Cách 3 (Dựa trên triết lý mã hóa)
 - Phương pháp nén thế hệ thứ nhất
 - Phương pháp nén thế hệ thứ hai

Mã loạt dài

- Thay bằng việc truyền đi một chuỗi “0” hoặc “1”, có thể truyền độ dài của chuỗi.
- Trong dữ liệu fax, 70%-80% không gian là các điểm ảnh có giá trị 0 (điểm ảnh trắng). Vì thế mã loạt dài hoạt động rất hiệu quả. Có thể thay thế độ dài chuỗi bằng một số nguyên có kích thước cố định.
- Mã loạt dài tỏ ra kém hiệu quả nếu tần suất xuất hiện của giá trị 1 (điểm ảnh màu đen) tăng.
- $111111111100000000000011111 \rightarrow 11,1,13,0,5,1$

Mã Huffman

- **Mã Huffman:** Phương pháp nén dự trên mô hình thống kê xem xét sắc xuất phân bố của ký tự. Mô hình xác định sắc xuất có thể là mô hình *tĩnh, thích nghi* hoặc *bán thích nghi (semi-adaptive)*.
- **Mô hình tĩnh** được tích hợp sẵn trong thiết bị nén và thiết bị giải nén.
- **Mô hình bán thích nghi** là mô hình cố định được xây dựng từ dữ liệu được nén.
- **Mô hình thích nghi** thay đổi trong quá trình nén.

Thuật toán cơ bản

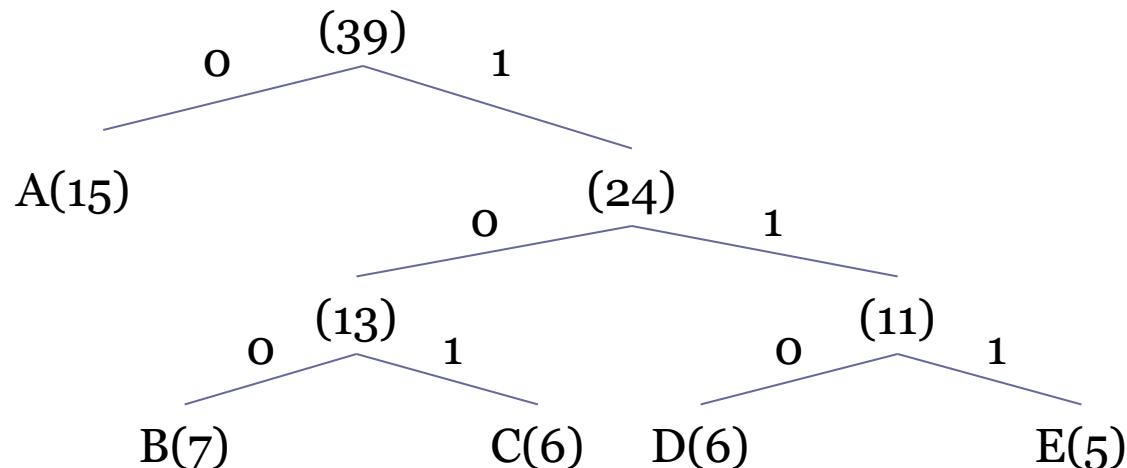
- Các ký tự có tần suất xuất hiện khác nhau.
- Các ký tự chiếm không gian biểu diễn khác nhau.
- Ký tự với độ dài từ mã nhị phân cố định trong bảng mã ASCII được thay bằng từ mã có độ dài thay đổi.
- Từ mã ngắn được gán cho các ký tự với tần suất xuất hiện cao trong văn bản hoặc tập văn bản.
- Bộ mã tạo ra có tính chất tiên tố.

Thuật toán cơ bản

- 1) Quét văn bản để thống kê tần suất xuất hiện của các ký tự.
- 2) Xây dựng cây mã Huffman dựa trên thống kê tần suất. Mỗi ký tự được biểu diễn bằng một nút lá. Nút lá xa gốc được gán cho ký tự ít xuất hiện.
- 3) Từ mã biểu diễn ký tự được lấy thông qua đường dẫn từ gốc đến nút lá tương ứng. Nhánh rẽ trái và nhánh rẽ phải được ký hiệu bởi bit 0 và 1.

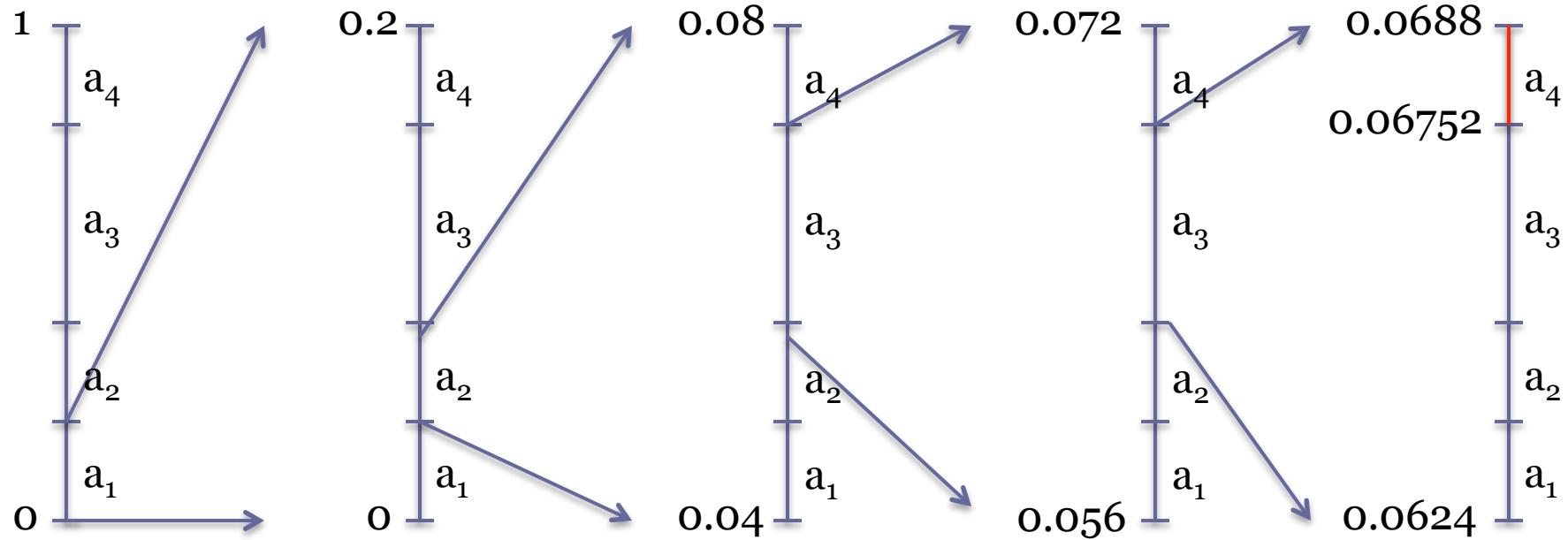
Minh họa

BCAACACDBDCADAEEEABACDBACACDBADABEABEAAA



Mã số học

- Chuỗi ký tự được gán bởi một từ mã số học duy nhất. Độ dài của từ mã tỷ lệ thuận với độ dài của chuỗi ký tự. Mã số học cần phải có bản tin đặc biệt để phân biệt giữa các từ mã.



Mã từ điển Lempel-Ziv

- Mã từ điển xây dựng từ mã mới cho một chuỗi các từ mã.
- **Ví dụ:** Bảng mã ASCII – mỗi từ mã có độ dài 8 bit, cho phép mã hóa 256 ký tự.
- Mã từ điển mở rộng thư viện với từ mã có độ dài từ 9 đến 12 bit. Từ mã mới là chuỗi các từ mã đã xuất hiện.
- Mã từ điển hoạt động không hiệu quả với văn bản gồm những chuỗi ký tự ngắn và khác nhau.

Minh họa

- Xét chuỗi ký tự sau: ABCBCABCABCD

Previous Input	Input	Output	Symbol	Index
NIL	A			
A	B	A	AB	256
B	C	B	BC	257
C	B	C	CB	258
B	C			
BC	A	BC	BCA	259
A	B			
AB	C	AB	ABC	260
C	A	C	CA	261
A	B			
AB	C			
ABC	D	ABC	ABCD	262
D	EOL	D		

Previous Input	Input	Output	Symbol	Index
NIL	A	A		
A	B	B	AB	256
B	C	C	BC	257
C	257	BC	CB	258
BC	256	AB	BCA	259
AB	C	C	ABC	260
C	260	ABC	CA	261
ABC	D	D	ABCD	262

Minh họa

- Mã từ điển được sử dụng trong nén ảnh graphic interchange format (GIF), tagged image file format (TIFF) và portable document format (PDF). Ảnh đa mức xám sử dụng 8 bit/pixel.

39	39	126	126
39	39	126	126
39	39	126	126
39	39	126	126

39-39-126-126-256-258-260-259-257-126

NÉN ẢNH FAX

Trần Quang Đức

Chuẩn né

- ITU-T Nhóm 1,2
 - Kỹ thuật điều chế: FM, AM, PM
 - Màu trắng: 1300 Hz, 1500 Hz
 - Màu đen: 2100 Hz, 2400 Hz
- ITU-T Nhóm 3: máy fax được thiết kế hoạt động trong mạng PSTN (9600 baud).
- ITU-T Nhóm 4: Máy fax được thiết kế hoạt động trong mạng ISDN (64K baud).

ITU-T Nhóm 3

- Hướng quét: Trái sang phải, Trên xuống dưới
- Độ rộng đường quét: 215 (255, 303)
- Số lượng pixel trên một đường: 1728 (2048, 2432)
- Mã hóa: RLC+MH (MMR - Nhóm 4, Tùy chọn nhóm 3)
- Tỷ lệ nén: 5%-20% của kích thước ban đầu (đến 95%)

Mã Huffman hiệu chỉnh

- Loạt điểm ảnh trắng có độ dài 1664 được thay thế bằng từ mã ngắn 011000.
- Mã Huffman được hiệu chỉnh để mã hóa những loạt điểm ảnh có độ dài là bội số của 64.
- Độ dài của loạt điểm ảnh có thể biểu diễn bằng một từ mã hoặc một vài từ mã (trong trường hợp loạt quá dài).

Mã Huffman hiệu chỉnh

Run Length	White Code Word	Black Code Word
0	00110101	0000110111
1	000111	010
2	0111	11
3	1000	10
4	1011	011
5	1100	0011
...
61	00110010	000001011010
62	00110011	000001100110
63	00110100	000001100111

Mã Huffman hiệu chỉnh

Run Length	White Code Word	Black Code Word
64	11011	0000001111
128	10010	000011001000
192	010111	000011001001
256	0110111	000001011011
320	00110110	000000110011
384	0110111	000000110100
...		
2432	000000011101	Same as white
2496	000000011110	Same as white
2560	000000011111	Same as white

Mã Huffman hiệu chỉnh

- Loại gồm 5 điểm ảnh trắng được mã hóa 1100
- Loại gồm 69 điểm ảnh trắng được mã hóa: 64+5
- Loại gồm 64 điểm ảnh trắng được mã hóa: 64+0
- Loại gồm 2561 điểm ảnh trắng được mã hóa: 2560+1
- Từ mã EOL: 000000000001

Mã Huffman hiệu chỉnh

- Không có loạt có độ dài 0. Tại sao vẫn có từ mã cho loạt gồm 0 điểm ảnh trắng và 0 điểm ảnh đen?
- Thường một dòng quét có kích thước 8.5 inch tương đương 1728 điểm ảnh, tại sao vẫn có mã cho những loạt có độ dài 2561?
- Loạt gồm 5 điểm ảnh đen có mã 0011, cũng là tiền tố của loạt gồm 61, 62 hoặc 63 điểm ảnh trắng. Giải thích lý do!
- Tìm tỷ số nén của ảnh gồm các điểm ảnh đen và trắng nằm xen kẽ nhau.

NÉN ẢNH JPEG

Trần Quang Đức

Giảm mẫu kênh màu



4:4:4

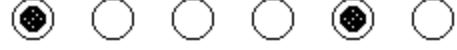


4:2:2

○ -- Pixel with only Y value

● -- Pixel with only Cr and Cb values

◎ -- Pixel with Y, Cr and Cb values



4:1:1



4:2:0

Giảm mẫu kênh màu

- **4:4:4 (1:1)**
 - Thiết bị quét phim cao cấp
 - Điện ảnh
- **4:2:2 (3:2)**
 - Định dạng video cao cấp
- **4:1:1 (2:1)**
 - DVCPRO (ví dụ NTSC, PAL)
- **4:2:0 (2:1)**
 - MPEG, mã hóa video H.26X
 - DVD, Blue-ray
 - JPEG, MJPEG

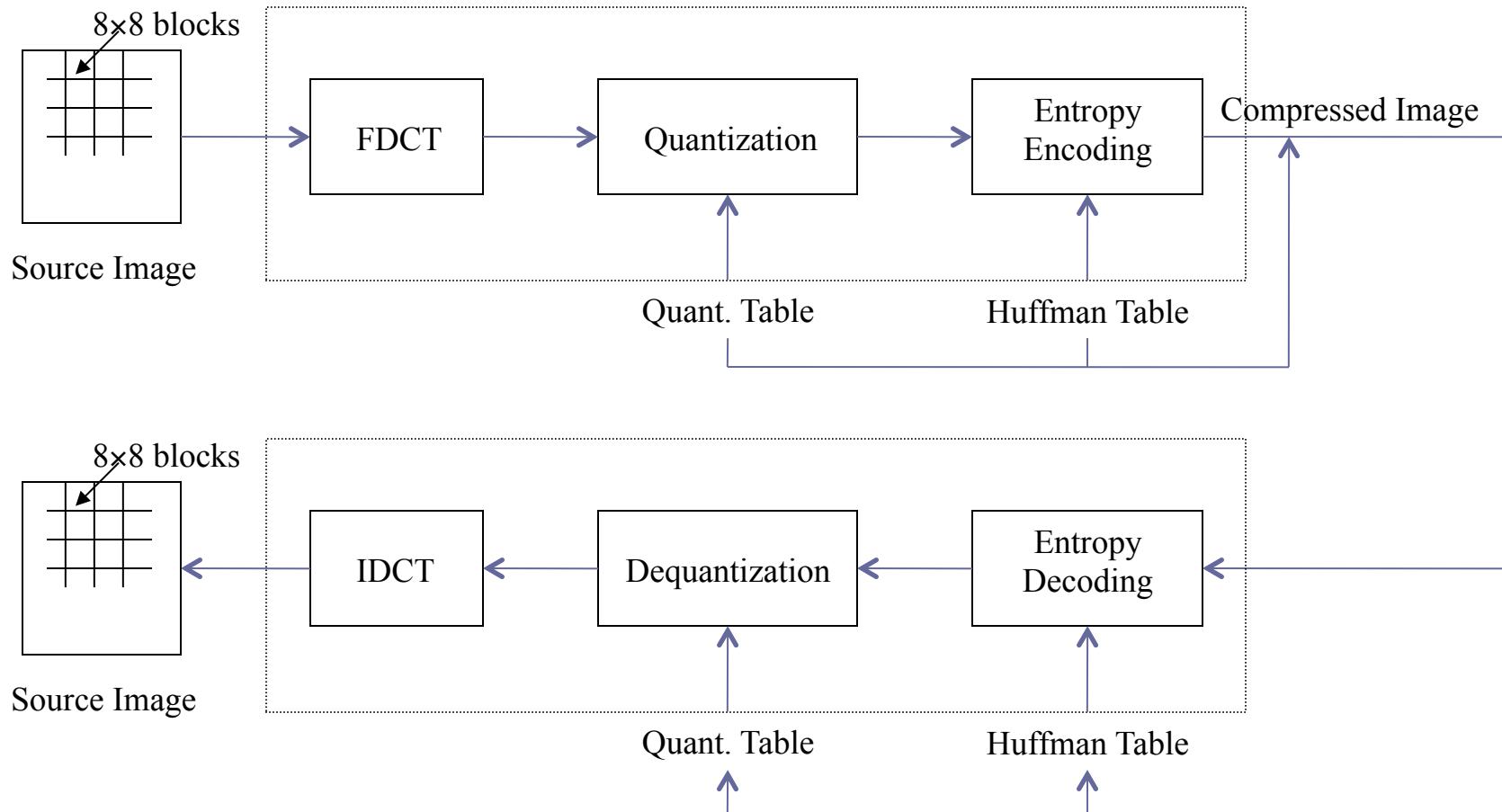
Chuẩn JPEG

- JPEG là chuẩn nén ảnh, được phát triển bởi “Joint Photographic Experts Group”.
- JPEG là định dạng thường được sử dụng trong lưu và truyền ảnh.
- JPEG nén được trên cả ảnh màu và đen trắng
- Tỷ lệ nén 1:10 nhưng ảnh hưởng nhiều đến cảm nhận của mắt người.

Chuẩn JPEG

- JPEG là kỹ thuật nén mất mát thông tin dựa trên biến đổi 2D-DCT (Discrete Cosine Transform).
 - **Nhận định 1:** Nội dung ảnh thay đổi tương đối chậm trên bề mặt của ảnh.
 - **Nhận định 2:** Mắt người nhạy cảm hơi với những mất mát thuộc về thành phần tần số thấp.
 - **Nhận định 3:** Mắt người cảm nhận tốt hơn sự thay đổi về độ chói so với màu sắc.

Chuẩn JPEG



Biến đổi DCT

- DCT biến đổi tín hiệu ảnh từ miền không gian sang miền tần số. Năng lượng ảnh tập chung chủ yếu ở các thành phần tần số thấp, nằm góc trên cùng bên trái của DCT. Các thành phần tần số cao thường có giá trị thấp trong biến đổi DCT nên có thể loại bỏ để tăng hiệu quả của quá trình nén.

$$F(u, v) = \frac{4C_k C_l}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f(x, y) \cos\left(\frac{\pi(2k+1)u}{2M}\right) \cos\left(\frac{\pi(2l+1)v}{2N}\right)$$

$$C_k = \begin{cases} 1/\sqrt{2} & If \ k = 0 \\ 1 & Otherwise \end{cases}$$

$$C_l = \begin{cases} 1/\sqrt{2} & If \ k = 0 \\ 1 & Otherwise \end{cases}$$

Biến đổi DCT

52	55	61	66	70	61	64	73
63	59	55	90	109	85	69	72
62	59	68	113	144	104	66	73
63	58	71	122	154	106	70	69
67	61	68	104	126	88	68	70
79	65	60	70	77	68	58	75
85	71	64	59	55	61	65	83
87	79	69	68	65	76	78	94

DCT

-145	-30	-61	27	56	-20	-2	0
4	-22	-61	10	13	-7	-9	5
--47	7	77	-25	-29	10	5	-6
-49	12	34	-15	-10	6	2	2
12	-7	-13	-4	-2	2	-3	3
-8	3	2	-6	-2	1	4	2
-1	0	0	-2	-1	-3	4	-1
0	0	-1	-4	-1	0	1	2

- Thành phần DC (i.e., $F(0,0)$) và AC (i.e., $F(u,v)$ ($u,v \neq 0$)) đều là số nguyên, có giá trị nằm trong khoảng từ -1024 đến 1023.

Lượng tử hóa

- Lượng tử hóa loại bỏ những thông tin thừa trong ảnh. Lượng tử hóa là nguyên nhân chủ yếu dẫn đến mất mát thông tin trong các kỹ thuật dựa trên biến đổi DCT.

$$F^Q(u,v) = \text{Integer Round } (F(u,v)/Q(u,v))$$

Quantization step size

- Bước lượng tử hóa được chọn dựa trên cảm nhận của mắt người. Nó là hàm với các biến gồm đặc trưng ảnh, đặc trưng hiển thị và khoảng cách hiển thị.
- Trong ứng dụng, bước lượng tử hóa thường được xác định dựa trên thực nghiệm.

Lượng tử hóa

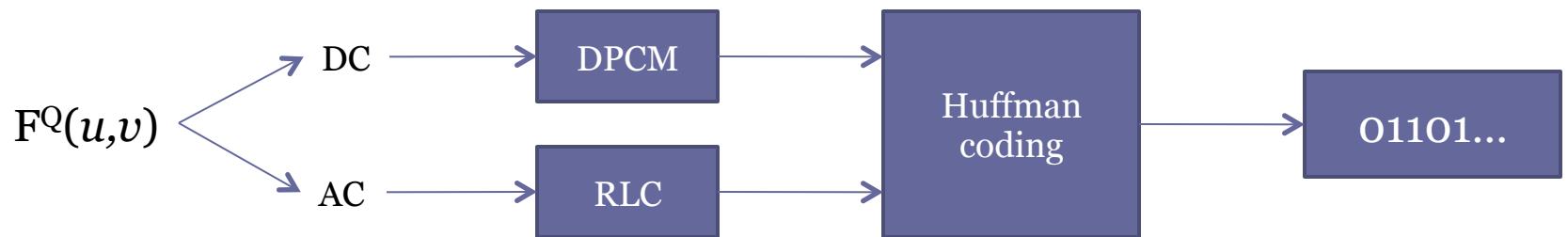
16	11	10	16	24	40	51	61
12	18	14	10	26	58	60	55
14	13					69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Tần số tăng dần

-26	-3	-6	2	2	-1	0	0
0	-2	-4	1	1	0	0	0
-3	1	5	-1	-1	0	0	0
-4	1	2	-1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

- Những thành phần tần số cao thường có giá trị bằng 0 hoặc rất nhỏ (kể cả giá trị âm và giá trị dương) nên có thể được biểu diễn bằng một vài bit.

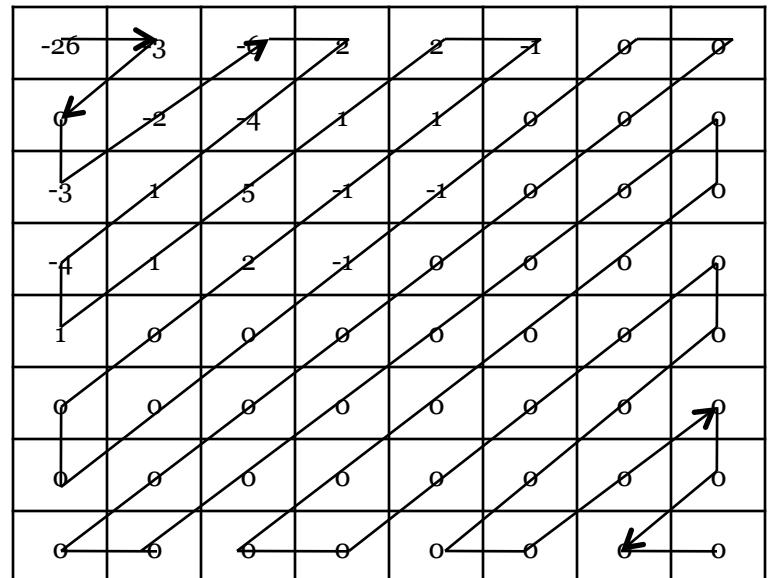
Mã hóa Entropy



- Mã hóa Entropy là dạng đặc biệt của nén không mất mát thông tin.
- Mã hóa Entropy gồm các bước sau: (1) sắp xếp các thành phần ảnh theo đường díc dắc; (2) mã hóa các thành phần ảnh bằng mã loạt dài (RLC) và Huffman.

Đường díc dắc

-26	-3	-6	2	2	-1	0	0
0	-2	-4	1	1	0	0	0
-3	1	5	-1	-1	0	0	0
-4	1	2	-1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0



- Tất cả các hệ số đều nằm trên đường díc dắc. Các thành phần tần số thấp (thường khác 0) được xếp trước các thành phần tần số cao.

DPCM và Mã loạt dài

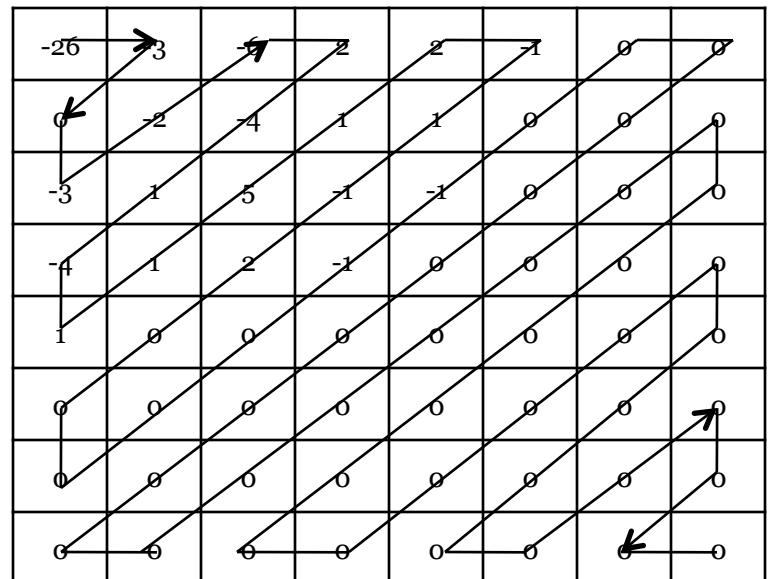
- Các thành phần DC của các khối ảnh lân cận có tính tương quan cao. Vì vậy, mã hóa sự sai khác giữa các thành phần trên làm tăng tính hiệu quả của quá trình nén.



- Mã loạt dài được sử dụng để mã hóa các thành phần AC

DPCM và Mã loạt dài

- Thành phần DC: $+3 \rightarrow (2), (3)$
- Thành phần AC:
 - $(0,2) (-3)$
 - $(1,2) (-3)$
 - $(0,2) (-2)$
 - $(0,3) (-6)$
 - $(0,2) (2)$
 - ...
 - $(0,0) \rightarrow \text{EOB}$
- Các thành phần AC được mô tả bởi một cặp ký hiệu (*loạt dài*, *kích thước*) (*biên độ*). Loạt dài được tính bằng số lượng 0 giữa hai thành phần khác 0. Kích thước biểu diễn số lượng bit dùng để mã hóa giá trị biên độ.



Mã Huffman

- Thực hiện tăng hiệu quả nén bằng việc thay thế những chuỗi ký tự dài bằng những từ mã ngắn.
- Độ dài của mỗi từ mã được xác định dựa trên tần suất xuất hiện của các ký tự.
- JPEG cung cấp bảng mã Huffman chuẩn, nhưng bảng mã Huffman cũng có thể được định nghĩa riêng tùy thuộc vào ứng dụng và đặc trưng của các ảnh sử dụng trong ứng dụng.

Mã Huffman

Category	Values	Bits for the value
1	-1,1	0,1
2	-3,-2,2,3	00,01,10,11
3	-7,-6,-5,-4,4,5,6,7	000,001,010,011,100,101,110,111
4	-15,...,-8,8,...,15	0000,...,0111,1000,...,1111
5	-31,...,-16,16,...31	00000,...,01111,10000,...,11111
6	-63,...,-32,32,...63	000000,...,011111,100000,...,111111
7	-127,...,-64,64,...,127	0000000,...,0111111,1000000,...,1111111
8	-255,...,-128,128,...,255	...
9	-511,...,-256,256,...,511	...
10	-1023,...,-512,512,...,1023	...
11	-2047,...,-1024,1024,...,2047	...

Giá trị và số lượng bit cho mỗi giá trị

Mã Huffman

Run, category	Code Length	Codeword
0,0	4	1010
0,1	2	00
0,2	2	01
...
0,10	16	1111111100000011
1,1	4	1100
1,2	5	11011
...
15,10	16	111111111111110

Bảng mã Huffman chuẩn cho các thành phần AC của kênh độ chói

Mã Huffman

Category	Code Length	Codeword
0	2	00
1	3	010
2	3	011
3	3	100
4	3	101
5	3	110
6	4	1110
7	5	11110
8	6	111110
9	7	1111110
10	8	11111110
11	9	111111110

Bảng mã chuẩn cho các thành phần DC của kênh độ chói

Mã Huffman

- Thành phần DC: +3 → (2),(3)
- Từ mã của thành phần DC: 011 11
- Thành phần AC: (0,2) (-3), (1,2) (-3)... EOB
- Từ mã của thành phần AC: 01 00 11011 00 ... 00

Minh họa



1



2



3



4



5

1. $Q = 100 - 83,2$ bytes
2. $Q = 50 - 15,1$ bytes
3. $Q = 25 - 9,5$ bytes
4. $Q = 10 - 4,7$ bytes
5. $Q = 5 - 1,5$ bytes

NÉN ẢNH JPEG2000

Trần Quang Đức

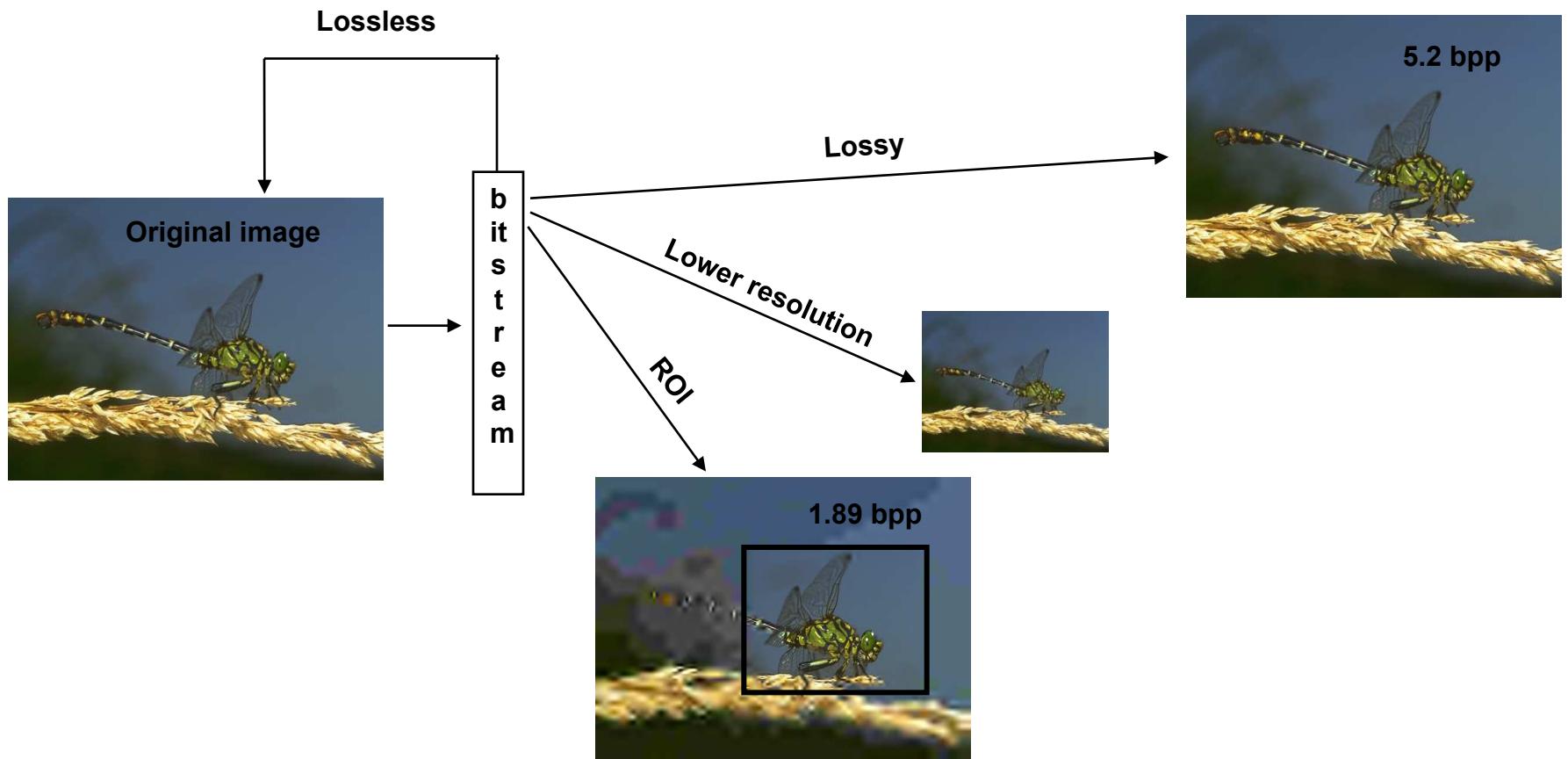
Chuẩn nén khác?

- Với tỷ lệ nén cao (ví dụ 0.25 bpp), méo của chuẩn JPEG là không thể chấp nhận.
- JPEG có 44 cách thức, rất nhiều trong số đó không được sử dụng trong bộ mã hóa JPEG.
- Chất lượng của ảnh giảm đi rõ rệt nếu xuất hiện lỗi trong quá trình truyền ảnh.
- JPEG được tối ưu hóa cho ảnh tự nhiên nhưng không phù hợp với ảnh đồ họa.
- JPEG không áp dụng được cho ảnh nhị phân.

Chuẩn JPEG2000

- **Đặc trưng**
 - Biến đổi sóng con Wavelet
 - Chất lượng tốt với tỷ lệ nén cao
 - Nén ảnh nhị phân
 - Nén mất mát và không mất mát thông tin
 - Linh hoạt với lỗi trong quá trình truyền
 - Có khả năng mở rộng
 - Chọn vùng quan tâm ROI
- **Một vài ứng dụng**
 - Website
 - Camera kỹ thuật số
 - Ảnh y học
 - Ảnh viễn thám

Chuẩn JPEG2000



Chuẩn JPEG2000



0.125 bpp

JPEG vs. JPEG2000



0.25 bpp



Chuẩn JPEG2000

- Nhược điểm
 - Thiết bị mã hóa và giải mã phức tạp, thời gian tính toán lâu.
 - Khi tỷ lệ nén khoảng 1:25, JPEG2000 so với JPEG có thể tạo ra ảnh nén ít bị chia ô nhưng giảm đang kể các chi tiết ảnh.