XỬ LÝ ẢNH

Khuất Thị Ngọc Ánh

Thông tin giảng viên

- Thông tin liên hệ
  - Bộ môn Hệ thống thông tin
  - Khoa Công Nghệ Thông Tin
  - Đại học Công nghệ Giao thông vận tải
  - E-mail: anhktn@utt.edu.vn
  - Điện thoại: (+84) 392120967

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

Khuất Thị Ngọc Ánh

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.

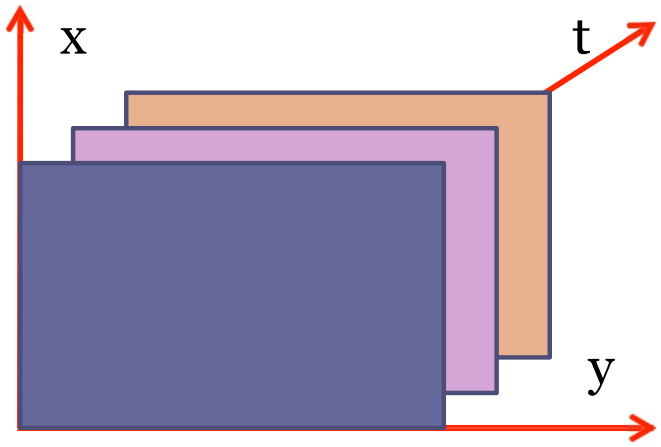
- Đối tượng của xử lý ảnh là xử lý các ảnh tự nhiên, ảnh chụp, dữ liệu ảnh có nguồn gốc từ tín hiệu ảnh đặc trưng bởi biên độ và dải tần số. Có sự phân biệt giữa xử lý ảnh với đồ họa.

- 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

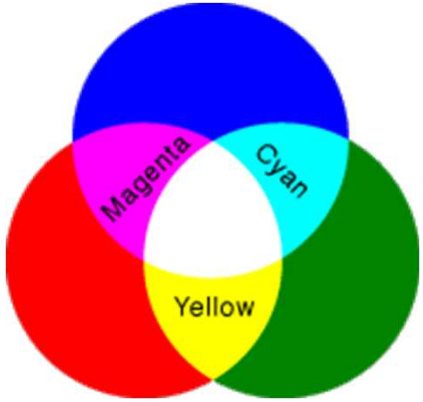
- **Ảnh tĩnh (Still Image)**: Biểu diễn bởi hàm độ chói của các biến toạn độ trong mặt phẳng ảnh I(x,y).

- **Chuỗi ảnh (Sequence of Images)**: Hàm độ chói của các biến tọa độ mặt phẳng và biến thời gian I(x,y,t).

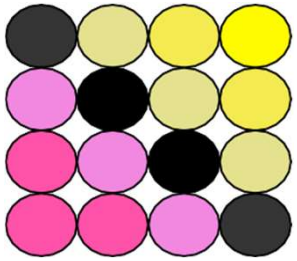


Ảnh tĩnh

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



Pixel (RGB)

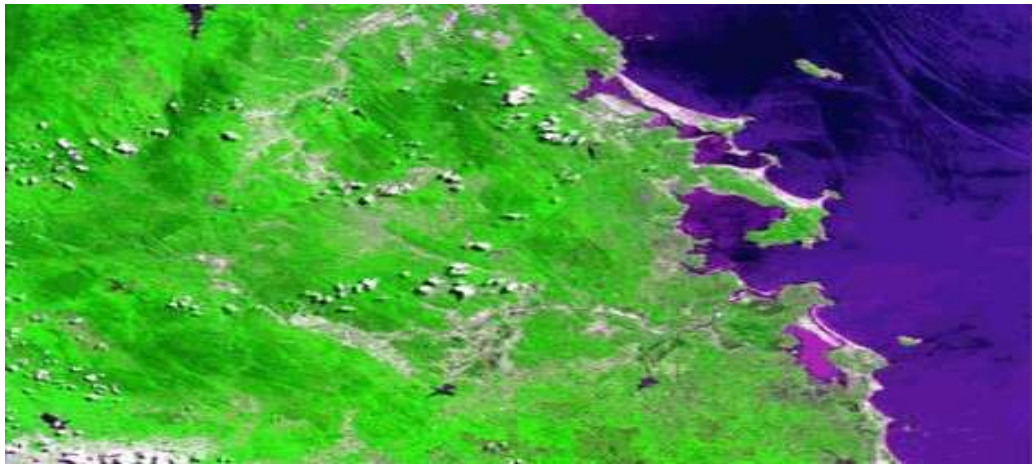


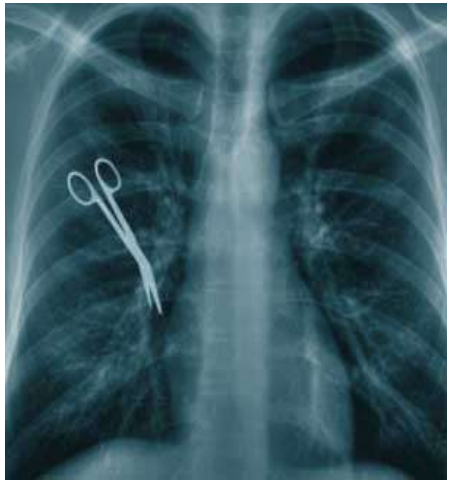
Chuỗi ảnh

- **Video:** Chuỗi các ảnh (khung hình), quan hệ thời gian giữa các khung hình biểu diễn ảnh động.

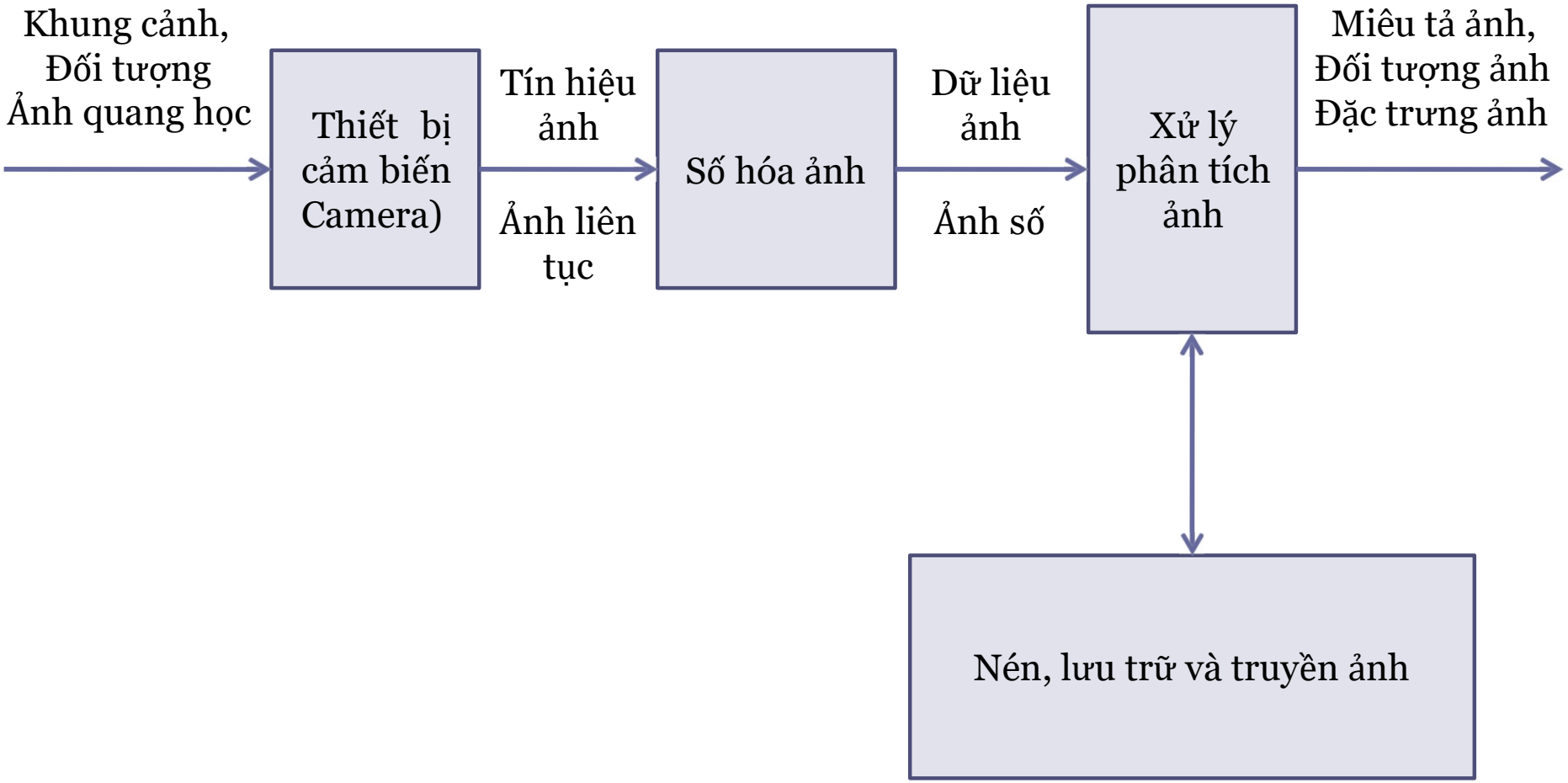
- Tấn số và Độ phân giải
  - NTSC (525 dòng, 30 khung hình/giây)
  - PAL (625 dòng, 25 khung hình/giây)
  - SECAM (625 dòng, 25 khung hình/giây)
  - Phim (24 khung hình/giây)
  - HDTV (16:9, 720 dòng, 60 khung hình/giây)
  - SVGA: 72 khung hình/giây, 1024x720 pixel

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

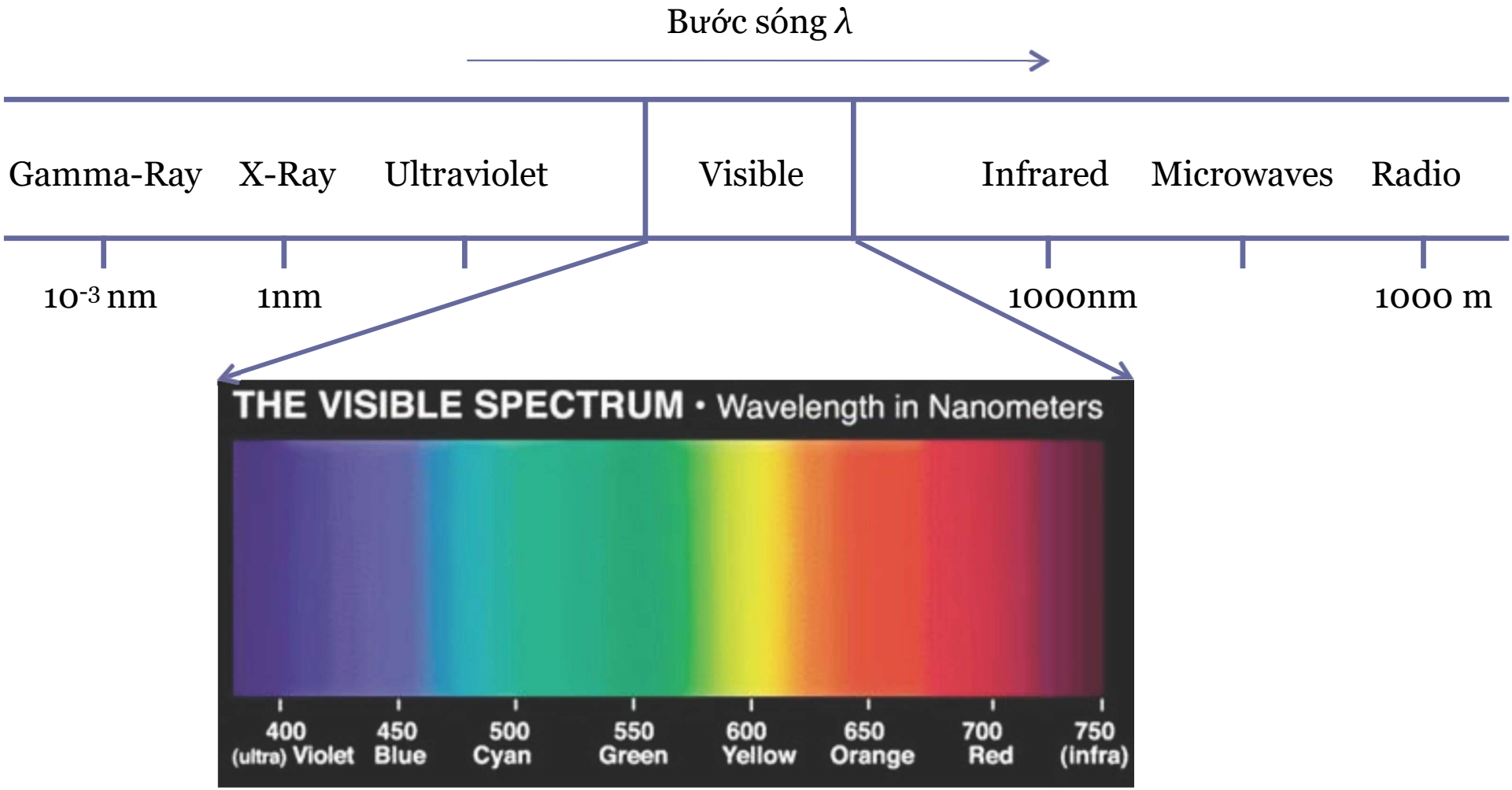
Ứng dụng của xử lý chuỗi ảnh

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

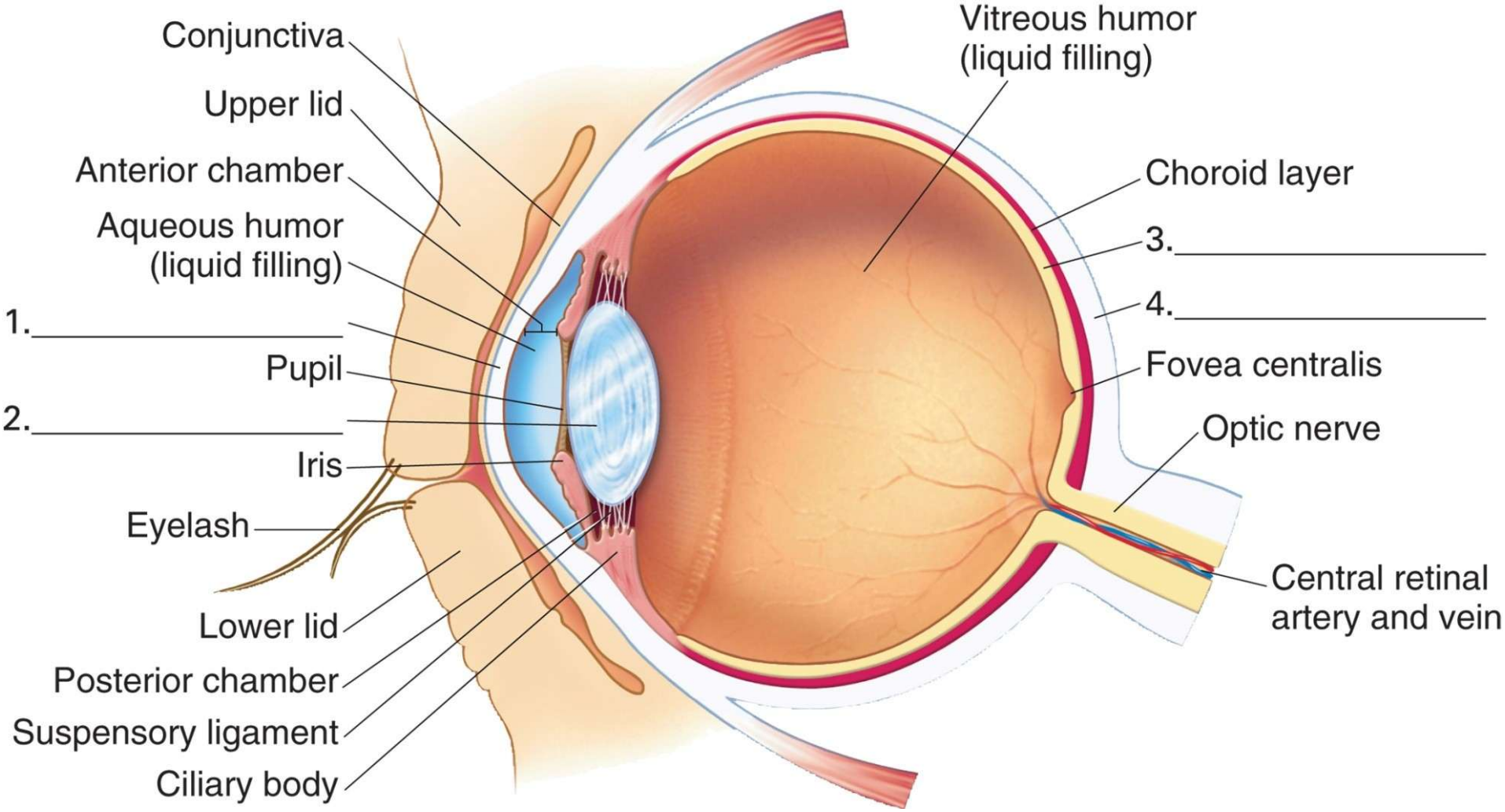
HÌNH THÀNH VÀ BIỂU DIỄN ẢNH

Khuất Thị Ngọc Ánh

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

Màu sắc

- Cảm nhận màu sắc phụ thuộc vào phổ của ánh sáng. Ánh sáng nhìn thầy 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.

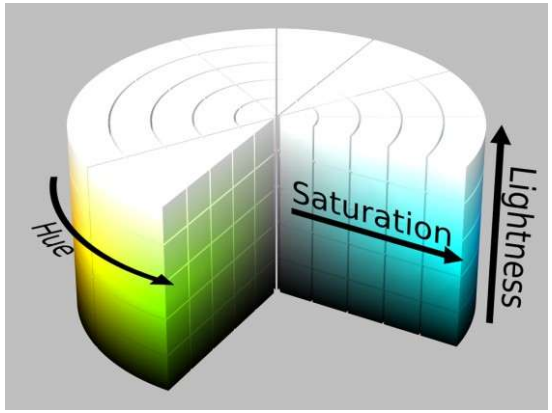
- Một màu có thể tạo nên bằng cách trộn 3 màu cơ bản tương đương với 3 dạng tế bào cảm nhận màu sắc (Đỏ, Lục, Lam). Phân biệt màu sắc dựa trên độ sáng, sắc độ và độ bão hòa.

Độ sáng, sắc độ và độ bão hòa

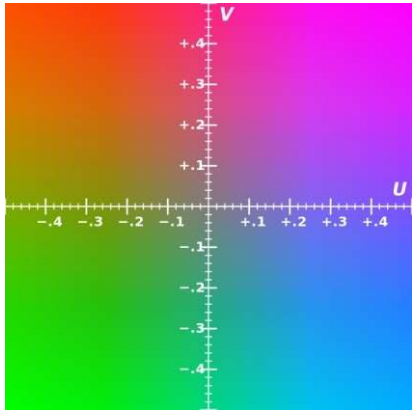
- **Độ sáng (Brightness)**
  - Thuộc tính chủ quan, đặc trưng cho khả năng cảm nhận độ rọi
  - Phụ thuộc vào độ rọi của môi trường xung quanh

- **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.

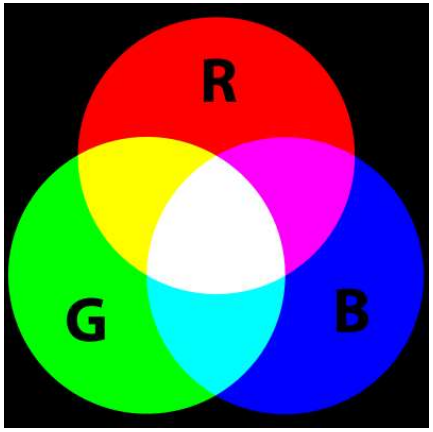
Các mô hình màu



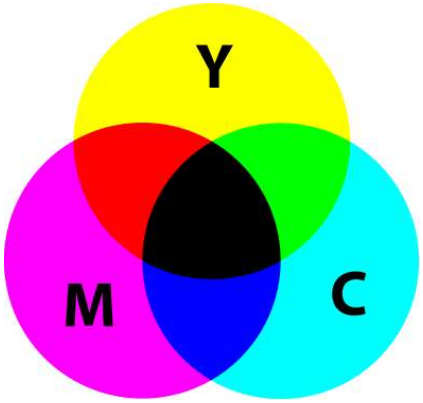
HSV



YCbCr (Y=0.5)



RGB



CMYK

Các mô hình màu

- **RGB** (Red-Green-Blue): Hệ màu RGB dựa trên việc kết hợp ba màu cơ bản là Đỏ (Red), Xanh lá cây (Green) và Xanh dương (Blue) để tạo ra tất cả các màu khác. Mỗi màu cơ bản trong hệ RGB có giá trị từ 0 đến 255. Khi kết hợp các mức màu cơ bản này, ta có thể tạo ra một loạt các màu sắc khác nhau

- **Hệ màu HSV** (Hue-Saturation-Value) là một cách biểu diễn màu sắc trong không gian màu dựa trên ba thành phần chính: Hue (sắc độ), Saturation (Độ bão hòa) và Value (Giá trị). Hệ màu HSV thường được sử dụng để mô tả màu sắc một cách trực quan và dễ hiểu hơn so với hệ màu RGB.

- **CMY(K)** (Cyan-Magenta-Yellow-(Black)): Hệ màu CMYK được sử dụng chủ yếu trong in ấn. Nó sử dụng bốn màu cơ bản là Cyan, Magenta, Yellow và Black để tạo ra màu sắc. Khi kết hợp các mức màu này, hệ thống tạo ra màu bằng cách loại bỏ ánh sáng từ màu cơ bản. Lớp màu đen (K) được thêm vào để cải thiện độ tương phản và định nghĩa màu

Các mô hình màu

- **Hệ màu HSL** (Hue-Saturation-Lightness) là một cách biểu diễn màu sắc trong không gian màu dựa trên ba thuộc tính chính: Hue (Màu sắc), Saturation (Độ bão hòa) và Lightness (Độ sáng). Hệ màu HSL cung cấp một cách tiếp cận khác để hiểu và điều chỉnh màu sắc so với hệ màu RGB hoặc hệ màu HSV.

- **Hệ màu YCbCr** (hoặc YUV) là một cách biểu diễn màu sắc trong không gian màu dựa trên ba thành phần chính: Y (Luma), Cb (Chrominance xanh dương), và Cr (Chrominance đỏ).

Hệ màu YCbCr thường được sử dụng trong việc nén ảnh và video, cũng như trong các ứng dụng truyền hình và phương tiện truyền thông. Nó cho phép tách luma (mức xám) và chrominance (màu sắc) của hình ảnh, cho phép tiết kiệm băng thông khi truyền tải ảnh và video.

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

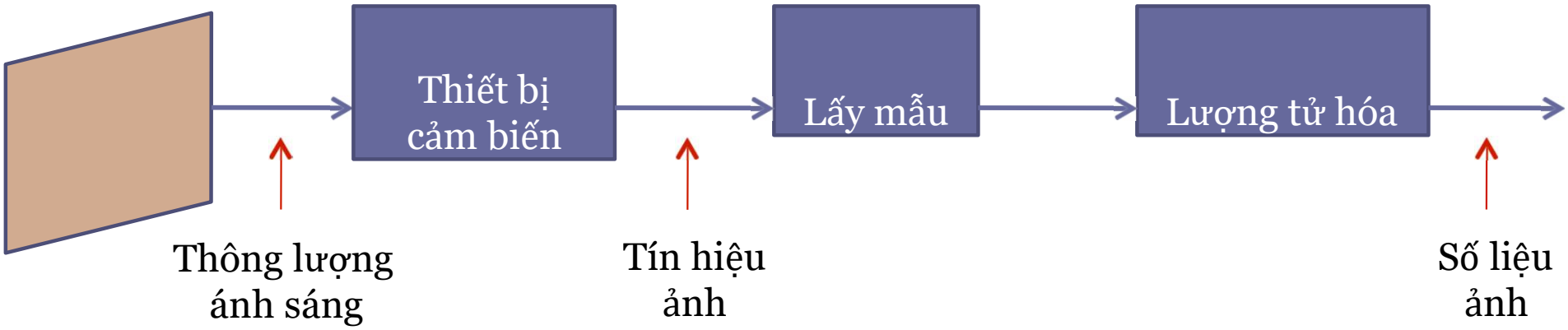
0	1	1
0	1	0
0	0	1

0	1	1
0	1	0
0	0	1

THU NHẬN VÀ SỐ HÓA ẢNH

Khuất Thị Ngọc Ánh

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

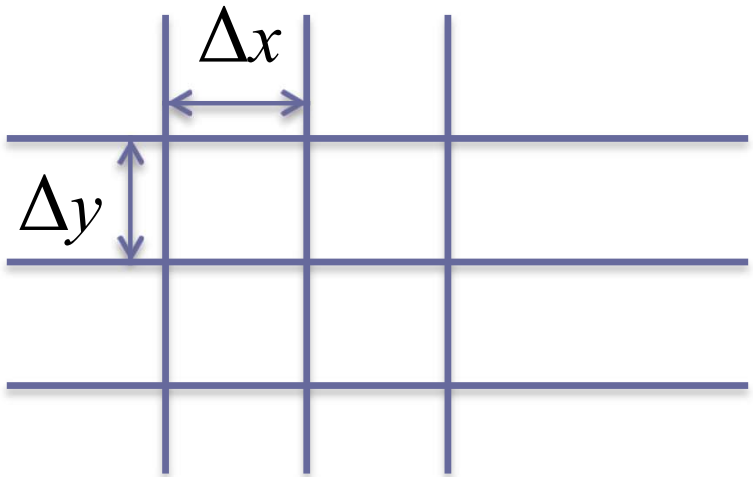
Lượng tử hóa: Lượng tử hóa Lloyd-Max

Lấy mẫu

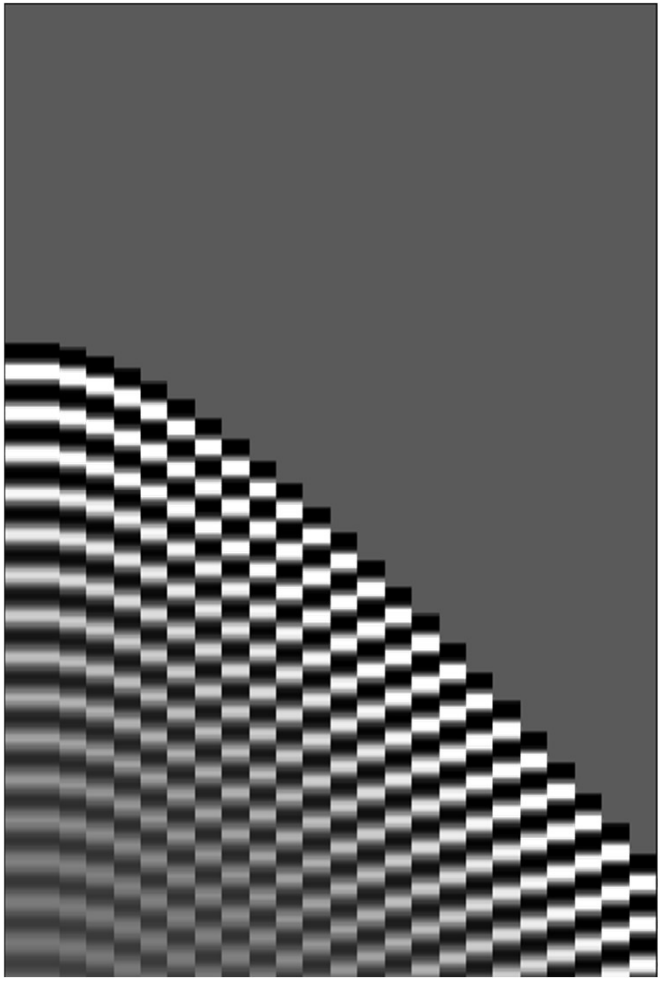
- Ảnh với dải giới hạn (Band limited images): Một hàm *f(x,y)* gọi là giải giới hạn nếu khai triên Fourier *F(u,v)* của nó là 0 bên ngoài miền bao, cụ thể *F(u,v)* = *0* nếu $|u|>u_o$ và $|v|>v_o$.

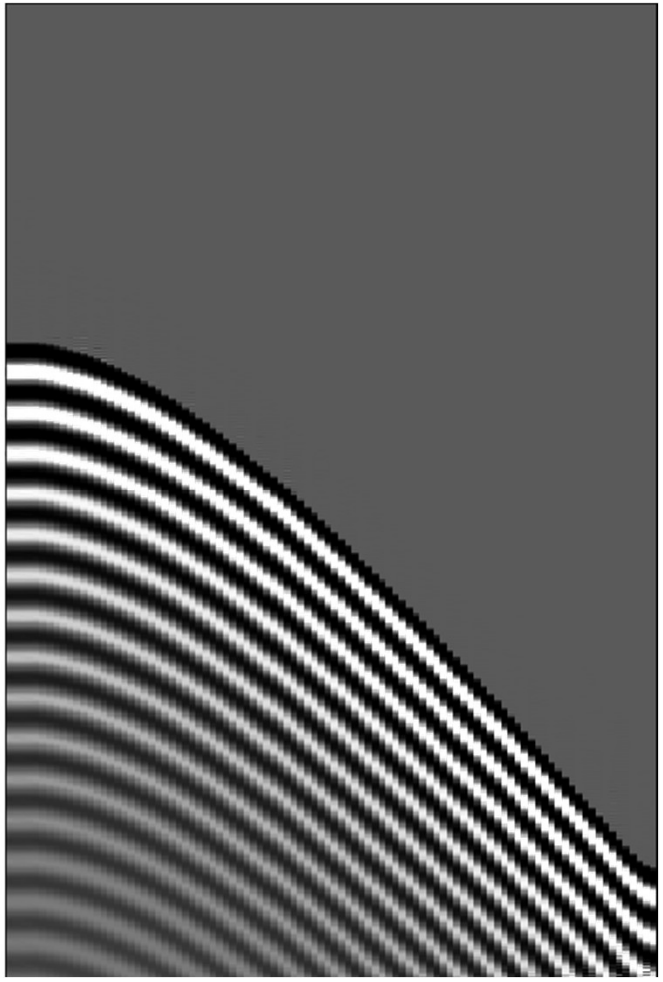
- Tỷ số Nyquist: Tần số lấy mẫu theo *x* và *y* tương đương hoặc lớn hơn 2 lần dải giới hạn u_o và v_o

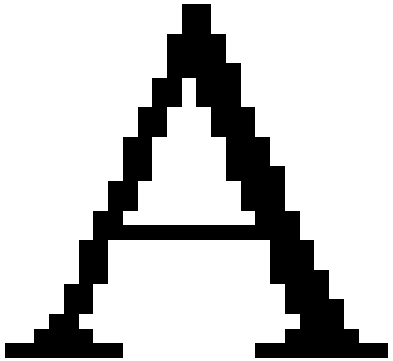
$$\frac{1}{\Delta x} \geq 2u_0 \qquad \frac{1}{\Delta y} \geq 2v_0$$

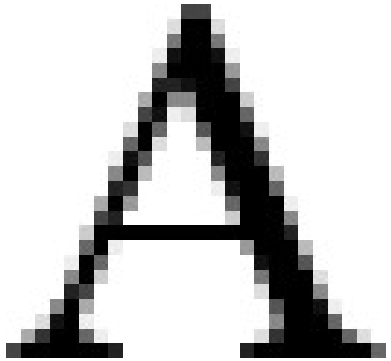


Hiện tượng răng cưa

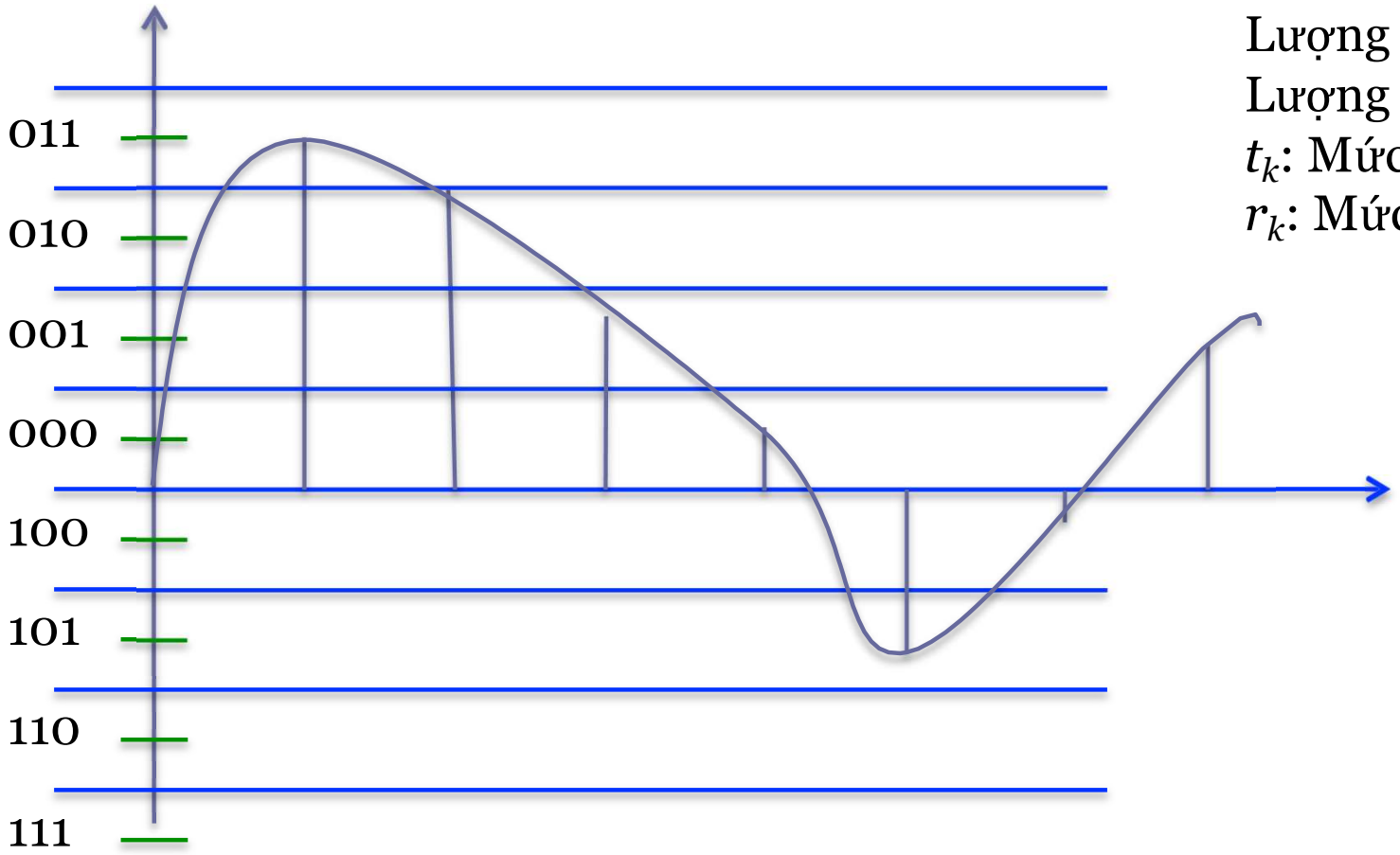








Lượng tử hóa



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.

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

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

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.

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

- Từ đó, ta có

$$t_k = \frac{r_k + r_{k-1}}{2} \qquad r_k = \frac{\int_{t_k}^{t_{k+1}} up(u)\, du}{\int_{t_k}^{t_{k+1}} p(u)\, du}$$

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

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

- Sai số lượng tử

$$\frac{q}{12} \qquad q = r_k - r_{k-1} = t_{k+1} - t_k$$

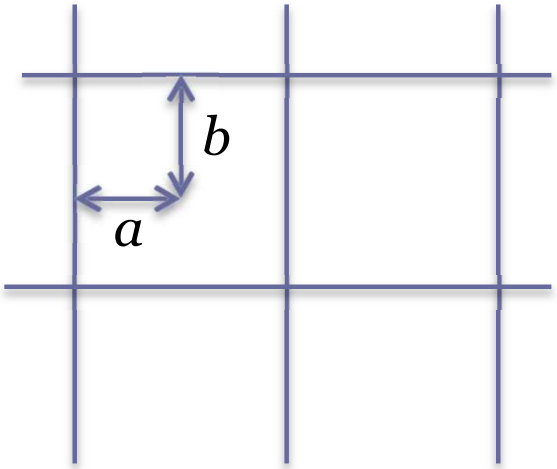
Thay đổi kích thước ảnh

- **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

Khuất Thị Ngọc Ánh

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 $f_2(m,n)$ tao nên tổ hợp tuyến tính của tín hiệu đầu ra $g_1(m,n)$ và $g_2(m,n)$.

$$T(\alpha_1 f_1(m,n) + \alpha_2 f_2(m,n)) = \alpha_1 T(f_1(m,n)) + \alpha_2 T(f_2(m,n))$$

$$= \alpha_1 g_1(m,n) + \alpha_2 g_2(m,n)$$

Tích chập

- Tích chập - Trường hợp liên tục

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

$$= \int_{\tau_1=-\infty}^{+\infty} \int_{\tau_2=-\infty}^{+\infty} f(\tau_1,\tau_2) h(x-\tau_1, y-\tau_2) d\tau_1 d\tau_2$$

- Tích chập - Trường hợp rời rạc

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

$$= \sum_{n_1=-\infty}^{+\infty} \sum_{n_2=-\infty}^{+\infty} f[n_1,n_2] h[x-n_1, y-n_2]$$

Tích chập

- Tính giao hoán

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

- Tính kết hợp

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

- Tính phân phối

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

Biến đổi Fourier

- Biến đổi Fourier rời rạc cho tín hiệu một chiều

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

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

- 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

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

- *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

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

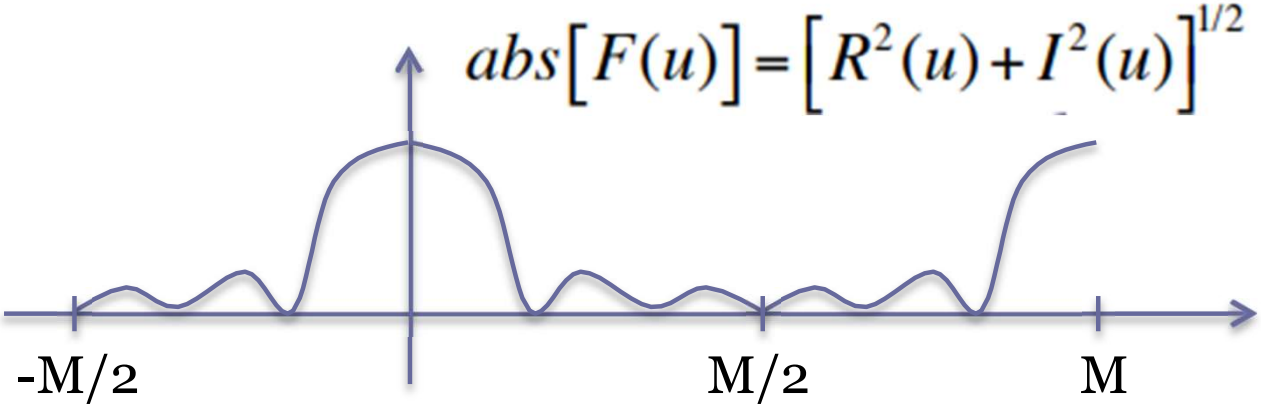
$$f(x,y)g(x,y) \Leftrightarrow F(u,v) * G(u,v)$$

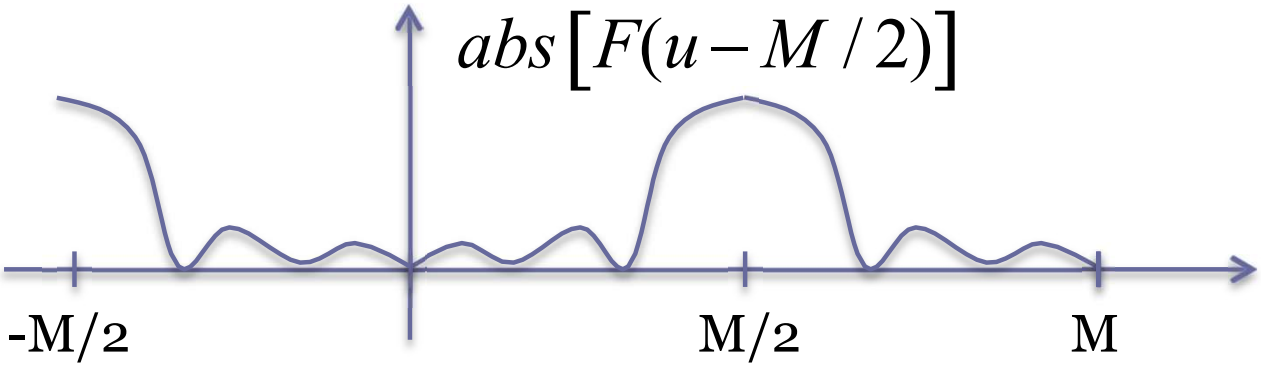
Biến đổi Fourier

$$P(u) = \left[F(u)\right]^2$$

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

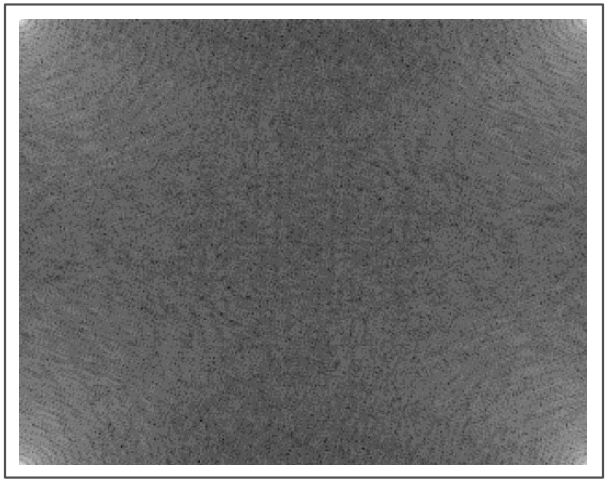
$$\Im\left[f(x)(-1)^x\right]$$



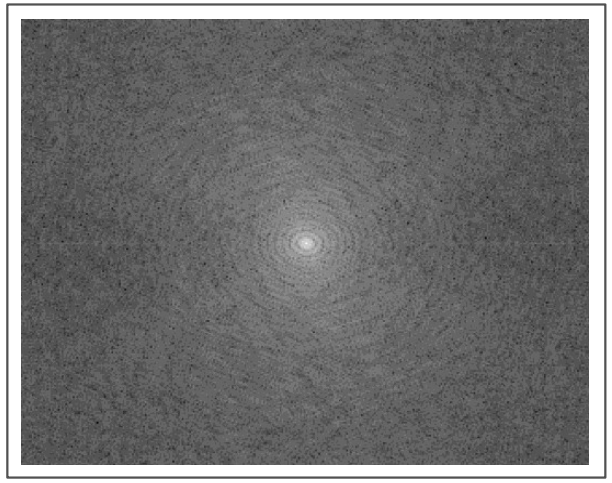


Biến đổi Fourier

(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)*

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

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 **x** là véc tơ chứa giá trị của các điểm ảnh trên

$$\mathbf{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{Bmatrix}$$

$$m_x = \mathrm{E}[\mathbf{x}]$$

Biến đổi KL

- Từ đó, ta có ma trận hiệp phương sai

$$C_x = \mathrm{E}\left[(\mathbf{x} - m_x)(\mathbf{x} - m_x)^T\right] = \mathrm{E}\left[\mathbf{x}\mathbf{x}^T\right] - \mathrm{E}\left[m_x m_x^T\right]$$

- Giá trị $c_{i,j}$ của C_x thể hiện mối tương quan giữa hai điểm ảnh x_i và x_j. Gọi **U** là ma trận của biến đổi tuyến tính $\mathbf{y}=\mathbf{U}^T(\mathbf{x}-m_x)$ 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 **y** có thể xác định như sau

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

Biến đổi KL

- Để giảm hoàn toàn tính tương quan giữa các điểm ảnh trong không gian biến đổi, **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

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

Biến đổi KL

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

$$C_y = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}$$

$$\begin{aligned} C_y &= \mathrm{E}\left[(\mathbf{y}-m_y)(\mathbf{y}-m_y)^T\right] = \mathrm{E}\left[\mathbf{y}\mathbf{y}^T\right] \\ &= \mathrm{E}\left[\mathbf{U}^T(\mathbf{x}-m_x)(\mathbf{x}-m_x)^T\mathbf{U}\right] \\ &= \mathbf{U}^T C_x \mathbf{U} \end{aligned}$$

- 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 **x** có thể khôi phục dựa trên phép biến đổi

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

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

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

- Với

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

Biến đổi KL

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

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

Biến đổi KL

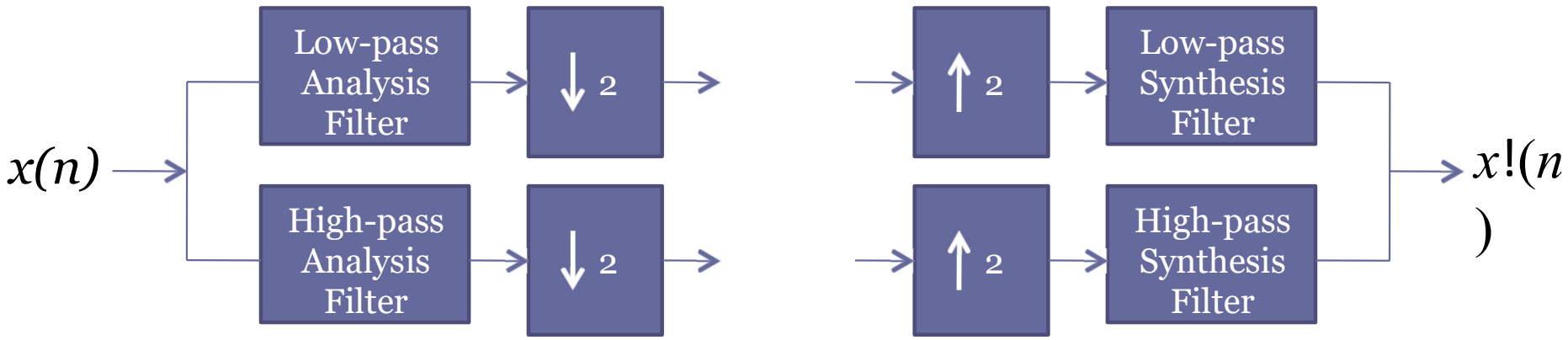
- Năng lượng của khối ảnh trước và sau khi biến đổi được bảo toàn.

- 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).



$h_0(n)$: Low-pass Analysis Filter
$h_1(n)$: High-pass Analysis Filter

$g_0(n)$: Low-pass Synthesis Filter
$g_1(n)$: 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

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

- Tăng mẫu 2 lần trên miền thời gian

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

Biến đổi Wavelet

- Giảm mẫu và tăng mẫu trên miên thời gian:

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

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

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

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

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

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

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 nhân 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)

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

CẢI THIỆN VÀ NÂNG CẤP ẢNH

Khuất Thị Ngọc Ánh

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
Số học/logic
Trừ ảnh
Trung bình ảnh

Toán tử trên miền không gian

Lọc làm tròn ảnh:
+ Lọc trung bình
+ Lọc gaussian
+ Lọc thống kê thứ tự:
  Lọc trung vị
  Lọc cực đại
  Lọc cực tiểu
Lọc tăng cường độ nét:
Laplacian, Robert, Sobel
Lọc High-boost

Toán tử trên miền tần số

Lọc thông thấp
+ lý tưởng
+ butterworth
+ gaussian
Lọc thông cao
+ lý tưởng
+ butterworth
+ gaussian
Lọc High-boost
Laplace trên miền tần số

TOÁN TỬ TRÊN ĐIỂM ẢNH

Khuất Thị Ngọc Ánh

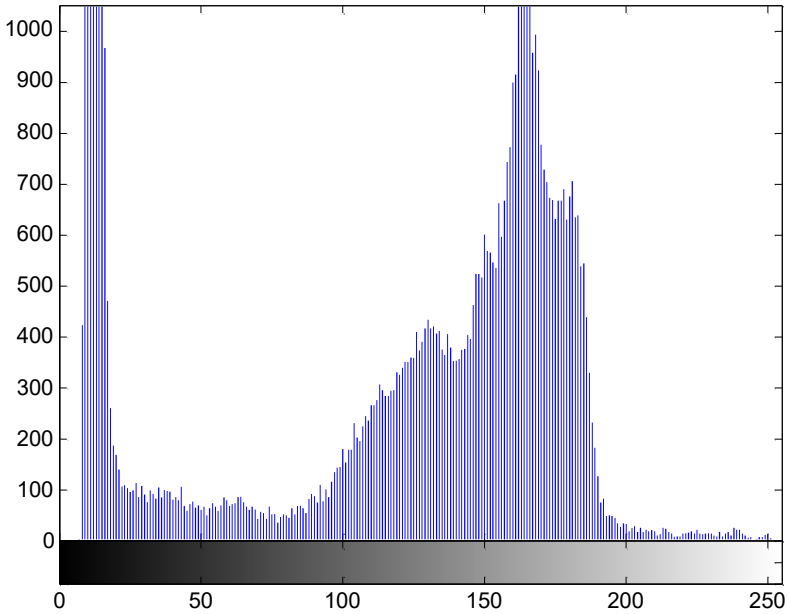
Phác họa nội dung

Cải thiện và nâng cấp ảnh

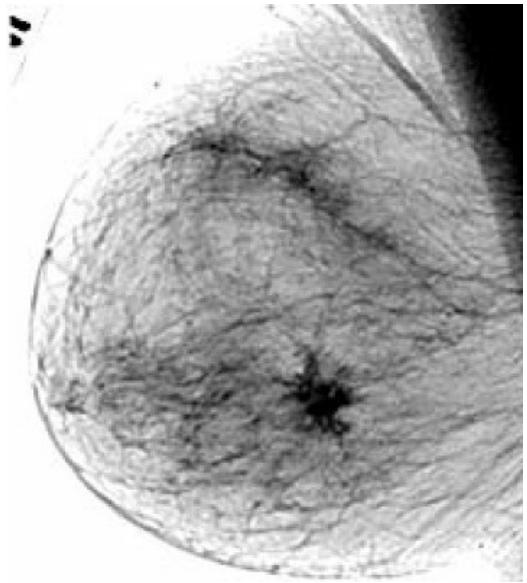
Toán tử trên điểm ảnh	Toán tử trên miền không gian	Toán tử trên miền tần số
Ảnh âm bản Biến đổi sử dụng hàm log Biến đổi sử dụng hàm mũ Dãn và co độ tương phản Cắt theo mức Trích chọn bit Cân bằng Histogram Số học/logic Trừ ảnh Trung bình ảnh	Lọc làm tròn ảnh: + Lọc trung bình + Lọc gaussian + Lọc thống kê thứ tự: Lọc trung vị Lọc cực đại Lọc cực tiểu Lọc tăng cường độ nét: Laplacian, Robert, Sobel Lọc High-boost	Lọc thông thấp + lý tưởng + butterworth + gaussian Lọc thông cao + lý tưởng + butterworth + gaussian 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 $n=\Sigma n_k$.



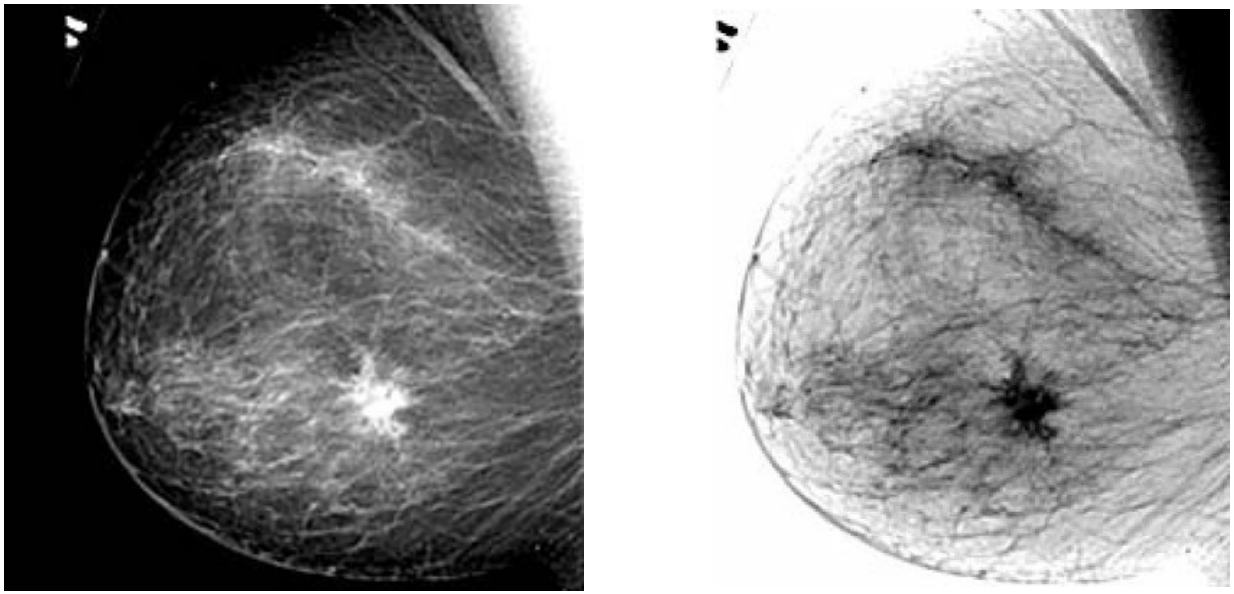
Phép biến đổi âm bản ảnh



$$g(m,n) = L - 1 - f(m,n)$$ trong đó:

- **g(m, n)** là giá trị mức xám sau phép biến đổi tại tọa độ (m, n).
- **L** là tổng số cấp xám khả dụng (thường là 256).
- **f(m, n)** là giá trị mức xám ban đầu tại tọa độ (m, n) trong khoảng [0, L-1].

Phép biến đổi âm bản ảnh

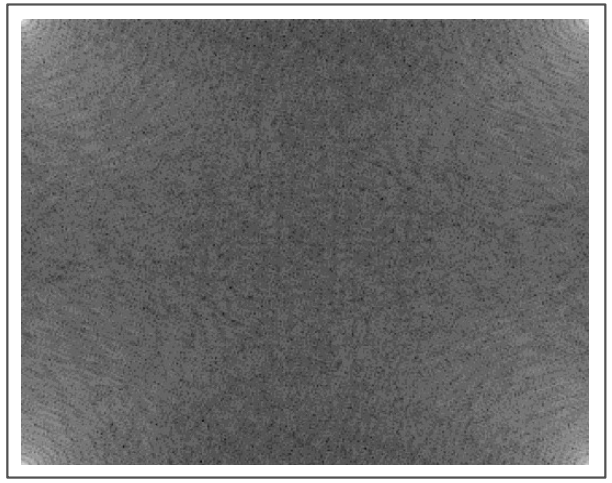


- Phép biến đổi âm bản là phép biến đổi đơn giản để tạo ra ảnh âm bản (ảnh đen trắng) từ một ảnh mức xám (ảnh dương bản) với các cấp xám nằm trong phạm vi [0, L-1]
- Phép biến đổi này hoạt động bằng cách lấy giá trị đối của mức xám f(m, n) và sau đó lùi ngược bằng cách trừ kết quả này từ giá trị tối đa của mức xám (L - 1). Kết quả là một phép biến đổi làm cho các vùng sáng trở thành đen và ngược lạ

Biến đổi sử dụng hàm log

(a)

(b)



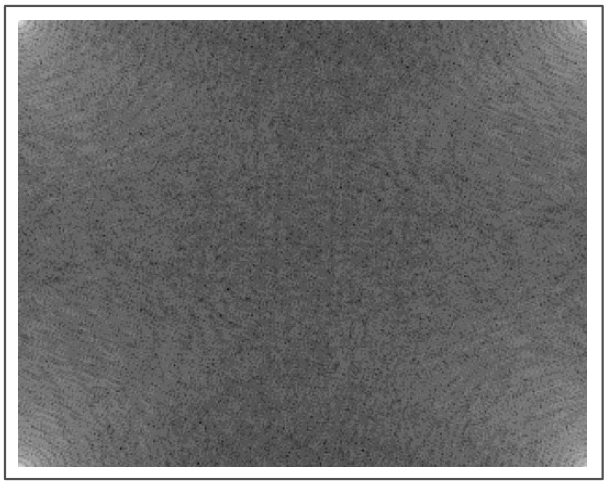
(c)

(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*)

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

Biến đổi sử dụng hàm log

(a)

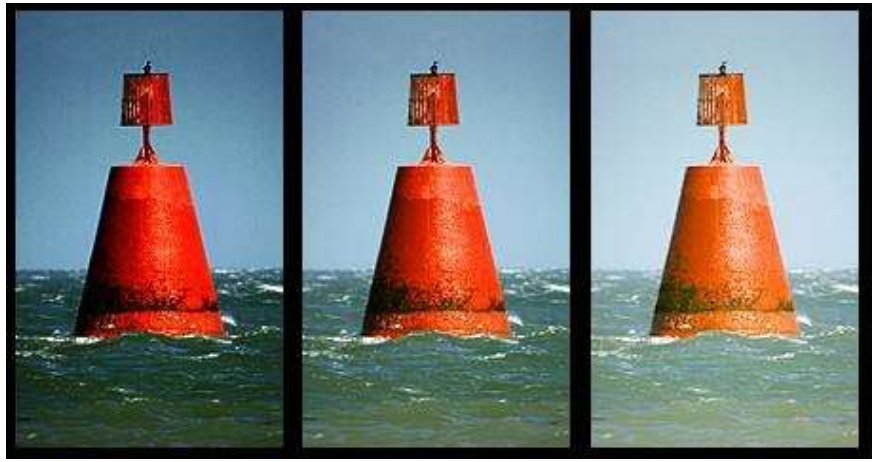


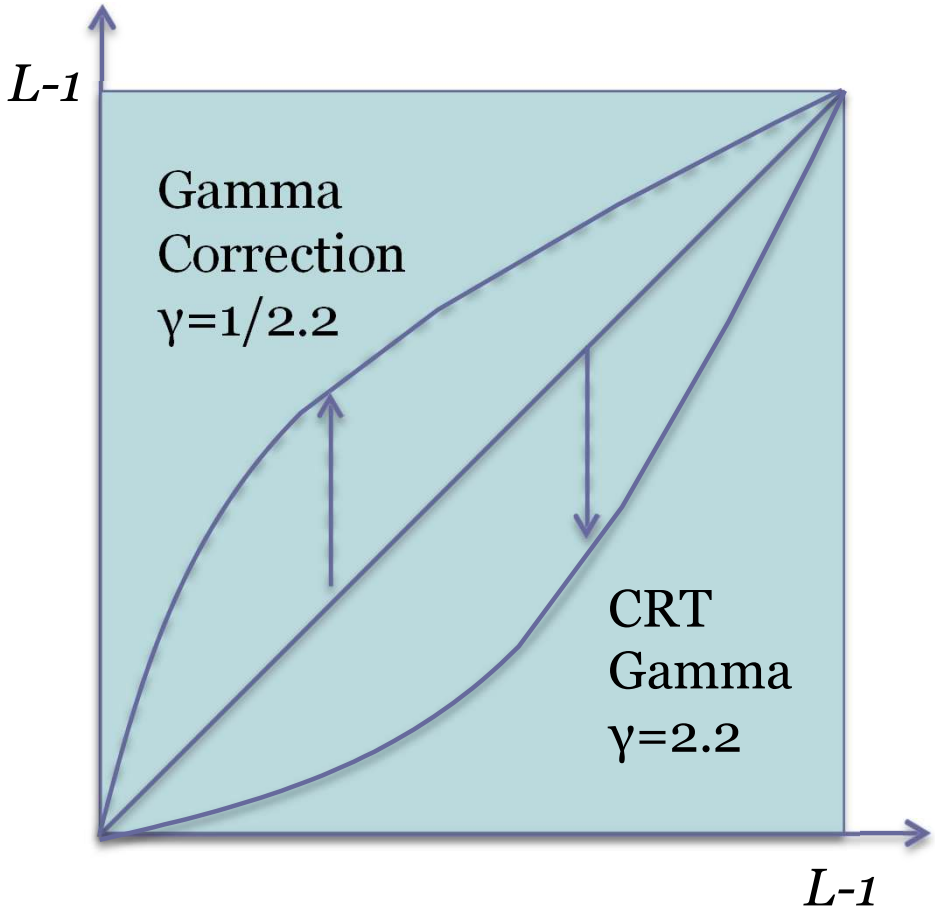
(b)

(a) Ảnh gốc *f(x,y)*
(b) Biến đổi sử dụng hàm log (*c=1*)

- Là một phương pháp biến đổi giúp tăng cường độ tương phản và tạo ra hiệu ứng đặc biệt cho ảnh. Phép biến đổi này thường được sử dụng để cải thiện chi tiết ở các vùng có độ sáng thấp trong ảnh, trong khi vẫn giữ nguyên các vùng có độ sáng cao
- c là hằng số để điều chỉnh mức độ của phép biến đổi (thường là một số dương).
- Thông thường để chọn c là thử nghiệm và điều chỉnh cho đến khi bạn đạt được kết quả mà bạn mong muốn.

Biến đổi sử dụng hàm mũ





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

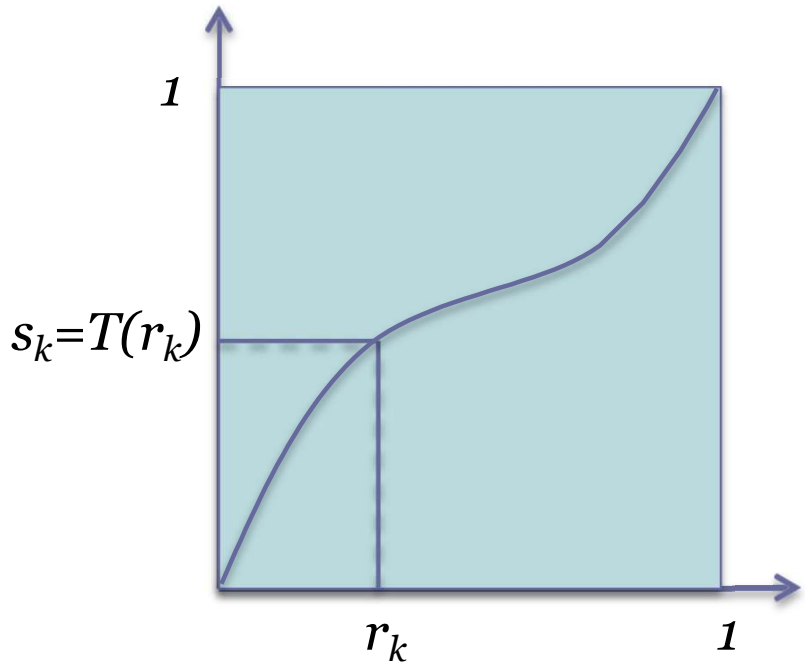
$$g(m,n) = cf(m,n)^{1/\gamma}$$

Cân bằng Histogram

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

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 $0 \le s \le 1$ và là hàm phân bố đều. Cân bằng histogram cho phép tạo ra histogram mới có phân bố đều.

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

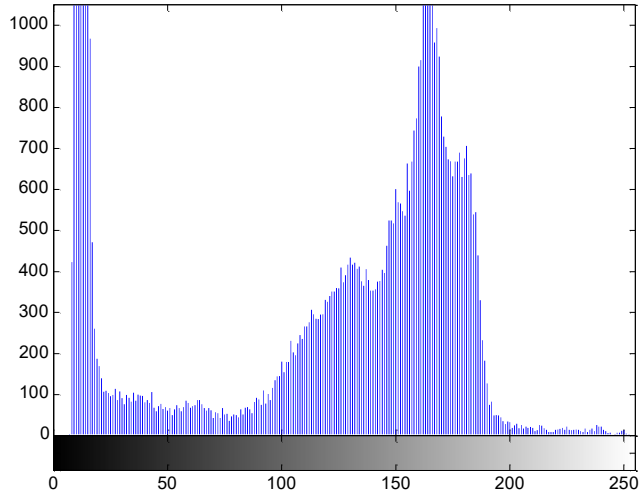
- Trên thực tế, giá trị mức xám mới được xác định như sau

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

Cân bằng Histogram

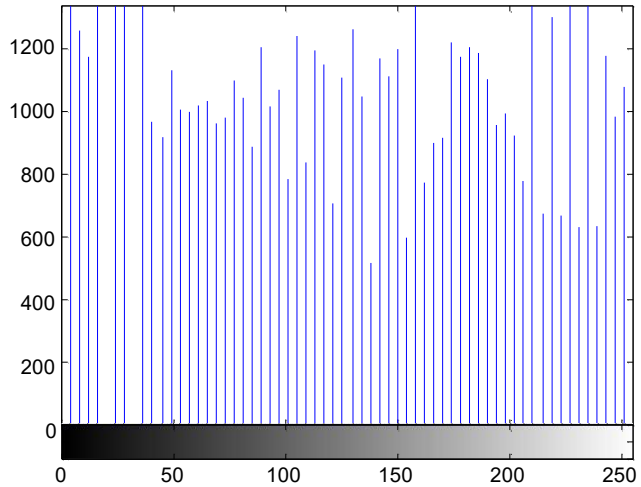
a

c



b

d



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

Cân bằng Histogram

- Ảnh I được gọi là cân bằng "lý tưởng" nếu với mọi mức xám g, g' ta có h(g) = h(g')

  Giả sử, ta có ảnh I ~ kích thước m × n

  new_level ~số mức xám của ảnh cân bằng

$$TB = \frac{m \times n}{new_level}$$ ~ số điểm ảnh trung bình của mỗi mức xám của ảnh cân bằng

$$t(g) = \sum_{i=0}^{g} h(i)$$ ~ số điểm ảnh có mức xám ≤ g

Xác định hàm f: g a f(g)

Sao cho: $$f(g) = \max\left\{0, round\left(\frac{t(g)}{TB}\right) - 1\right\}$$

Bài tập Histogram

1. Cân bằng ảnh sau với new_level= 4

$$I = \begin{pmatrix} 1 & 2 & 4 & 6 & 7 \\ 2 & 1 & 3 & 4 & 5 \\ 7 & 2 & 6 & 9 & 1 \\ 4 & 1 & 2 & 1 & 2 \end{pmatrix}$$

g	h(g)	t(g)	f(g)
1	5	5	0
2	5	10	1
3	1	11	1
4	3	14	2
5	1	15	2
6	2	17	2
7	2	19	3
9	1	20	3

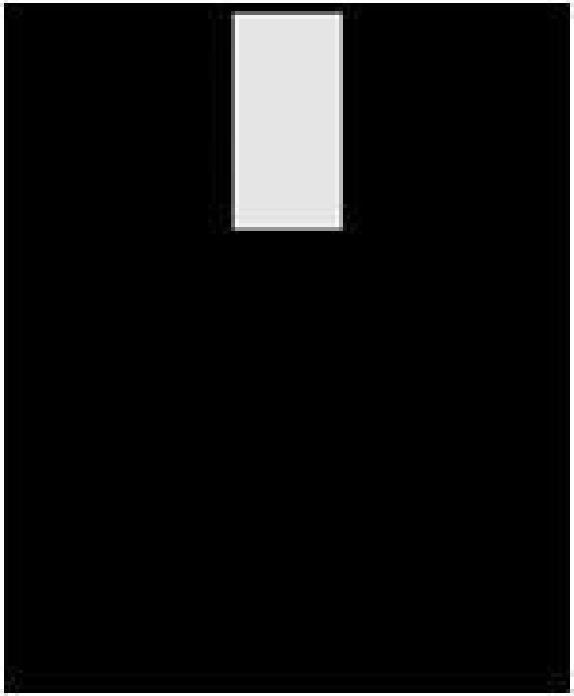
$$I_{kq} = \begin{pmatrix} 0 & 1 & 2 & 2 & 3 \\ 1 & 0 & 1 & 2 & 2 \\ 3 & 1 & 2 & 3 & 0 \\ 2 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Các phép biến đổi ảnh dựa trên các phép toán số học/logic

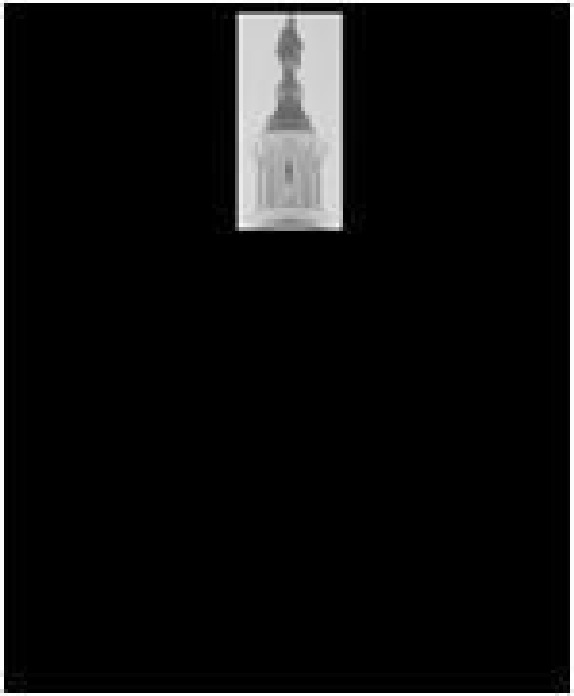
- Phép AND ảnh
- Phép OR ảnh
- Phép trừ ảnh
- Phép cộng ảnh

Phép AND ảnh

Ảnh gốc



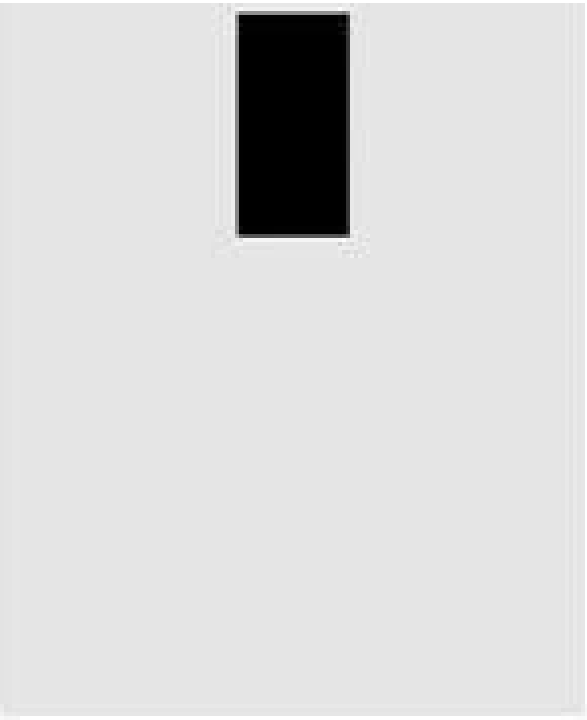
Mặt nạ AND



Phép AND ảnh

Phép OR ảnh

Ảnh gốc

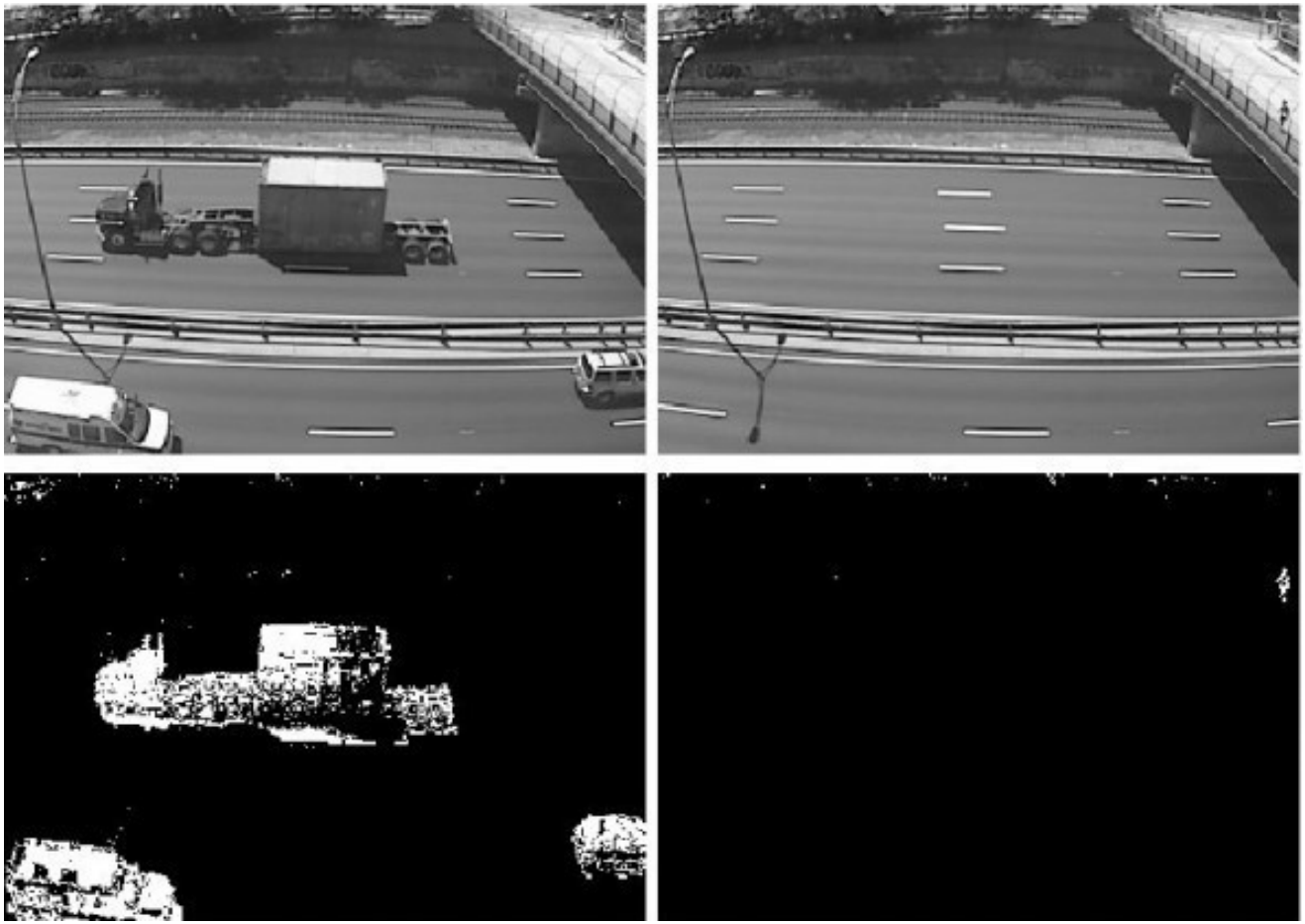


Mặt nạ OR

Phép OR ảnh

Trừ ảnh

a b



c d

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

(a) f(x,y)
(b) h(x,y)
(c) *g(x,y)=/f(x,y)-h(x,y)/*

Cộng ảnh

TOÁN TỬ TRÊN MIỀN KHÔNG GIAN

Khuất Thị Ngọc Ánh

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
Trừ 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: Laplacian, Robert, Sobel
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 đều, 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.

$$p(z) = \begin{cases} \dfrac{1}{b-a} & a \le z \le b \\ 0 & OtherwiseA \end{cases}$$

- Nhiễu lượng tử có đặc trưng sau

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

Nhiễu trắng

- Nhiễu trắng là nhiễu có phổ năng lượng không đổi. Nếu nhiễu có phổ năng lượng nhiều hơn ở một vài tấn 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 mẫu 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. Nhiễu 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 muốn tiêu** xuất hiệu khi ảnh bị bão hòa bởi nhiễu xung.

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

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 bình

Kỹ thuật lọc trung bình

Giả sử ta có ảnh I, điểm ảnh P, cửa sổ W(P) và ngưỡng θ. Khi đó kỹ thuật lọc trung bình phụ thuộc không gian bao gồm các bước cơ bản sau:

+ ***Bước 1:*** Tìm trung bình

$$\{I(q) \mid q \in W(P)\} \rightarrow AV(P)$$

+ ***Bước 2:*** Gán giá trị

$$I(P) = \begin{cases} I(P) & |I(P) - AV(P)| \le \theta \\ AV(P) & Nguoclai \end{cases}$$

Lọc trung bình

Ví dụ: Lọc ảnh sau với W (3x3) và ngưỡng = 2

$$I = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 4 & 16 & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix} \Rightarrow \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 2 \\ \hline 4 & 4 & 2 & 1 \\ \hline 4 & 2 & 1 & 1 \\ \hline 2 & 1 & 2 & 1 \\ \hline \end{array}$$

Ví dụ 2: Lọc ảnh sau với W (3x3) và ngưỡng = 80

10	41	9	3	123
232	186	79	21	134
1	43	56	1	12
88	98	34	4	2
134	76	89	104	5
88	54	12	231	1
74	12	42	111	0

Lọc trung bình theo K giá trị gần nhất

Giả sử ta có ảnh I, điểm ảnh P, cửa sổ W(P), ngưỡng θ và số k. Khi đó, lọc trung bình theo k giá trị gần nhất bao gồm các bước sau:

+ ***Bước 1***: Tìm K giá trị gần nhất

$$\{I(q) \mid q \in W(p)\} \rightarrow \{k \sim \text{giá trị gần } I(P) \text{ nhất}\}$$

+ ***Bước 2:*** Tính trung bình

$$\{k \sim \text{giá trị gần } I(P) \text{ nhất}\} \rightarrow AV_k(P)$$

+ ***Bước 3:*** Gán giá trị

$$I(P) = \begin{cases} I(P) & |I(P) - AV_k(P)| \le \theta \\ AV_k(P) & Nguoclai \end{cases}$$

Ví dụ:

$$I = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 4 & 16 & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix}$$

W(3 × 3); θ = 2; k = 3

Lọc thống kê thứ tự

- Lọc trung vị
- Lọc cực đại
- Lọc cực tiểu

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

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

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

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

0	0	0	8	89	8	8	7	8

0	0	0	7	8	8	8	8	89

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