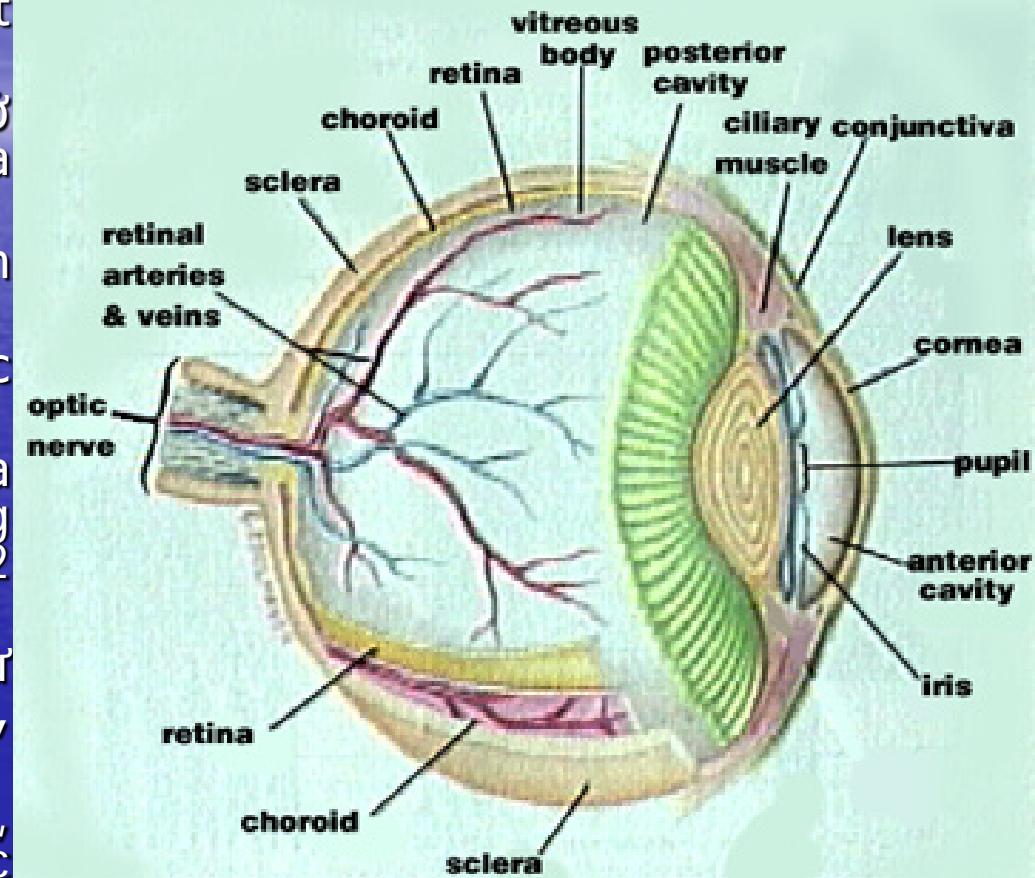


# Hệ thống thị giác của con người

- Việc hiểu hệ thống thị giác của con người giúp ích rất nhiều đến việc thiết kế, tính chính xác và hiệu quả của hệ thống phân tích, xử lý ảnh.
- Hệ thống thị giác của con người bao gồm các phần chính:
  - Mắt (cảm biến hình ảnh hoặc camera)
  - Thần kinh thị giác (đường truyền tín hiệu)
  - Và não (đơn vị xử lý thông tin hình ảnh hay máy tính).

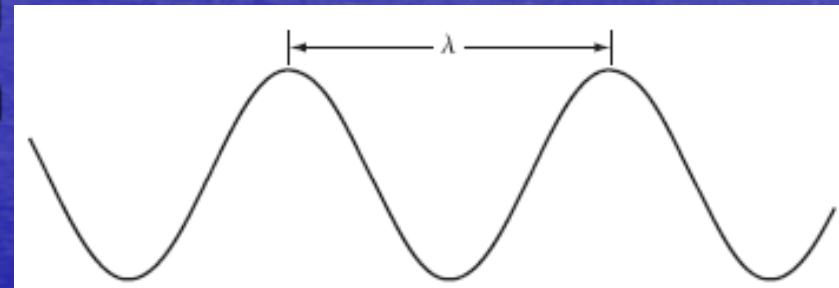
# Mắt người

- Đường kính nhăn cầu khoảng 20 mm.
- Giác mạc (Cornea) - màng trong suốt dai ở ngoài cùng, bao phủ bề mặt phía trước.
- Màng cứng (Sclera) - màng mờ dai ở ngoài cùng, bao phủ phần còn lại của quả cầu thị giác.
- Màng bao (Choroid) - chứa các mạch máu, cung cấp dinh dưỡng
- Mống mắt(Iris) - Ở phần phía trước của choroid, quyết định màu mắt.
- Đồng tử(Pupil) – Vùng mở giữa của Iris, điều khiển số lượng ánh sáng đến mắt (đường kính thay đổi từ 2 đến 8 mm).
- Thủy tinh thể (Lens) – được tạo từ các lớp đồng tâm gồm các tế bào sợi, chứa 60-70% nước.
- Võng mạc (Retina) – lớp trong cùng, màng tạo ảnh dựa trên lens, gồm các tế bào cảm quang (các tế bào này rất nhạy cảm ánh sáng).



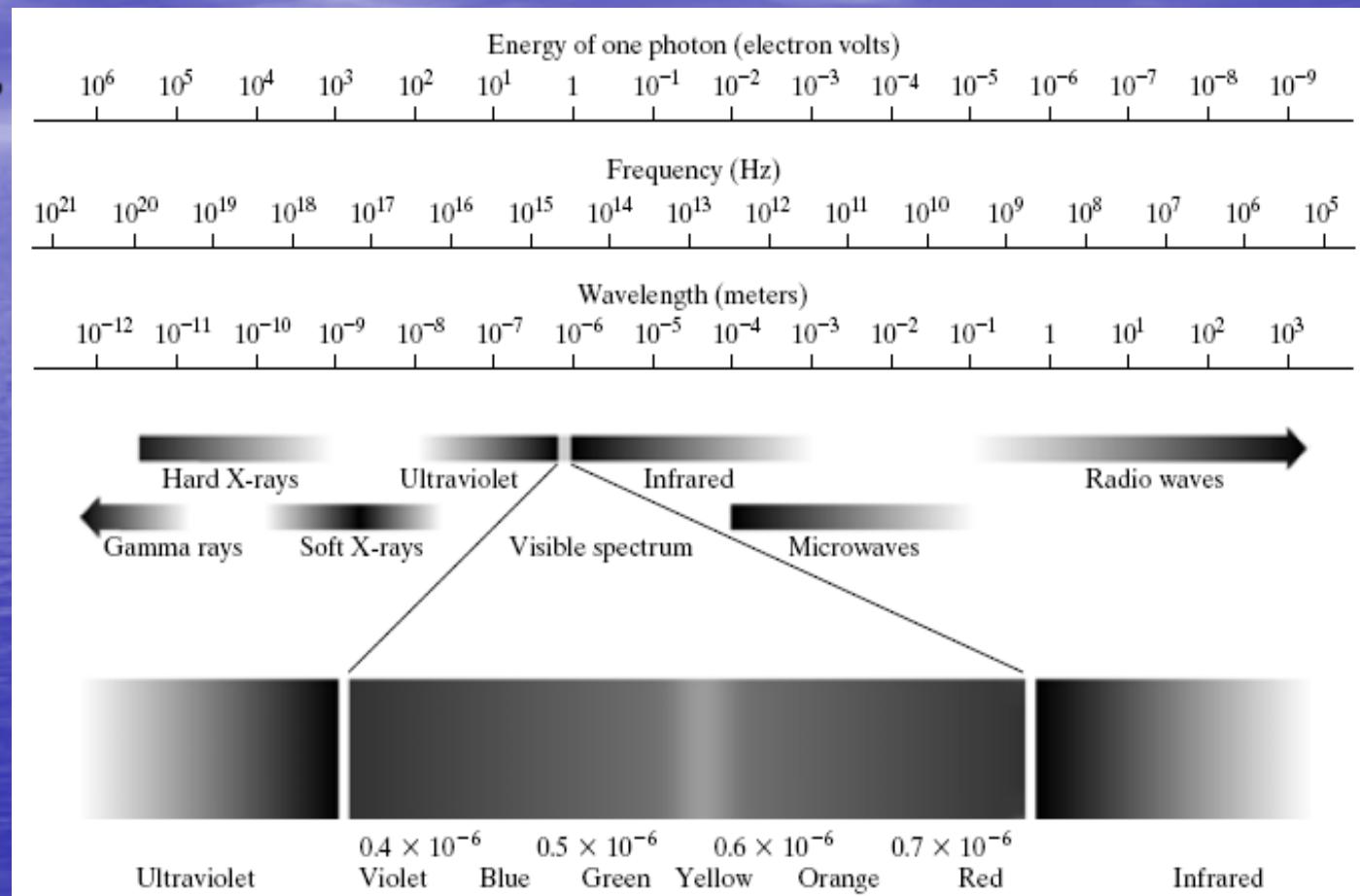
# Ánh sáng và phổ điện tử

- Sóng điện tử hoặc bức xạ có dạng hình sin, với bước sóng  $\lambda$  và tần số  $\nu$  thì  $c=\lambda\nu$ ,  $c$  là vận tốc ánh sáng ( $2.998 \times 10^8$ m/s).
- Năng lượng của phổ điện tử được tính là  $E=h\nu$ ,  $h$  là hằng số Planck.



# Ánh sáng và phổ điện tử

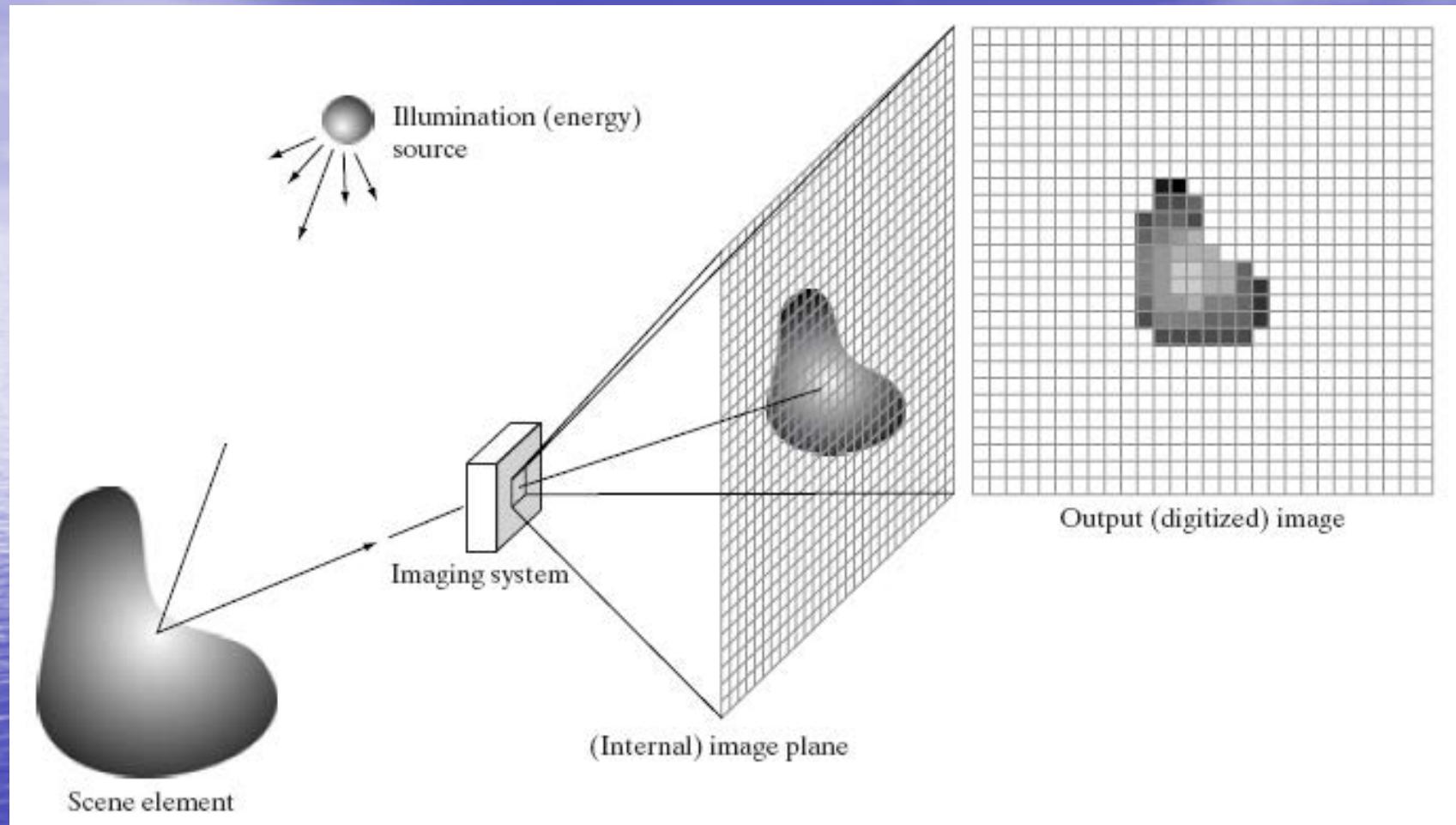
- Ánh sáng là một dạng của bức xạ điện tử mà mắt người có thể nhìn thấy. Phạm vi nhìn thấy của bước sóng là từ 0.43 đến 0.79  $\mu\text{m}$ .
- Các vùng khác nhau của vùng phổ thấy được sẽ tương ứng các màu khác nhau.



# Ánh sáng và phổ điện tử

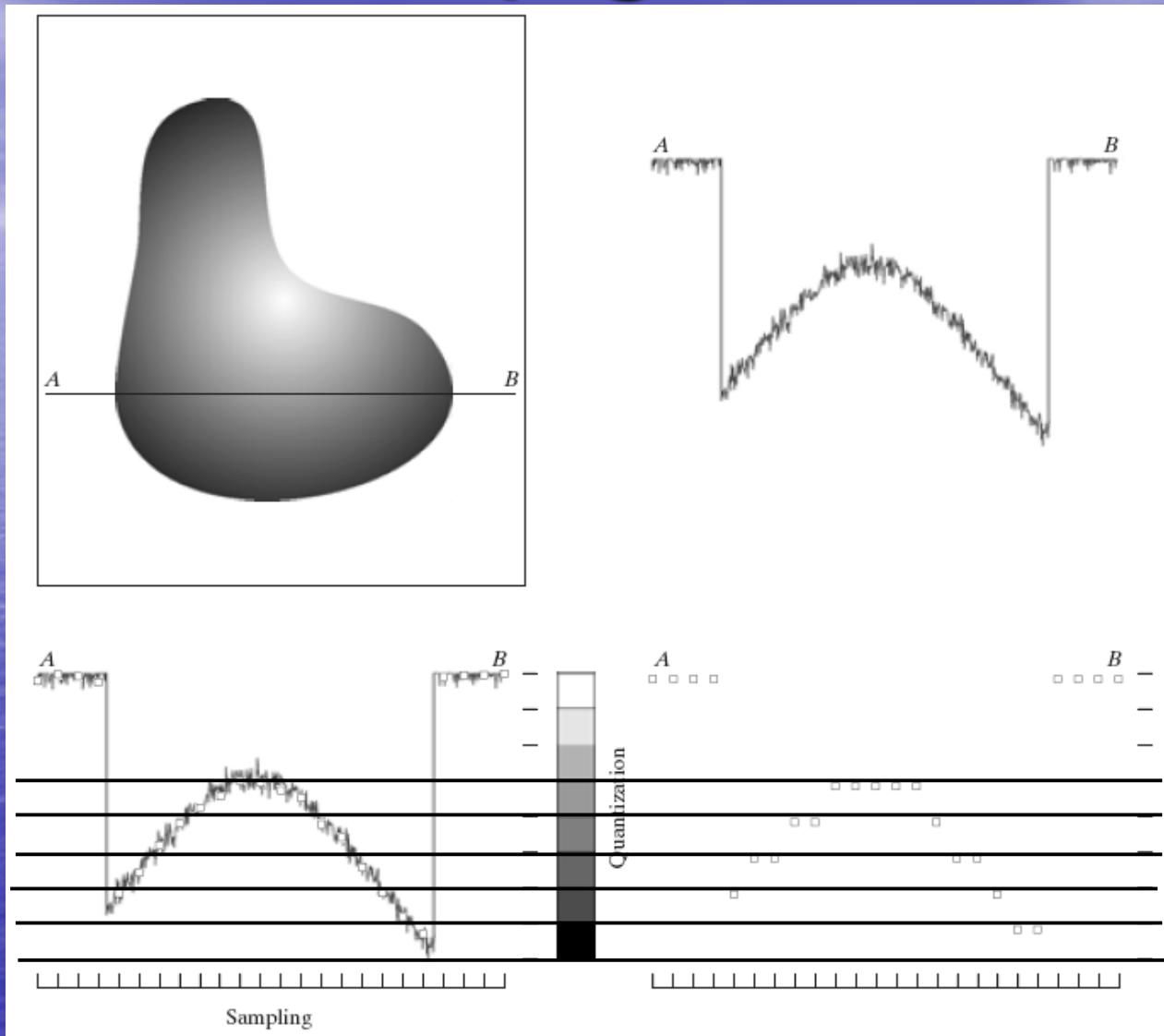
- Ánh sáng không màu hay đơn sắc.
- Cường độ của ánh sáng được biểu hiện bởi một giá trị xám hay mức xám. Màu trắng tương ứng mức xám cao nhất và màu đen là mức xám thấp nhất.
- Có ba thuộc tính thường được dùng để mô tả nguồn ánh sáng không màu:
  - **Độ chói (Radiance)** là tổng số lượng năng lượng từ nguồn, được tính bằng Watt (W).
  - **Độ sáng (Luminance)** là mức độ năng lượng ánh sáng mà nhận được, được tính bằng lumens (lm).
  - **Mức sáng (Brightness)** là mô tả chủ quan về mức độ ánh sáng (cảm nhận của con người).
- Bước sóng dùng sẽ phụ thuộc vào ứng dụng, tùy mức độ “thấy được”.
- Ngoài sóng điện tử, còn các loại năng lượng khác như: âm thanh (hình ảnh sóng siêu âm), chùm electron (ảnh dưới kính hiển vi điện tử), ...

# Thu ảnh số

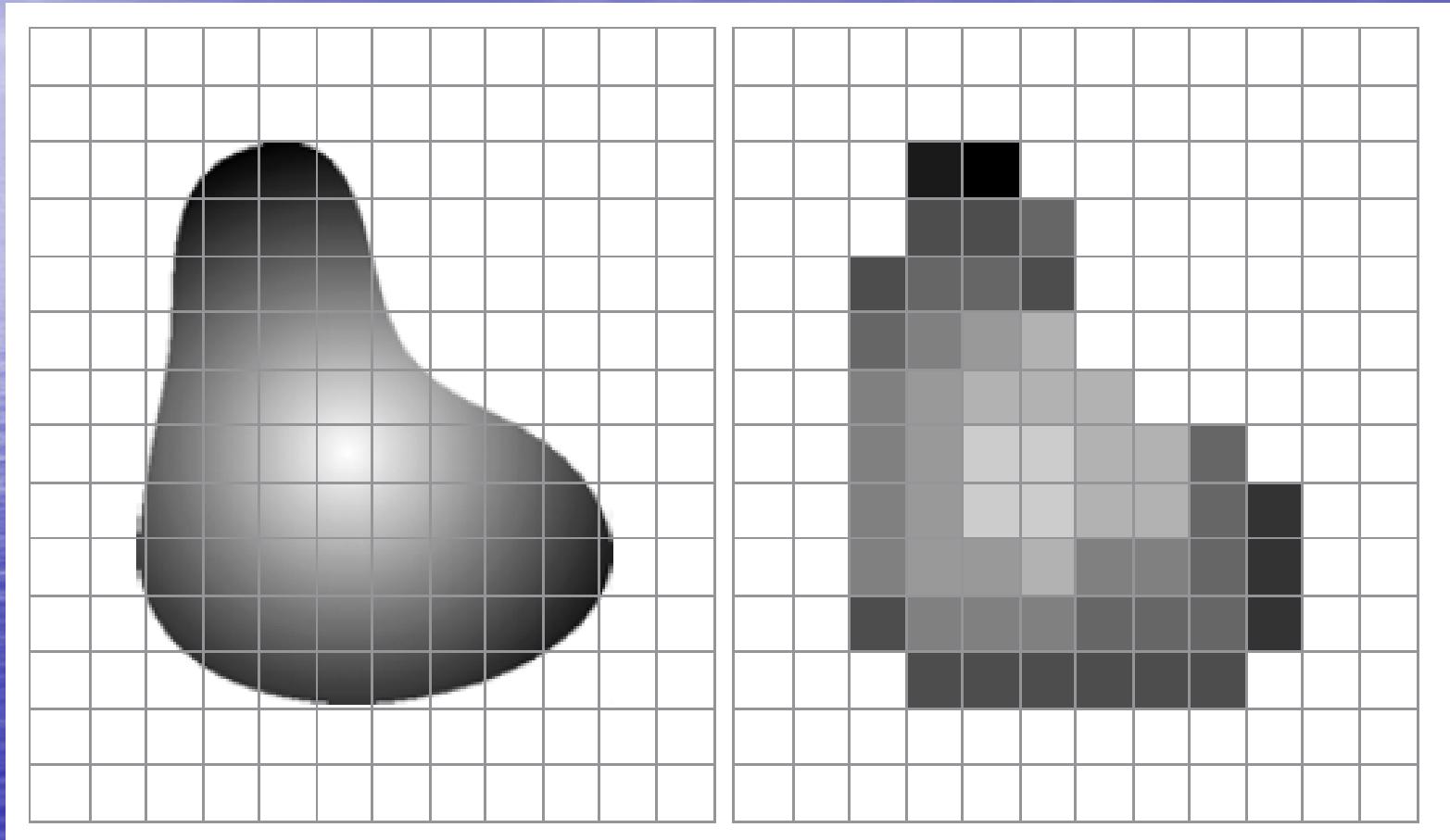


Phụ thuộc vào loại tia, một màn chắn thích hợp sẽ được dùng để tiếp nhận tia.

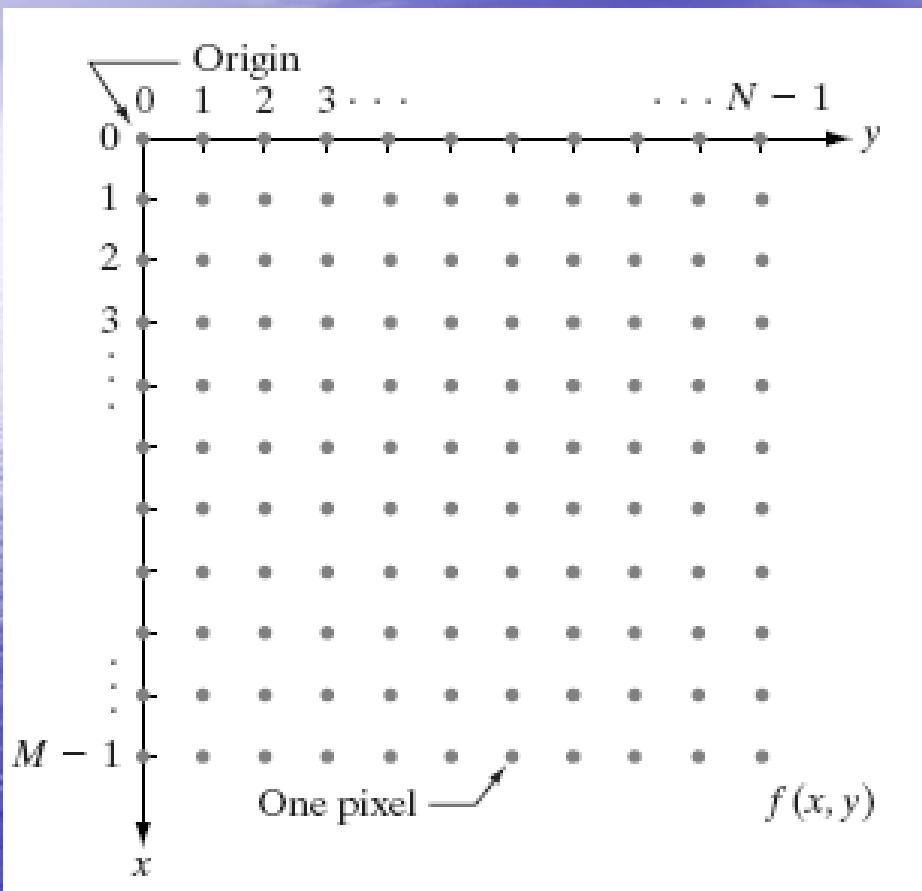
# Lấy mẫu và lượng hóa



# Kết quả lượng tử hóa



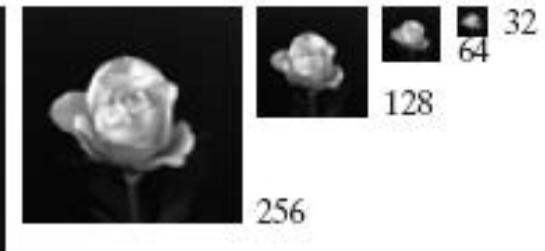
# Biểu diễn ảnh



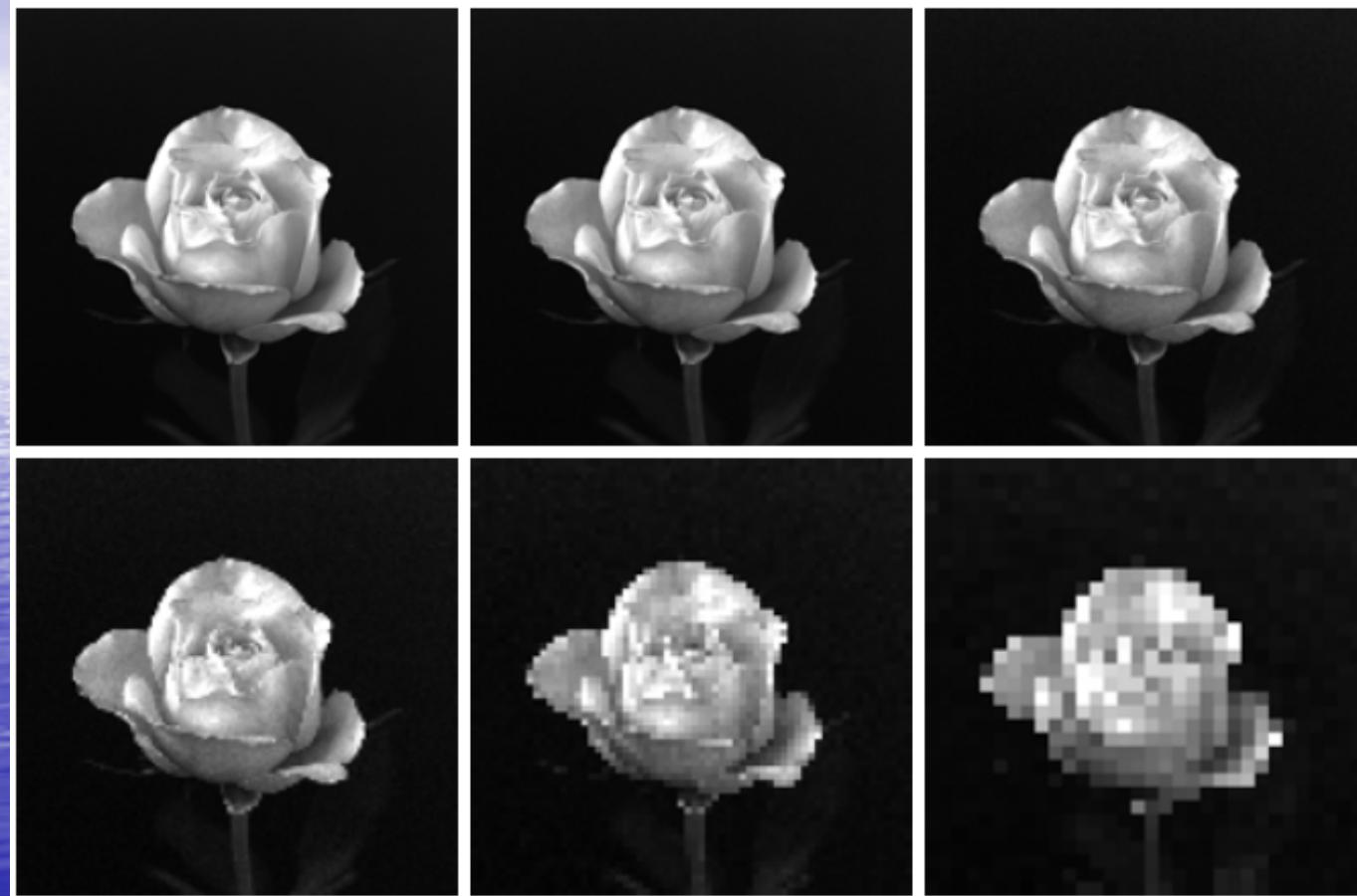
- Tọa độ  $(x,y)$  mô tả vị trí trong không gian và giá trị của hàm  $f(x,y)$  là cường độ ánh sáng tại điểm này.
- Điểm  $(x,y)$  là điểm ảnh trong ảnh và giá trị  $f(x,y)$  xem là giá trị xám (hoặc mức xám) của ảnh  $f$  tại  $(x,y)$ .

$$f(x, y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N-1) \\ f(1,0) & f(1,1) & \cdots & f(1, N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1, N-1) \end{bmatrix}$$

# Ảnh ở các độ phân giải khác nhau



# Quan hệ giữa kích thước và độ phân giải

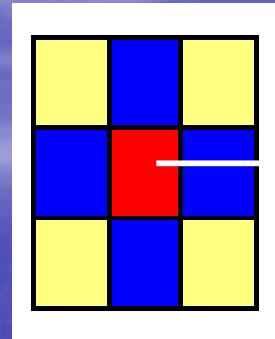


- Kích thước ảnh 1024x1024, 8 bit.
- Độ phân giải khác nhau:
  - ✓ 1024x1024
  - ✓ 256x256
  - ✓ 128x128
  - ✓ 64x64
  - ✓ 32x32

# Quan hệ giữa cơ bản các điểm ảnh

## 1. Láng giềng: xét điểm $p(x,y)$

- 4 lảng giềng của  $p$ ,  $N_4(p)$   
 $(x+1,y), (x-1,y), (x,y+1), (x,y-1)$
- 8 lảng giềng của  $p$ ,  $N_8(p)$   
gọi  $N_D(p) = (x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)$   
 $N_8(p) = N_4(p) + N_D(p)$



## 2. Kề

Định nghĩa  $V$  là tập các giá trị mức xám cần xét (ảnh nhị phân  
 $V=\{1\}$ ), mức xám  $=[0..255]$

- a. 4-kề: 2 điểm ảnh  $p$  và  $q$  là 4-kề theo tập  $V$  nếu  $q \in N_4(p)$
- b. 8-kề: 2 điểm ảnh  $p$  và  $q$  là 8-kề theo tập  $V$  nếu  $q \in N_8(p)$
- c. m-kề: 2 điểm ảnh  $p$  và  $q$  là m-kề theo tập  $V$  nếu:
  - i.  $q \in N_4(p)$  hoặc
  - ii.  $q \in N_D(p)$  và  $N_4(p) \cap N_4(q) = \{\phi\}$

### 3. Liên thông, Vùng và Biên

- Đường đi giữa hai điểm  $p$  và  $q(s,t)$  là dãy các pixel  $(x_0, y_0), \dots, (x_n, y_n)$ , với  $(x_0, y_0) = (x, y)$  và  $(x_n, y_n) = (s, t)$  và các pixel còn lại phải kế nhau, chiều dài đường đi là  $n$ .
- Nếu  $(x_0, y_0) = (x_n, y_n)$  thì đây là đường đóng.
- Gọi  $S$  là một tập con các pixel của ảnh, 2 pixel  $p$  và  $q$  được gọi là liên thông (nối nhau) trong  $S$  nếu tồn tại một đường đi trong  $S$  giữa  $p$  và  $q$ .
- $\forall p \in S$ , tập các pixel mà liên kết nhau trong  $S$  được gọi là thành phần liên thông của  $S$ .

– Nếu  $S$  chỉ có một thành phần liên thông thì  $S$  được gọi là tập liên thông.

–  $R$  là một tập các pixel của ảnh,  $T$  được gọi là một vùng nếu  $R$  là tập liên thông.

– Biên của vùng  $R$  là tập các pixel trong vùng mà các pixel này có một hoặc nhiều láng giềng không thuộc  $R$ .



# Độ đo khoảng cách

- Xét 3 pixel có tọa độ  $(x,y)$ ,  $(s,t)$ , và  $(v,w)$ , hàm khoảng cách phải thỏa tính chất:

- $D(p,q) \geq 0$  ( $D(p,q)=0$  khi và chỉ khi  $p=q$ )
- $D(p,q) = D(q,p)$ , và
- $D(p,z) \leq D(p,q) + D(q,z)$

1. Khoảng cách Euclid

2. Khoảng cách city-block

3. Khoảng cách chessboard

$$D_e(p,q) = \sqrt{(x-s)^2 + (y-t)^2}$$

$$D_4(p,q) = |x-s| + |y-t|$$

$$D_8(p,q) = \max(|x-s|, |y-t|)$$

```

function [rt, f, g] = twodsin(A,u0, v0, M, N)
% TWODSIN Compares for loops vs. vectorization.
% The comparison is based on implementing the function
%  $f(x,y) = A \sin(u_0 x + v_0 y)$  for  $x=0,1,2,\dots,M-1$  and
%  $y=0,1,2,\dots,N-1$ . The inputs to the function are
% M and N and the constants in the function.
% First implement using for loops
tic % start timing
for r=1:M
    u0x=u0*(r-1);
    for c=1:N
        v0y=v0*(c-1); f(r,c)=A*sin(u0x+v0y);
    end
end
t1=toc; % End timing
% Now implement using vectorization
tic % start timing
r=0:M-1;c=0:N-1;
[C,R]=meshgrid(c,r);
% special MATLAB function for fast 2F function evaluations
% creates all the (x,y) pairs for function evaluation
g=A*sin(u0*R+v0*C);
t2=toc; %End timing
%compute the ratio of the two times
rt=t1/(t2+eps); %use eps in case t2 is close to zero.

```

# Image processing toolbox

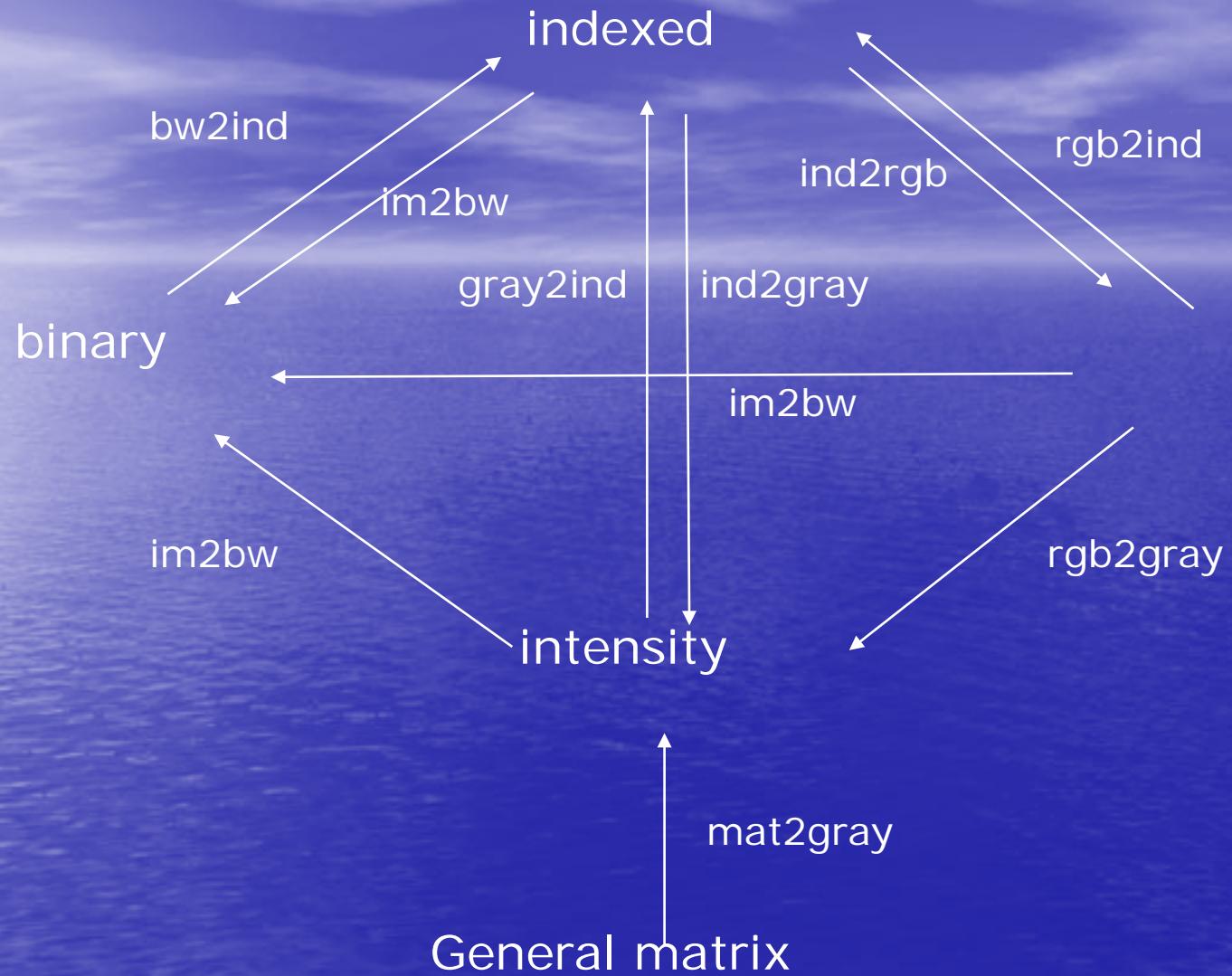
```

>> [rt,f,g]=twodsin(1, 1/(2*pi), 1/(4*pi), 512, 512);
>>g=mat2gray(g);
>> imshow(g)

```

Hàm thường dùng:

- `imshow`, xem ảnh
- `imwrite`, lưu ảnh
- `imfinfo`, thông tin file ảnh
- `Imread`, đọc file ảnh



# Tăng cường ảnh

- Xử lý một ảnh nghĩa là làm cho ảnh đó tốt hơn ban đầu trong một ứng dụng cụ thể nào đó.
- Tăng cường (Enhancement) sẽ phụ thuộc rất nhiều vào vấn đề/ảnh cụ thể.
- Việc tăng cường được xem xét trong hai loại:
  - Miền không gian: thao tác trên điểm ảnh và lân cận  
$$g(x,y) = T[f(x,y)]$$
  - Miền tần số: thao tác trên tần số

## Kỹ thuật tăng cường ảnh

- Âm bản
  - Tương phản
  - Nén theo vùng linh động
  - Cắt nhỏ mức xám
  - Bù ánh
  - Trung bình ảnh
  - Histogram
- Trên điểm**
- Làm tròn
  - Lọc Median
  - Nỗi bật
  - Đạo hàm
  - Histogram
- Lọc BT thấp
  - Lọc BT cao
  - Lọc theo bước
  - Lọc theo đồng hình dạng
  - Histogram
- Xử lý màu theo màu sắc

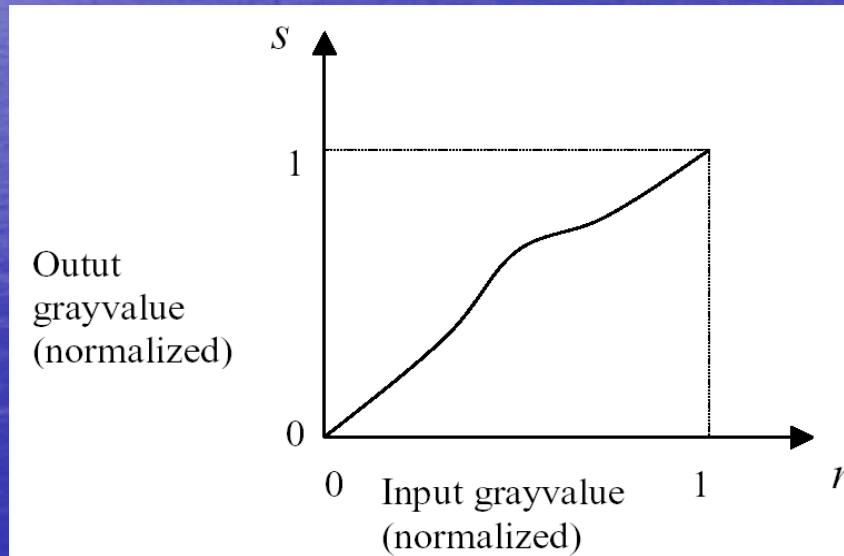
**Dùng mặt nạ**

**Biến đổi**

**Màu sắc**

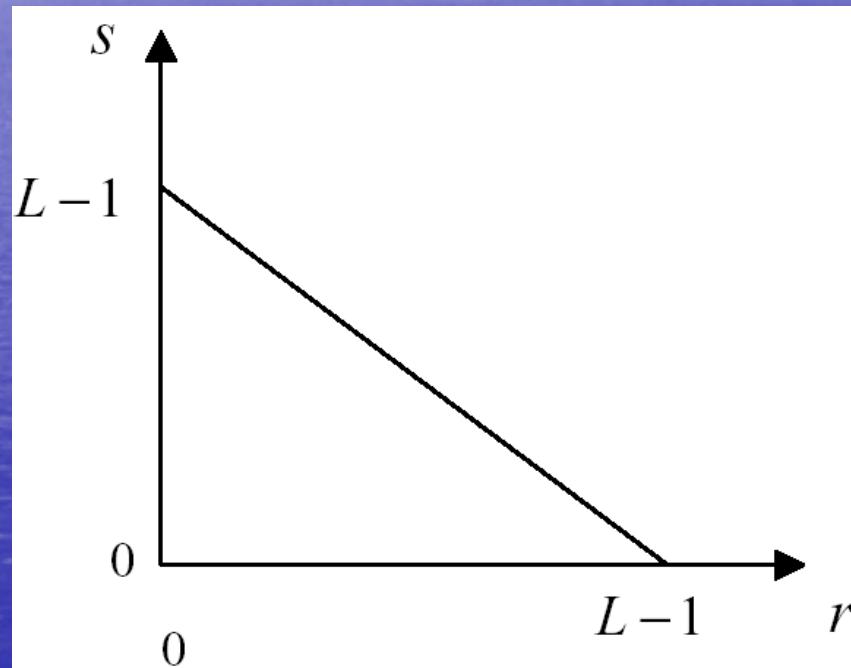
# Thao tác trên điểm

- Giá trị  $g(x,y)$  sẽ phụ thuộc vào  $f(x,y)$  tại  $(x,y)$  (và không phụ thuộc vào các pixel lân giềng).
- Chúng ta thường viết  $s = T(r)$  với  $s$  là giá trị sau khi thao tác và  $r$  là giá trị pixel cần biến đổi.



- $T$  là hàm tăng ánh xạ từ  $[0,1]$  vào  $[0,1]$ .

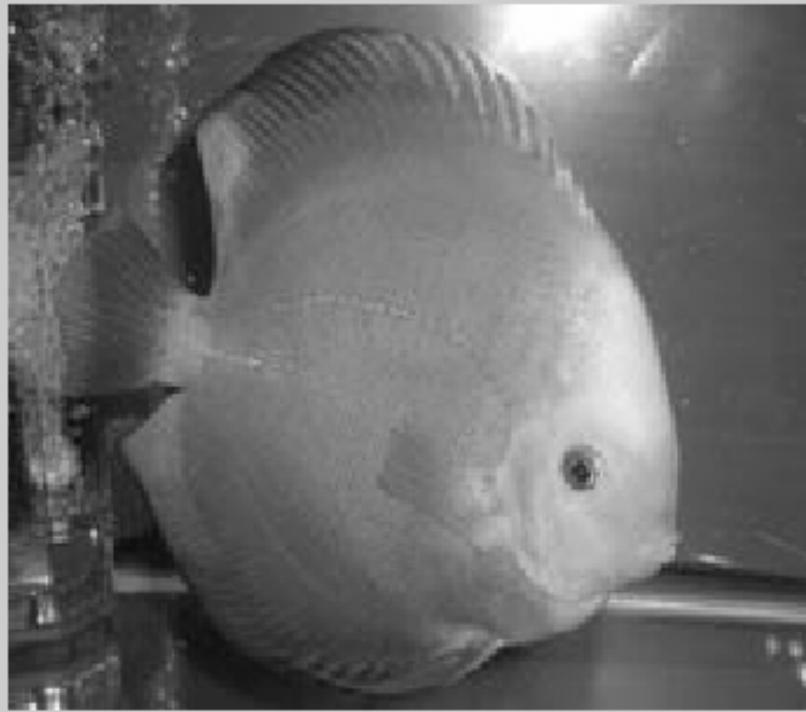
# Ảnh âm bản



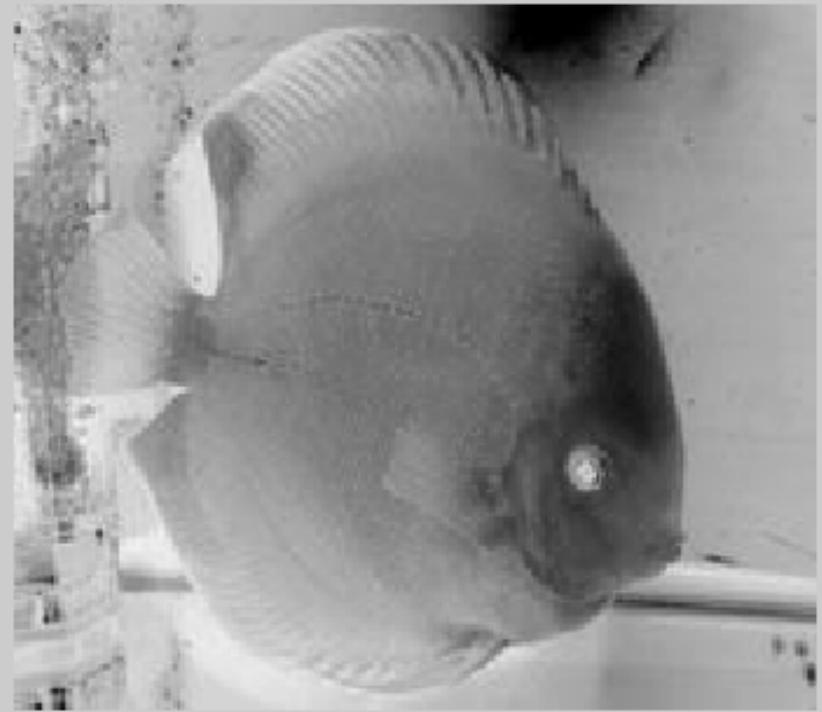
$$T(r) = s = L - 1 - r, \quad L: \text{Giá trị xám lớn nhất}$$

# Âm bản

Origin

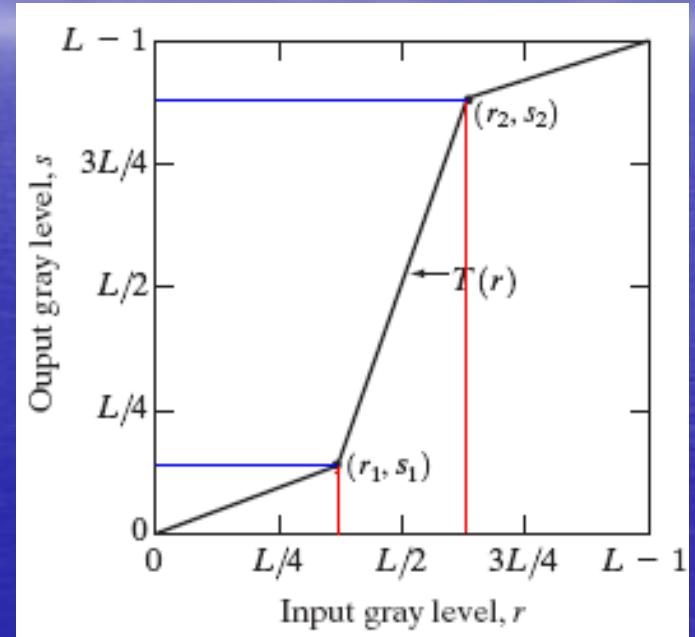


Negative



# Độ tương phản

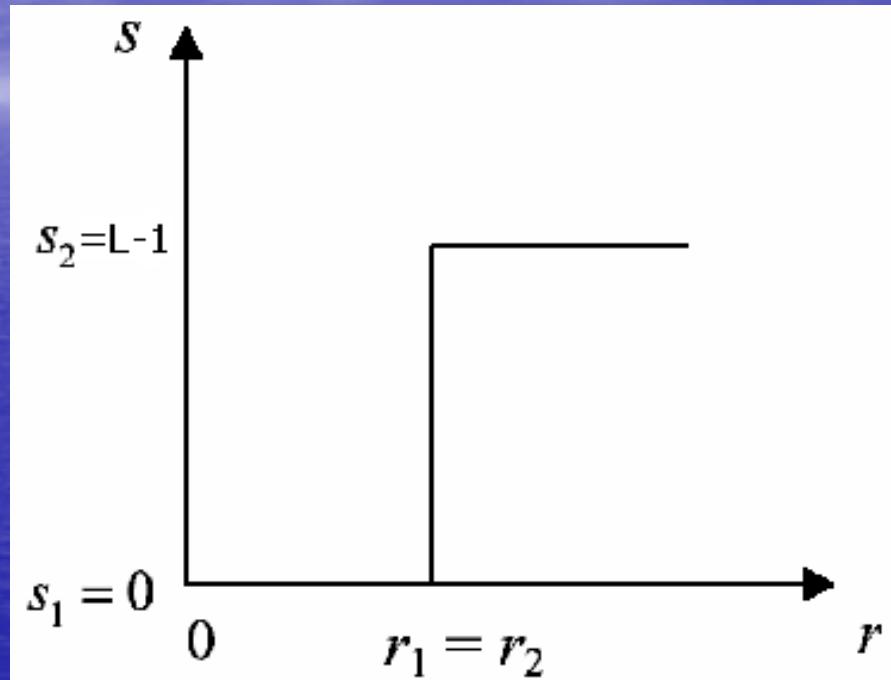
- Tăng vùng linh hoạt của giá trị xám trong ảnh ban đầu.
- Giả sử vùng cần mở rộng là đoạn  $[r_1, r_2]$ . Chúng ta sẽ mở rộng thành  $[s_1, s_2]$ .
- Với  $(r_1 - r_2) < (s_1 - s_2)$ .



# Độ tương phản

## • Trường hợp đặc biệt 1:

- **Nếu**  $r_1=s_1$  và  $r_2=s_2$ , không thay đổi
- **Nếu**  $r_1=r_2$ ,  $s_1=0$  và  $s_2=L-1$ , biến đổi này trở thành hàm ngưỡng hay nhị phân hóa.
- Hữu dụng khi chúng ta chỉ quan tâm đến hình dạng của đối tượng và ở mức xám.



# Độ tương phản

Low-contrast



Contrast stretching

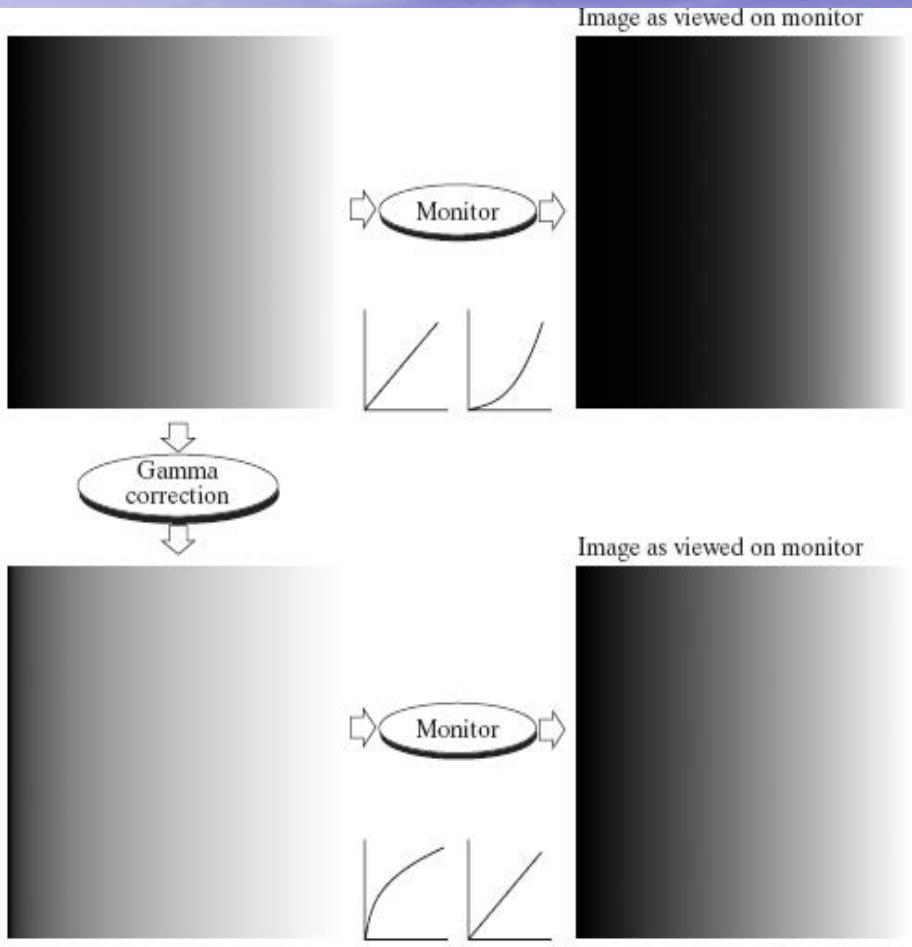


Binary

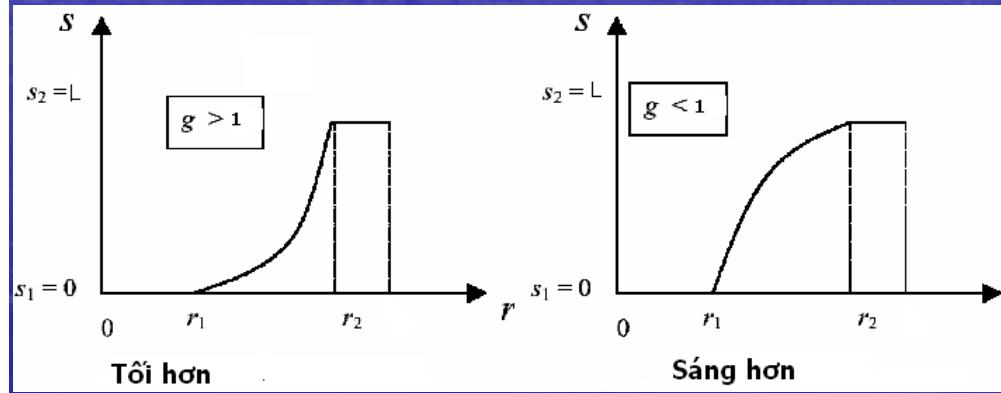


# Độ tương phản

- Trường hợp đặc biệt 2:
  - Gamma correction:  $S_1 = 0, S_2 = 1$  và



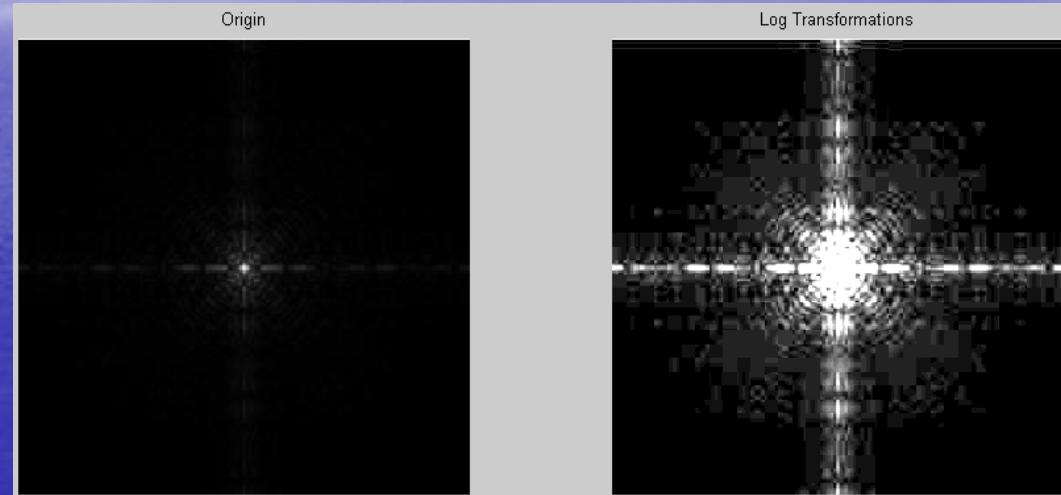
$$T(r) = \begin{cases} 0, & r < r_1 \\ \left(\frac{r - r_1}{r_2 - r_1}\right)^g, & r_1 \leq r \leq r_2 \\ L, & r > r_2 \end{cases}$$



# Nén vùng linh hoạt

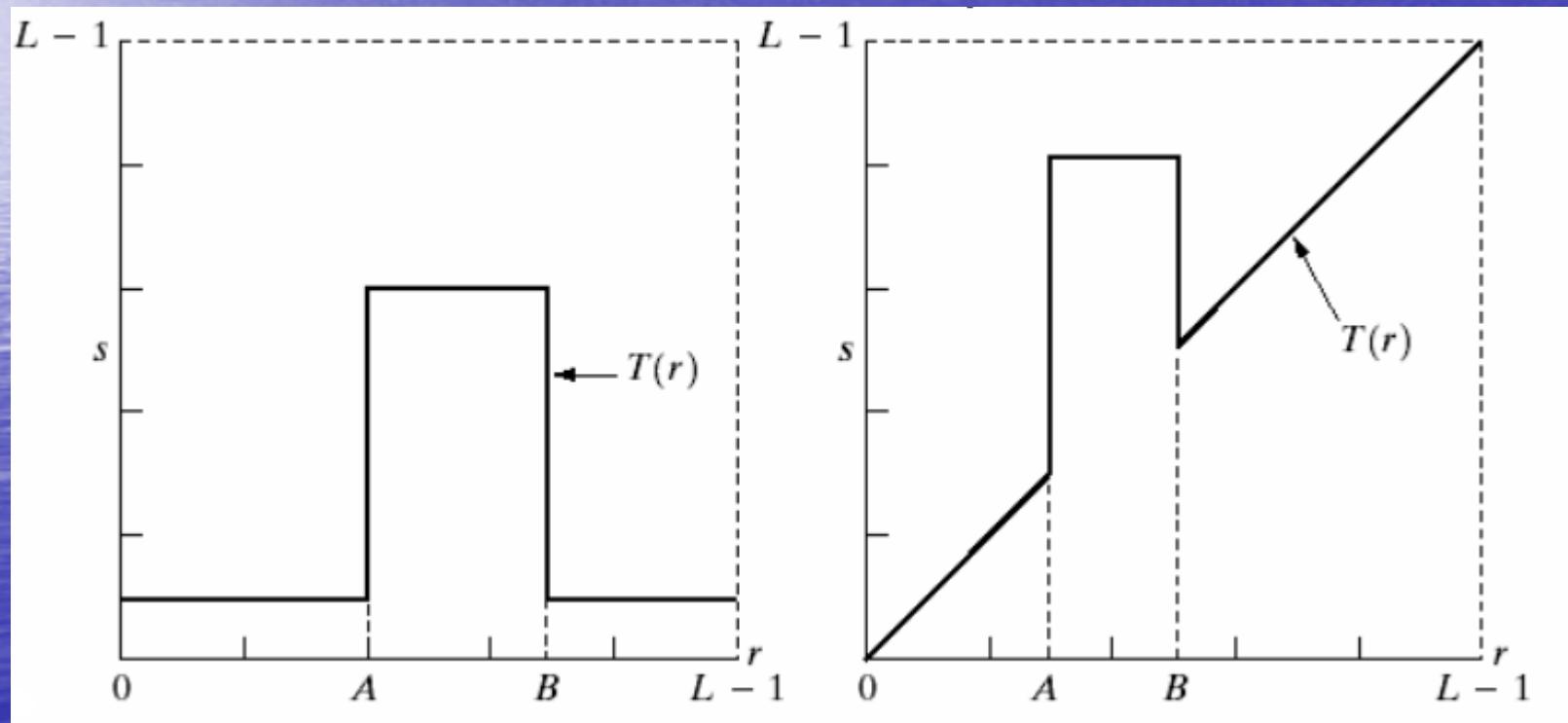
- Khi vùng linh hoạt lớn, chúng ta cần "nén" vùng này lại, ví dụ biến đổi Fourier về cường độ.
- Thông thường dùng tỷ lệ log.

$$s = T(r) = c \log(1 + |r|)$$



# Nén vùng linh hoạt

- Chia nhỏ mức xám: Làm nổi bật một vùng đặc biệt của mức xám.



# Nén vùng linh hoạt

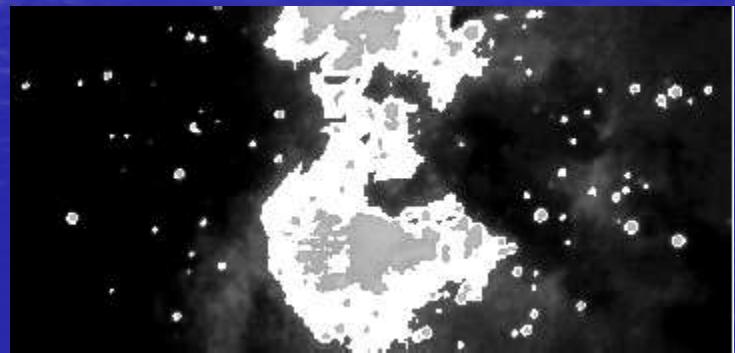
- **Ví dụ:**



Ảnh ban đầu



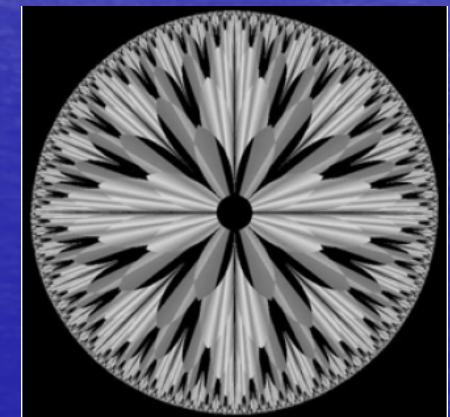
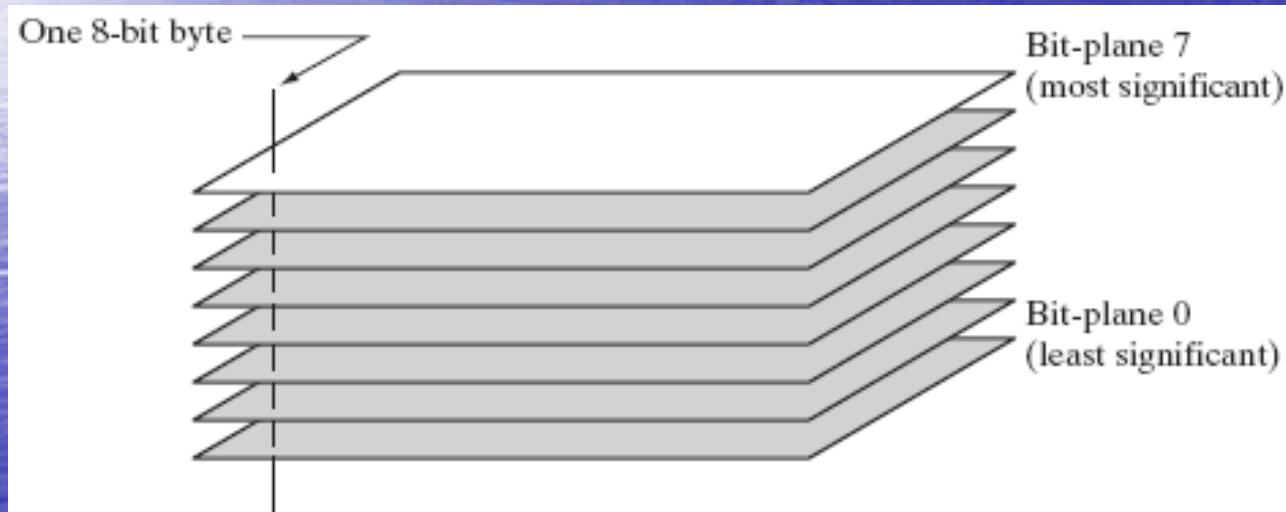
Ảnh được làm nổi bật (không nền)



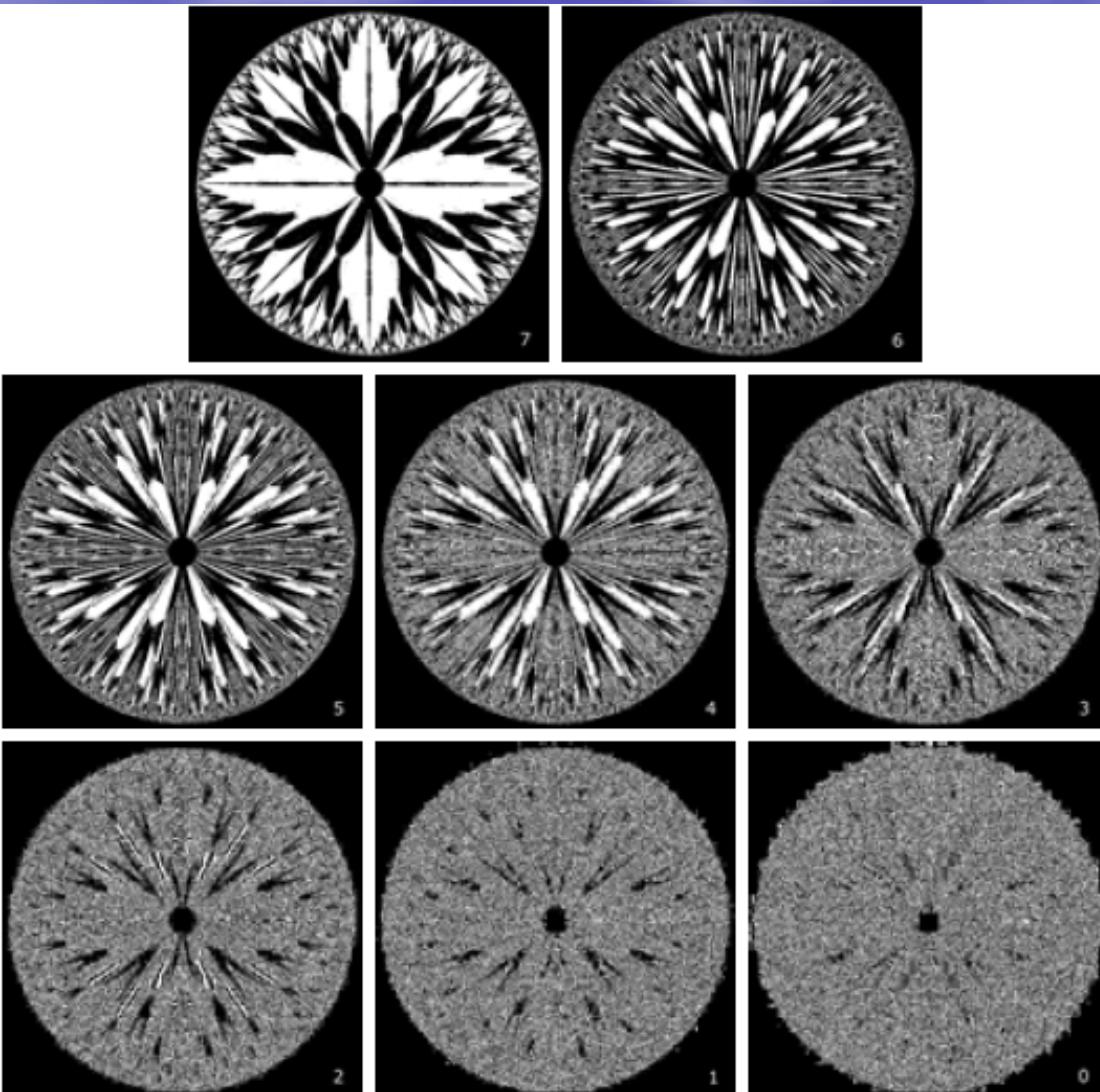
Ảnh được làm nổi bật (có nền)

# Nén vùng linh hoạt

- **Chia nhỏ mặt phẳng Bit:** Hiển thị các bit khác nhau các ảnh nhị phân riêng biệt.



Tám mặt  
phẳng bit



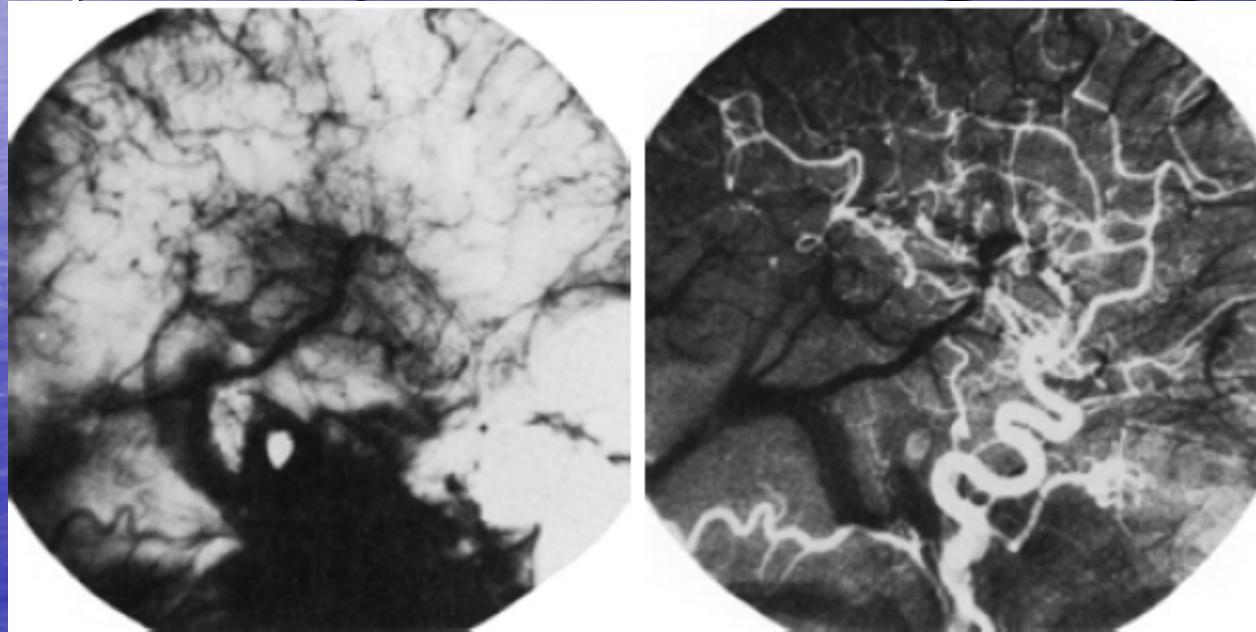
# Trừ ảnh

- Trong trường hợp độ sai khác của hai bức ảnh “tương đồng” theo công thức:

$$g(m, n) = f_1(m, n) - f_2(m, n)$$

- Dùng để phân đoạn ảnh và làm tăng cường ảnh.

- $f_1(m, n)$ :  
Ảnh trước  
khi đưa hóa  
chất vào
- $f_2(m, n)$ :  
Ảnh sau khi  
đưa hóa  
chất vào



$g(m, n)$ : Ảnh  
sau khi đưa  
hóa chất  
vào , sau  
khi trừ ảnh

X-quang não người bệnh

# Dùng trung bình ảnh để giảm nhiễu

- Nhiễu là hiện tượng ngẫu nhiên (không dự đoán trước được), làm hư ảnh.
- Nhiễu xảy ra trong quá trình:
  - Thu ảnh
  - Truyền ảnh
  - Lưu trữ ảnh
- Nhiễu có thể mô hình hóa như tiến trình:

$$g(m,n) = f(m,n) + \eta(m,n)$$

← Anh nhiễu      ← Anh không nhiễu      ← Nhiễu

# Dùng trung bình ảnh để giảm nhiễu

- Nhiễu  $h(m,n)$  tại pixel  $(m,n)$  được mô hình như một biến ngẫu nhiên.
- Thông thường,  $h(m,n)$  xem như là **zero-mean** và các giá trị nhiễu tại các pixel khác nhau thì **không có tương quan**.
- Giả sử chúng ta có  $M$  quan sát  $\{g_i(m, n)\}$ ,  $i=1,2,\dots,M$ , chúng ta có thể giảm nhiễu (một phần) bằng "tính trung bình của ảnh".

$$\bar{g}(m, n) = \frac{1}{M} \sum_{i=1}^M g_i(m, n)$$

# Dùng trung bình ảnh để giảm nhiễu

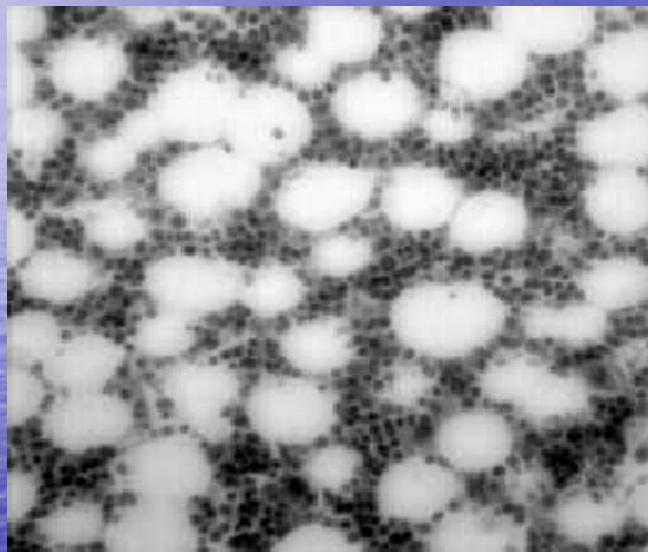
- Trong trường hợp này:

$$E[\bar{g}(m, n)] = f(m, n)$$

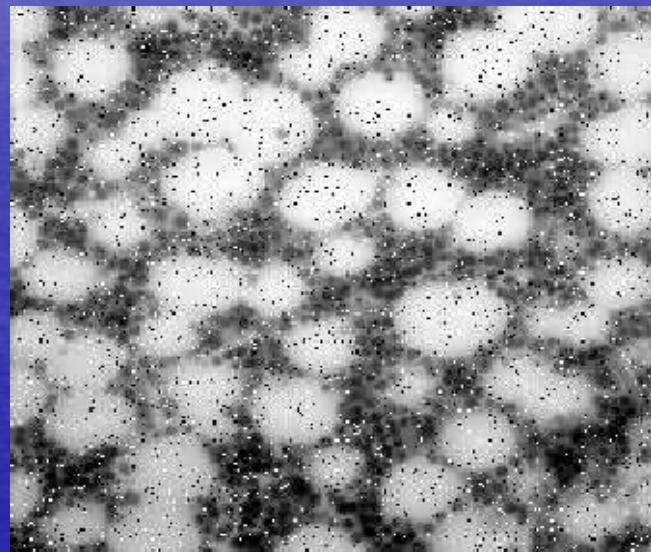
$$\text{Var}[\bar{g}(m, n)] = \frac{1}{M} \text{Var}[\eta(m, n)]$$

- Hơn nữa, nếu số lượng quan sát tăng ( $M \rightarrow \infty$ ), tác động nhiễu sẽ giảm về zero.

# Ví dụ



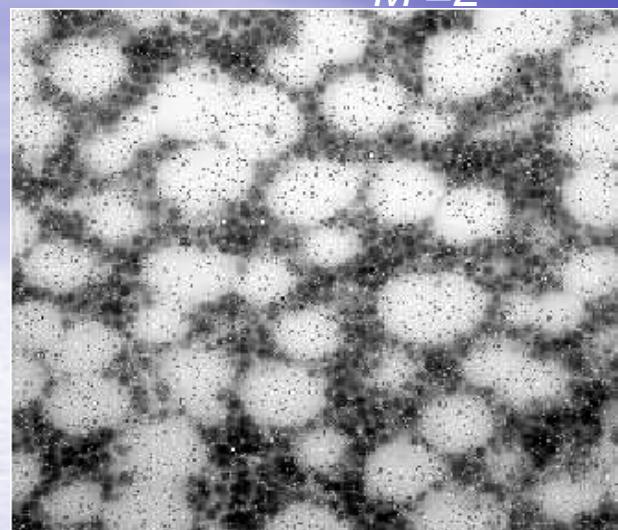
Ảnh không nhiễu



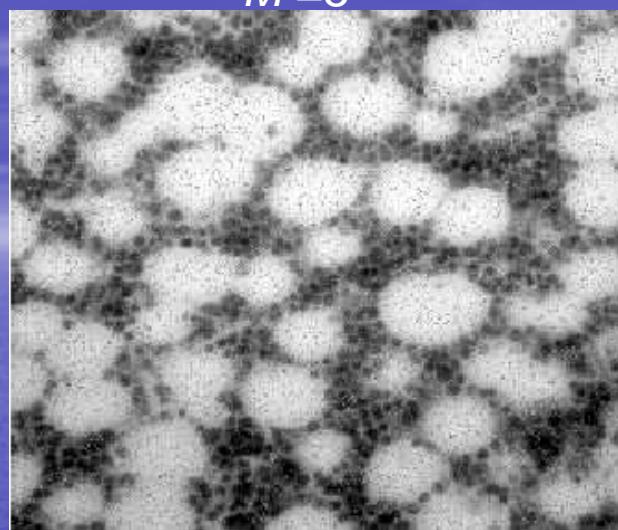
Ảnh nhiễu  
 $\text{Variance} = 0.05$

# Ví dụ

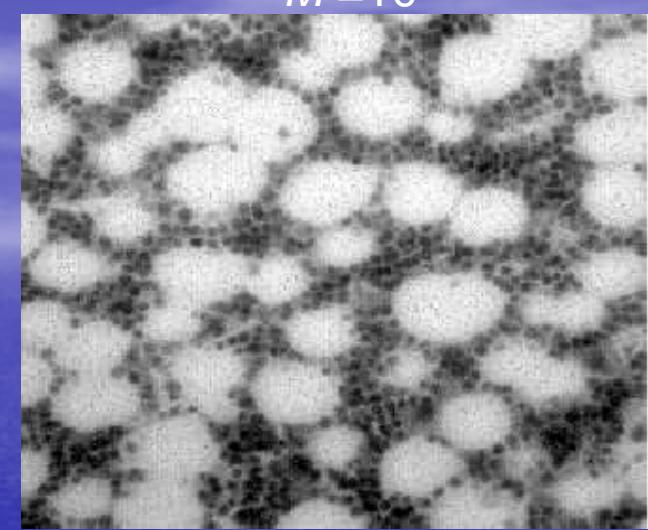
$M = 2$



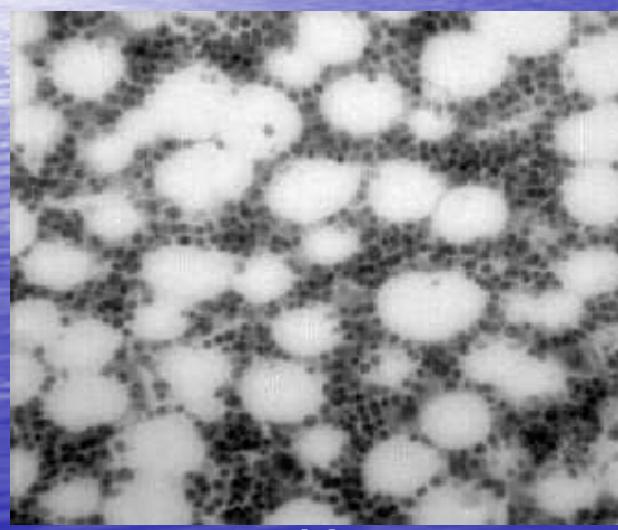
$M = 5$



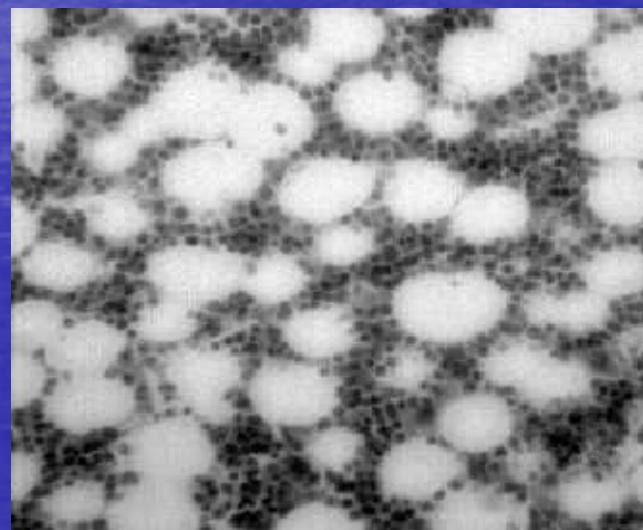
$M = 10$



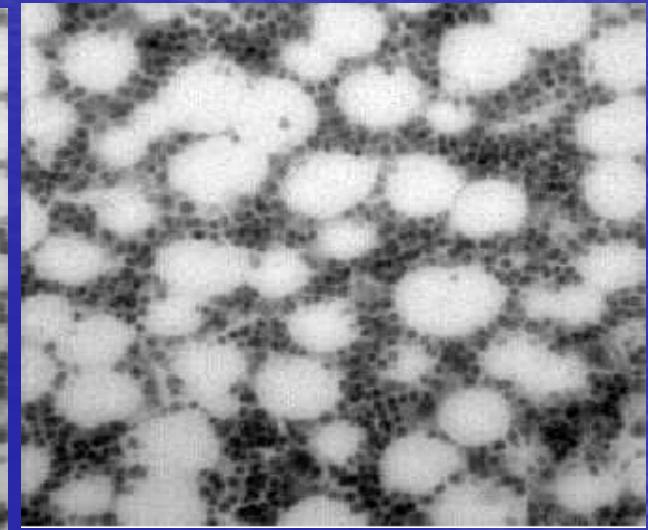
$M = 25$



$M = 50$



$M = 100$



# Vài bộ lọc trung bình

$$\frac{1}{9} \times \begin{matrix} & & \\ \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix} & & \end{matrix} \quad \frac{1}{16} \times \begin{matrix} & & \\ \begin{matrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{matrix} & & \end{matrix}$$

# Histogram

- Histogram của ảnh xám, có các giá trị xám  $r_0, r_1, \dots, r_{L-1}$  là hàm rời rạc

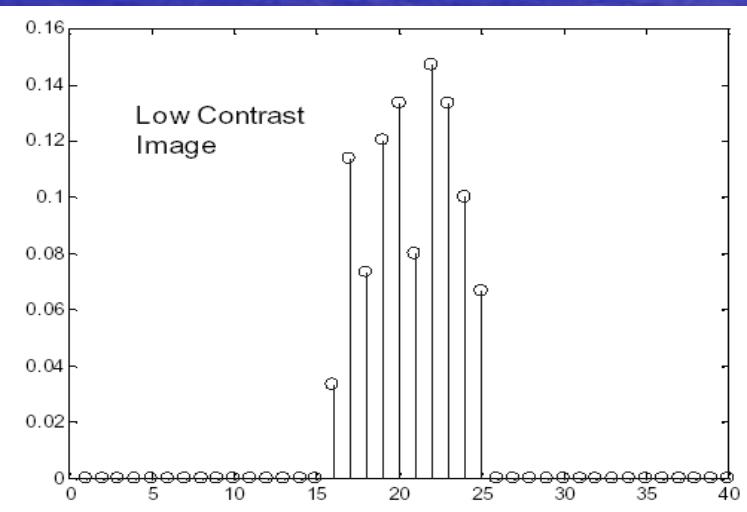
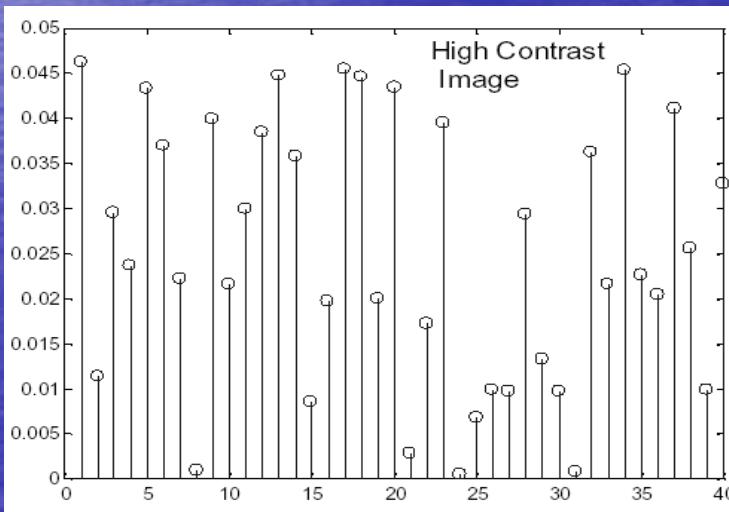
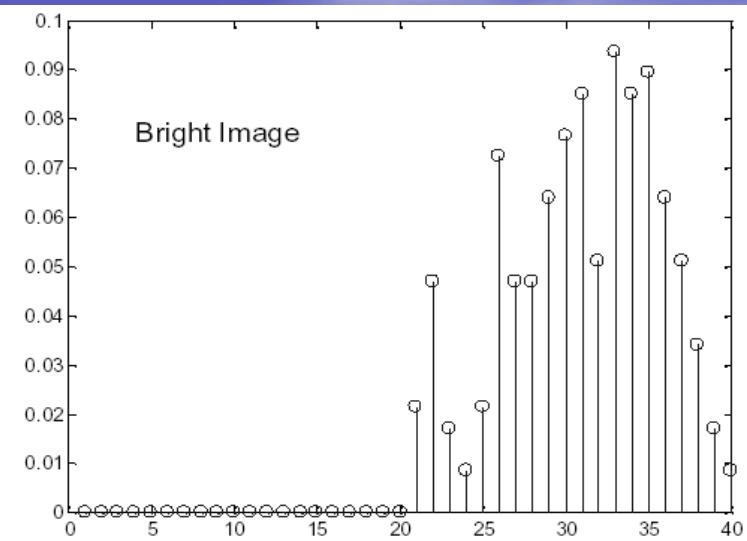
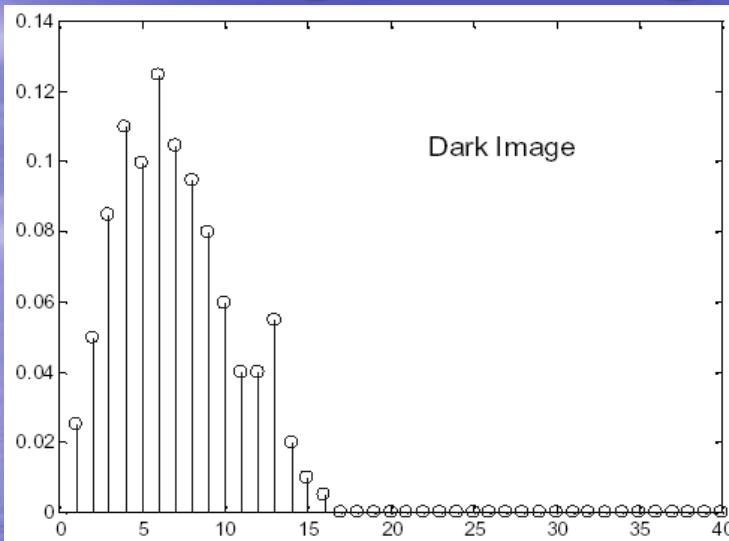
$$p(r_k) = \frac{n_k}{n}, \text{ với } \begin{cases} n_k \text{ là số lượng pixel có giá trị xám } r_k \\ n \text{ là tổng số pixel của ảnh} \end{cases}$$

- Hàm  $p(r_k)$  mô tả tỷ lệ tổng số pixel có mức xám  $r_k$  và tổng số pixel của ảnh.
- Histogram cung cấp một mô tả tổng thể phân bố của ảnh.
- Nếu chúng ta xem mức xám trong ảnh như biến ngẫu nhiên  $R$ , mật độ xác suất, thì histogram cung cấp như một xấp xỉ của mật độ xác suất.

$$\Pr[R=r_k] \approx p(r_k)$$

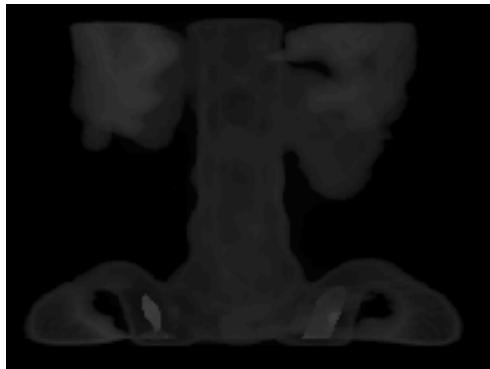
# Vài hình dạng Histogram

- Hình dạng của một histogram cung cấp thông tin hữu dụng cho hiệu chỉnh độ tương phản, làm tăng cường.

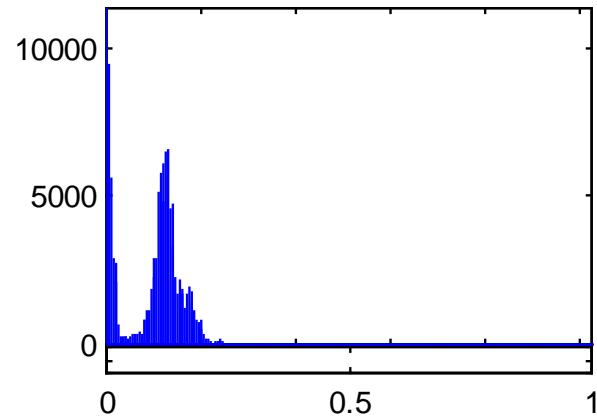


# Ví dụ kéo dãn histogram

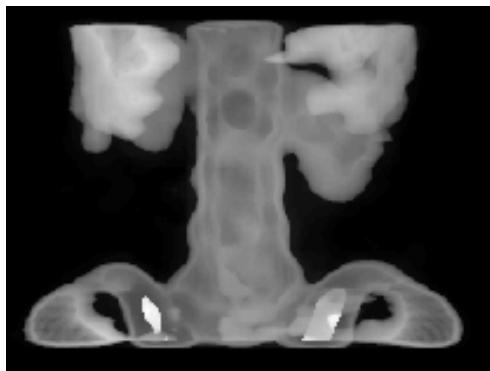
The original image displayed



The histogram of the original image



The stretched image



The stretched histogram

