

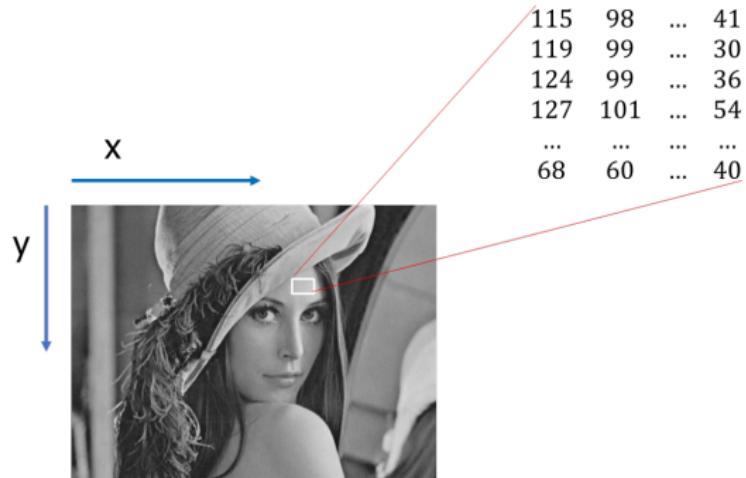
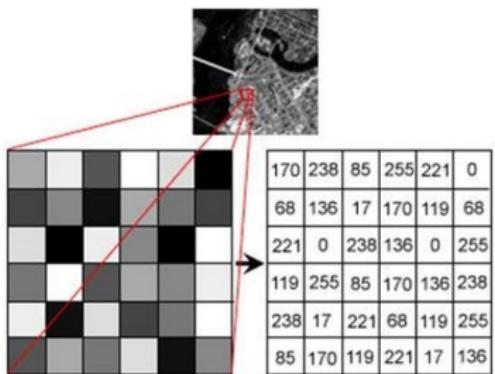
# CƠ SỞ DỮ LIỆU ẢNH

## Image Databases

### Hệ cơ sở dữ liệu đa phương tiện

HK1, 2023 - 2024

# Ảnh và lưu trữ



# Ảnh và lưu trữ

- Ảnh là tín hiệu 2D (x,y)

- Về mặt toán học:

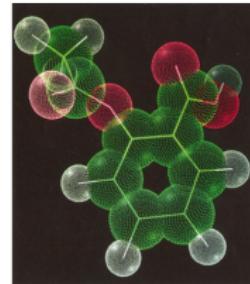
Ảnh là một ma trận biểu diễn tín hiệu

- Đối với người dùng:

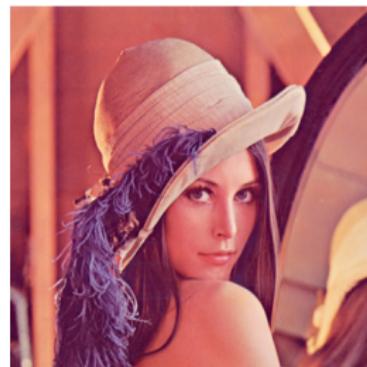
Ảnh chứa các thông tin ngữ nghĩa

# Ảnh và lưu trữ

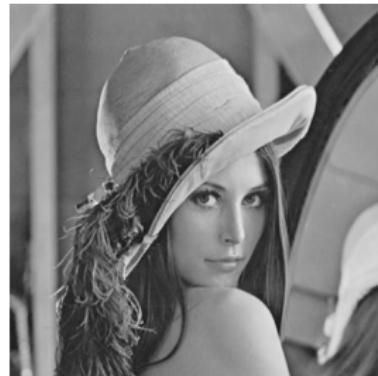
- **Ảnh tự nhiên:** ảnh thu nhận từ các thiết bị như
  - Camera, microscope, tomography, satellite, ...
- **Ảnh nhân tạo:** tạo ra từ các phần mềm chuyên biệt
  - Đồ họa máy tính (computer graphics), thực tại ảo (virtual reality)



# Ảnh và lưu trữ



Ảnh màu  
 $I_R(x, y), I_G(x, y), I_B(x, y)$

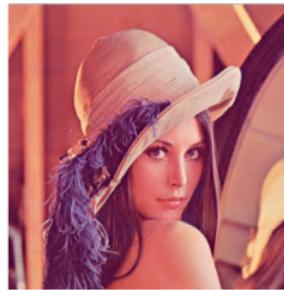


Ảnh mức xám  
 $I(x, y) \in [0, \dots 255]$



Ảnh nhị phân  
 $I(x, y) \in \{0, 1\}$

# Ảnh và lưu trữ



Ảnh màu trong hệ tọa độ RGB

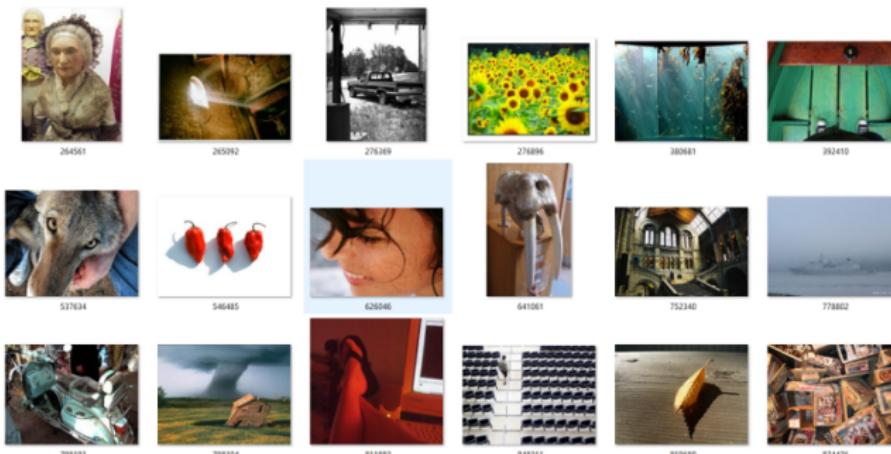
# Ảnh và lưu trữ



**Ảnh màu trong một số hệ tọa độ khác**

# Cơ sở dữ liệu ảnh

- **CSDL chung (general database)**: photo collection, Internet
  - Nội dung đa dạng và không đồng nhất
  - Biểu diễn dựa trên các đặc trưng chung (shape, color, texture...)
  - Vùng ứng dụng rộng



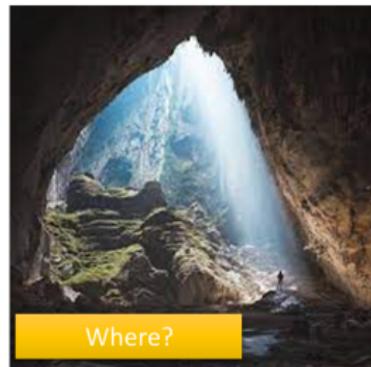
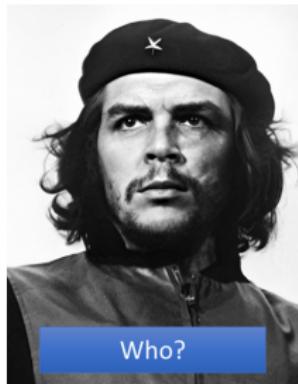
# Cơ sở dữ liệu ảnh

- **CSDL đặc thù (specific database)**: face database, medical database
  - Nội dung khá thuần nhất
  - Thường dành cho các ứng dụng đặc biệt (vd: ảnh sinh trắc học – biometry, ảnh y khoa – medical images)
  - Đặc trưng được trích chọn tùy thuộc vào mỗi lĩnh vực ứng dụng



# Tìm kiếm hình ảnh

- Tìm cách bức ảnh “Vietnam”?
- Đây là hình chụp của ai? Đây là cái gì? Hình này chụp ở đâu?



# Kiểu truy vấn: dựa trên từ khóa

- **Kiểu truy vấn dựa trên từ** (text-based image retrieval)

- Tìm các bức hình của “bill gates”?
- Được sử dụng rộng rãi: Google, Youtube, ...



# Kiểu truy vấn: dựa trên từ khóa

## Hạn chế

- Từ khóa được lấy từ
  - Văn bản xung quanh ảnh (Google Image, Google Video, ...)
  - Đánh dấu thủ công → giá thành cao
- Nhập nhằng: chọn từ khóa như thế nào?
  - Tính chủ quan của người đánh dấu keyword-image
  - Phụ thuộc ngôn ngữ
  - Phụ thuộc ngữ cảnh

# Kiểu truy vấn: dựa trên nội dung

**IRMA Query Demo 3.3**



Restart

Found: 100 (1-12) - Query Time: 0.08 s

					
Select	Select	Select	Select	Select	Select
					
Select	Select	Select	Select	Select	Select

Navigation icons: back, forward, search, etc.

# Kiểu truy vấn: dựa trên nội dung



# Kiểu truy vấn: dựa trên nội dung

- Content-Based Image Retrieval

- Kỹ thuật **sử dụng nội dung trực quan (visual content)** để tìm kiếm hình ảnh từ cơ sở dữ liệu hình ảnh lớn tùy theo nhu cầu người dùng

- Giúp tìm được hình ảnh mong muốn

- Đặc điểm

- Truy vấn **chứa hình ảnh**

- Mục đích tìm kiếm: “**tương tự**” – đối sánh không chính xác

# Kiểu truy vấn: dựa trên nội dung

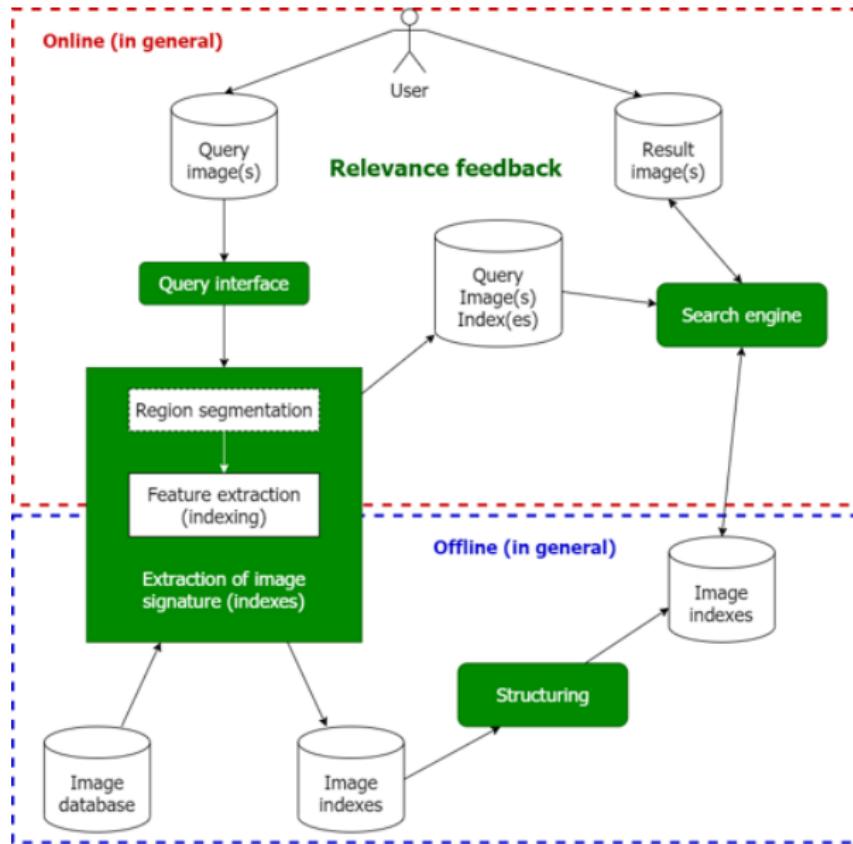
- Trích chọn tự động các đặc trưng và tìm kiếm dựa trên **độ tương tự giữa các đặc trưng**
- Lĩnh vực nghiên cứu đang được quan tâm
- Truy vấn:
  - Bằng ví dụ
  - Bộ phận
  - Bằng hình vẽ (sketch)

# Kiểu truy vấn: dựa trên nội dung

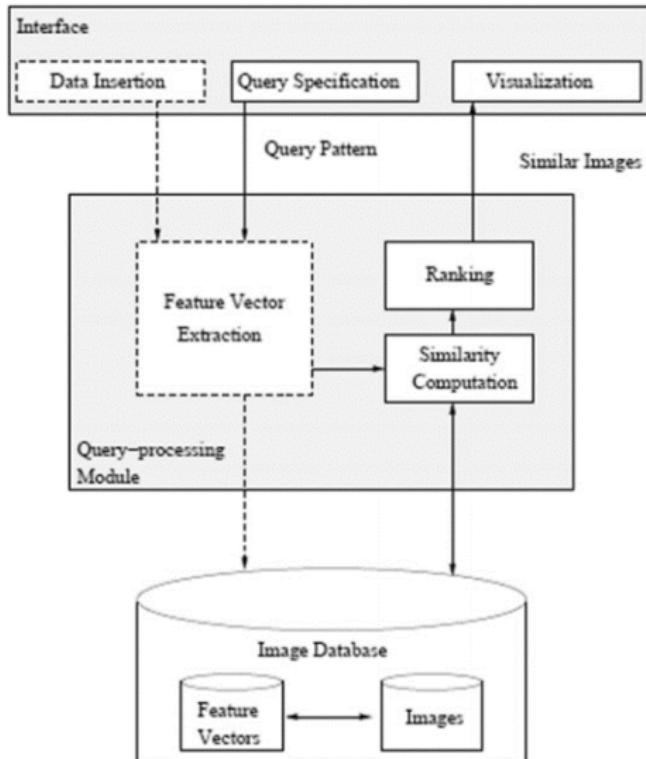
## Vấn đề

- Nội dung đa dạng
- Khối lượng dữ liệu cần xử lý lớn và thường trùng lặp: *Numeric gap*
- Vấn đề về ngữ nghĩa: *Semantic gap*  
Làm thế nào để biết 1 người đang vui/buồn khi nhìn vào hình?

# Kiểu truy vấn: dựa trên nội dung



# Kiểu truy vấn: dựa trên nội dung



# Kiểu truy vấn: dựa trên nội dung

## Các cấp độ truy vấn (query)

- **Cấp độ 1:** truy xuất theo **đặc trưng ban đầu** như màu sắc (color), kết cấu (texture), hình dạng (shape), vị trí không gian (spatial location) của các thành phần trong hình ảnh
- **Cấp độ 2:** truy xuất các đối tượng thuộc loại đã cho được xác định bởi các **thuộc tính được trích ra**
- **Cấp độ 3:** truy xuất các **thuộc tính trừu tượng** liên quan đến một lượng lớn suy luận cấp cao về mục đích của các đối tượng hoặc cảnh được mô tả

# Kiểu truy vấn: dựa trên nội dung

## CBIR – các bước chính

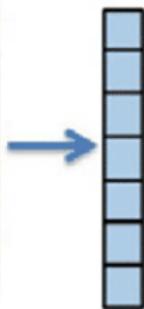
- Trích chọn đặc trưng (biểu diễn ảnh)
- Đánh chỉ mục để tăng hiệu năng tìm kiếm
- Lựa chọn thước đo độ tương tự

# Kiểu truy vấn: dựa trên nội dung

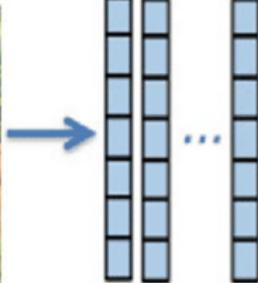
Biểu diễn ảnh – trích chọn đặc trưng

- Đặc trưng toàn cục
  - Màu sắc, hình dáng, kết cấu, ...
- Đặc trưng cục bộ
  - Phân vùng
  - Key points (character points)
- Đặc trưng riêng: tùy thuộc vào từng ứng dụng cụ thể
  - Nhận dạng vân tay: minuties (các điểm đầu, điểm rẽ nhánh)
  - Nhận dạng mặt: EigenFaces

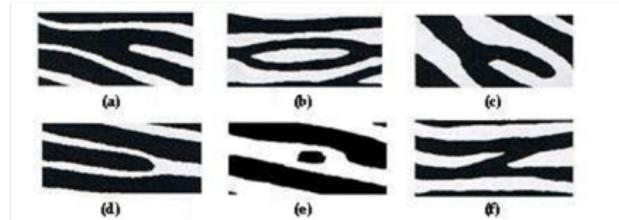
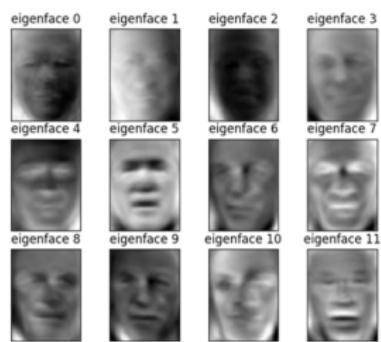
# Kiểu truy vấn: dựa trên nội dung



Global feature representation



Local feature representation



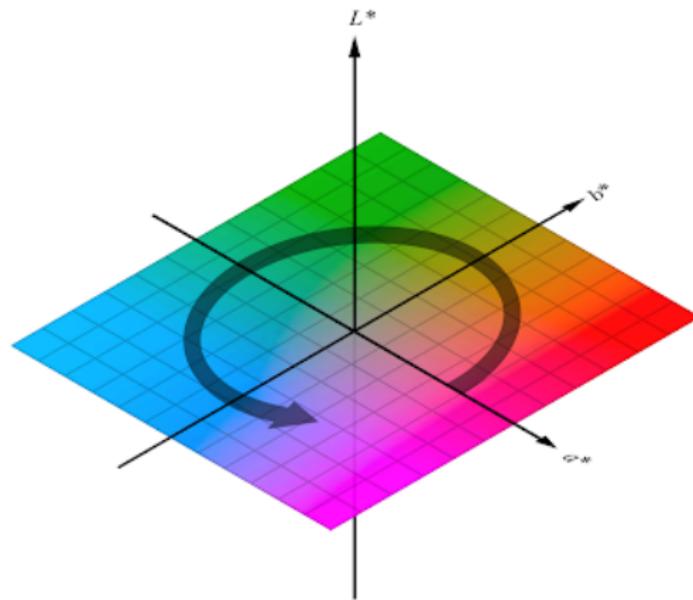
# Đặc trưng màu sắc

Không gian màu sắc: RGB, Munsell, CIE L\*a\*b\*, CIE L\*u\*v\*, HSV (hay HSL, HSB),...

- **RGB**: sử dụng rộng rãi để hiển thị hình ảnh
- **CMY**: sử dụng cho in ấn (bao gồm bộ 3 màu cyan, magenta và yellow)
- **CIE L\*a\*b\***, **CIE L\*u\*v\***: không phụ thuộc vào thiết bị; bao gồm thành phần độ sáng (L) và hai thành phần màu
- **HSV**: được sử dụng rộng rãi trong đồ họa máy tính, là một cách trực quan để mô tả màu sắc; bao gồm 03 thành phần màu là màu (hue), độ bảo hòa (saturation) và giá trị (value, độ đậm nhạt hay độ sáng)

# Đặc trưng màu sắc

CIE L\*a\*b\*

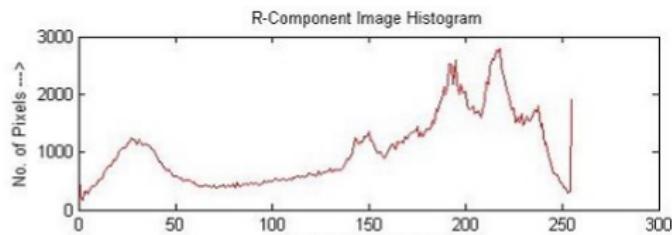


# Histogram

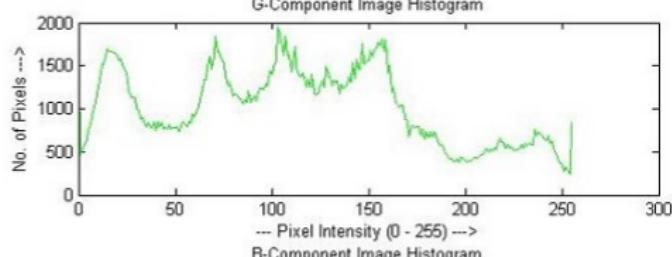
- Thể hiện sự phân bố màu sắc trên ảnh
  - Thể hiện hiệu quả nội dung màu của hình ảnh nếu mẫu màu là duy nhất so với phần còn lại của tập dữ liệu
  - Cách tính:
    - Lấy mẫu màu của ảnh (bin)
    - Tính số điểm ảnh tương ứng với mỗi bin
      - Chuẩn hóa: chia cho số điểm ảnh
- $$H = (h_1, h_2, \dots, h_n) \text{ với } h_i \in [0, 1] \quad \text{và } \sum h_i = 1$$

# Histogram

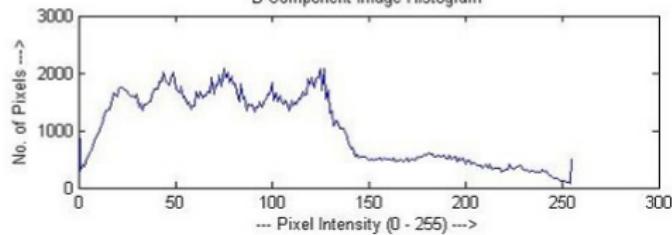
R-Component Image



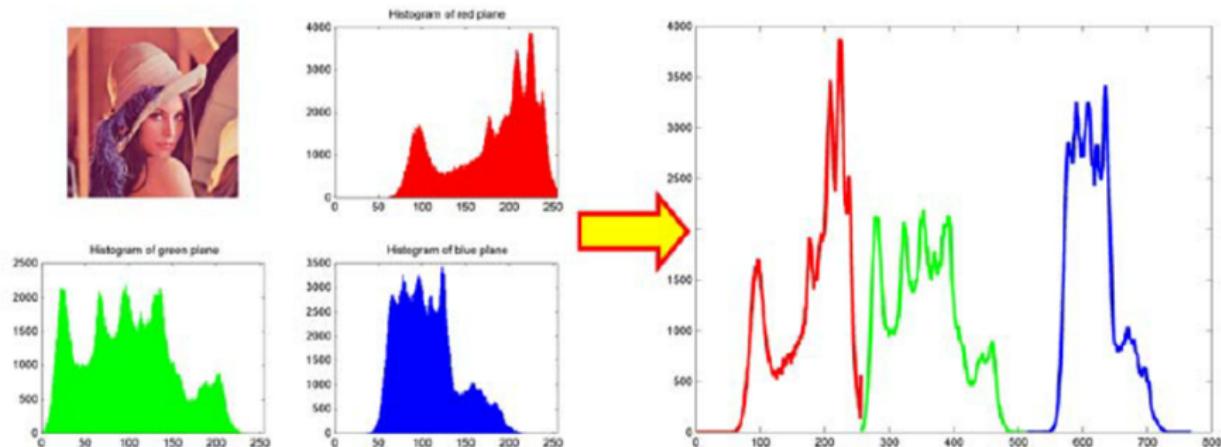
G-Component Image



B-Component Image



# Histogram



# Histogram

## Khoảng cách dựa trên sự phân bố màu

- Thường sử dụng khoảng cách L1 hay L2 (Euclidean)

$$d_{L1}(H, G) = \sum_{i=1}^N |h_i - g_i|$$

- Giao giữa 2 histogram

$$\cap(H, G) = \frac{\sum_i \min(h_i, g_i)}{\sum_i g_i}$$

# Histogram

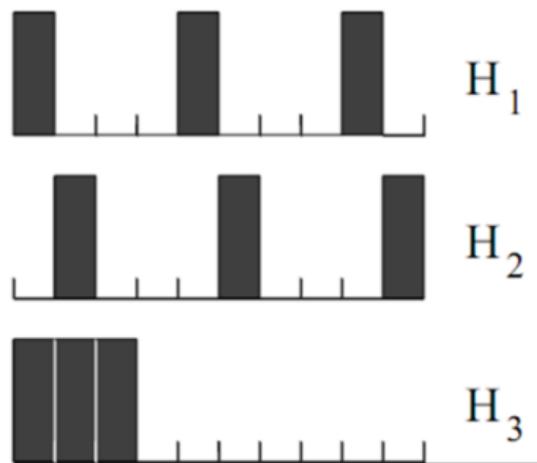
Ví dụ: có 03 hình kích thước  $8 \times 8$ , mỗi pixel thuộc về 1 trong 8 màu  $C_1$  đến  $C_8$

- Hình 1: mỗi màu gồm 8 pixel
  - Hình 2: màu  $C_1$  đến  $C_4$  có 7 pixel, màu  $C_5$  đến  $C_8$  có 9 pixel
  - Hình 3:  $C_1$  và  $C_2$  có 2 pixel,  $C_3$  đến  $C_8$  có 10 pixel
- 2 hình nào có mức độ tương tự cao nhất?

# Histogram

## Vấn đề gấp phải với phân bố màu

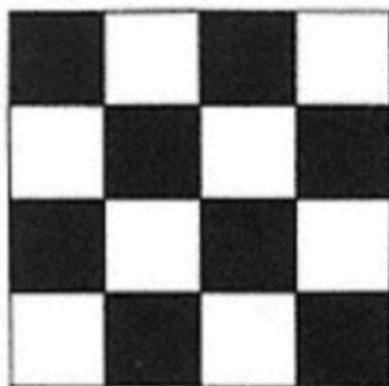
- Không tính đến sự gần giống giữa các màu trong các bin gần nhau



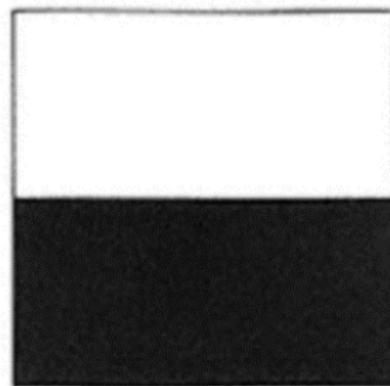
# Histogram

## Vấn đề gấp phải với phân bố màu

- Việc phân bố trong không gian các màu điểm ảnh: 2 ảnh khác nhau, histogram có thể giống nhau



(a)

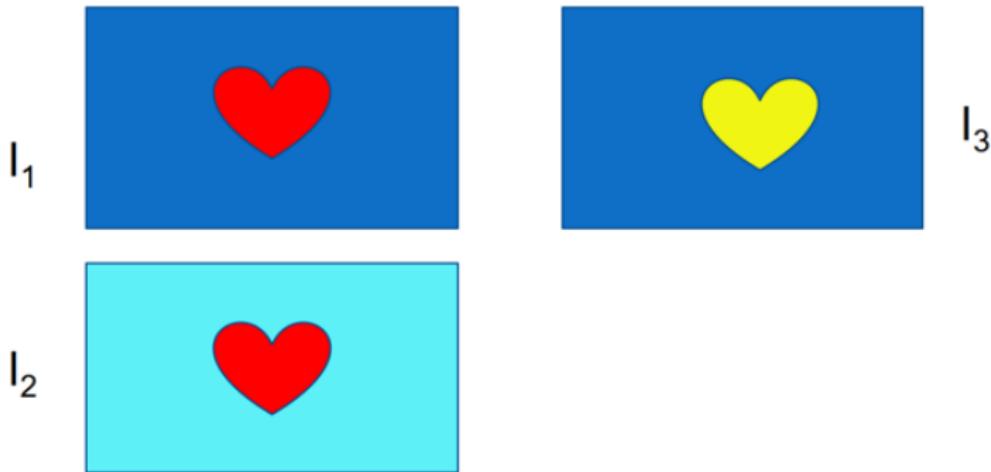


(b)

# Histogram

## Vấn đề gấp phai với phân bố màu

- Ảnh hưởng của nền:  $d(l_1, l_2) \neq d(l_1, l_3)$



# Histogram

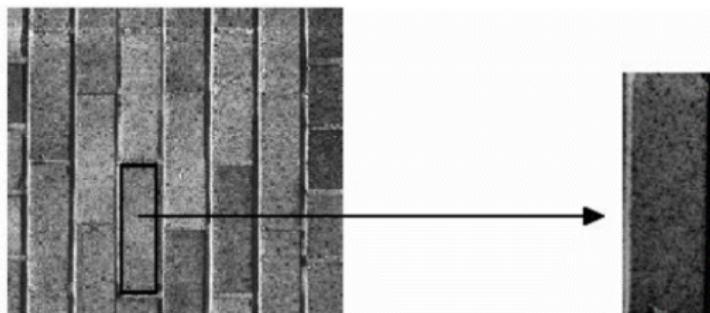
## Vấn đề gấp phải với phân bố màu

- Phụ thuộc cách biểu diễn màu (không gian màu)
- Phụ thuộc thiết bị
- ...

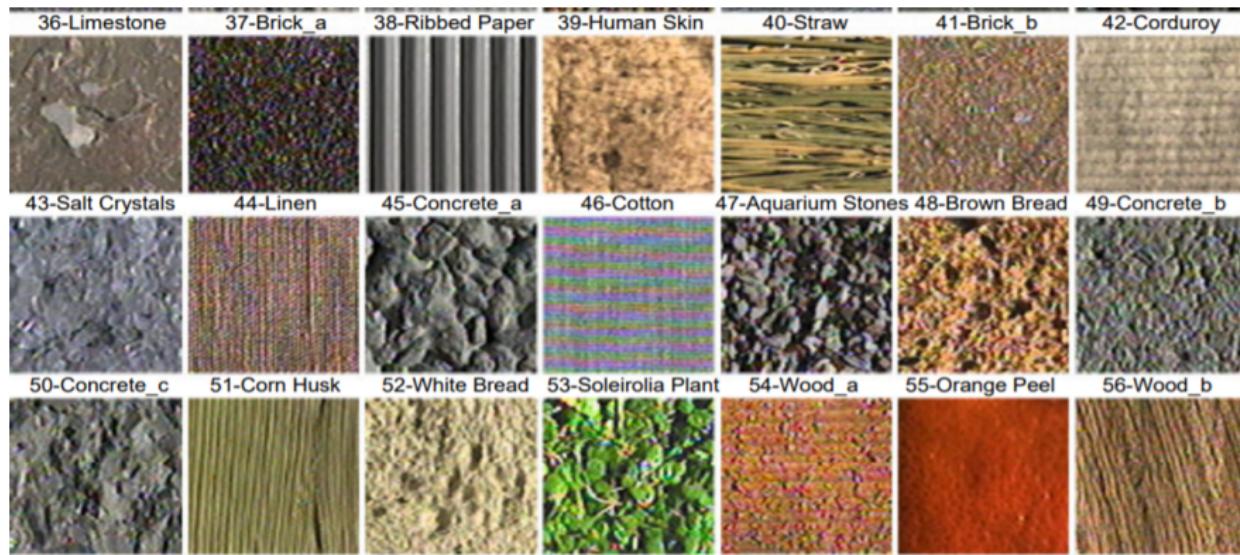
# Texture

## Đặc trưng kết cấu (texture)

- Tập các primitive (mẫu) được sắp xếp theo quy tắc sắp chỗ nào đó
- Primitive: tập 1 hoặc nhiều điểm liên thông với nhau có độ xám (màu sắc) gần giống nhau
- Kết cấu: có thể có chu kỳ hoặc không có chu kỳ

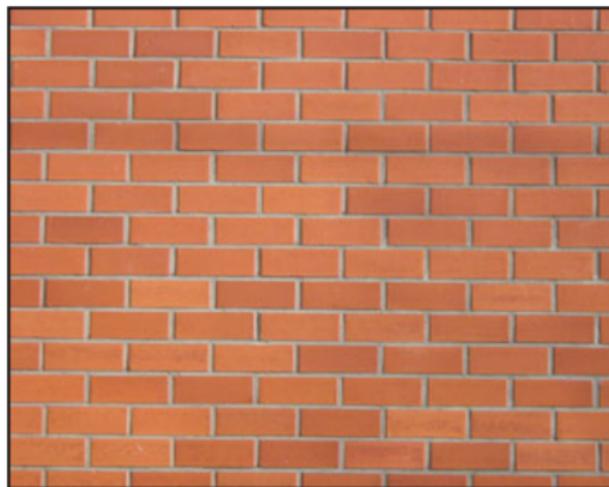


# Texture



# Texture

Kết cấu có thể phân hai nhóm chính: kết cấu có cấu trúc (structural texture) và kết cấu ngẫu nhiên (stochastic texture)

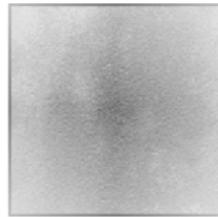


# Texture

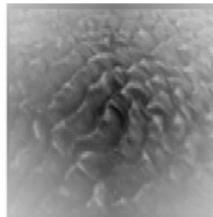
6 đặc trưng mô tả kết cấu (theo Tamura)

- Tính thô ráp (coarseness)
- Tính tương phản (contrast)
- Tính có hướng (directionality)
- Tính giống đường (line likeness)
- Tính trật tự (regularity)
- Tính gồ ghề (roughness)

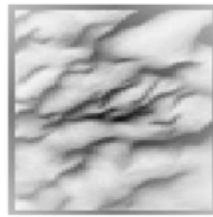
# Texture



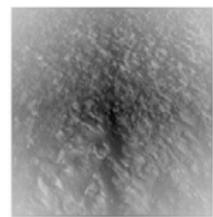
Fine



Coarse



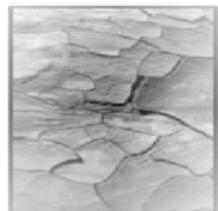
High contrast



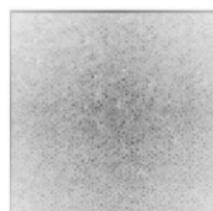
Low contrast



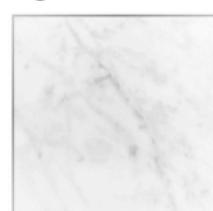
Directional



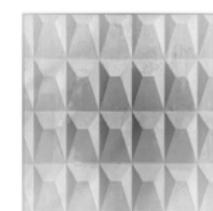
Non directional



Rough



Smooth

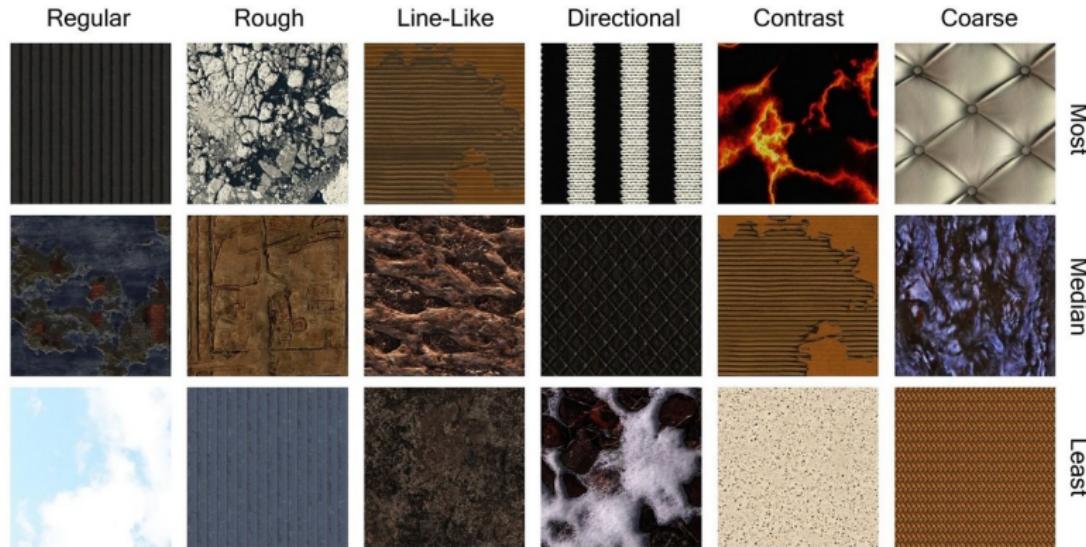


Regular



Random

# Texture



# Texture

6 đặc trưng mô tả kết cấu (theo Tamura)

- **Tính thô ráp (coarseness)**

- Đặc điểm kết cấu cơ bản nhất. Các thành phần hình ảnh có sự khác biệt càng lớn thì hình ảnh càng thô
  - ⇒ Hình ảnh phóng to sẽ “thô” hơn hình ảnh gốc

- **Tính tương phản (contrast)**

- Chất lượng ảnh
- Các tham số đo độ tương phản
  - Phạm vi động của các cấp độ xám
  - Phân cực của sự phân bố màu đen và trắng trên biểu đồ mức xám hoặc tỷ lệ của các vùng đen và trắng
  - Độ sắc nét của các cạnh
  - Khoảng lặp lại các mẫu

# Texture

6 đặc trưng mô tả kết cấu (theo Tamura)

- **Tính có hướng (directionality)**

- Tập trung vào hình dạng của thành phần kết cấu và mối quan hệ của các thành phần (tổ chức của các thành phần kết cấu trong một vùng cụ thể)
- Tamura: đo lường tổng số hướng khác nhau của 1 phần tử kết cấu
- Ví dụ: 02 mẫu kết cấu khác nhau về hướng thì có cùng giá trị về tính có hướng

- **Tính giống đường (line likeness)**

- Hình dạng của 1 phần tử kết cấu
- Được định nghĩa là sự trùng của các hướng cạnh (các góc định hướng) cùng xảy ra trong các cặp pixel cách nhau một khoảng  $d$  đọc theo hướng cạnh của mỗi pixel

# Texture

6 đặc trưng mô tả kết cấu (theo Tamura)

- **Tính trật tự (regularity)**

- Đo lường sự thay đổi của quy tắc vị trí phần tử.
- Liên quan đến việc kết cấu là đều đặn hay không đều.
- Hình dạng phần tử khác nhau làm giảm tính đều đặn. Một kết cấu tốt có xu hướng được coi là đều đặn.

- **Tính gồ ghề (roughness)**

- Đo lường xem kết cấu là thô hay mịn
- Liên quan đến thô ráp và tương phản.

# Texture

Kết cấu có thể mô tả dựa trên

- Phân tích thống kê
  - Tham số tính từ biểu đồ phân bố màu (histogram)
  - Phân tích ma trận đồng xuất hiện (Co-occurrence matrix) ⇒ đặc trưng Haralick
- Phân tích trên miền tần số
  - Biến đổi Fourier
  - Biến đổi Gabor...
- Mô hình

# Ma trận đồng xuất hiện mức xám (GLCM)

## Gray Level Co-occurrence Matrix (GLCM)

- Ma trận đồng xuất hiện xác định tần xuất hai giá trị pixel sẽ xuất hiện trong một mối quan hệ không gian xác định, cách nhau một khoảng cách nhất định
- Từ đó, đặc trưng kết cấu được xác định phân tích thống kê của các ma trận này
- Thường 4 hướng chính:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$

