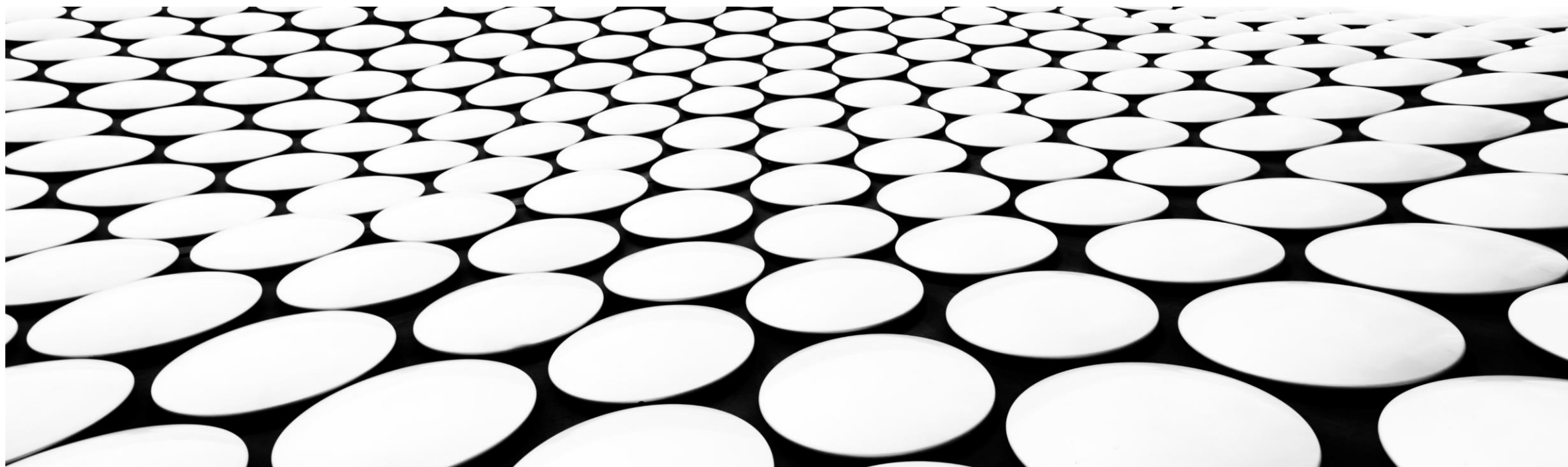


---

# CƠ SỞ DỮ LIỆU ẢNH IMAGE DATABASE (TT)

HỆ CSDL ĐA PHƯƠNG TIỆN



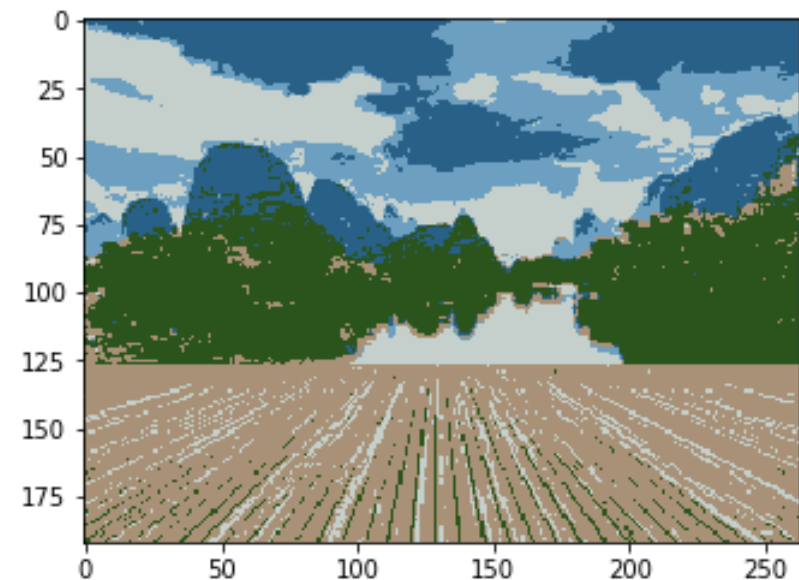
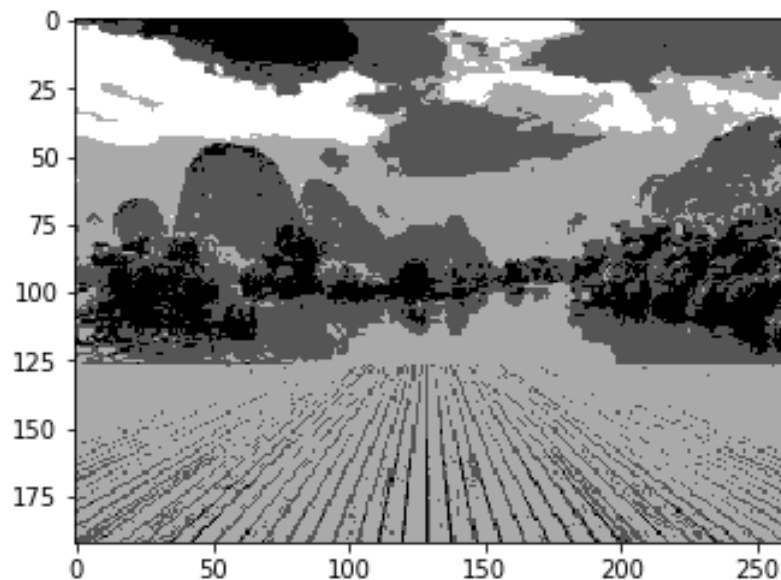
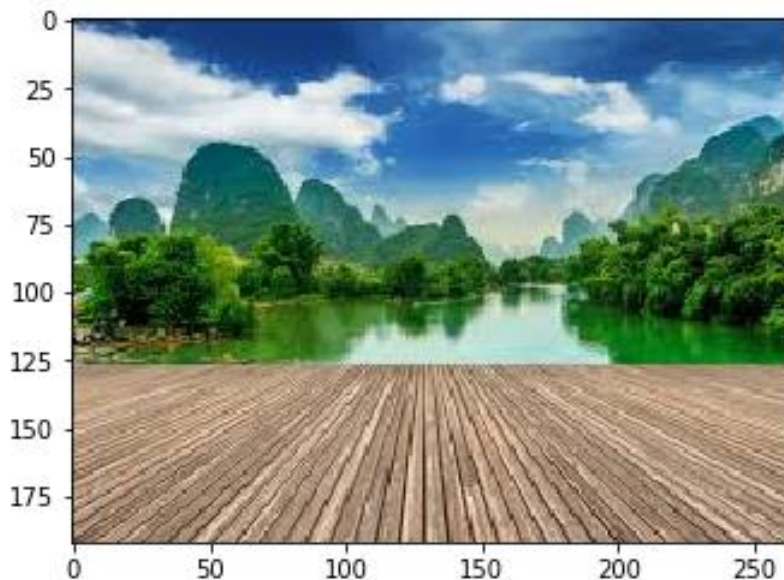
# PHÂN ĐOẠN ẢNH

- ❖ Phân đoạn (segmentation) chia một hình ảnh thành các vùng hoặc các đối tượng trong ảnh.
- ❖ Phân đoạn cho phép trích xuất các đối tượng trong ảnh
- ❖ Hữu ích:
  - Sau khi phân đoạn thành công, đường bao của đối tượng được trích xuất
  - Hình dạng của đối tượng được mô tả
  - Dựa trên hình dạng, kết cấu và màu sắc, các đối tượng được xác định
  - Kỹ thuật phân đoạn ảnh được sử dụng trong tìm kiếm tương tự

# PHÂN ĐOẠN ẢNH

## ❖ Mục tiêu:

- Chia ảnh thành các vùng tương ứng với các đối tượng trong ảnh
- Sau đó xác định các đặc trưng cho đối tượng



# PHÂN ĐOẠN ẢNH

- ❖ Thuật toán phân đoạn dựa trên một trong hai thuộc tính cơ bản của màu sắc, các giá trị xám hoặc kết cấu là tính **không liên tục (discontinuity)** và **tương tự (similarity)**
  - Sự không liên tục: phân vùng dựa trên sự thay đổi về cường độ
    - ✓ Các biên (edge, contour)
  - Sự tương tự: phân chia hình ảnh thành các vùng tương tự nhau theo tiêu chí xác định
    - ✓ Ngưỡng biểu đồ (histogram thresholding)...

# PHÂN ĐOẠN ẢNH

## ❖ Các cách tiếp cận

- Tiếp cận điểm
- Tiếp cận vùng: tìm các vùng đồng nhất trong ảnh
- Tiếp cận biên: phát hiện sự không liên tục trong ảnh
- Kết hợp biên - vùng

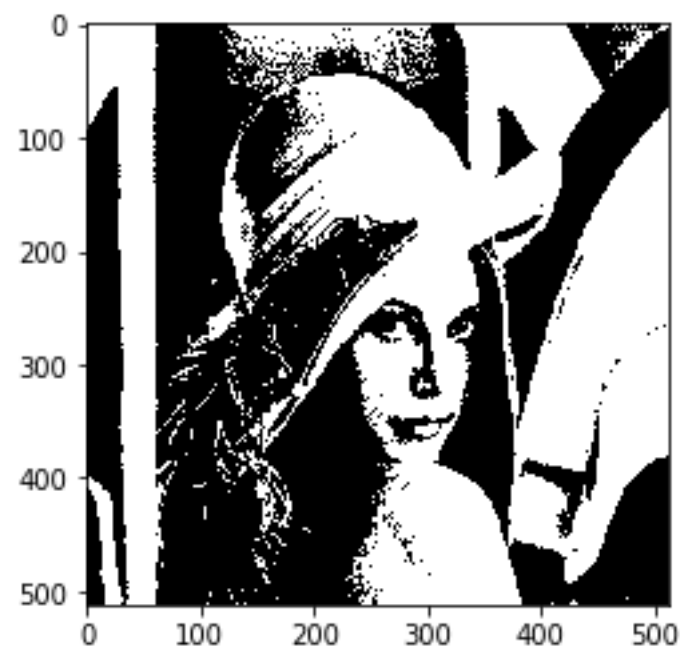
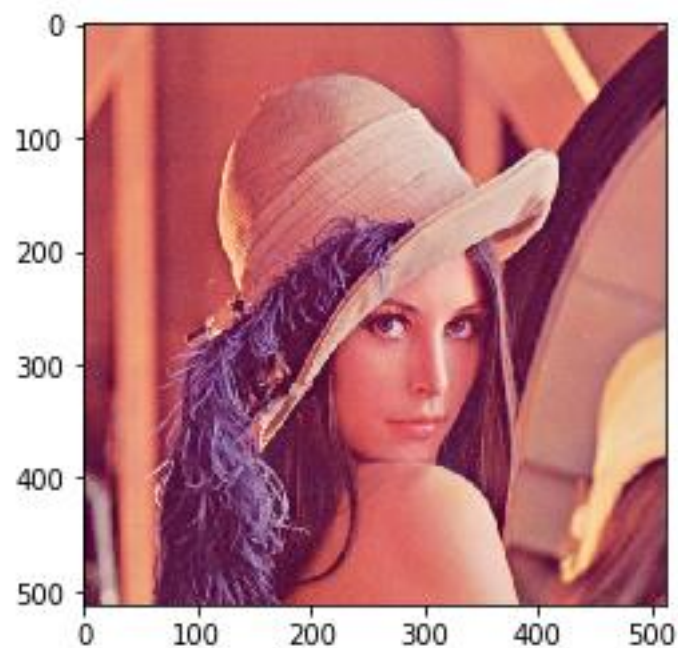
# PHÂN ĐOẠN BẰNG NGƯỠNG

## ❖ Phân đoạn bằng ngưỡng (Segmentation by Thresholding)

- Giả sử biểu đồ mức xám (gray-level histogram) ứng với hình ảnh  $f(x, y)$  gồm các đối tượng tối trên nền sáng
  - ✓ Pixel của đối tượng và background có mức xám được nhóm vào trong 2 chế độ
    - If  $f(x, y) \geq \mathcal{T}$  then  $f(x, y) = 1$
    - If  $f(x, y) < \mathcal{T}$  then  $f(x, y) = 0$

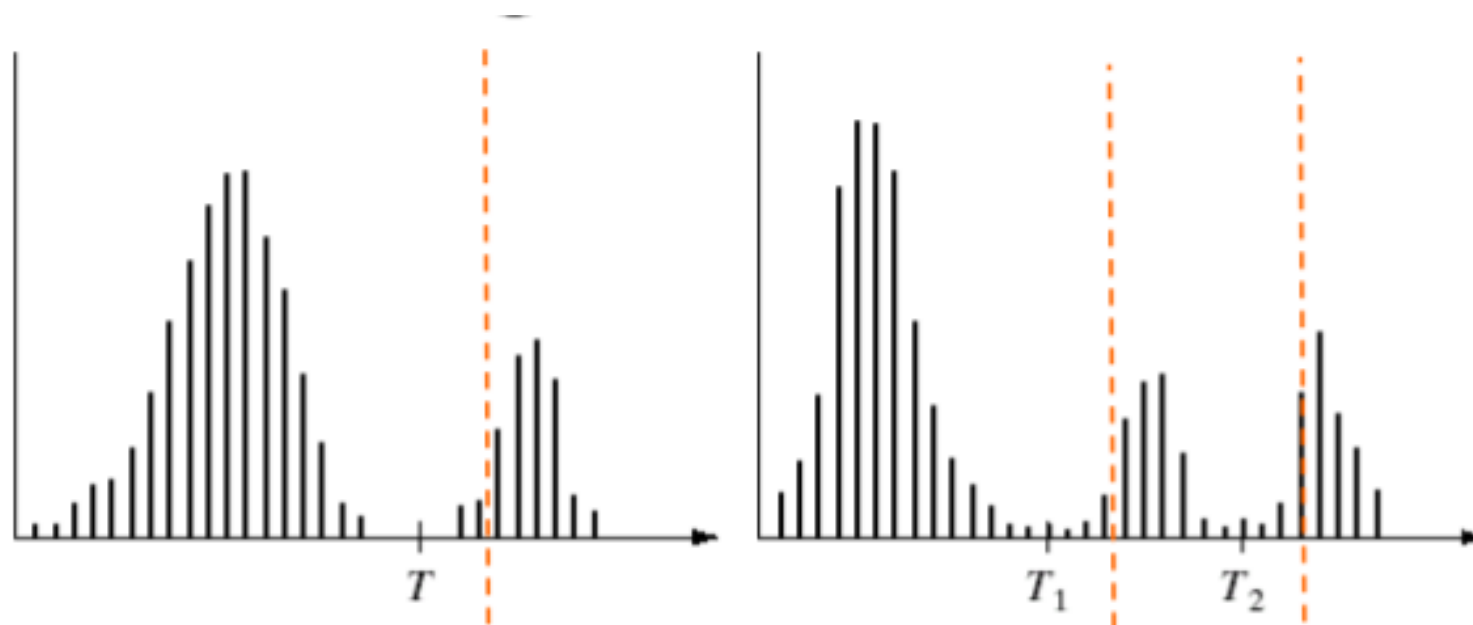
# PHÂN ĐOẠN BẰNG NGƯỠNG

❖ Không thuộc vào cách tiếp cận vùng



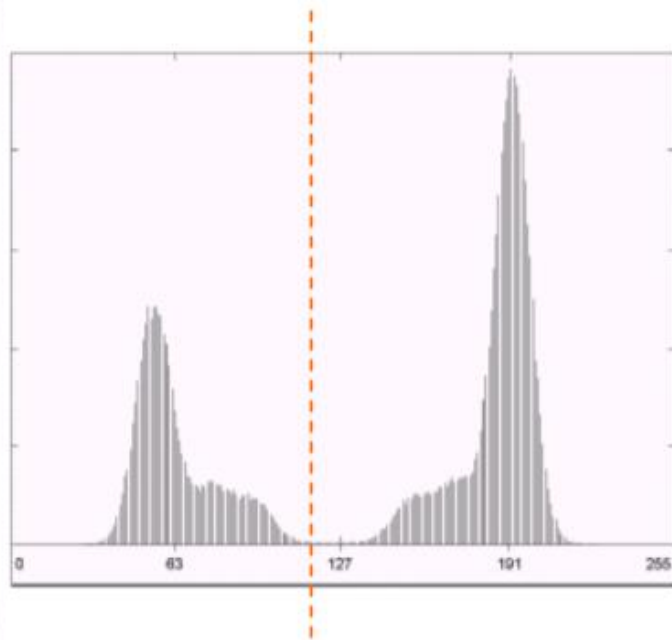
# VẤN ĐỀ VỚI NGƯỠNG GIÁ TRỊ ĐƠN

- ❖ Ngưỡng giá trị đơn (single value thresholding) phù hợp với biểu đồ hai phương thức (bimodal histogram)





# VÍ DỤ PHÂN ĐOẠN BẰNG NGƯỠNG



# PHÂN ĐOẠN BẰNG NGƯỠNG

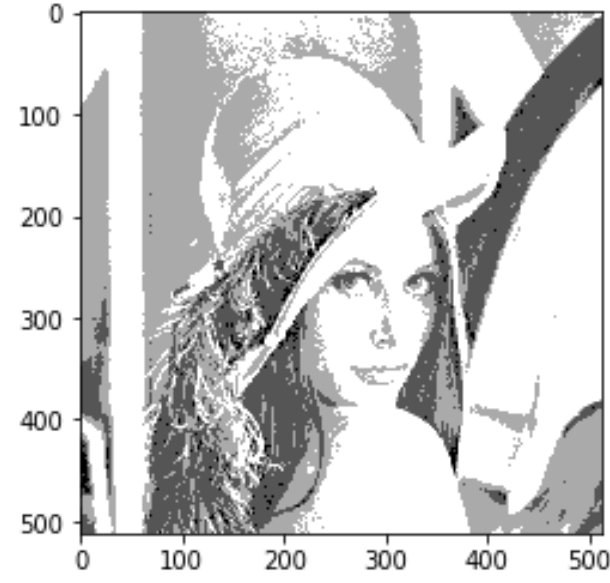
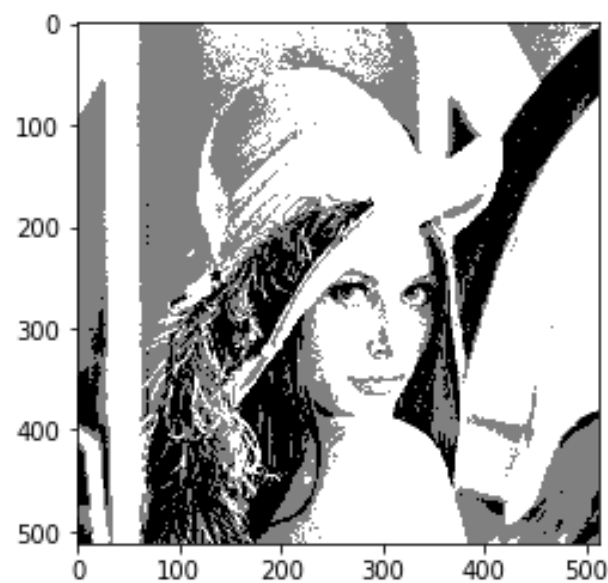
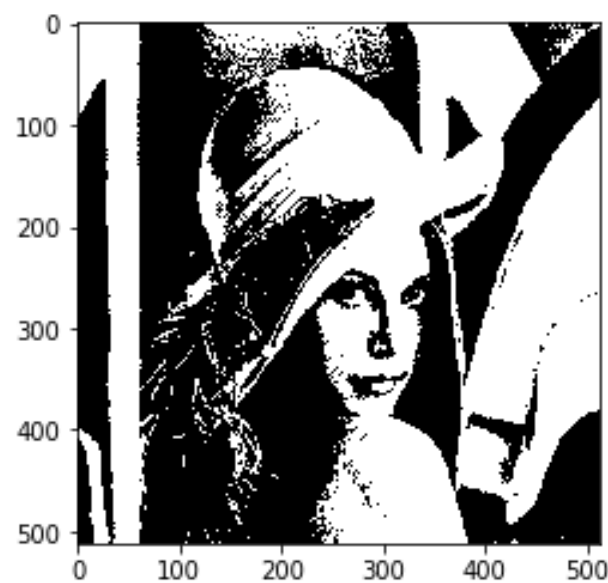
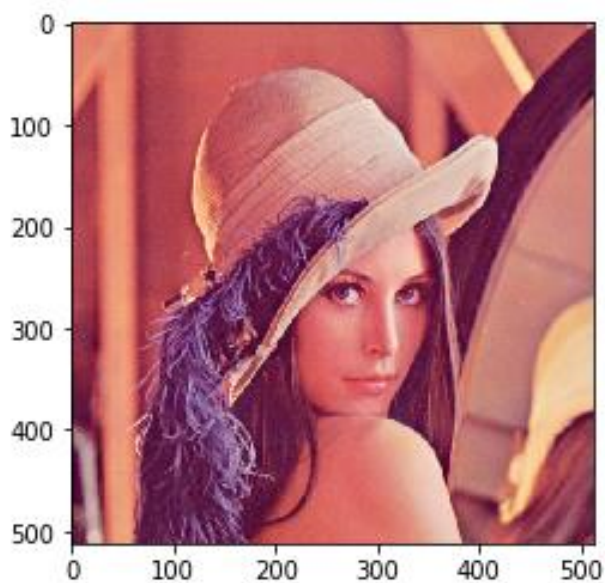
❖ Có thể dùng  $n$  ngưỡng  $\rightarrow (n + 1)$  vùng

❖ Ngưỡng:

- Toàn cục (global threshold): một ngưỡng cho toàn bộ ảnh
- Cục bộ (local threshold): một ngưỡng cho 1 vùng của ảnh
- Thích nghi (adaptive threshold): ngưỡng được lựa chọn phù hợp với từng ảnh hoặc từng vùng của ảnh

# PHÂN ĐOẠN BẰNG NGƯỠNG

❖ Các ngưỡng khác nhau



# NGƯỠNG DỰA TRÊN PHÂN BỐ

## ❖ Ngưỡng toàn cục:

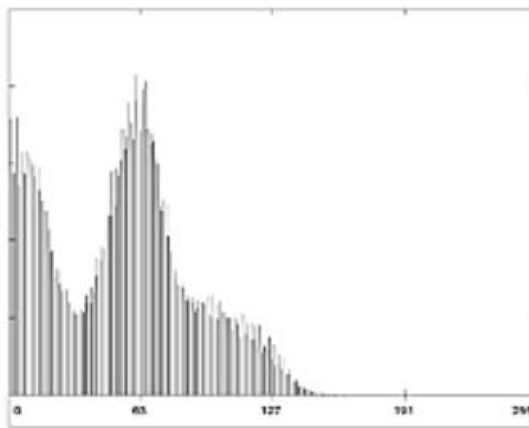
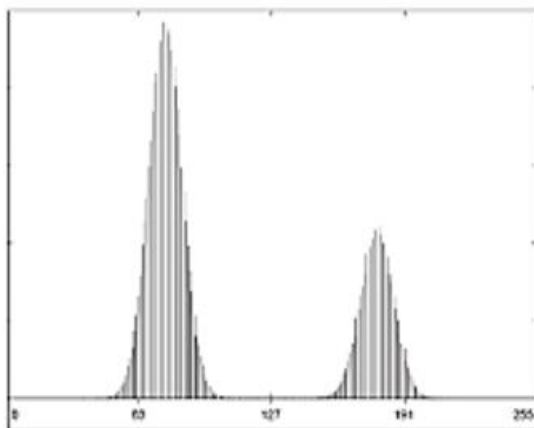
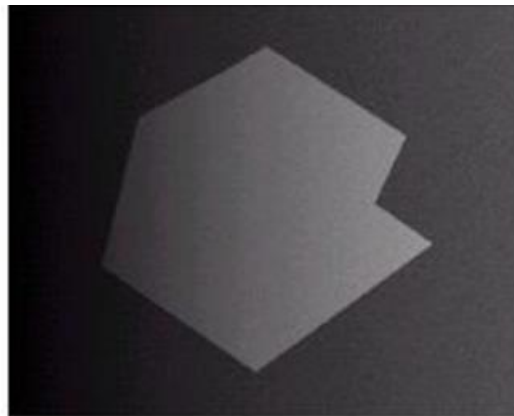
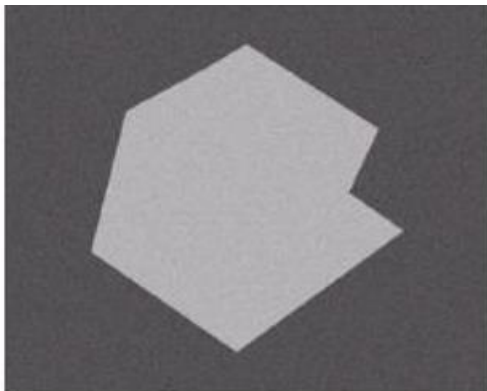
- Phù hợp cho môi trường ổn định
  - Đơn giản, nhanh
  - Nhược điểm:
    - ✓ Cần biết số lớp
    - ✓ Không có ràng buộc về không gian
  - Xác định giá trị ngưỡng
    - ✓ Bảng thử nghiệm, trung bình mức xám, trung bình max min
    - ✓ Giá trị cho phép cân bằng giữa các vùng của phân bố
- Tự động: T sao cho lỗi là nhỏ nhất

# NGƯỠNG DỰA TRÊN PHÂN BỐ

❖ Lựa chọn ngưỡng tối ưu lặp lại (Iterative optimal threshold selection)

1. Chọn giá trị ước tính đầu cho  $T$
2. Phân đoạn hình ảnh bằng cách sử dụng  $T \rightarrow$  tạo 2 nhóm: gồm  $G_1$  pixel có giá trị  $\geq T$  và  $G_2$  pixel có giá trị  $< T$
3. Tính  $\mu_1$  và  $\mu_2$  là giá trị trung bình của  $G_1$  và  $G_2$  pixel
4. Ngưỡng mới:  $T = \frac{1}{2} (\mu_1 + \mu_2)$
5. Lặp lại bước 2 – 4 cho đến khi  $T$  ổn định

# VẤN ĐỀ VỚI NGƯỠNG TOÀN CỤC



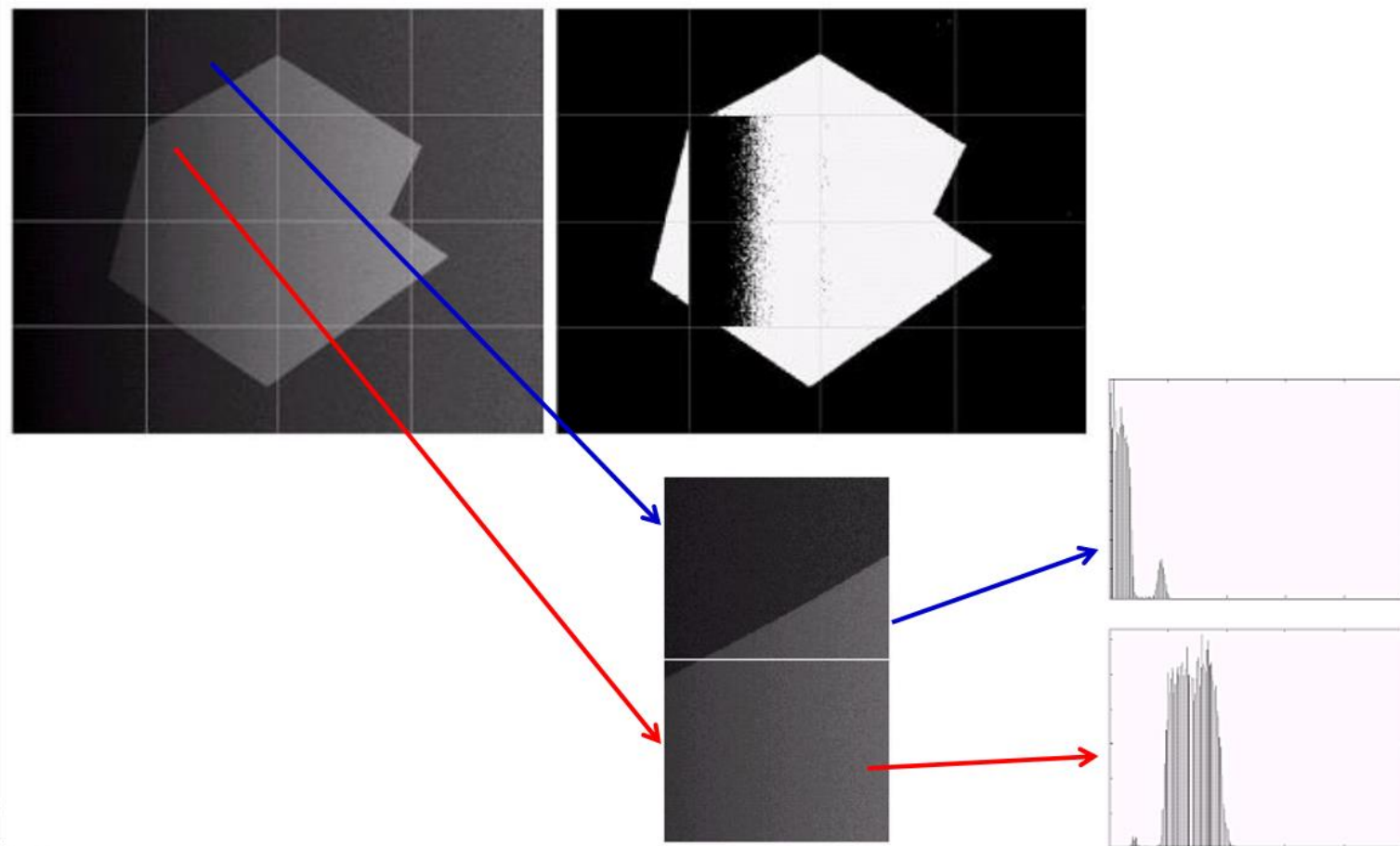
# VẤN ĐỀ VỚI NGƯỠNG TOÀN CỤC

❖ Sử dụng ngưỡng cục bộ. Ví dụ:

- Chia một ảnh thành các ảnh nhỏ và lựa chọn ngưỡng cho từng ảnh
- Xác định kích thước của ảnh nhỏ
- Trước khi phân đoạn, kiểm tra giá trị phương sai để xác định ảnh được chia có ít nhất 2 vùng

Ví dụ: nếu  $\text{variance} < 100$  thì không xác định ngưỡng

# VẤN ĐỀ VỚI NGUỒN TOÀN CỤC





# ĐA NGƯỠNG

❖ Sử dụng  $n$  ngưỡng  $\rightarrow n + 1$  vùng

- If  $f(x, y) < T_1$  then  $f(x, y) = 0$
- If  $f(x, y) \geq T_1 \ \&\& \ f(x, y) < T_2$  then  $f(x, y) = 1$
- ...
- If  $f(x, y) \geq T_n$  then  $f(x, y) = n$

# ĐA NGŨƠNG



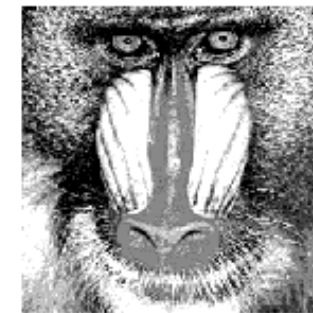
*Baboon (n=1)*



*Baboon (n=2)*



*Baboon (n=3)*



*Baboon (n=4)*



*Cameraman (n=1)*



*Cameraman (n=2)*



*Cameraman (n=3)*



*Cameraman (n=4)*



*Lena (n=1)*



*Lena (n=2)*



*Lena (n=3)*

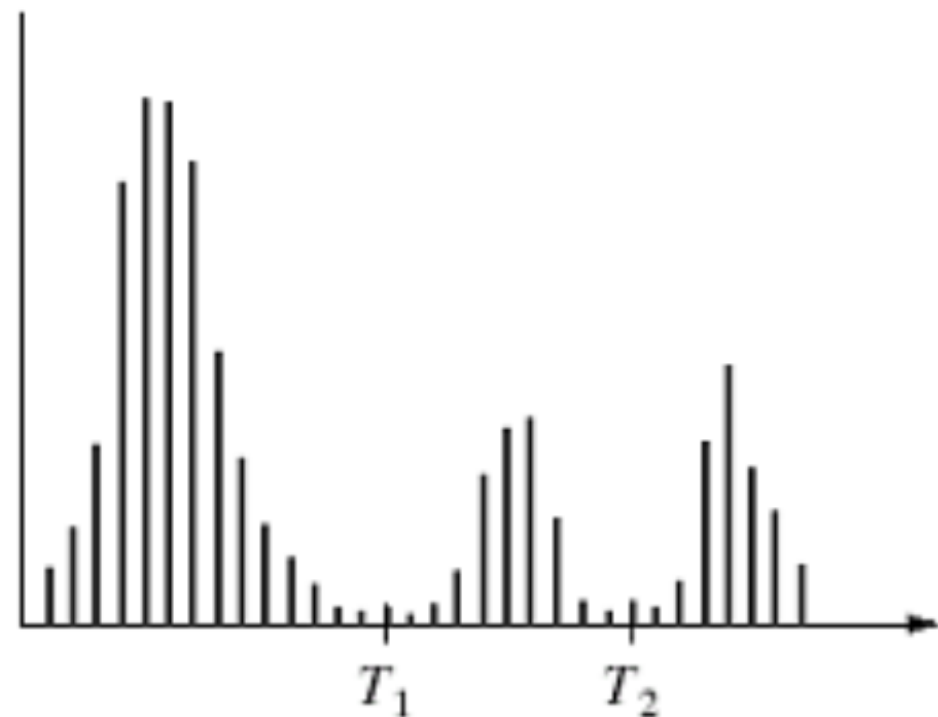


*Lena (n=4)*

Figure 5. The segmentation effect of PSO in the case that thresholds number

# ĐA NGUỒN DỰA TRÊN HISTOGRAM

$$\diamond g(x, y) = \begin{cases} 2 & \text{if } f(x, y) \geq T_2 \\ 1 & \text{if } f(x, y) < T_2 \wedge f(x, y) \geq T_1 \\ 0 & \text{if } f(x, y) < T_1 \end{cases}$$

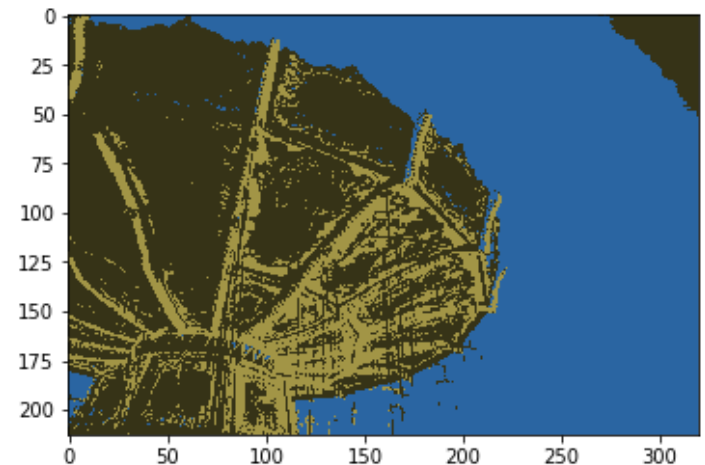
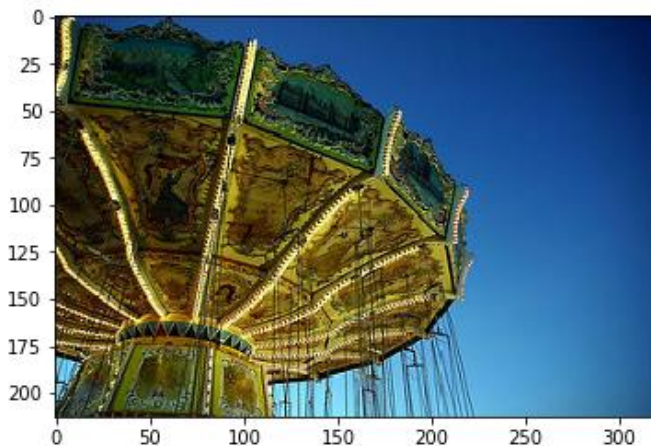


# PHÂN CỤM TRONG KHÔNG GIAN MÀU

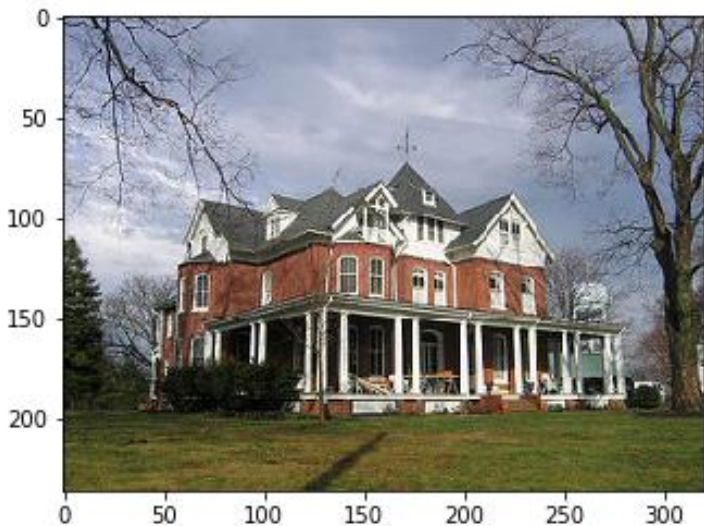
- ❖ Mỗi điểm ảnh được ánh xạ với 1 điểm trong không gian màu

$$\text{Img}(i, j) = (R(i, j), G(i, j), B(i, j))$$

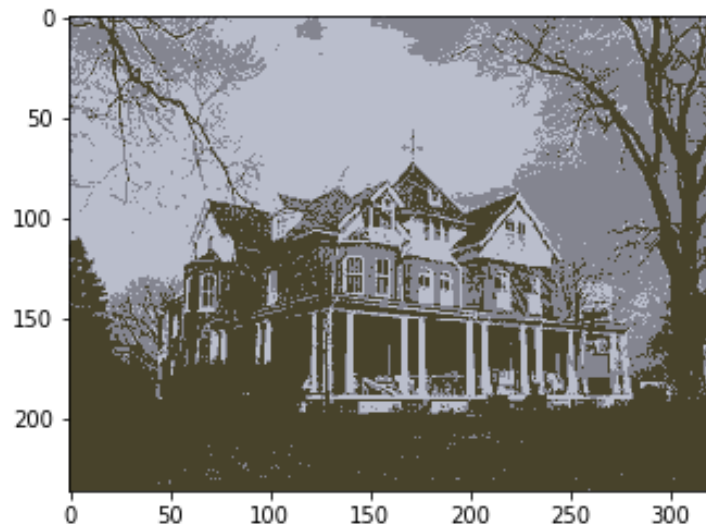
- ❖ Các điểm trong không gian màu sẽ được nhóm theo cụm
- ❖ Các cụm sẽ được ánh xạ trở lại các vùng trong hình ảnh



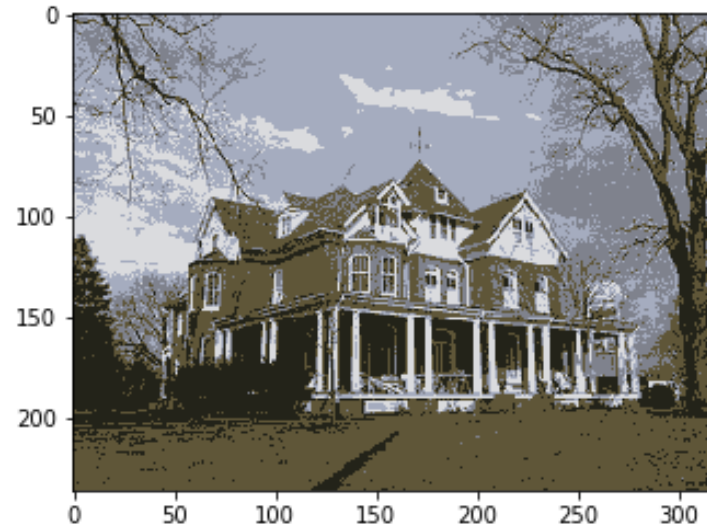
# PHÂN CỤM TRONG KHÔNG GIAN MÀU



**Hình gốc**



**Số cụm: 3**



**Số cụm: 5**

# GIẢI THUẬT K-MEANS

- ❖ Phương pháp để phân đoạn ảnh thành nhiều vùng ( $k$  vùng)
  - Các điểm được phân vào  $k$  nhóm (clusters)
  - Tâm của nhóm: giá trị trung bình của tất cả các phần tử trong nhóm
- ❖ Số lớp  $k$  biết trước và cố định
- ❖ Dựa trên điểm ảnh, không tính đến phân bố điểm ảnh trong không gian

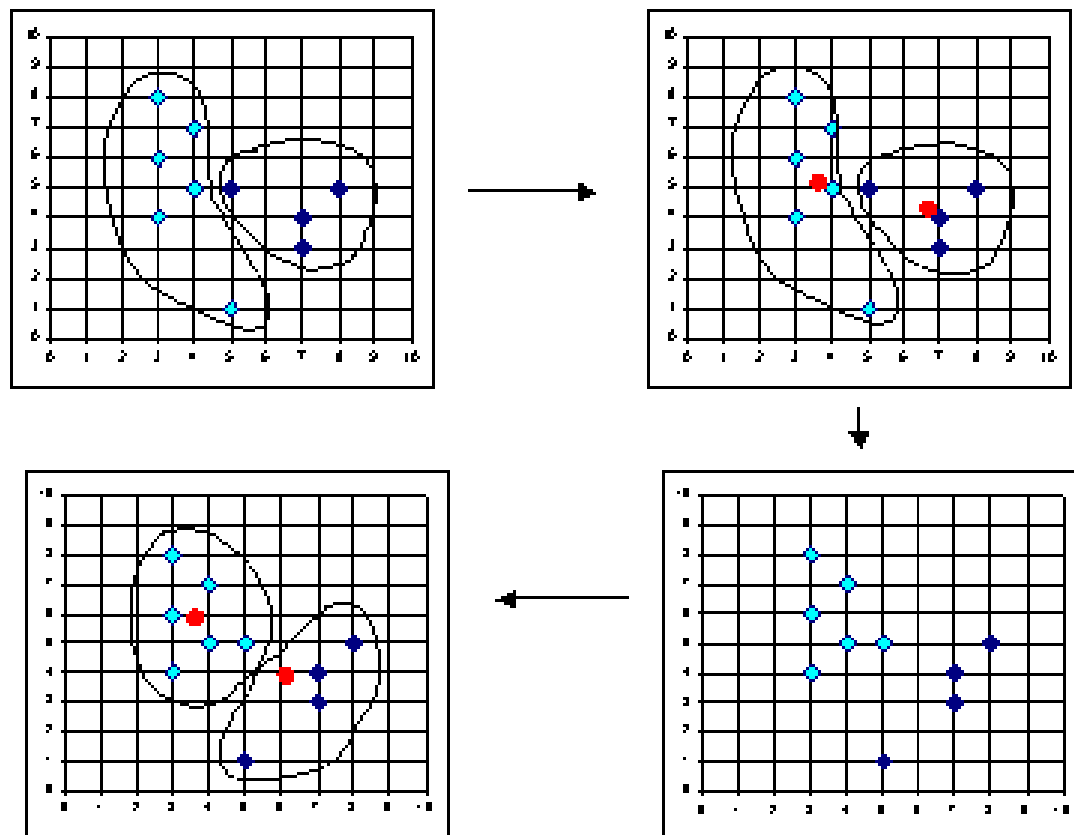
# GIẢI THUẬT K-MEANS

## ❖ 4 bước

- 1. Chia dữ liệu thành  $k$  tập không rỗng
- 2. Tính tâm cho mỗi nhóm
- 3. Sắp xếp lại dữ liệu vào các nhóm gần nhất
- 4. Quay lại 2

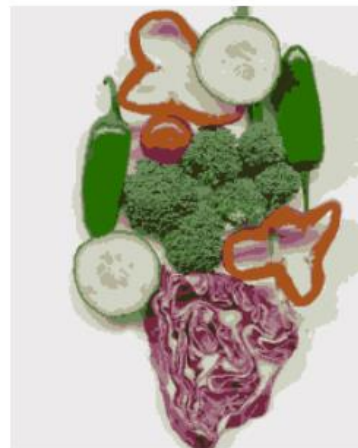
Giải thuật dừng khi tất cả các nhóm ổn định

# GIẢI THUẬT K-MEANS





# GIẢI THUẬT K-MEANS



## PHÂN ĐOẠN DỰA TRÊN ĐIỂM ẢNH

- ❖ Phân đoạn chỉ dựa trên giá trị điểm ảnh (cường độ sáng, màu sắc,...)
- ❖ Không tính đến mối tương quan vị trí trong không gian (2D) của điểm ảnh
- ❖ Các vùng thu được có thể không liên thông
- ❖ Xử lý: xóa bỏ các điểm ảnh rời rạc

# PHÂN ĐOẠN ẢNH DỰA TRÊN VÙNG

- ❖ Dựa trên sự đồng nhất trong các vùng
- ❖ Tính đến phân bố trong không gian của điểm ảnh
- ❖ Theo Pavlidis
  - Một vùng  $S$  là một phân vùng của hình ảnh  $I$  thành một tập hợp các vùng  $S_i$  thỏa mãn
    - ✓  $\cup S_i = S$
    - ✓  $S_i \cap S_j = \emptyset, i \neq j$
    - ✓  $\forall S_i, P(S_i) = \text{true}$
    - ✓  $P(S_i \cup S_j) = \text{false}, i \neq j, S_i$  liền kề  $S_j$

# PHÁT TRIỂN VÙNG (REGION GROWING)

## ❖ Ý tưởng:

- Bắt đầu với tập điểm ảnh và lựa chọn các điểm ảnh lân cận để gộp lại theo tiêu chuẩn nào đó
- Điểm ảnh khởi tạo được gọi là nhân/hạt giống (seed)
- Vùng sẽ được phát triển từ nhân
- Các hiểu biết về ảnh cần phân đoạn được sử dụng để lựa chọn nhân tốt

## ❖ Phát triển vùng phù hợp đối với các ảnh nhiễu khi các cạnh khó được nhận dạng

# PHÁT TRIỂN VÙNG (REGION GROWING)

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

$$|f(seed) - f(neighbors)| \leq 1$$

a	a	d	d	d	d	d
a	a	a	d	d	d	a
b	a	a	a	a	a	a
b	b	a	a	a	a	b
b	b	b	c	b	b	b
b	b	c	b	c	b	b
c	c	c	c	c	b	b

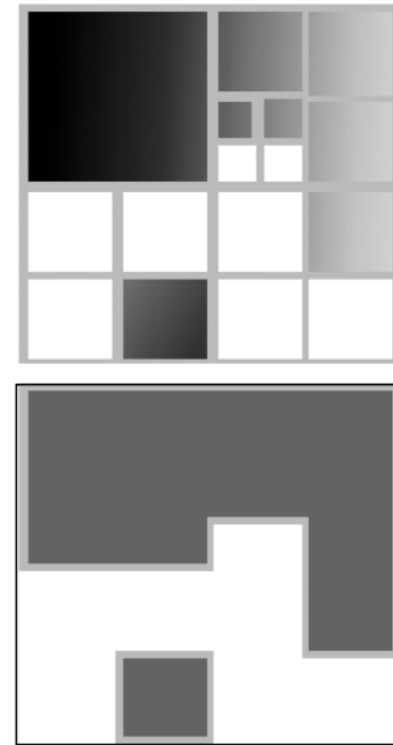
# SPLIT-AND-MERGE (CHIA VÀ HỢP)

## ❖ Bước 1: Chia

- Chia đệ quy các vùng không đồng nhất theo một tiêu chuẩn nào đó (phương sai, max, min) thành các vùng nhỏ hơn
- Chia một vùng thành 4 vùng con
- Tính chất của vùng con được tính toán

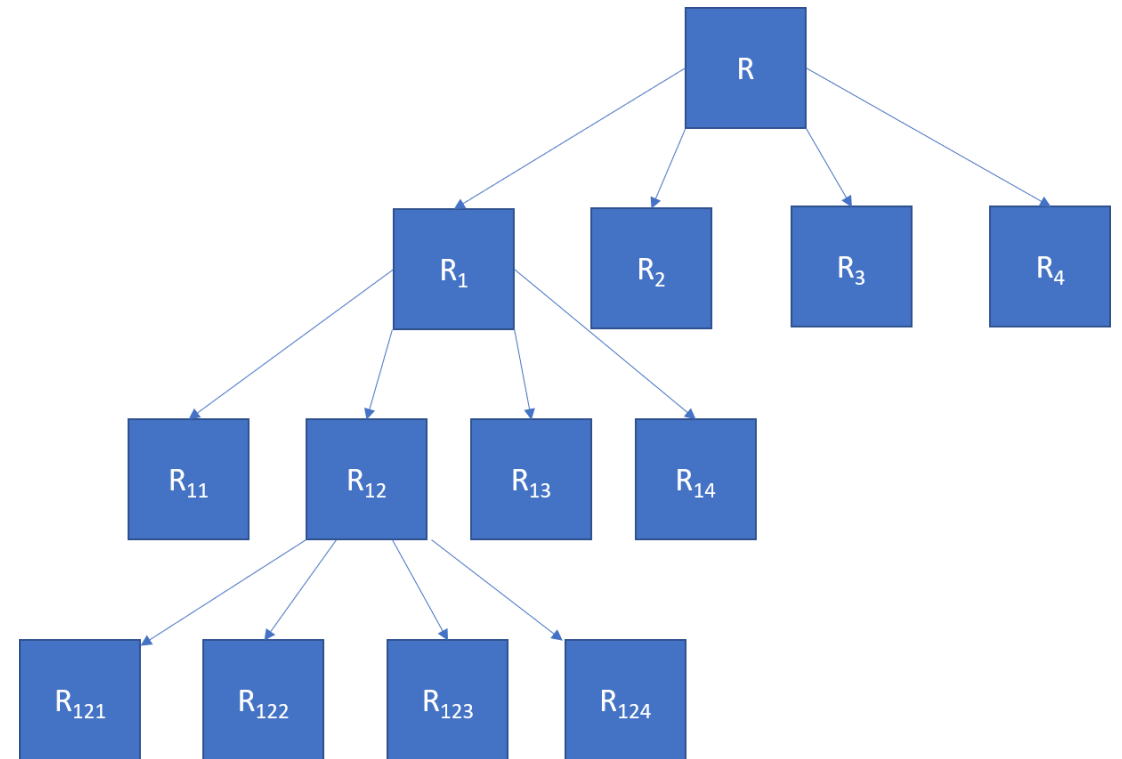
## ❖ Bước 2: Hợp

- Gộp các vùng đồng nhất lân cận nhau theo một tiêu chuẩn nào đó



# SPLIT-AND-MERGE (CHIA VÀ HỢP)

- ❖ Ảnh có thể biểu diễn thành 1 cây
  - Nút gốc là toàn bộ ảnh
- ❖ Mỗi nút  $F$  không đồng nhất được chia thành 4 phần
  - 4 phần sẽ trở thành nút con của  $F$
- ❖ Giải thuật tiếp tục cho đến khi không còn vùng không đồng nhất



# SPLIT-AND-MERGE (CHIA VÀ HỢP)

0	1	0	0	7	7	7	7
1	0	2	2	7	7	7	7
0	2	2	2	7	7	7	7
4	4	2	2	7	7	7	7
0	0	1	1	3	3	7	7
1	1	2	2	3	7	7	7
2	4	3	0	5	7	7	7
2	3	3	5	5	0	7	7

Đầu vào

0	1	0	0	7	7	7	7
1	0	2	2	7	7	7	7
0	2	2	2	7	7	7	7
4	4	2	2	7	7	7	7
0	0	1	1	3	3	7	7
1	1	2	2	3	7	7	7
2	4	3	0	5	7	7	7
2	3	3	5	5	0	7	7

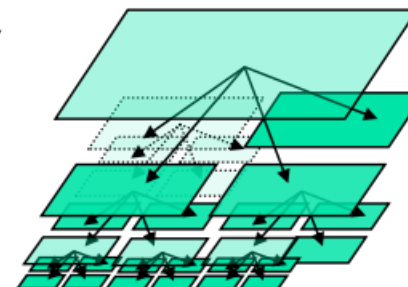
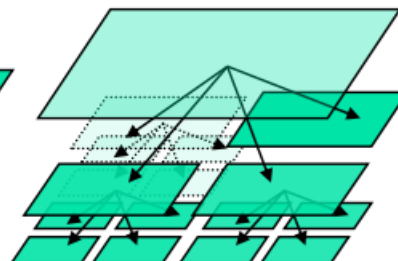
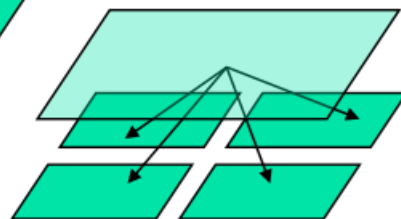
1

0	1	0	0	7	7	7	7
1	0	2	2	7	7	7	7
0	2	2	2	7	7	7	7
4	4	2	2	7	7	7	7
0	0	1	1	3	3	7	7
1	1	2	2	3	7	7	7
2	4	3	0	5	7	7	7
2	3	3	5	5	0	7	7

2

0	1	0	0	7	7	7	7
1	0	2	2	7	7	7	7
0	2	2	2	7	7	7	7
4	4	2	2	7	7	7	7
0	0	1	1	3	3	7	7
1	1	2	2	3	7	7	7
2	4	3	0	5	7	7	7
2	3	3	5	5	0	7	7

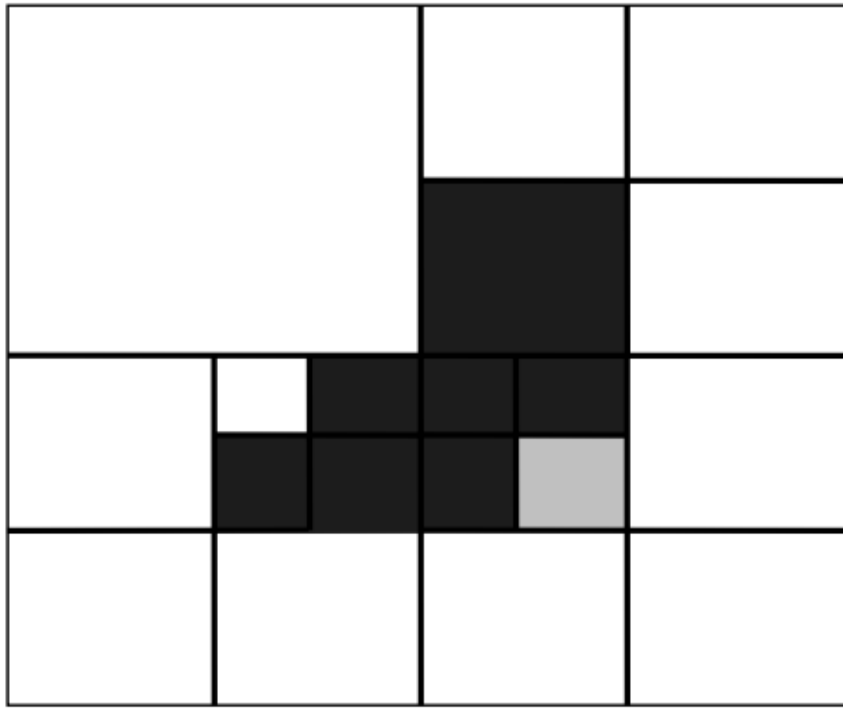
3



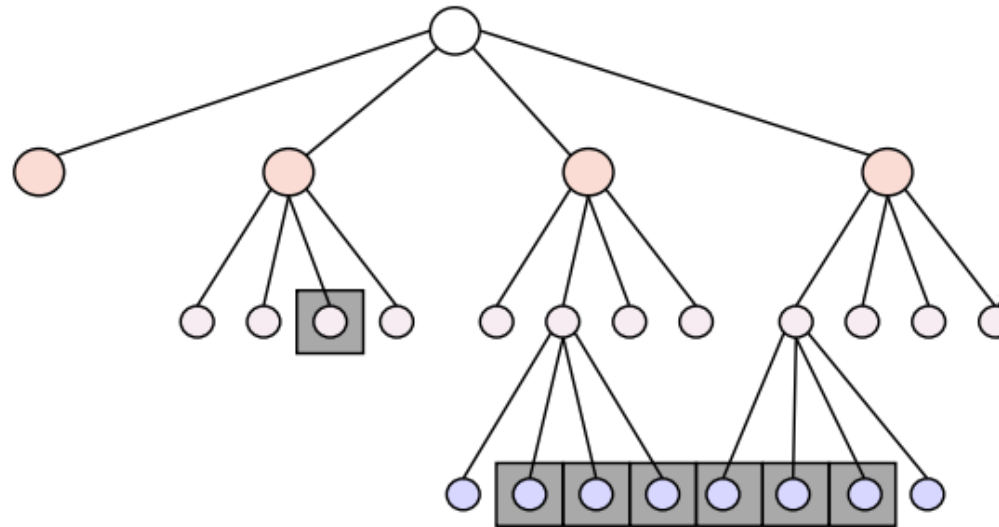


# SPLIT-AND-MERGE (CHIA VÀ HỢP)

*Nối các vùng lân cận đồng nhất*



**Cây tứ phân**



# SPLIT-AND-MERGE (CHIA VÀ HỢP)

1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	0
3	1	4	9	9	8	1	0
1	1	8	8	8	4	1	0
1	1	6	6	6	3	1	0
1	1	5	6	6	3	1	0
1	1	5	6	6	2	1	0
1	1	1	1	1	1	0	0

1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	0
3	1	4	9	9	8	1	0
1	1	8	8	8	4	1	0
1	1	6	6	6	3	1	0
1	1	5	6	6	3	1	0
1	1	5	6	6	2	1	0
1	1	1	1	1	1	0	0

# SPLIT-AND-MERGE (CHIA VÀ HỢP)

1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	0
3	1	4	9	9	8	1	0
1	1	8	8	8	4	1	0
1	1	6	6	6	3	1	0
1	1	5	6	6	3	1	0
1	1	5	6	6	2	1	0
1	1	1	1	1	1	0	0

1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	0
3	1	4	9	9	8	1	0
1	1	8	8	8	4	1	0
1	1	6	6	6	3	1	0
1	1	5	6	6	3	1	0
1	1	5	6	6	2	1	0
1	1	1	1	1	1	0	0

# SPLIT-AND-MERGE (CHIA VÀ HỢP)

1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	0
3	1	4	9	9	8	1	0
1	1	8	8	8	4	1	0
1	1	6	6	6	3	1	0
1	1	5	6	6	3	1	0
1	1	5	6	6	2	1	0
1	1	1	1	1	1	0	0

1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	0
3	1	4	9	9	8	1	0
1	1	8	8	8	4	1	0
1	1	6	6	6	3	1	0
1	1	5	6	6	3	1	0
1	1	5	6	6	2	1	0
1	1	1	1	1	1	0	0

# MỘT SỐ LƯU Ý

- ❖ Tất cả các giải thuật đều cần tiêu chuẩn, tham số
  - Không có tham số chung, phải lựa chọn tùy vào ứng dụng
  - Ngưỡng
    - ✓ Giá trị cố định (hằng số)
    - ✓ Giá trị tương đối (% , trung bình...)
    - ✓ Giá trị được tính toán bằng thuật toán

# MỘT SỐ LƯU Ý

- ❖ Không có thuật toán tối ưu cho bài toán phân đoạn
- ❖ Không có thuật toán chung cho tất cả các bài toán → tùy thuộc ứng dụng
- ❖ Nên thực hiện các phép tiền xử lý để phân đoạn đạt hiệu quả
- ❖ Xác định rõ mục đích của phân đoạn
  - Các phần muốn nhìn thấy trong ảnh
  - Kết quả phân đoạn tăng hiệu quả bài toán truy vấn

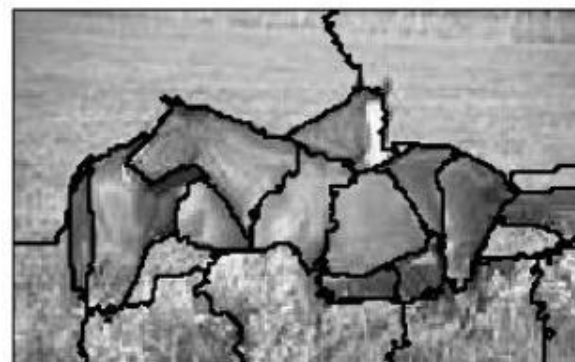
# HẠN CHẾ CỦA VIỆC PHÂN VÙNG

❖ Phân vùng không cho phép xác định tất cả các đối tượng



# HẠN CHẾ CỦA VIỆC PHÂN VÙNG

❖ 1 vùng có thể không chứa 01 đối tượng





# TRUY VẤN DỰA TRÊN SỰ TƯƠNG TỰ (SIMILARITY-BASED RETRIEVAL)

## ❖ Phương pháp đo trực tiếp (direct metric approach)

- Hàm tính khoảng cách được định nghĩa cho các hình ảnh
- Tìm láng giềng ( $k$ ) của hình truy vấn
- Hàm khoảng cách ngây thơ (naïve distance function) cho hình ảnh  $m \times n$ :
  - ✓ L1: tổng các khoảng cách Euclidean theo từng cặp của các pixel RGB
  - ✓ L2: tổng khoảng cách Euclidean trong không gian ( $m \times n \times 3$ )

# TRUY VẤN DỰA TRÊN SỰ TƯƠNG TỰ (SIMILARITY-BASED RETRIEVAL)

## ❖ Phương pháp đo dựa trên đặc trưng (feature-based metrics)

- Sử dụng đặc trưng để giảm kích thước
- Hàm khoảng cách “thực”  $d$  của ảnh/phân đoạn (segment) và hàm khoảng cách  $d'$  của vector đặc trưng được trích xuất phải thỏa

$$d(a, b) < d(a, c) \Rightarrow d'(a, b) < d'(a, c)$$

- Sử dụng chỉ mục để tăng tốc độ truy xuất:
  - ✓ Lập chỉ mục đa chiều (multidimensional indexing) cho các nhóm đặc trưng
  - ✓ Lập chỉ mục ngược (inverted indexing) các đặc trưng riêng
  - ✓ Kết hợp cả 2 cách lập chỉ mục

# TRUY VẤN DỰA TRÊN SỰ TƯƠNG TỰ (SIMILARITY-BASED RETRIEVAL)

## ❖ Cách tiếp cận chuyển đổi:

- Ý tưởng: sự khác biệt tỷ lệ thuận với chi phí tối thiểu để chuyển đổi hình ảnh (phân đoạn) này sang hình ảnh khác
- Chọn hình ảnh ít khác biệt với hình ảnh truy vấn
- Ví dụ về các phép chuyển đổi: dịch chuyển (translation), xoay (rotation), thay đổi kích thước (scaling),... Mỗi phép toán sẽ gắn với một hàm chi phí
- Chi phí chuyển đổi của ảnh sẽ bằng tổng chi phí các phép toán thành phần
- Linh hoạt hơn phương pháp theo phép đo
- Phương pháp dựa trên phép đo hỗ trợ lập chỉ mục tốt hơn

# CẤU TRÚC CSDL ẢNH

❖ Mỗi ảnh có thể xem như từng file riêng

## (a) Biểu diễn quan hệ (relational representation)

- Quan hệ hình ảnh (image relation): id ảnh và thuộc tính mức ảnh (toàn cục)
- Quan hệ đối tượng (object relation): đối tượng (phân đoạn, khung bao) trong ảnh; trích thủ công hoặc tự động

Thuộc tính bao gồm: id ảnh, id đối tượng, tọa độ đối tượng, đặc trưng

- Khái quát: quan hệ xác suất (probabilistic relation)
- Truy vấn: áp dụng kỹ thuật CSDL thông thường sử dụng giá trị đặc trưng trong điều kiện truy vấn

# CẤU TRÚC CSDL ẢNH

## (b) Biểu diễn không gian (spatial representation)

- ❖ Xây dựng cấu trúc cây (như R-tree, R\*-tree) cho tất cả hình ảnh trong CSDL
- ❖ Nút lá chứa tập các đối tượng gần nhau và danh sách trỏ đến ảnh gốc
- ❖ Mỗi phần tử chứa các thuộc tính bổ sung của đối tượng
- ❖ Các chỉ mục riêng biệt có thể được xây dựng cho các thuộc tính không gian khác

# CẤU TRÚC CSDL ẢNH

## ❖ Tổng quát:

- Ở khía cạnh phi không gian, hình ảnh thường được xem là 1 tài liệu (document) và được truy xuất bằng cách sử dụng các kỹ thuật được phát triển để truy xuất thông tin (information retrieval)
- Sử dụng kết hợp các tiêu chí không gian và phi không gian trong truy xuất được thực hiện đơn giản bằng cách kết hợp (hợp, giao...) các danh sách con trở từ các chỉ mục liên quan.