



PHÂN VÙNG ẢNH - ssđf

nhập môn công nghệ thông tin (Đại học Công nghiệp Thực phẩm Thành phố Hồ Chí Minh)



Scan to open on Studocu

PHÂN VÙNG ẢNH

5.1. GIỚI THIỆU

Xử lý ảnh bao gồm các bước: thu nhận ảnh, tiền xử lý, phân đoạn ảnh, biểu diễn và giải thích, nhận dạng và mô tả (tuy nhiên không phải ứng dụng xử lý ảnh nào cũng phải có đầy đủ tất cả các bước trên). Trong các bước đó, bước phân đoạn ảnh là bước quan trọng nhất nhưng và cũng khó khăn nhất. Mặt khác, có nhiều ứng dụng quan trọng cần đến kỹ thuật phân đoạn: ảnh màu hoặc đa mức xám thì cần phải phân ngưỡng; muốn nhận dạng được các đối tượng thì trước hết cần phải phân ảnh thành các vùng khác nhau chứa các đối tượng trong ảnh: tách ra vùng là chữ, số trên bì thư hoặc fax trong phân loại thư, fax tự động; tách ra các vùng là chữ, hình ảnh, bảng biểu ... trong các ứng dụng nhập dữ liệu tự động; tách ra vùng bị bóng để đánh giá phần trăm diện tích bóng trên da; tách ra vùng rừng bị cháy trong ảnh chụp từ máy bay, vệ tinh để phát hiện cháy rừng; tách ra ao, hồ, cây xanh từ ảnh chụp thành phố từ vệ tinh để đánh giá tỷ lệ ao hồ, cây xanh, nhà cửa của thành phố v.v..

Hình dáng của một đối tượng có thể được miêu tả hoặc bởi các tham số của đường biên hoặc các tham số của vùng mà nó chiếm giữ. Sự miêu tả hình dáng dựa trên thông tin đường biên yêu cầu việc phát hiện biên. Sự mô tả hình dáng dựa vào vùng đòi hỏi việc phân đoạn ảnh thành một số vùng đồng nhất. Như vậy, phát hiện biên và phân vùng là hai cách tiếp cận đối ngẫu trong việc phân tích ảnh. Các vùng ảnh yêu cầu phải có các đặc tính đồng nhất (ví dụ như: cường độ, kết cấu) giúp phân biệt được từng vùng. Các đặc tính này tạo nên các vector đặc trưng (feature vectors) để phân biệt một vùng với các vùng khác. Các đặc trưng được sử dụng trong suốt quá trình phân đoạn theo các nguyên tắc kiểm tra tính đồng nhất của vùng.

Chúng ta hãy giả sử rằng một miền ảnh X phải được phân đoạn thành N vùng khác nhau: R_1, \dots, R_N và nguyên tắc phân đoạn là một vị từ của công thức $P(R)$. Cả miền ảnh X và các vùng của nó R_1, \dots, R_N có thể được miêu tả một cách tiện lợi bởi các tập con của plane ảnh (image plane) Z_2 . Việc phân đoạn ảnh chia tập X thành các tập con $R_i, i = 1..N$ có những tính chất sau:

$$X = \bigcup_{i=1}^N R_i$$

$$R_i \cap R_j = \emptyset \text{ với } i \neq j$$

$$P(R_i) = \text{TRUE với } i = 1..N$$

$$P(R_i \cup R_j) = \text{FALSE với } i \neq j$$

Những vùng $R_i, i=1..N$ phải lấp kín hoàn toàn ảnh, nó được thể hiện trong tính chất (5.1). Tính chất (5.2) đảm bảo rằng hai vùng khác nhau là những tập hợp rời nhau. Vị từ $P(R_i)$ phải là TRUE trên mỗi vùng R_i để bảo đảm tính đồng nhất của vùng, có thể được

thấy ở 5.3. Cuối cùng, vị từ $P(R_i \cap R_j)$ ($i \neq j$) phải là FALSE, bởi vì $R_i \cap R_j$, $i \neq j$ tương ứng với một vùng ảnh không đồng nhất.

Công thức của vị từ phân đoạn P và những đặc trưng mà nó sử dụng đóng vai trò quan trọng trong kết quả phân đoạn. P thường là một vị từ của công thức $P(R, x, t)$, trong đó x là vector đặc trưng gắn với một điểm ảnh và t là một tập hợp các tham số (thường là các ngưỡng). Các đặc trưng cấu trúc thường được chứa trong vector đặc trưng. Trong trường hợp đơn giản nhất, vector đặc trưng x chỉ có cường độ ảnh $f(k, l)$ và vector ngưỡng chỉ gồm một ngưỡng T. Một nguyên tắc phân đoạn đơn giản có công thức:

$$P(R): f(k, l) < T \quad (5.5)$$

Trong trường hợp các ảnh màu, vector đặc trưng x có thể là ba thành phần ảnh RGB $[f_R(k, l), f_G(k, l), f_B(k, l)]T$. Một nguyên tắc phân đoạn ảnh đơn giản có thể có công thức:

$$P(R, x, t): ((f_R(k, l) < T_R) \text{ và } (f_G(k, l) < T_G) \text{ và } (f_B(k, l) < T_R)) \quad (5.6)$$

Trong nhiều ứng dụng, sự liên thông của vùng đóng vai trò quan trọng trong việc phân đoạn ảnh. Một vùng R được gọi là liên thông nếu bất kỳ hai điểm (x_A, y_A) và (x_B, y_B) thuộc vào R có thể được nối bởi một đường $(x_A, y_A) \dots (x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1}) \dots (x_B, y_B)$, mà các điểm (x_i, y_i) thuộc vào R và bất kỳ điểm (x_i, y_i) nào đều kề sát với điểm trước (x_{i-1}, y_{i-1}) và điểm tiếp theo (x_{i+1}, y_{i+1}) trên đường đó. Một điểm (x_k, y_k) được gọi là kề với điểm (x_l, y_l) nếu (x_l, y_l) thuộc vào láng giềng trực tiếp của (x_k, y_k) . Chúng ta có thể định nghĩa hai loại láng giềng. 4-láng giềng của một điểm (x, y) là một tập hợp bao gồm láng giềng dọc và láng giềng ngang của nó:

$$N_4((x, y)) = \{(x+1, y), (x-1, y), (x, y+1), (x, y-1)\} \quad (5.7)$$

8-láng giềng của (x, y) là một tập cha của 4-láng giềng và bao gồm láng giềng ngang, dọc và chéo:

$$N_8((x, y)) = N_4((x, y)) \cup \{(x+1, y+1), (x-1, y-1), (x+1, y-1), (x-1, y+1)\} \quad (5.8)$$

Các đường được xác định bằng cách sử dụng 4-láng giềng gồm có các vết dọc và ngang với $\Delta x = \Delta y = 1$. Những đường sử dụng 8-láng giềng có thêm các vết chéo với khoảng cách $\sqrt{2}$.

Các kỹ thuật phân đoạn có thể chia thành ba lớp khác nhau. Các kỹ thuật cục bộ (Local techniques) dựa vào các thuộc tính cục bộ của các điểm và láng giềng của nó. Các kỹ thuật toàn thể (global) phân ảnh dựa trên thông tin chung của toàn bộ ảnh (ví dụ bằng cách sử dụng lược đồ xám của ảnh – image histogram). Các kỹ thuật tách (split), hợp (merge) và growing sử dụng cả khái niệm đồng nhất và gần về hình học. Hai vùng có thể được hợp lại nếu chúng giống nhau (nếu $P(R_i \cup R_j) = \text{TRUE}$) và kề nhau. Một vùng không đồng nhất có thể bị chia thành những vùng nhỏ hơn. Một vùng có thể mở rộng bằng cách thêm các điểm sao cho nó vẫn đồng nhất, $P(R_i) = \text{TRUE}$. Các chương tiếp theo sẽ trình bày một số phương pháp phân đoạn ảnh cụ thể.

5.2. PHÂN VÙNG ẢNH THEO NGUỖNG BIÊN ĐỘ

Trong hầu hết các trường hợp, ngưỡng được chọn từ lược đồ độ sáng của vùng hay ảnh cần phân đoạn. Có rất nhiều kỹ thuật chọn ngưỡng tự động xuất phát từ lược đồ xám $\{h[b] \mid b = 0, 1, \dots, B-1\}$ đã được đưa ra, B là số mức xám của ảnh, với ảnh grayscale

B=256. Những kỹ thuật phổ biến sẽ được trình bày dưới đây. Những kỹ thuật này có thể tận dụng những lợi thế do sự làm tròn dữ liệu lược đồ ban đầu mang lại nhằm loại bỏ những dao động nhỏ về độ sáng. Tuy nhiên các thuật toán làm tròn cần phải cẩn thận, không được làm dịch chuyển các vị trí đỉnh của lược đồ. Nhận xét này dẫn đến thuật toán làm tròn dưới đây:

$$h_{\text{CNOOTH}}[b] = \frac{1}{W} \sum_{w=-(M-1)/2}^{(M-1)/2} h_{\text{raw}}[b - w] \quad (5.9)$$

Trong đó,

- W thường được chọn là số lẻ (3 hoặc 5).
- $h_{\text{smooth}}[b]$ là giá trị tần xuất của độ sáng b sau khi làm tròn
- $h_{\text{raw}}[b]$ là giá trị tần xuất của độ sáng b gốc

5.2.1. Thuật toán đẳng liệu

Đây là kỹ thuật chọn ngưỡng theo kiểu lặp do Ridler và Calvard đưa ra.

- Trước hết, lược đồ sẽ được phân đoạn thành hai phần bằng một giá trị ngưỡng khởi động với $\Theta_0 = B/2$ tức là bằng phân nửa thang độ xám động của ảnh.
- Sau đó, các trung bình mẫu ($m_{f,0}$) của những điểm ảnh thuộc đối tượng và ($m_{b,0}$) của những điểm ảnh nền sẽ được tính toán.
- Một giá trị ngưỡng mới Θ_1 sẽ được tính kế đó bằng cách lấy giá trị trung bình của hai trung bình mẫu nói trên.
- Quá trình này cứ thế sẽ được tiếp tục với ngưỡng mới cho đến khi nào giá trị ngưỡng không thay đổi nữa thì dừng lại.

Nếu biểu diễn dưới dạng công thức toán học, chúng ta có:

$$N = \frac{\sum_{b=B_{\text{min}}}^{B_{\text{max}}} b \times h(b)}{\sum_{b=B_{\text{min}}}^{B_{\text{max}}} h(b)}$$

với

- $m_{f,k+1}$ thì $B_{\text{min}} = 0$ và $B_{\text{max}} = \theta_k$
- $m_{b,k+1}$ thì $B_{\text{min}} = \theta_k$ và $B_{\text{max}} = B - 1$

$$\theta_{k+1} = \frac{N_{f,k} + N_{b,k}}{2}$$

Thực hiện cho tới khi $\theta_k = \theta_{k+1}$

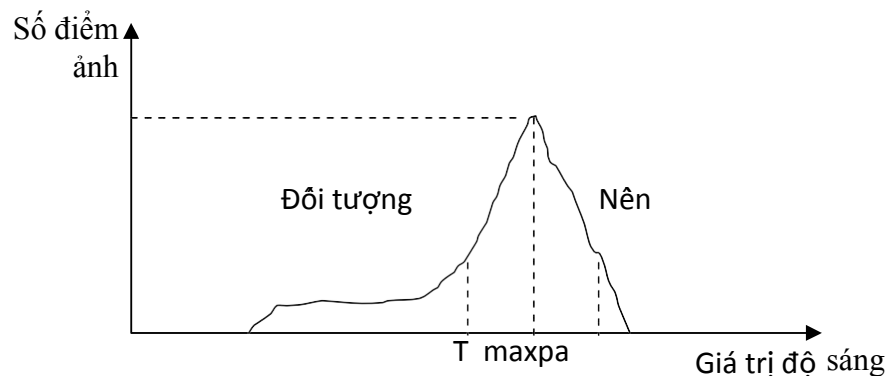
(5.10)

5.2.2. Thuật toán đối xứng nền

Kỹ thuật này dựa trên sự giả định là tồn tại hai đỉnh phân biệt trong lược đồ nằm đối xứng nhau qua đỉnh có giá trị lớn nhất trong phần lược đồ thuộc về các điểm ảnh nền. Kỹ thuật này có thể tận dụng ưu điểm của việc làm trơn được mô tả trong phương trình (5.9). Đỉnh cực đại $maxp$ tìm được nhờ tiến hành tìm giá trị cực đại trong lược đồ. Sau đó thuật toán sẽ được áp dụng ở phía không phải là điểm ảnh thuộc đối tượng ứng với giá trị cực đại đó nhằm tìm ra giá trị độ sáng a ứng với giá trị phần trăm $p\%$ mà: $P(a) = p\%$, trong đó $P(a)$ là hàm phân phối xác suất về độ sáng.

Định nghĩa: [Hàm phân phối xác suất về độ sáng]

Hàm phân phối xác suất $P(a)$ thể hiện xác suất chọn được một giá trị độ sáng từ một vùng ảnh cho trước, sao cho giá trị này không vượt quá một giá trị sáng cho trước a . Khi a biến thiên từ $-\infty$ đến $+\infty$, $P(a)$ sẽ nhận các giá trị từ 0 đến 1. $P(a)$ là hàm đơn điệu không giảm theo a , do vậy $\frac{dP}{da} \geq 0$.



Hình 5.1 Minh hoạ thuật toán đối xứng nền

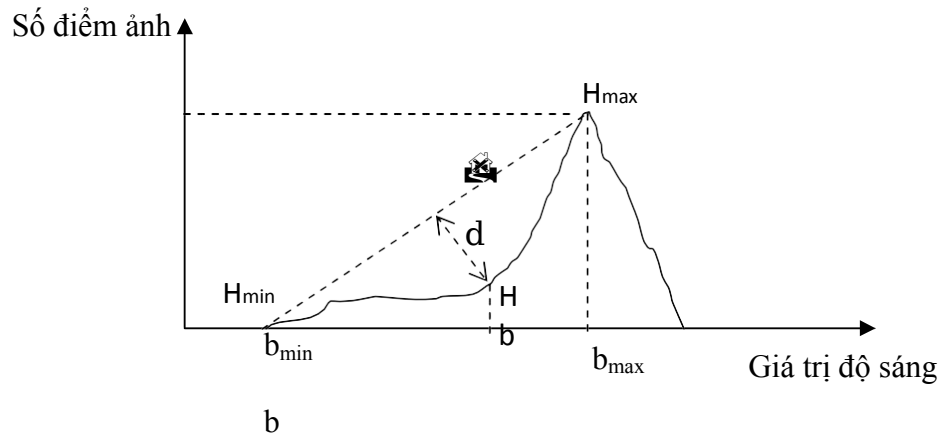
Ở đây ta đang giả thiết là ảnh có các đối tượng tối trên nền sáng. Giả sử độ chắc chắn là 95%, thì có nghĩa là ta phải ở bên phải đỉnh $maxp$ một giá trị a sao cho $P(a)=95\%$. Do tính đối xứng đã giả định ở trên, chúng ta lấy đối xứng qua $maxp$ để có được ngưỡng T :

$$T = maxp - (a - maxp) \quad (5.11)$$

Kỹ thuật này thực hiện ngược lại với tình huống ảnh có các đối tượng sáng trên một nền tối.

5.2.3. Thuật toán tam giác

Thuật toán này do Zack đề xuất. Hình 5.2 minh họa thuật toán này. Các bước của thuật toán như sau: Nối đường thẳng Δ từ điểm H_{\max} của lược đồ (điểm có histogram lớn nhất, có mức xám b_{\max}) đến điểm H_{\min} của lược đồ (điểm ứng với độ sáng nhỏ nhất b_{\min}). Với mỗi độ sáng b trong khoảng $[b_{\min}, b_{\max}]$, chúng ta đi tính khoảng cách d từ điểm H_b của lược đồ (ứng với giá trị độ sáng b) đến Δ . Giá trị b_0 ứng với khoảng cách d lớn nhất sẽ được chọn làm giá trị ngưỡng T . Kỹ thuật này đặc biệt hiệu quả khi các điểm ảnh thuộc đối tượng tạo nên một đỉnh yếu trong lược đồ ảnh.



Hình 5.2. Minh họa thuật toán tam giác

5.2.4. Chọn ngưỡng đối với Bimodal Histogram

Nếu ảnh chứa một đối tượng và một nền có cường độ đồng nhất, nó thường có một histogram hai mode (bimodal histogram) như một trường hợp trình bày trong hình 5.3.

Ngưỡng T được chọn ở tại vị trí cực tiểu địa phương của histogram nằm giữa hai đỉnh của histogram. Điểm cực đại địa phương của histogram có thể dễ dàng được phát hiện bằng cách sử dụng biến đổi chóp mũ (top hat) do Meyer đưa ra: Phụ thuộc vào tình huống chúng ta đang phải làm việc là với nhưng đối tượng sáng trên nền tối hay đối tượng tối trên nền sáng mà phép biến đổi top hat sẽ có một trong hai dạng sau:

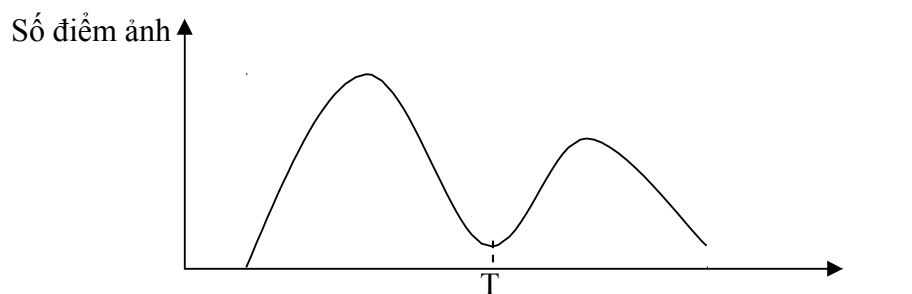
a) Các đối tượng sáng:

$$TopHat(A, B) = A - (A \circ B) = A - \max_B(\min_B(A)) \quad (5.12)$$

b) Các đối tượng tối:

$$TopHat(A, B) = A - (A \circ B) = A - \min_B(\max_B(A)) \quad (5.13)$$

Việc tính toán giá trị cực tiểu địa phương của histogram thì khó nếu histogram nhiều. Do đó, trong trường hợp này nên làm trơn histogram, ví dụ sử dụng thuật toán (5.9).



Hình 5.3. Bimodal histogram

Trong một số ứng dụng nhất định, cường độ của đối tượng hay nền thay đổi khá chậm. Trong trường hợp này, histogram ảnh có thể không chứa hai thùy phân biệt rõ ràng, vì vậy có thể phải dùng ngưỡng thay đổi theo không gian. Hình ảnh được chia thành những khối hình vuông, histogram và ngưỡng được tính cho mỗi khối tương ứng. Nếu histogram cục bộ không phải là bimodal histogram thì ngưỡng được tính bằng cách nội suy ngưỡng của các khối láng giềng. Khi ngưỡng cục bộ đã có thì áp dụng thuật toán phân ngưỡng ở hình 5.3 cho khối này.

5.3. PHÂN VÙNG THEO MIỀN ĐỒNG NHẤT

5.3.1. Giới thiệu

Kỹ thuật phân đoạn ảnh thành các miền đồng nhất dựa vào các thuộc tính quan trọng nào đó của miền. Mỗi một thuộc tính khi sử dụng thì có một tiêu chuẩn phân đoạn tương ứng. Một số thuộc tính tiêu biểu là: mức xám, màu sắc (đối với ảnh màu), kết cấu sợi...

Ta có thể dùng logic vị từ để làm tiêu chuẩn đánh giá phân đoạn. Giả sử ảnh X phải phân thành n vùng khác nhau: Z_1, Z_2, \dots, Z_n và logic vị từ có dạng $P(Z)$. Việc phân vùng phải thoả mãn các tính chất sau:

Kết quả của việc phân đoạn ảnh phụ thuộc vào dạng của vị từ P và các đặc tính biểu diễn bởi vector đặc tính. Thường vị từ P có dạng $P(Z, X, t)$, với X là véc tơ đặc tính, t là ngưỡng. Trường hợp đơn giản nhất, véc tơ đặc tính chỉ chứa giá trị mức xám của ảnh $I(k, t)$ và ngưỡng chỉ đơn thuần là giá trị T.

$$P(Z): I(k, I) < T. \quad (5.15)$$

Với ảnh màu, véc tơ đặc tính X có thể là thành phần ba màu R, G, B và $I_R(k, l)$, $I_G(k, l)$, $I_B(k, l)$ là các thành phần tương ứng. Lúc đó luật phân ngưỡng có dạng:

$$P(Z, X, t): I_R(k, l) < T_R \text{ và } I_G(k, l) < T_G \text{ và } I_B(k, l) < T_B \quad (5.16)$$

Có ba cách tiếp cận chủ yếu trong phân vùng ảnh theo miền đồng nhất và độc lập với tiêu chuẩn lựa chọn tính đồng nhất:

- Phương pháp phân tách – cây tứ phân (split – quad trees)
- Phương pháp hợp (merge).
- Phương pháp tách - hợp (split – merge).

5.3.2. Phương pháp tách cây tứ phân (quad tree)

Về nguyên tắc phương pháp này kiểm tra tính hợp thức của tiêu chuẩn đồng nhất một cách tổng thể trên miền lớn. Nếu tiêu chuẩn được thỏa việc phân đoạn coi như kết thúc. Trong trường hợp ngược lại ta chia miền đang xét thành 4 miền nhỏ hơn, ta lại áp dụng đệ quy bằng phương pháp trên cho mỗi miền nhỏ hơn cho đến khi tất cả các miền đều thỏa mãn. Phương pháp này có thể mô tả bằng thuật toán sau:

```

Procedure PhanDoan(Mien)
Begin
  If miền đang xét không thỏa Then
    Begin
      Chia miền đang xét thành 4 miền:  $Z_1, Z_2, Z_3, Z_4$ 
      For i=1 to 4 Do PhanDoan( $Z_i$ )
    End
  Else Exit
End;
```

Nếu miền đang xét không thỏa mãn tiêu chuẩn thì không phải là miền đồng nhất và sẽ được chia thành 4 phần.

Thuật toán kiểm tra tiêu chuẩn theo (5.19) có thể được viết:

```

Function Examin_Criteria(I, N1, M1, N2, M2, T)
/* Giả sử là ảnh có tối đa 255 mức xám.
(N1, M1), (N2, M2) là tọa độ điểm đầu và điểm cuối của miền, T là ngưỡng. */
Begin
  1. Max = 0; Min = 255;
  2. For i = N1 to N2 do
    For j = M1 to M2 do
      If  $I(i, j) < Min$  Then  $Min = I(i, j)$ ;
      If  $I(i, j) > Max$  Then  $Max = I(i, j)$ ;
    End
  3. If  $ABS(max - min) < T$  Then Examin_Criteria = 0
  Else Examin_Criteria = 1;
```


End

Thuật toán kiểm tra tiêu chuẩn theo (5.20) được viết:

```
Fuction Examin_Criteria(I,N1,M1,N2, M2, T)
/* Giả sử là ảnh có tối đa 255 mức xám.
(N1,M1),(N2, M2) là toạ độ điểm đầu và điểm cuối của miền, T là ngưỡng.*/
Begin
1. xtb = 0; e=0; n = (N2-N1)*(M2-M1)
2. For i = N1 to N2 do
    For j = N1 to N2 do
        xtb = xtb + I(i,j)/n
3. For i = N1 to N2 do
    For j = N1 to N2 do
        e = e + sqr(I(i,j)-xtb)/n
4. If e < T Then Examin_Criteria = 0
    Else Examin_Criteria = 1;
End
```

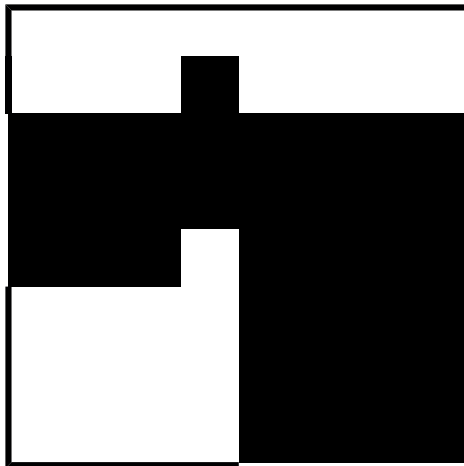
Nếu hàm trả về giá trị 0, có nghĩa vùng đang xét là đồng nhất, trường hợp ngược lại nghĩa là mức không đồng nhất. Trong giải thuật trên, khi miền là đồng nhất cần tính lại giá trị trung bình và cập nhật lại ảnh đầu ra. Giá trị trung bình được tính bởi:

Tổng giá trị mức xám/ tổng số điểm.

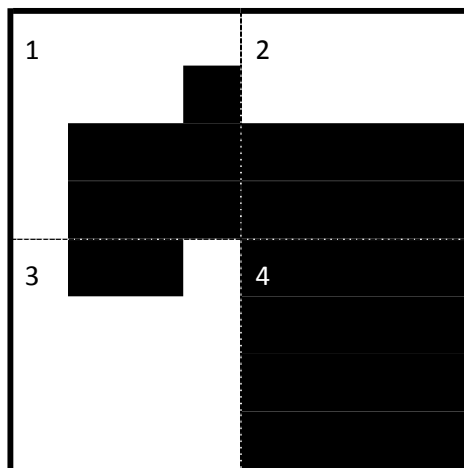
Thuật toán này tạo nên một cây mà mỗi nút cha có 4 nút con ở mọi mức, trừ mức ngoài cùng. Vì thế cây này có tên là cây tứ phân.

Một vùng thoả tiêu chuẩn tạo nên một nút lá, nếu không sẽ tạo nên một nút trong và 4 nút con tương ứng của việc chia làm 4 vùng. Mỗi nút lá của cây biểu diễn một vùng đã phân chia theo tiêu chuẩn.

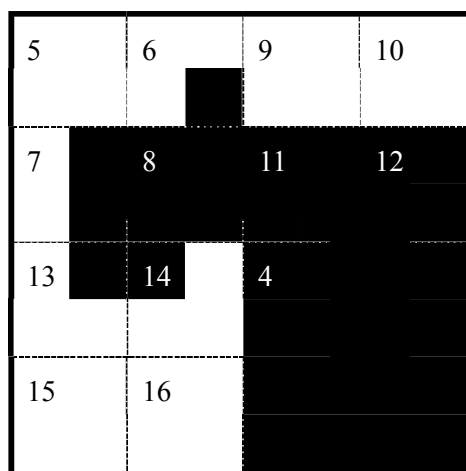
Hình 5.4 a-e minh họa thuật toán tách cây tứ phân: ảnh gốc (a) được chia thành 4 phần được kết quả phân mức 1 (b), tiếp tục thực hiện đối với các phần nhỏ, ta được phân mức 2, 3.



a) Ảnh gốc



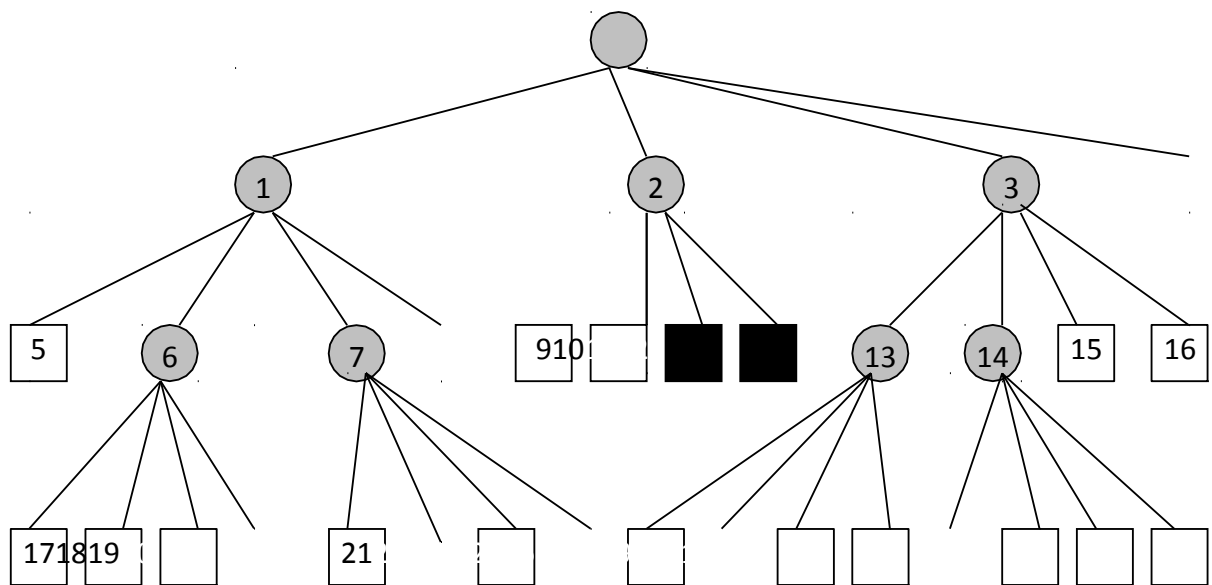
b) Phân mức 1



c) Phân mức 2

5	17	18	9	10
	19	20		
21	22	8	11	12
23	24			
25	26	29	30	4
27	28	31	32	
15	16			

d) Phân mức 3



e) Cây tương ứng

Hình 5.4. Phương pháp tách cây tứ phân

Tiêu chuẩn phân vùng ở đây là màu sắc. Nếu mọi điểm của vùng là màu trắng sẽ tạo nên một nút lá trắng và tương tự như vậy với nút lá đen. Nút màu ghi có nghĩa là vùng không thuần nhất và phải tiếp tục chia.

5.3.3. Các phương pháp phân vùng bởi hợp

Ý tưởng của phương pháp này là xem xét ảnh từ các miền nhỏ nhất rồi hợp chúng lại nếu thỏa mãn tiêu chuẩn để được một miền đồng nhất lớn hơn. Ta lại tiếp tục với miền thu được cho đến khi không thể hợp được nữa. Số miền còn lại cho ta kết quả phân đoạn ảnh. Miền nhỏ nhất của bước xuất phát là điểm ảnh. Việc hợp 2 vùng được thực hiện theo nguyên tắc sau:

- Hai vùng phải đáp ứng tiêu chuẩn, như cùng màu hay cùng mức xám.
- Chúng phải kề cận nhau

Trong xử lý ảnh người ta dùng khái niệm liên thông để xác định kề cận. Với 4 liên thông một điểm ảnh $I(x,y)$ sẽ có 4 kề cận theo hướng x, y. Trong khi đó với 8 liên thông, điểm ảnh $I(x,y)$ sẽ có 4 liên thông theo hai hướng x và y và 4 liên thông khác theo hướng chéo 45° .

5.3.4. Phương pháp tách hợp (split-merge)

Hai phương pháp vừa xét ở trên có một số nhược điểm.

- Phương pháp tách tạo nên một cấu trúc phân cấp và thiết lập mối quan hệ giữa các vùng. Tuy nhiên nó thực hiện việc chia quá chi tiết.
- Phương pháp hợp cho phép làm giảm số miền liên thông xuống tối thiểu nhưng cấu trúc hàng ngang dần trái, không cho ta thấy mối liên hệ giữa các vùng.

Chính vì nhược điểm này mà ta nghĩ đến phương pháp phối hợp cả 2 phương pháp. Trước tiên dùng phương pháp tách để tạo nên cây tứ phân, phân đoạn theo hướng từ gốc đến lá. Tiếp theo tiến hành duyệt cây theo chiều ngược lại và hợp các vùng có cùng tiêu chuẩn. Với phương pháp này ta thu được miêu tả cấu trúc của ảnh với các miền liên thông có kích thước tối đa.

Giải thuật trên gồm một số bước sau:

1. Kiểm tra tiêu chuẩn đồng nhất
 - i. Nếu không thỏa và số điểm trong vùng lớn hơn một điểm, tách làm 4 vùng (trên, dưới, trái, phải) bằng cách gọi đệ quy. Nếu kết quả tách xong và không tách được nữa chuyển sang bước ii.
 - ii. Nếu tiêu chuẩn đồng nhất là thỏa thì tiến hành hợp vùng và cập nhật giá trị trung bình cho vùng.
2. Hợp vùng Cần kiểm tra 4 lân cận đã nêu trên. Có thể có nhiều vùng thỏa mãn khi đó ta chọn vùng tối ưu rồi tiến hành hợp.

Phương pháp này thu được kết quả số vùng là nhỏ hơn phương pháp tách và ảnh được làm trơn hơn.

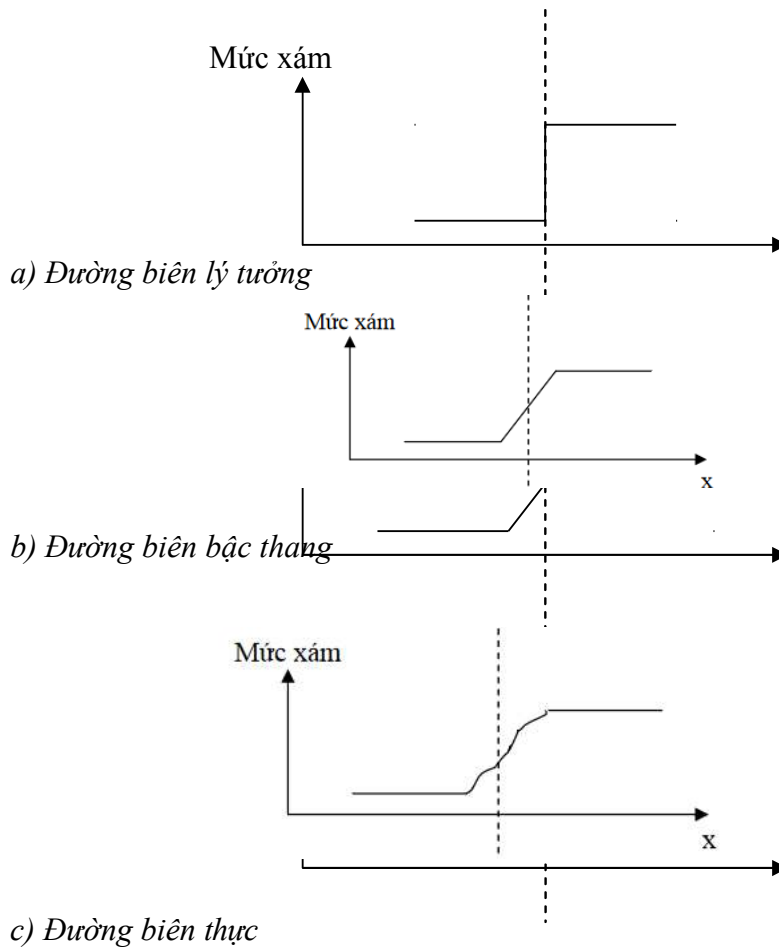
5.4. PHÂN VÙNG DỰA THEO ĐƯỜNG BIÊN

Việc phân đoạn ảnh dựa vào biên được tiến hành qua một số bước:

- Phát hiện biên và làm nổi biên
- Làm mảnh biên
- Nhị phân hoá đường biên
- Miêu tả đường biên

5.4.1. Phát hiện biên

Về mặt toán học, người ta coi điểm biên của ảnh là điểm có sự biến đổi đột ngột về độ xám, như trong hình 5.5 sau đây:



Hình 5.5. Đường biên của ảnh

Có nhiều định nghĩa về đường biên, điển hình có ba loại đường biên chính:

- Đường biên lý tưởng (Hình 5.5 a) được định nghĩa là sự thay đổi giá trị cấp xám tại một vị trí xác định. Vị trí của đường biên chính là vị trí thay đổi cấp xám.
- Đường biên bậc thang (Hình 5.5 b) xuất hiện khi sự thay đổi cấp xám trải rộng qua nhiều điểm ảnh. Vị trí của đường biên được xem như vị trí chính giữa của
- đường nối giữa cấp xám thấp và cấp xám cao.
- Đường biên thực (Hình 5.5 c): Đó là sự thay đổi cấp xám tại nhiều điểm nhưng không trơn.

Định nghĩa toán học của biên ở trên là cơ sở cho các kỹ thuật phát hiện biên. Điểm quan trọng là biến thiên giữa các điểm ảnh thường là nhỏ, trong khi đó biến thiên độ sáng của điểm biên (khi qua biên) lại khá lớn.

5.4.2. Làm mảnh biên

Làm mảnh biên là việc làm nổi biên với độ rộng chỉ 1 pixel. Trong phần trên, ta thấy, kỹ thuật Laplace dùng trong việc phát hiện biên cho kết quả trực tiếp biên ảnh với độ rộng 1 pixel. Còn với các kỹ thuật khác, ví dụ như kỹ thuật Gradient thì không như vậy, ta phải có bước làm mảnh biên để thu được biên với độ rộng 1 pixel.

Khi thực hiện đạo hàm một ảnh, ta thu được những điểm cực trị cục bộ. Theo kỹ thuật Gradient, những điểm cực trị cục bộ có thể coi như biên. Do vậy cần tách biệt những điểm cực trị đó để xác định chính xác biên ảnh và để giảm độ rộng biên ảnh. Một trong những phương pháp hay dùng đó là phương pháp “Loại bỏ các điểm không cực đại”.

5.4.3. Nhị phân hóa đường biên

Nhị phân hóa đường biên là giai đoạn then chốt trong quá trình trích chọn vì nó xác định đường bao nào thực sự cần và đường bao nào có thể loại bỏ. Nói chung, người ta thường nhị phân hóa đường biên theo cách thức làm giảm nhiễu hoặc tránh hiện tượng kéo sợi trên ảnh. Điều này cũng giải thích tại sao phân đoạn dựa theo biên có hiệu quả khi ảnh có độ tương phản tốt. Trong trường hợp ngược lại, có thể sẽ bị mất một phần đường bao hay đường bao có chân, không khép kín, v.v., do đó sẽ bất lợi cho biểu diễn sau này. Một phương pháp hay được dùng là chọn ngưỡng thích nghi. Với cách chọn này, ngưỡng sẽ phụ thuộc vào hướng của gradient nhằm làm giảm sự xoắn của biên. Đầu tiên, người ta định ra một ngưỡng nào đó và sau đó sử dụng một hệ số sinh thích nghi thông qua lời giải toán tử đạo hàm theo hướng tìm được để tinh chỉnh.

5.4.4. Miêu tả đường biên

Khi đã có bản đồ biên ảnh, ta cần phải biểu diễn nó dưới dạng thích hợp phục vụ cho việc phân tích và làm giảm lượng thông tin dùng để miêu tả, lưu trữ đối tượng. Người ta thường thực hiện theo nguyên tắc: tách riêng từng biên và gán cho mỗi biên một mã.

Có nhiều phương pháp miêu tả đường biên khác nhau, mỗi phương pháp thích hợp với một loại ứng dụng. Việc tách các đường bao có thể phải bổ xung thêm các điều kiện nhằm loại bỏ các đường bao không khép kín, hoặc bỏ đi các chân rết bám theo các đường bao kín.

Có nhiều cách mã hóa đường bao, có thể biểu diễn chính xác đường bao hay xấp xỉ nhờ nội suy. Một số cấu trúc cơ sở mã hóa đường bao thường dùng là: điểm, đoạn thẳng, cung, đường cong. Có một số phương pháp mã hóa đường bao hay dùng như: mã hóa theo tọa độ Đề các, mã hóa Freeman, xấp xỉ bởi đoạn thẳng, xấp xỉ đa thức.

5.5. PHÂN VÙNG THEO KẾT CẤU BỀ MẶT

Kết cấu thường được nhận biết trên bề mặt của các đối tượng như gỗ, cát, vải

vóc... Kết cấu là thuật ngữ phản ánh sự lặp lại của các phần tử sợi (texel) cơ bản. Sự lặp lại này có thể ngẫu nhiên hay có tính chu kỳ hoặc gần chu kỳ. Một texel chứa rất nhiều điểm ảnh. Trong phân tích ảnh, kết cấu được chia làm hai loại chính là: loại thống kê và loại cấu trúc.

5.4.1. Phương pháp thống kê

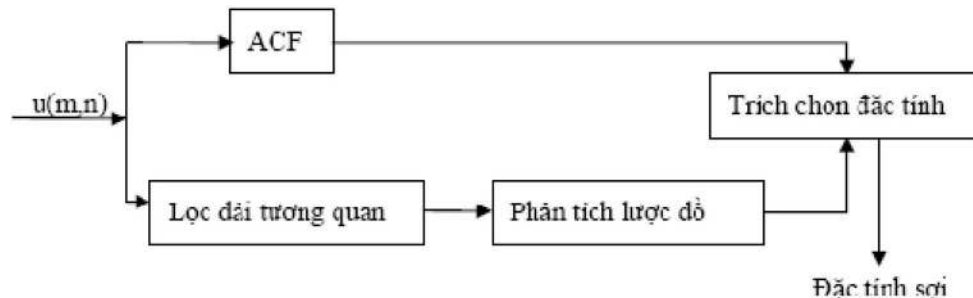
Tính kết cấu ngẫu nhiên rất phù hợp với các đặc trưng thống kê. Vì vậy, người ta có thể dùng các đặc trưng ngẫu nhiên để đo nó như: Hàm tự tương quan (AutoCorrelation Function- ACF), các biến đổi mật độ gờ, ma trận tương tranh,...

Theo cách tiếp cận bằng hàm tự tương quan, độ thô của kết cấu sợi tỉ lệ với độ rộng của ACF, được biểu diễn bởi khoảng cách x_0, y_0 sao cho $r(x_0, 0) = r(0, y_0) = 1$. Người ta cũng dùng cách đo nhánh của ACF nhờ hàm khởi sinh moment:

$$M(k, l) = \sum_m \sum_n (m - \mu_1)^k (n - \mu_2)^l r(m, n) \quad (5.25)$$

với: $\mu_1 = \sum_m \sum_n m r(m, n)$ và $\mu_2 = \sum_m \sum_n n r(m, n)$

Các đặc trưng của kết cấu sợi như độ thô, độ mịn hay hướng có thể ước lượng nhờ các biến đổi ảnh bằng kỹ thuật lọc tuyến tính. Một mô hình đơn giản trong trường hợp ngẫu nhiên cho việc phân tích tính kết cấu được mô tả trong hình dưới đây:



Hình 5.6. Phân tích kết cấu sợi bằng dải tương quan

Trong mô hình này, trường kết cấu sợi trước tiên được giải chấp bởi bộ lọc lấy từ đầu ra của ACF. Như vậy, nếu $r(m, n)$ là ACF thì:

$$u(m, n) \otimes a(m, n) = \varepsilon(m, n) \quad (5.26)$$

là trường ngẫu nhiên không tương quan.

Lưu ý rằng, bộ lọc là không duy nhất, có thể là nhân quả, bán nhân quả hay không nhân quả. Các ACF hay dùng như $M(0, 2)$, $M(2, 0)$, $M(1, 1)$, $M(2, 2)$. Các đặc trưng của lược đồ bậc một của $\varepsilon(m, n)$ chẳng hạn như trung bình m_1 , độ phân tán 2μ cũng hay được sử dụng.

Ngoài các đặc trưng trên, có thể đưa thêm một số khái niệm và định nghĩa các đại lượng dựa trên đó như: lược đồ mức xám (Histogram Grey Level Difference), ma trận

xuất hiện mức xám (Grey Level Occurrence Matrices).

Lược đồ hiệu mức xám:

Lược đồ hiệu mức xám dùng để mô tả các thông tin mang tính không gian và được định nghĩa như sau. Cho $d=(d_1, d_2)$ là vecto dịch chuyển giữa 2 điểm ảnh và $g(d)$ là hiệu mức xám với khoảng cách d :

$$g(d) = |f(k, l) - f(k+d_1, l+d_2)| \quad (5.27)$$

với hàm $f(k, l)$ cho giá trị mức xám tại tọa độ (k, l) . Gọi $h_g(g, d)$ là lược đồ của hiệu mức xám khoảng cách d . Với mỗi khoảng cách d ta có một lược đồ mức xám riêng.

Với một miền ảnh có kết cấu thô, lược đồ $h_g(g, d)$ có khuynh hướng tập trung xung quanh $g=0$ với khoảng cách d nhỏ. Trái lại, với một miền ảnh có kết cấu mịn, $h_g(g, d)$ sẽ phân nhánh đều với vecto dịch chuyển d khá nhỏ. Dựa trên lược đồ này, người ta định nghĩa lại một số đại lượng:

- Trung bình:
$$\mu_d = \sum_{k=1}^N g_k h_g(g_k, d) \quad (5.28)$$

- Phương sai:
$$\sigma_d^2 = \sum_{k=1}^N (g_k - \mu_d)^2 h_g(g_k, d) \quad (5.29)$$

- Độ tương phản:
$$c_d = \sum_{k=1}^N g_k^2 h_g(g_k, d) \quad (5.30)$$

Phương sai đo độ tản mát của hiệu mức xám tại một khoảng cách d nào đấy. Kết cấu tất định thường có phương sai σ tương đối nhỏ. Độ tương phản c_d chính là mômen của lược đồ $h_g(g, d)$ xung quanh $g=0$ và đo độ tương phản của hiệu mức xám.

Người ta sử dụng entropy để đo độ đồng nhất của lược đồ h_g :

$$H_g = - \sum_{k=1}^N h_g(g_k, d) \ln(h_g(g_k, d)) \quad (5.31)$$

Ưu điểm cơ bản của lược đồ hiệu mức xám là tính toán đơn giản. Ngoài ra còn có khả năng cho ta tổ chức kết cấu không gian.

Ma trận xuất hiện liên hiệp mức xám

Giả sử $P(k, l, d)$ là xác suất liên hiệp của hai điểm ảnh f_k và f_l với các mức xám k, l tương ứng cách nhau một khoảng d . Xác suất này dễ dàng tính được nhờ việc tính số lần xuất hiện $n_{k,l}$ của cặp điểm ảnh (f_k, f_l) có mức xám k và l với khoảng cách d . Gọi n là tổng số cặp liên hiệp có thể với khoảng cách d trong ảnh. Các phần tử $c_{k,l}$ của ma trận xuất hiện liên hiệp mức xám cd được tính như sau: $cd = (c_{k,l})$

Và
$$c_{k,l} = P(k, l, d) = \frac{n_{k,l}}{n} \quad (5.32)$$

Ma trận xuất hiện liên hiệp mức xám Cd là ma trận vuông $N \times N$ phần tử (N là số mức xám của ảnh). Ma trận này chứa các thông tin hữu ích về tổ chức kết cấu không

gian. Nếu

kết cấu tương đối thô thì các phần tử của ma trận tập trung xung quanh đường chéo chính. Ngược lại, nếu kết cấu bề mặt mịn, giá trị các phần tử của cd sẽ phân rải tương đối rõ.

Dựa trên khái niệm này người ta định nghĩa về một số độ đo:

▪ Xác suất cực đại:
$$P_d = \max_{(k,l)} C_{k,l} \quad (5.33)$$

▪ Entropy:
$$H_d = -\sum_{k=1}^N \sum_{l=1}^N C_{k,l} \ln(C_{k,l}) \quad (5.34)$$

Dễ dàng thấy được entropy cực đại khi xác suất liên hiệp $P(k,l,d)$ có phân phối đều.

Mô men bậc m:
$$I_d = \sum_{k=1}^N \sum_{l=1}^N |k-l|^m C_{k,l} \quad (5.35)$$

I_d cực tiểu khi các phần tử của ma trận C tập trung trên đường chéo chính vì khoảng cách $|k-l|^m$ rất nhỏ, I_d nhỏ có nghĩa là kết cấu khá thô. Người ta cũng còn đưa vào một số độ đo khác như hàm tự tương quan, phổ năng lượng.

Để áp dụng cách tiếp cận này, cần cài đặt các giải thuật tính các đại lượng đo trên.

5.4.2. Phương pháp cấu trúc

Kết cấu sợi có cấu trúc thuần nhất là những texels xác định, mà sự xuất hiện lặp đi lặp lại tuân theo một luật tất định hay ngẫu nhiên nào đấy. Một texel về thực tế là một nhóm các điểm ảnh có cùng một số tính chất bất biến lặp trên ảnh. Một texel cũng có định nghĩa theo mức xám, theo bề mặt hay tính đồng nhất đối với một số các tính chất như kích thước, hướng, lược đồ bậc hai (ma trận tương tranh).

Với các texel được phân bố ngẫu nhiên, tính kết cấu sợi tương ứng của nó được coi là yếu (Weak) ngược với qui luật phân bố tất định gọi là khỏe (Strong). Khi tính kết cấu sợi là yếu, luật phân bố có thể đo bởi:

- Mật độ gờ
- Các loạt dài của các texel liên thông tối đa
- Mật độ cực trị tương đối; số pixel trên một đơn vị diện tích có mức xám cực trị địa phương so với các lân cận.

Ngoài hai cách tiếp cận trên, người ta còn dùng cách tiếp cận khác bằng cách lấy tổ hợp 2 cách trên và gọi là kỹ thuật mosaic. Mô hình này biểu diễn các quá trình học ngẫu nhiên, ví dụ như khảm ngẫu nhiên hay đều của một mặt phẳng vào các đường cong lồi sẽ làm nổi lên tính kết cấu tế bào.

5.4.3. Tiếp cận theo tính kết cấu

Khi đối tượng xuất hiện trên một nền có tính kết cấu cao, việc phân đoạn dựa vào tính kết cấu trở nên quan trọng. Nguyên nhân là kết cấu sợi thường chứa mật độ cao các gờ (edge) làm cho phân đoạn theo biên kém hiệu quả, trừ phi ta loại tính kết cấu.

Việc phân đoạn dựa vào miền đồng nhất cũng có thể áp dụng cho các đặc trưng kết cấu và có thể dùng để phân đoạn các miền có tính kết cấu.

Nhìn chung, việc phân loại và phân vùng dựa vào kết cấu là một vấn đề phức tạp. Ở đây, tài liệu chỉ mang tính chất giới thiệu. Có thể giải quyết vấn đề này trong thực tế nếu ta biết trước các loại kết cấu (dựa vào quy luật hay các phân bố của nó).