

Luận văn

Tìm hiểu phương pháp phân đoạn ảnh

MỞ ĐẦU

Xử lý ảnh (XLA) là một trong những chuyên ngành quan trọng và lâu đời của Công nghệ thông tin. XLA được áp dụng trong nhiều lĩnh vực khác nhau như y học, vật lý, hoá học, tìm kiếm tội phạm,... Mục đích chung của việc XLA thường là: (1) xử lý ảnh ban đầu để có được một bức ảnh mới theo một yêu cầu cụ thể; (2) phân tích ảnh để thu được các thông tin đặc trưng trên ảnh nhằm hỗ trợ cho việc phân loại và nhận biết ảnh; (3) phân đoạn ảnh (image segmentation) để nhận diện được các thành phần trong ảnh nhằm hiểu được kết cấu của bức ảnh ở mức độ cao hơn. Để xử lý được một bức ảnh thì phải trải qua nhiều bước, nhưng bước quan trọng và khó khăn nhất đó là phân đoạn ảnh. Nếu bước phân đoạn ảnh không tốt thì dẫn đến việc nhận diện sai lầm về các đối tượng có trong ảnh.

Trong khoảng 30 năm trở lại đây đã có rất nhiều các thuật toán được đề xuất để giải quyết bài toán phân đoạn ảnh. Các thuật toán hầu hết đều dựa vào hai thuộc tính quan trọng của mỗi điểm ảnh so với các điểm lân cận của nó, đó là: sự khác (dissimilarity) và giống nhau (similarity). Các phương pháp dựa trên sự khác nhau của các điểm ảnh được gọi là các phương pháp biên (boundary-based methods), còn các phương pháp dựa trên sự giống nhau của các điểm ảnh được gọi là phương pháp miền (region-based methods). Tuy nhiên, cho đến nay các thuật toán theo cả hai hướng này đều vẫn chưa cho kết quả phân đoạn tốt, vì cả hai loại phương pháp này đều chỉ nắm bắt được các thuộc tính cục bộ (local) của ảnh. Do đó, trong thời gian gần đây, việc tìm ra các thuật toán nắm bắt được các thuộc tính toàn cục (global) của bức ảnh đã trở thành một xu hướng.

Mục đích chính của em là tìm hiểu và hệ thống lại các phương pháp phân đoạn ảnh đã có theo các hướng: như phân đoạn theo ngưỡng, phân đoạn theo đường biên và theo miền đồng nhất. Ngoài ra, trong đồ án này em cũng tìm hiểu và trình bày thêm một phương pháp được đánh giá là hiệu quả hơn các phương pháp trước đây. Phương pháp này dựa vào việc coi một bức ảnh như một đồ thị, sau đó định

nghĩa một tính chất để so sánh giữa các cặp miền của ảnh. Thuật toán này tuân theo một chiến lược tham lam, có thời gian chạy gần như tuyến tính, nhưng vẫn đảm bảo được việc phân đoạn chính xác và hiệu quả.

Ngoài phần mở đầu và kết luận, luận văn được chia làm 4 chương, cụ thể nội dung các chương như sau:

Chương 1 Trình bày sơ lược về XLA, giới thiệu các giai đoạn xử lý trong một hệ thống XLA, trong đó có bước phân đoạn ảnh. Một số khái niệm, thuật ngữ trong XLA, như điểm ảnh, mức xám, biên,... được trình bày như là các khái niệm.

Chương 2 Hệ thống lại một số thuật toán phân đoạn ảnh theo các hướng: phân đoạn theo ngưỡng, phân đoạn theo đường biên và phân đoạn theo miền đồng nhất. Trong mỗi loại phương pháp này chúng tôi trình bày ngắn gọn phương pháp và ưu nhược điểm của chúng.

Chương 3 Trình bày một thuật toán phân đoạn dựa trên đồ thị :Thuật toán coi mỗi pixel là một đỉnh của đồ thị, sự khác nhau giữa hai điểm ảnh là trọng số của cạnh nối hai đỉnh tương ứng với nhau. Thuật toán dựa theo chiến lược tham lam, nhưng có thể nắm bắt được các thuộc tính non-local của bức ảnh. Một số định lý và hệ quả liên quan đến thuật toán được trình bày và chứng minh ngắn gọn.

Chương 4 đưa ra các đoạn mã chương trình (code) bằng C++ mã hoá một số thuật toán được trình bày trong luận văn.

Khi viết báo cáo này em đã cố gắng hết sức để hoàn thành công việc được giao, song điều kiện thời gian và trình độ còn hạn chế nên không tránh khỏi thiếu sót. Em mong nhận được sự góp ý của thầy giáo hướng dẫn, thầy cô giáo và bạn bè trong khoa Công nghệ thông tin để em có được những kinh nghiệm thực tế và bổ ích để sau này có thể xây dựng được một chương trình hoàn thiện hơn.

CHƯƠNG 1 : TỔNG QUAN VỀ XỬ LÝ ẢNH VÀ PHÂN ĐOẠN ẢNH

Xử lý ảnh ngày nay đã trở thành một ngành khoa học lớn và có mặt trong nhiều lĩnh vực của cuộc sống. Điều này hoàn toàn có thể lý giải được từ một định nghĩa đơn giản về ngành khoa học này: Xử lý ảnh là ngành khoa học nghiên cứu các quá trình xử lý thông tin dạng hình ảnh^{Error! Reference source not found.}, mà hình ảnh là một trong những dạng thông tin phong phú nhất đối với chúng ta.. Trong quá trình xử lý ảnh bước quan trọng nhất và cũng là có khấn nhất là bước phân đoạn ảnh. Phân đoạn nhằm mục đích phân tách các đối tượng cấu thành nên ảnh thô để có thể sử dụng cho các ứng dụng về sau.

1.1 TỔNG QUAN VỀ XỬ LÝ ẢNH

1.1.1 Giới thiệu về Xử lý ảnh

Trong xã hội loài người, ngôn ngữ là một phương tiện trao đổi thông tin phổ biến trong quá trình giao tiếp. Bên cạnh ngôn ngữ, hình ảnh cũng là một cách trao đổi thông tin mang tính chính xác, biểu cảm khá cao và đặc biệt không bị cảm giác chủ quan của đối tượng giao tiếp chi phối. Thông tin trên hình ảnh rất phong phú, đa dạng và có thể xử lý bằng máy tính. Chính vì vậy, trong những năm gần đây sự kết hợp giữa ảnh và đồ hoạ đã trở nên rất chặt chẽ trong lĩnh vực xử lý thông tin.

Cũng như xử lý dữ liệu hình ảnh bằng đồ hoạ, việc XLA số là một lĩnh vực của tin học ứng dụng. Việc xử lý dữ liệu bằng đồ hoạ đề cập đến những ảnh nhân tạo, các ảnh này được xem xét như là những cấu trúc dữ liệu và được tạo ra bởi các chương trình. XLA số thao tác trên các ảnh tự nhiên thông qua các phương pháp và kỹ thuật mã hoá. Ảnh sau khi được thu nhận bằng các thiết bị thu nhận ảnh sẽ được biến đổi thành ảnh số theo các phương pháp số hoá được nhúng trong các thiết bị kỹ thuật khác nhau và được biểu diễn trong máy tính dưới dạng ma trận 2 chiều hoặc 3 chiều.

Mục đích của việc XLA được chia làm hai phần

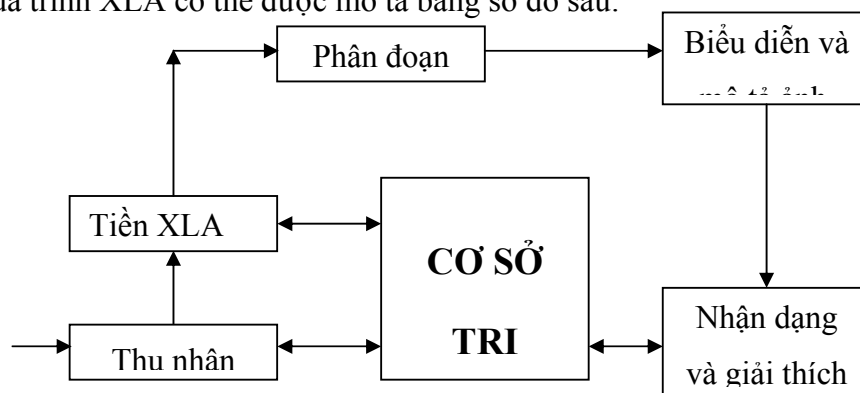
- Biến đổi ảnh làm tăng chất lượng ảnh
- Tự động nhận dạng, đoán ảnh, đánh giá nội dung của ảnh.

Phương pháp biến đổi ảnh được sử dụng trong việc xử lý các ảnh chụp từ không trung (chương trình đo đạc từ máy bay, vệ tinh và các ảnh vũ trụ) hoặc xử lý các ảnh trong y học (ảnh chụp cắt lát, ảnh siêu âm, vv...). Một ứng dụng khác của việc biến đổi ảnh là mã hoá ảnh, trong đó các ảnh được xử lý để rồi lưu trữ hoặc truyền đi.

Các phương pháp nhận dạng ảnh được sử dụng khi xử lý tế bào, nhiễm sắc thể, nhận dạng chữ vv... Thực chất của công việc nhận dạng chính là sự phân loại đối tượng thành các lớp đối tượng đã biết hoặc thành những lớp đối tượng chưa biết. Bài toán nhận dạng ảnh là một bài toán lớn, có rất nhiều ý nghĩa thực tiễn và ta cũng có thể thấy rằng để công việc nhận dạng trở nên dễ dàng thì ảnh phải được tách thành các đối tượng riêng biệt – đây là mục đích chính của bài toán phân đoạn ảnh. Nếu phân đoạn ảnh không tốt sẽ dẫn đến sai lầm trong quá trình nhận dạng ảnh, bởi vậy người ta xem công đoạn phân đoạn ảnh là vấn đề then chốt trong quá trình xử lý ảnh nói chung.

1.1.2 Quá trình XLA

Quá trình XLA có thể được mô tả bằng sơ đồ sau:



Hình 1. Quá trình xử lý ảnh

- Thu nhận ảnh: Đây là công đoạn đầu tiên mang tính quyết định đối với quá trình XLA. Ảnh đầu vào sẽ được thu nhận qua các thiết bị như camera, sensor, máy scanner, vv ... và sau đó các tín hiệu này sẽ được số hoá. Các thông số quan trọng ở bước này là độ phân giải, chất lượng màu, dung lượng bộ nhớ và tốc độ thu nhận ảnh của các thiết bị.
- Tiền xử lý: Ở bước này, ảnh sẽ được cải thiện về độ tương phản, khử nhiễu, khử bóng, khử độ lệch, v.v.. với mục đích làm cho chất lượng ảnh trở nên tốt hơn nữa và thường được thực hiện bởi các bộ lọc.
- Phân đoạn ảnh: Phân đoạn ảnh là *bước then chốt* trong XLA. Giai đoạn này nhằm phân tích ảnh thành những thành phần có cùng tính chất nào đó dựa theo biên hay các vùng liên thông. Tiêu chuẩn để xác định các vùng liên thông có thể là cùng màu, cùng mức xám hay cùng độ nhám vv ... Mục đích của phân đoạn ảnh là để có một miêu tả tổng hợp về nhiều phần tử khác nhau cấu tạo nên ảnh thô. Vì lượng thông tin chứa trong ảnh rất lớn – trong khi trong đa số các ứng dụng chúng ta chỉ cần trích chọn một vài đặc trưng nào đó, do vậy cần có một quá trình để giảm lượng thông tin khổng lồ ấy. Quá trình này bao gồm **phân vùng ảnh** và **trích chọn đặc tính** chủ yếu.
- Biểu diễn và mô tả ảnh: Kết quả của bước phân đoạn ảnh thường được cho dưới dạng dữ liệu điểm ảnh thô, trong đó hàm chứa biên của một vùng ảnh, hoặc tập hợp tất cả các điểm ảnh thuộc về chính vùng ảnh đó. Trong cả hai trường hợp, sự chuyển đổi dữ liệu thô này thành một dạng thích hợp hơn cho việc xử lý trong máy tính là rất cần thiết. Để chuyển đổi chúng, câu hỏi đầu tiên cần phải trả lời là nên biểu diễn một vùng ảnh *dưới dạng biên* hay *dưới dạng một vùng hoàn chỉnh* gồm tất cả những điểm ảnh thuộc về nó. Biểu diễn dạng biên cho một vùng phù hợp với những ứng dụng chỉ quan tâm chủ yếu đến các đặc trưng hình dạng bên ngoài của đối tượng, ví dụ như các góc cạnh và điểm uốn trên biên chẳng hạn. Biểu diễn dạng vùng lại thích hợp cho những ứng dụng khai thác các tính chất bên trong của đối tượng, ví dụ như vân ảnh hoặc cấu trúc xương của

nó. Sự chọn lựa cách biểu diễn thích hợp cho một vùng ảnh chỉ mới là một phần trong việc chuyển đổi dữ liệu ảnh thô sang một dạng thích hợp hơn cho các xử lý về sau. Chúng ta còn phải đưa ra một phương pháp mô tả dữ liệu đã được chuyển đổi đó sao cho những tính chất cần quan tâm đến sẽ được làm nổi bật lên, thuận tiện cho việc xử lý chúng.

- Nhận dạng và giải thích: Đây là bước cuối cùng trong quá trình XLA. Nhận dạng ảnh (image recognition) có thể được nhìn nhận một cách đơn giản là việc *gán nhãn* cho các đối tượng trong ảnh. Giải thích là công đoạn *gán nghĩa* cho một tập các đối tượng đã được nhận biết.

Chúng ta cũng có thể thấy rằng, không phải bất kỳ một ứng dụng XLA nào cũng bắt buộc phải tuân theo tất cả các bước xử lý đã nêu ở trên, ví dụ như các ứng dụng chỉnh sửa ảnh nghệ thuật chỉ dừng lại ở bước tiền xử lý. Một cách tổng quát thì những chức năng xử lý bao gồm nhận cả nhận dạng và giải thích thường chỉ có mặt trong hệ thống phân tích ảnh tự động hoặc bán tự động, được dùng để rút trích ra những thông tin quan trọng từ ảnh, ví dụ như các ứng dụng nhận dạng ký tự quang học, nhận dạng chữ viết tay vv...

1.2. TỔNG QUAN VỀ PHÂN ĐOẠN ẢNH

Để phân tích các đối tượng trong ảnh, chúng ta cần phải phân biệt được các đối tượng cần quan tâm với phần còn lại của ảnh, hay còn gọi là nền ảnh. Những đối tượng này có thể tìm ra được nhờ các *kỹ thuật phân đoạn ảnh*, theo nghĩa tách phần tiền cảnh ra khỏi hậu cảnh trong ảnh. Mỗi một đối tượng trong ảnh được gọi là một vùng hay miền, đường bao quanh đối tượng ta gọi là đường biên. Mỗi một vùng ảnh phải có các đặc tính đồng nhất (ví dụ: màu sắc, kết cấu, mức xám vv...). Các đặc tính này tạo nên một véc tơ đặc trưng riêng của vùng (**feature vectors**) giúp chúng ta phân biệt được các vùng khác nhau.

Như vậy, 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. Có thể thấy kỹ thuật phát hiện biên và phân vùng ảnh là hai bài toán đối ngẫu của nhau. Thực vậy, dò biên để thực hiện phân lớp đối tượng và một khi đã phân lớp xong cũng có nghĩa là đã phân vùng được ảnh. Ngược lại, khi đã phân vùng, ảnh được phân lập thành các đối tượng, ta có thể phát hiện biên.

Có rất nhiều kỹ thuật phân đoạn ảnh, nhưng nhìn chung chúng ta 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 techniques) 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 các khái niệm đồng nhất và gần về hình học.

1.3. MỘT SỐ KHÁI NIỆM CƠ BẢN

1.3.1 Điểm ảnh - Pixel

Ảnh trong thực tế là một ảnh liên tục về không gian và về giá trị độ sáng. Để có thể XLA bằng máy tính cần phải tiến hành số hoá ảnh. Trong quá trình số hoá, người ta biến đổi tín hiệu liên tục sang tín hiệu rời rạc thông qua quá trình lấy mẫu (rời rạc hoá về không gian) và lượng hoá thành phần giá trị mà về nguyên tắc bằng mắt thường không phân biệt được hai điểm kề nhau. Trong quá trình này người ta sử dụng khái niệm *Picture element* mà ta quen gọi là Pixel - phần tử ảnh. Như vậy, một ảnh là một tập hợp các *Pixel*

1.3.2 Mức xám – Gray level

Mức xám là kết quả *sự mã hoá* tương ứng một cường độ sáng của mỗi điểm ảnh với một giá trị số - kết quả của quá trình lượng hoá. Cách mã hoá kinh điển

thường dùng 16, 32 hay 64 mức. Phổ dụng nhất là mã hoá ở mức 256, ở mức này mỗi Pixel sẽ được mã hoá bởi 8 bit.

1.3.3 Biên

Biên là một đặc tính rất quan trọng của đối tượng trong ảnh, nhờ vào biên mà chúng ta phân biệt được đối tượng này với đối tượng kia. Một điểm ảnh có thể gọi là *điểm biên* nếu ở đó có sự thay đổi đột ngột về mức xám. Tập hợp các điểm biên tạo thành *biên* hay còn gọi là đường bao ảnh.

1.3.4 Láng giềng

Trong XLA có một khái niệm rất quan trọng, đó là khái niệm *láng giềng*. Có hai loại láng giềng: *4-láng giềng* và *8-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:

$$N_4((x,y)) = \{(x+1,y), (x-1,y), (x,y+1), (x,y-1)\} \quad (1.1)$$

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

$$N_8((x,y)) = N_4((x,y)) \cup \{(x+1,y+1), (x-1,y-1), (x+1,y-1), (x-1,y+1)\} \quad (1.2)$$

1.3.5 Vùng liên thông

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) .

CHƯƠNG 2 : MỘT SỐ KỸ THUẬT PHÂN ĐOẠN ẢNH

Phân đoạn (segmentation) là một quá trình chia ảnh ra các vùng con khác nhau mà trong mỗi vùng chứa các thực thể có ý nghĩa cho việc phân lớp - mỗi thực thể được xem là một đối tượng mang những thông tin đặc trưng riêng. Có rất nhiều kỹ thuật phân đoạn ảnh, trong chương này chúng tôi giới thiệu một số kỹ thuật tiêu biểu như: Phân đoạn dựa vào ngưỡng, phân đoạn dựa vào biên, phân đoạn theo miền đồng nhất. Cũng có thể thấy rằng không có một kỹ thuật phân đoạn nào là vạn năng – theo nghĩa là có thể áp dụng cho mọi loại ảnh và cũng không có một kỹ thuật phân đoạn ảnh nào là hoàn hảo.

2.1 PHÂN ĐOẠN DỰA VÀO NGƯỠNG

2.1.1 Giới thiệu chung

Biên độ của các tính chất vật lý của ảnh (như là độ phản xạ, độ truyền sáng, màu sắc ...) là một đặc tính đơn giản và rất hữu ích. Nếu biên độ đủ lớn đặc trưng cho ảnh thì chúng ta có thể dùng ngưỡng biên độ để phân đoạn ảnh. Thí dụ, biên độ trong bộ cảm biến hồng ngoại có thể phản ánh vùng có nhiệt độ thấp hay vùng có nhiệt độ cao. Đặc biệt, kỹ thuật phân ngưỡng theo biên độ rất có ích đối với ảnh nhị phân như văn bản in, đồ họa, ảnh màu hay ảnh X-quang.

Việc chọn ngưỡng trong kỹ thuật này là một bước vô cùng quan trọng, thông thường người ta tiến hành theo các bước chung như sau:

- ✓ Xem xét lược đồ xám của ảnh để xác định và khe. Nếu ảnh có nhiều đỉnh và khe thì các khe có thể sử dụng để chọn ngưỡng.
- ✓ Chọn ngưỡng T sao cho một phần xác định trước η của toàn bộ số mẫu là thấp hơn T.
- ✓ Điều chỉnh ngưỡng dựa trên xét lược đồ xám của các điểm lân cận.

- ✓ Chọn ngưỡng bằng cách xem xét lược đồ xám của những điểm thoả tiêu chuẩn đã chọn.

Một thuật toán đơn giản trong kỹ thuật này là: giả sử rằng chúng ta đang quan tâm đến các đối tượng sáng (object) trên nền tối (background), một tham số T - gọi là ngưỡng độ sáng, sẽ được chọn cho một ảnh $f[x,y]$ theo cách:

$$\begin{aligned} \text{If } f[x,y] \geq T & \quad f[x,y] = \text{object} = 1 \\ \text{Else} & \quad f[x,y] = \text{Background} = 0. \end{aligned}$$

Ngược lại, đối với các đối tượng tối trên nền sáng chúng ta có thuật toán sau:

$$\begin{aligned} \text{If } f[x,y] < T & \quad f[x,y] = \text{object} = 1 \\ \text{Else} & \quad f[x,y] = \text{Background} = 0. \end{aligned}$$

Vấn đề chính là chúng ta nên chọn ngưỡng T như thế nào để việc phân vùng đạt được kết quả cao nhất?.

Có rất nhiều thuật toán chọn ngưỡng: ngưỡng cố định, dựa trên lược đồ, sử dụng Entropy, sử dụng tập mờ, chọn ngưỡng thông qua sự không ổn định của lớp và tính thuần nhất của vùng vv... Ở đây chúng tôi đề cập đến hai thuật toán chọn ngưỡng đó là chọn ngưỡng cố định và chọn ngưỡng dựa trên lược đồ.

2.1.2 Chọn ngưỡng cố định

Đây là phương pháp chọn ngưỡng độc lập với dữ liệu ảnh. Nếu chúng ta biết trước là chương trình ứng dụng sẽ làm việc với các ảnh có độ tương phản rất cao, trong đó các đối tượng quan tâm rất tối còn nền gần như là đồng nhất và rất sáng thì việc chọn ngưỡng $T = 128$ (xét trên thang độ sáng từ 0 đến 255) là một giá trị chọn khá chính xác. Chính xác ở đây hiểu theo nghĩa là số các điểm ảnh bị phân lớp sai là cực tiểu.

2.1.3 Chọn ngưỡng dựa trên lược đồ (Histogram)

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 $\{h[b] \mid b = 0, 1, \dots, 2^B-1\}$ đã được đưa ra. 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 trơn dữ liệu lược đồ ban đầu mang lại nhằm loại bỏ những dao động nhỏ về độ sáng. Tuy nhiên các thuật toán làm trơn cần phải cẩn thận, không được làm dịch chuyển các vị trí đỉnh của lược đồ. Nhận xét này dẫn đến thuật toán làm trơn dưới đây:

$$h_{smooth}[b] = \frac{1}{W} \sum_{w=-(W-1)/2}^{(W-1)/2} h_{raw}[b-w] \quad W \text{ lẻ} \quad (2.1)$$

Trong đó, W thường được chọn là 3 hoặc 5

2.1.3.1 Thuật toán đẳng liệu

Đây là kỹ thuật chọn ngưỡng theo kiểu lặp do Ridler và Calvard đưa ra. Thuật toán được mô tả như sau:

- B1: Chọn giá trị ngưỡng khởi động $\theta_0 = 2^{B-1}$
- B2: Tính các trung bình mẫu ($m_{f,0}$) của những điểm ảnh thuộc đối tượng và ($m_{b,0}$) của những điểm ảnh nền.
- B3: Tính các ngưỡng trung gian theo công thức:

$$\theta_k = \frac{m_{f,k-1} + m_{b,k-1}}{2} \quad \text{với } k = 1, 2, \dots \quad (2.2)$$

- B4: Nếu $\theta_k = \theta_{k-1}$: Kết thúc, dừng thuật toán.

Ngược lại : Lặp lại bước 2.

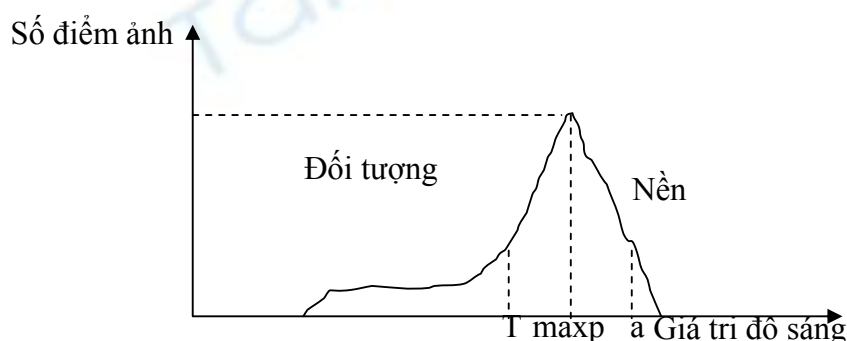
2.1.3.2 Thuật toán đối xứng nền

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 (2.1). Đỉnh cực đại $\max p$ 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 được định nghĩa như sau:

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 độ 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 $dP/da \geq 0$.



Hình 2. Minh họa thuật toán đối xứng nền

Ở đây ta đang giả thiết là ảnh có các đối tượng tối trên nền sáng. Giả sử mức là 5%, thì có nghĩa là ta phải ở bên phải đỉnh $\max p$ một giá trị a sao cho $P(a)=95\%$. Do tính đối xứng đã giả định ở trên, chúng ta sử dụng độ dịch chuyển về phía trái của điểm cực đại tìm giá trị ngưỡng T :

$$T = \max p - (a - \max p) \quad (2.3)$$

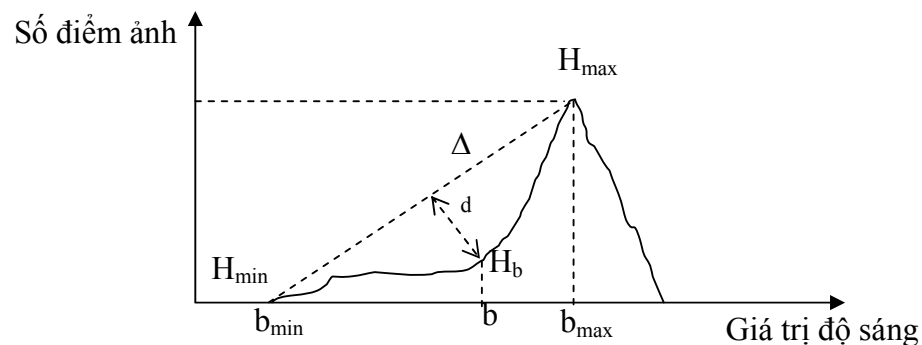
Kỹ thuật này dễ dàng điều chỉnh được cho phù hợp với tình huống ảnh có các đối tượng sáng trên một nền tối.

2.1.3.3 Thuật toán tam giác

Khi một ảnh có các điểm ảnh thuộc đối tượng tạo nên một đỉnh yếu trong lược đồ ảnh thì thuật toán tam giác hoạt động rất hiệu quả. Thuật toán này do Zack đề xuất và được mô tả như sau:

- B1: Xây dựng đường thẳng Δ là đường nối hai điểm (H_{\max}, b_{\max}) và (H_{\min}, b_{\min}) , trong đó H_{\max} là điểm có Histogram lớn nhất ứng với mức xám b_{\max} và H_{\min} là điểm có Histogram ứng với độ sáng nhỏ nhất b_{\min} .
- B2: Tính khoảng cách d từ H_b của lược đồ (ứng với điểm sáng b) đến Δ . Trong đó, $b \in [b_{\max}, b_{\min}]$.
- B3: Chọn ngưỡng $T = \text{Max} \{H_b\}$

Minh hoạ thuật toán tam giác bởi hình vẽ như sau:



Hình 3. Minh hoạ thuật toán tam giác

2.1.3.4 Chọn ngưỡng đối với Bimodal Histogram

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:

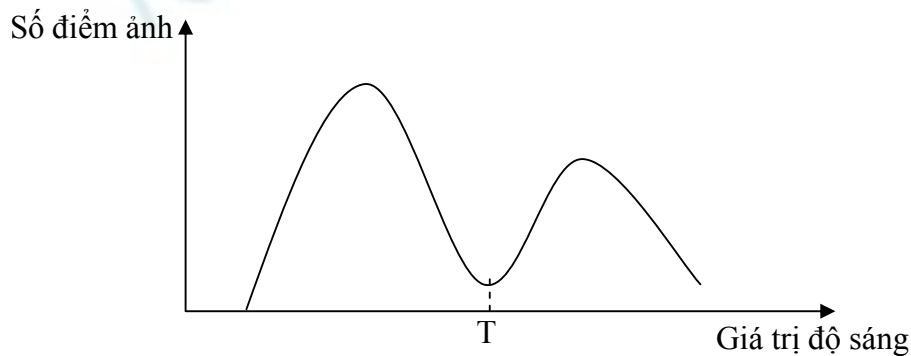
a/ Các đối tượng sáng:

$$TopHat(A, B) = A - (A \circ B) = A - \max_B(\min_B(A)) \quad (2.4)$$

b/ Các đối tượng tối:

$$TopHat(A, B) = A - (A \circ B) = A - \min_B(\max_B(A)) \quad (2.5)$$

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 (2.1).



Hình 4. Bimodal Histogram

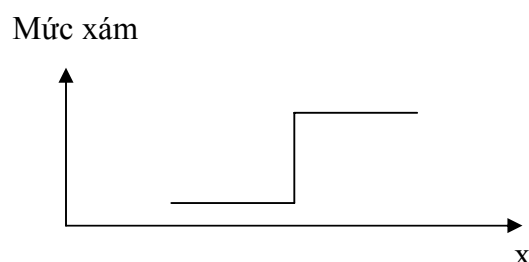
Trong một số ứng dụng nhất định, cường độ của đối tượng hay nền thay đổi khá chậm. Trong trường hợp này, histogram ảnh có thể không chứa hai thùy phân biệt rõ ràng, vì vậy có thể phải dùng ngưỡng thay đổi theo không gian. Hình ảnh được chia thành những khối hình vuông, histogram và ngưỡng được tính cho 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 2.1 cho khối này.

2.2 PHÂN ĐOẠN DỰA THEO ĐƯỜNG BIÊN

2.2.1 Giới thiệu chung

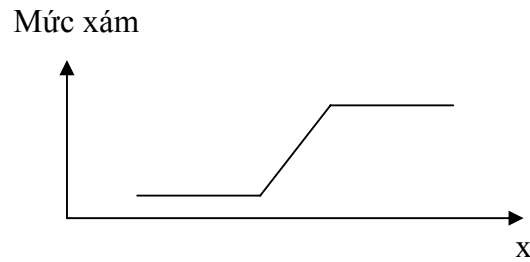
Như chúng ta đã biết, Biên là một đặc tính rất quan trọng để phân vùng các đối tượng. Có thể hình dung tầm qua trọng của biên thông qua ví dụ sau: Khi một người hoạ sĩ vẽ một cái bàn gỗ, chỉ cần phác thảo vài nét về hình dáng như cái mặt bàn, cái chân bàn mà không cần thêm các chi tiết khác, người xem đã có thể nhận ra đó là cái bàn. Vài nét phác thảo của người hoạ sĩ chính là đường biên bao quanh đối tượng. Nếu ứng dụng của ta là phân lớp nhận diện các đối tượng thì coi như nhiệm vụ đã hoàn thành. Tuy nhiên, nếu đòi hỏi thêm các chi tiết khác như vân gỗ, màu sắc, kích thước vv ... thì chừng ấy thông tin là chưa đầy đủ.

Trong toán học, người ta đưa ra khái niệm đường biên lý tưởng như sau: **Đường biên lý tưởng 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.** Thể hiện của định nghĩa là hình vẽ 2



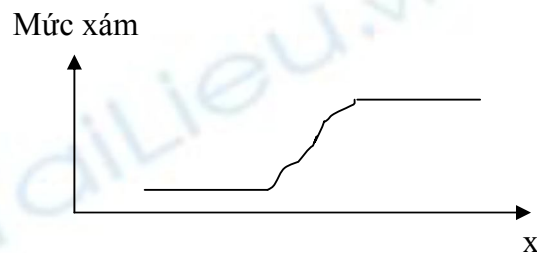
Hình 5. Đường biên lý tưởng

Một loại đường biên nữa - được gọi là đường biên bậc thang: **Đường biên bậc thang 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.**



Hình 6. Đường biên bậc thang

Trong thực tế đường biên của chúng ta thường có dạng như sau:



Hình 7. Đường biên thực

Như đã nói ở trên, biên là một trong những đặc trưng quan trọng của ảnh, chính vì vậy mà trong nhiều ứng dụng người ta sử dụng cách phân đoạn dựa theo biên. Việc phân đoạn ảnh dựa vào biên được tiến hành qua các bước:

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

2.2.2 Phát hiện biên

Phát hiện biên một cách lý tưởng là xác định được tất cả các đường bao trong các đối tượng. Có nhiều phương pháp phát hiện biên, thông thường chúng ta sử dụng phương pháp phát hiện biên trực tiếp. Phương pháp này nhằm làm nổi biên dựa vào sự biến thiên về giá trị độ sáng của điểm ảnh. Kỹ thuật chủ yếu dùng ở đây là kỹ thuật đạo hàm. Nếu lấy đạo hàm bậc nhất của ảnh ta có phương pháp

Gradient, nếu lấy đạo hàm bậc hai ta có kỹ thuật Laplace. Phương pháp này có ưu điểm là ít chịu ảnh hưởng của nhiễu, song nếu sự biến thiên của độ sáng không đột ngột thì hiệu quả đạt được là rất kém.

2.2.2.1 Kỹ thuật Gradient

Phương pháp Gradient là phương pháp dò biên cục bộ dựa vào cực đại của đạo hàm. Theo định nghĩa, Gradient là một vectơ có các thành phần biểu thị tốc độ thay đổi giá trị của điểm ảnh theo hai hướng x và y. Các thành phần của gradient được tính theo công thức:

$$\begin{aligned}\frac{\partial f(x, y)}{\partial x} &= f_x \approx \frac{f(x + dx, y) - f(x, y)}{dx} \\ \frac{\partial f(x, y)}{\partial y} &= f_y \approx \frac{f(x, y + dy) - f(x, y)}{dy}\end{aligned}\quad (2.6)$$

trong đó, dx là khoảng cách giữa các điểm theo hướng x (khoảng cách tính bằng số điểm), dy là khoảng cách giữa các điểm theo hướng y. Thực tế, người ta hay dùng với $dx = dy = 1$.

Với một ảnh liên tục $f(x, y)$, các đạo hàm riêng của nó cho phép xác định vị trí cực đại cục bộ theo hướng của biên. Thực vậy, một ảnh liên tục được biểu diễn bởi một hàm $f(x, y)$ dọc theo r với góc φ (toạ độ cực):

$$f(x, y) = f(r \cdot \cos \varphi, r \cdot \sin \varphi) \quad (2.7)$$

gradient được định nghĩa:

$$\begin{aligned}\frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = f_x \cos \varphi + f_y \sin \varphi \\ \frac{\partial f}{\partial \varphi} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial \varphi} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \varphi} = -f_x r \sin \varphi + f_y r \cos \varphi\end{aligned}\quad (2.8)$$

φ là hướng của biên khi:

$$\begin{aligned}\frac{\partial f}{\partial \varphi} &= 0 \\ \Leftrightarrow -f_x r \sin \varphi + f_y r \cos \varphi &= 0 \\ \Leftrightarrow \operatorname{tg} \varphi &= \frac{f_x}{f_y} \\ \Leftrightarrow \varphi &= \operatorname{arctg} \left(\frac{f_x}{f_y} \right)\end{aligned}$$

Thực ra, đạo hàm của ảnh là không tồn tại vì $f(x,y)$ không liên tục. Ở đây, ta chỉ sử dụng mô phỏng theo ý nghĩa của đạo hàm, việc tính toán là xấp xỉ đạo hàm bằng kỹ thuật nhân chập. Trong phương pháp gradient, người ta chia nhỏ thành hai kỹ thuật (tương ứng với hai toán tử khác nhau):

- + Kỹ thuật gradient dùng toán tử gradient, lấy đạo hàm theo một hướng;
- + Kỹ thuật la bàn dùng toán tử la bàn, lấy đạo hàm theo tám hướng: Bắc, Nam, Đông, Tây, và Đông Bắc, Tây Bắc, Đông Nam, Tây Nam.

2.2.2.2 Kỹ thuật Gradient

Kỹ thuật gradient sử dụng một cặp mặt nạ H_1, H_2 trực giao (theo hai hướng vuông góc). Nếu định nghĩa g_x, g_y là gradient tương ứng theo hai hướng x, y thì biên độ của gradient tại điểm (i,j) - ký hiệu là $g(i,j)$ được tính theo công thức:

$$g(i, j) = A_0 = \sqrt{g_x^2(i, j) + g_y^2(i, j)} \quad (2.9)$$

Góc φ :

$$\varphi_r(i, j) = \arctan\left(\frac{g_x(i, j)}{g_y(i, j)}\right) \quad (2.10)$$

Có nhiều toán tử đạo hàm khác nhau đã được áp dụng. Em xin trình bày một số toán tử tiêu biểu (tương ứng là các mặt nạ khác nhau) như Toán tử Robert, toán tử Sobel, Toán tử Prewitt ...

+/- Toán tử Robert (Do Robert đề xuất năm 1965): Toán tử này là áp dụng trực tiếp của công thức đạo hàm tại điểm (x,y). Chọn cặp mặt nạ H_1, H_2 như sau:

$$H_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Với mỗi điểm ảnh $I(x,y)$ của I , gọi g_x, g_y tương ứng là các đạo hàm theo các hướng x và y, ta có:

$$\begin{cases} g_x(i, j) = I(i+1, j) - I(i, j) \\ g_y(i, j) = I(i, j+1) - I(i, j) \end{cases} \quad (2.11)$$

Điều bày tương đương với việc chập ảnh với hai mặt nạ H_1, H_2 :

$$\begin{cases} g_x(i, j) = I(i, j) \otimes H_1 \\ g_y(i, j) = I(i, j) \otimes H_2 \end{cases} \quad (2.12)$$

Người ta gọi H_1, H_2 là mặt nạ Robert.

Trong trường hợp tổng quát, giá trị gradient biên độ g và gradient hướng ϕ_r được tính bởi công thức (2.9), (2.10). Ngoài ra, để giảm thời gian tính toán ta cũng có thể dùng các chuẩn sau để tính $g(i,j)$:

$$A_1 = |g_x(i, j)| + |g_y(i, j)| \quad (2.13)$$

$$\text{Hoặc } A_2 = \text{Max}(|g_x(i, j)|, |g_y(i, j)|) \quad (2.14)$$

Một điểm nữa là: khi di chuyển mặt nạ trên ảnh, trường hợp gặp các điểm biên, thì coi các điểm ứng với mặt nạ ở bên ngoài ảnh có giá trị 0.

+/- Toán tử Solbel:

Toán tử Solbel sử dụng hai mặt nạ H_1, H_2 như sau:

$$H_1 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (2.15)$$