

LỜI CẢM ƠN

Lời đầu tiên em xin bày tỏ lòng biết ơn sâu sắc tới thầy PGS.TS Ngô Quốc Tạo, Viện Công Nghệ Thông Tin – Viện Khoa Học Công Nghệ Việt Nam, người đã tận tình hướng dẫn, chỉ bảo, giúp đỡ em hoàn thành đồ án tốt nghiệp này. Em xin gửi lời cảm ơn chân thành đến Thầy PGS.TS Đỗ Năng Toàn, Viện Công Nghệ Thông Tin - Viện Khoa Học & Công Nghệ Việt Nam.

Em cũng xin gửi lời cảm ơn chân thành đến các thầy cô giáo trong bộ môn Công nghệ thông tin trường Đại học Dân Lập Hải Phòng đã nhiệt tình giảng dạy, truyền đạt kiến thức cho em trong suốt 4 năm học qua.

Cuối cùng em xin gửi lời cảm ơn đến gia đình, bạn bè những người đã động viên giúp đỡ em rất nhiều trong suốt quá trình học tập tại trường cũng như quá trình làm đồ án này.

Hải Phòng , tháng 7 năm 2010

Sinh viên

Vũ Văn Thành

MỤC LỤC

LỜI CẢM ƠN	1
LỜI MỞ ĐẦU	4
CHƯƠNG 1. TỔNG QUAN XỬ LÝ ẢNH	6
1.1. Giới thiệu về xử lý ảnh	6
1.2 Quá trình xử lý ảnh	7
1.3. Tổng quan về phân đoạn ảnh	9
1.4 Một số khái niệm cơ bản.....	10
1.4.1 Điểm ảnh – Pixel	10
1.4.2 Mức xám - Gray level	10
1.4.3 Biên	10
1.4.4 Láng giềng.....	11
1.4.5 Vùng liên thông	11
1.4.6 Biểu diễn ảnh.....	11
1.4.7 Tăng cường và khôi phục ảnh	12
1.4.8 Biến đổi ảnh.....	12
1.4.9. Phân tích ảnh	12
1.4.10 Nhận dạng ảnh.....	12
1.4.11 Nén ảnh.....	12
1.5 Các định dạng cơ bản trong xử lý ảnh	12
CHƯƠNG 2. PHÂN ĐOẠN ẢNH DỰA VÀO NGƯỠNG.....	13
2.1 Giới thiệu chung	13
2.2 Chọn ngưỡng cố định.....	14
2.3 Chọn ngưỡng dựa trên lược đồ (Histogram).....	15
2.3.1 Thuật toán đẳng liệu	15
2.3.2 Thuật toán đối xứng nền.....	15
2.3.3 Thuật toán tam giác	17
2.3.4 Chọn ngưỡng đối với Bimodal Histogram	17
2.4 Phân ngưỡng tối ưu dựa trên sự không ổn định của lớp và tính đồng nhất của vùng.....	19
2.4.1 Giới thiệu	19
2.4.2 Cơ sở lý thuyết và thuật toán.....	20
CHƯƠNG 3. PHÂN ĐOẠN THEO MIỀN ĐỒNG NHẤT	33
3.1 Giới thiệu	33
3.2 Phương pháp tách cây tứ phân.....	34
3.3 Phương pháp phân vùng hợp	37
3.4 Phương pháp tách hợp (Split- Meger).....	38
3.5 Nhận xét.....	39
CHƯƠNG 4. PHÂN ĐOẠN DỰA VÀO ĐỒ THỊ.....	40
4.1 Giới thiệu	40
4.2 Phân đoạn dựa vào đồ thị	41
4.3 Tính chất của so sánh cặp miền.....	42
4.4 Thuật toán và các tính chất	43
4.5 Nhận xét.....	49
CHƯƠNG 5. CÀI ĐẶT THỬ NGHIỆM.....	50
5.1 Thuật toán Đẳng liệu :	50
5.2 Thuật toán Tam giác :	54

5.3 Thuật toán GraphBased :	57
5.4 Kết quả đạt được	60
KẾT LUẬN	62
TÀI LIỆU THAM KHẢO	65

LỜI MỞ ĐẦU

Cùng với sự phát triển ngày càng mạnh mẽ của khoa học kỹ thuật trong một vài thập kỷ gần đây, xử lý ảnh tuy là một ngành khoa học còn tương đối mới mẻ so với nhiều ngành khoa học khác nhưng hiện nay nó đang là một trong những lĩnh vực phát triển rất nhanh và thu hút sự quan tâm đặc biệt từ các nhà khoa học, thúc đẩy các trung tâm nghiên cứu, ứng dụng về lĩnh vực hấp dẫn này. Điều này hoàn toàn có thể lý giải được từ một định nghĩa đơn giản: Xử lý ảnh là ngành khoa học nghiên cứu các quá trình xử lý thông tin dạng hình ảnh, mà hình ảnh là một dạng thông tin vô cùng phong phú, đa dạng và là phương tiện giao tiếp, trao đổi chủ yếu của con người. Thông tin hình ảnh ngày nay có thể được xử lý dễ dàng bằng máy tính, chính vì thế, trong những năm gần đây sự kết hợp giữa ảnh và đồ họa đã trở nên rất chặt chẽ trong lĩnh vực xử lý thông tin. Mục tiêu chính của xử lý ảnh thường là:

- Xử lý ảnh ban đầu để có được ảnh mới theo một yêu cầu xác định (ví dụ như ảnh mờ cần xử lý để được ảnh rõ hơn)
- Phân tích ảnh để thu được các thông tin đặc trưng giúp cho việc phân loại, nhận biết ảnh (ví dụ phân tích ảnh vân tay để trích chọn các đặc trưng vân tay)
- Hiểu ảnh đầu vào để có những mô tả về ảnh ở mức cao hơn, sâu hơn (ví dụ từ ảnh một tai nạn giao thông phức tạp hiện trường tai nạn).

Qua đó, ta có thể thấy xử lý ảnh đóng vai trò quan trọng như thế nào trong các ứng dụng thực tế về khoa học kỹ thuật cũng như trong cuộc sống thường ngày. Những ứng dụng này dường như là vô hạn cùng với sự khám phá của con người và sự phát triển như vũ bão của công nghệ số hóa, chẳng hạn, trong các lĩnh vực như: sản xuất và kiểm tra chất lượng, sự di chuyển của Robot, các phương tiện đi lại tự trị, công cụ hướng dẫn cho người mù, an ninh và giám sát, nhận dạng đối tượng, nhận dạng mặt, các ứng dụng trong y học, sản xuất, hiệu chỉnh Video, và chinh phục vũ trụ...

Để xử lý được một bức ảnh thì phải trải qua nhiều khâu khác nhau tùy theo mục đích của việc xử lý, nhưng khâu quan trọng và khó khăn nhất đó là phân đoạn ảnh. Trong một số lượng lớn các ứng dụng về xử lý ảnh và hiển thị máy tính, phân đoạn đóng vai trò chính yếu như là bước đầu tiên trước khi áp dụng các thao tác xử lý ảnh mức cao hơn như: nhận dạng, giải thích ngữ nghĩa, và biểu diễn ảnh. Nếu bước phân đoạn ảnh không tốt thì dẫn đến việc nhận diện sai lầm về các đối tượng có trong ảnh.

Phân đoạn ảnh đã và đang là một trong những vấn đề nhận được nhiều sự quan tâm trong lĩnh vực xử lý ảnh. Trong khoảng 30 năm trở lại đây đã có rất nhiều các thuật toán được đề xuất để giải bài toán này. 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 nhau (dissimilarity) và giống nhau (similarity) giữa chúng. 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 của ảnh. Do đó, trong thời gian gần đây, việc tìm ra các thuật toán nắm bắt được các thuộc tính toàn cục của bức ảnh đã trở thành một xu hướng phổ biến.

Nhận thấy, xử lý ảnh là một lĩnh vực hay và khó. Được sự khuyến khích và hỗ trợ của thầy giáo hướng dẫn, em đã chọn đề tài nghiên cứu và hệ thống một số phương pháp phân đoạn ảnh để làm luận văn tốt nghiệp.

CHƯƠNG 1. TỔNG QUAN XỬ LÝ ẢNH

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 họa đã trở nên rất chặt chẽ trong lĩnh vực xử lý thông tin.

Cũng như xử lý dữ liệu hình ảnh bằng đồ họa, việc xử lý ảnh số là một lĩnh vực của tin học ứng dụng. Việc xử lý dữ liệu bằng đồ họa 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. Xử lý ảnh 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 xử lý ảnh được chia làm 2 phần:

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

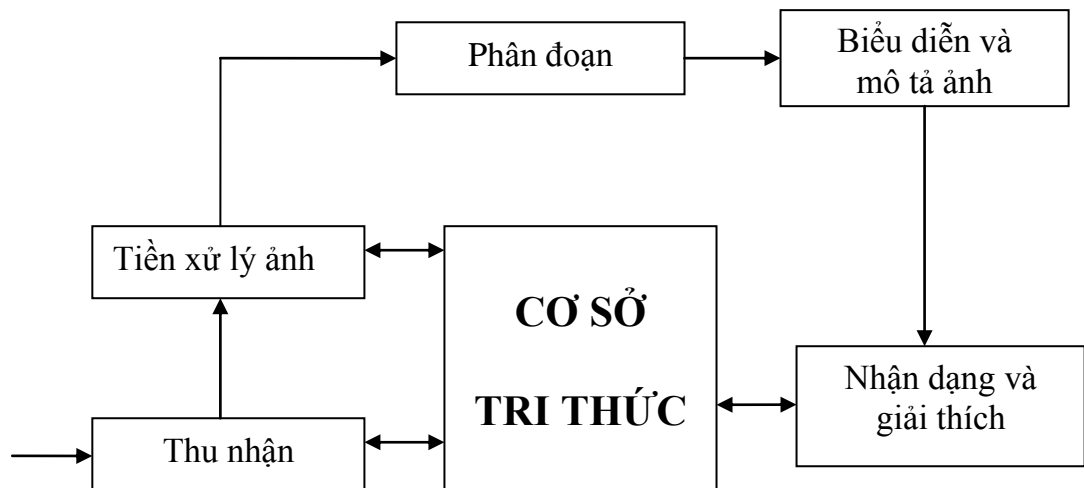
Phương pháp biến đổi ảnh được sử dụng trong việc xử lý các ảnh chụp từ không trung (Chương trình đo đạc từ máy bay, vệ tinh và các ảnh vũ trụ) hoặc xử lý các ảnh trong y học (ảnh siêu âm, ảnh chụp cắt lát, vv...). Một ứng dụng khác của biến đổi ảnh đó là mã hóa ảnh, trong đó các ảnh được xử lý để lưu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

Các phương pháp nhận dạng ảnh được sử dụng khi xử lý tế bào, nhiễm sắc thể, nhận dạng chữ viết, 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 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 quá trình then chốt trong quá trình xử lý ảnh nói chung.

1.2 Quá trình xử lý ảnh

Quá trình xử lý ảnh 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 xử lý ảnh. Ảnh đầu vào sẽ được thu nhận qua các thiết bị như camera, sensor, máy quét, vv... và sau đó các tín hiệu này sẽ được số hóa. 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, vv... với mục đích làm cho chất lượng ảnh trở nên tốt hơn và thường được thực hiện bởi các bộ lọc.

- Phân đoạn ảnh: Phân đoạn ảnh là bước then chốt trong xử lý ảnh, 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 từ nhiều phần tử khác nhau cấu tạo nên ảnh thô. Vì lượng thông tin chứa trong ảnh rất lớn, trong khi đó 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 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. 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 hay 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 những 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 xử lý ảnh.

Nhận dạng ảnh 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 có thể thấy rằng, không phải bất kỳ một ứng dụng xử lý ảnh nào cũng bắt buộc phải tuân theo 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 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 kí tự quang học, nhận dạng chữ viết tay vv...

1.3. 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 kĩ 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 (như màu sắc, kết cấu, mức xám vv...). Các đặc tính này tạo nên một vector đặc trưng riêng của vùng 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 kĩ 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 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.

1.4 Một số khái niệm cơ bản

1.4.1 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.

1.4.2 Mức xám - Gray level

Mức xám là kết quả của sự mã hóa tương ứng một cường độ sáng của mỗi điểm ảnh với một giá trị sáng, kết quả của quá trình lượng tử hóa. Cách mã hóa kinh điển thường dùng là 16, 32 hay 64 mức. Phổ dụng nhất là mã hóa ở mức 256, ở mức này mỗi Pixel sẽ được mã hóa bởi 8 bit.

1.4.3 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 gọi là biên hay còn gọi là đường bao ảnh.

1.4.4 Láng giềng

Trong xử lý ảnh 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:

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

1.4.5 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 $(x_A, y_B) \dots (x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1}) \dots (x_B, y_B)$, 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) .

1.4.6 Biểu diễn ảnh

Trong biểu diễn ảnh, người ta dùng các phần tử đặc trưng của ảnh là pixel. Có thể xem một hàm hai biến chứa các thông tin như biểu diễn của ảnh, việc xử lý ảnh số yêu cầu ảnh phải được mã hóa và lượng tử hóa. Việc lượng tử hóa ảnh là chuyển đổi tín hiệu tương tự sang tín hiệu số của một ảnh đã lấy mẫu sang một số hữu hạn mức xám.

Một số mô hình thường được dùng trong xử lý ảnh, mô hình toán, mô hình thống kê.

1.4.7 Tăng cường và khôi phục ảnh

Tăng cường ảnh là bước quan trọng tạo tiền đề cho xử lý ảnh, gồm một loạt các kỹ thuật như: lọc độ tương phản, khử nhiễu, nổi màu...

Khôi phục ảnh là nhằm loại bỏ các suy giảm trong ảnh.

1.4.8 Biến đổi ảnh

Trong thuật ngữ biến đổi ảnh thường được dùng để nói đến một lớp các ma trận đơn vị và các kỹ thuật dùng để biến đổi ảnh. Một số loại biến đổi được dùng như: biến đổi Fourier, Sin, Cosin, Hadamard, tích Kronecker, biến đổi Karhunen Loeve...

1.4.9. Phân tích ảnh

Liên quan đến việc xác định các độ đo định lượng của một ảnh để đưa ra một mô tả đầy đủ về ảnh. Các kỹ thuật được sử dụng ở đây nhằm mục đích xác định biên của ảnh.

1.4.10 Nhận dạng ảnh

Là quá trình liên quan đến việc mô tả các đối tượng mà người ta muốn đặc tả nó. Quá trình nhận dạng thường đi sau quá trình trích chọn các đặc tính chủ yếu của đối tượng.

Có hai kiểu mô tả đối tượng đó là: mô tả tham số (nhận dạng theo tham số) và mô tả theo cấu trúc (nhận dạng theo cấu trúc).

1.4.11 Nén ảnh

Dữ liệu ảnh cũng như các dữ liệu khác cần phải lưu trữ hay truyền đi trên mạng, lượng thông tin để biểu diễn cho một ảnh là rất lớn . Do đó làm giảm lượng thông tin hay nén dữ liệu là một nhu cầu cần thiết.

1.5 Các định dạng cơ bản trong xử lý ảnh

Trong quá trình xử lý ảnh, một ảnh thu nhận vào máy tính phải được mã hóa. Hình ảnh khi lưu trữ dưới dạng tệp tin sẽ được số hóa. Một số dạng ảnh đã được chuẩn hóa như: ảnh GIF, BMP, PCX, IMG, TIFF...

- Ảnh IMG: Là ảnh đen trắng, phần đầu của ảnh có 16 byte chứa các thông tin cần thiết, ảnh IMG được nén theo từng dòng. Mỗi dòng bao gồm các gói (pack). Các dòng giống nhau cũng nén thành một gói.

- Ảnh PCX: Định dạng ảnh PCX là một trong những định dạng ảnh cổ điển nhất, nó thường được dùng để lưu trữ ảnh, nó sử dụng phương pháp mã loại dài RLE (Run-Length-Encoded) để nén dữ liệu ảnh, quá trình nén và giải nén được thực hiện trên từng dòng ảnh.

- Ảnh TIFF: Là ảnh mà dữ liệu chứa trong tệp thường được tổ chức thành các nhóm dòng (cột) quét của dữ liệu ảnh.

- Ảnh GIF (Graphics Interchanger Format): Với định dạng ảnh GIF những vướng mắc mà các định dạng khác gặp phải khi số trong ảnh tăng lên không còn nữa. Dạng ảnh GIF cho chất lượng cao độ phân giải đồ họa cũng đạt cao, cho phép hiển thị trên hầu hết các phần cứng.

CHƯƠNG 2. PHÂN ĐOẠN ẢNH DỰA VÀO NGƯỠNG

2.1 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

- 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.

2.2 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.

2.3 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:

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

Trong đó w thường được chọn là 3 hoặc 5

2.3.1 Thuật toán đẳng liệu

Đây là kỹ thuật chọn ngưỡng theo kiểu lặp do Ridler và Calvart 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 thì lặp tiếp bước 2.

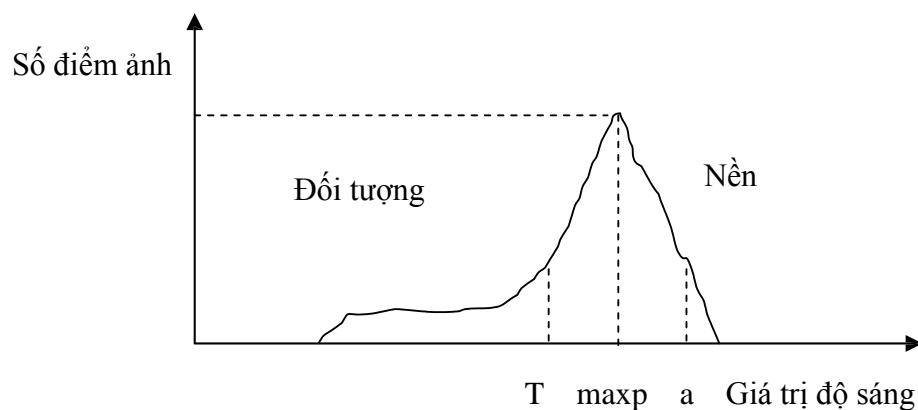
2.3.2 Thuật toán đối xứng nền

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 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 $dp/da \geq 0$.



Hình2. 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 :

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

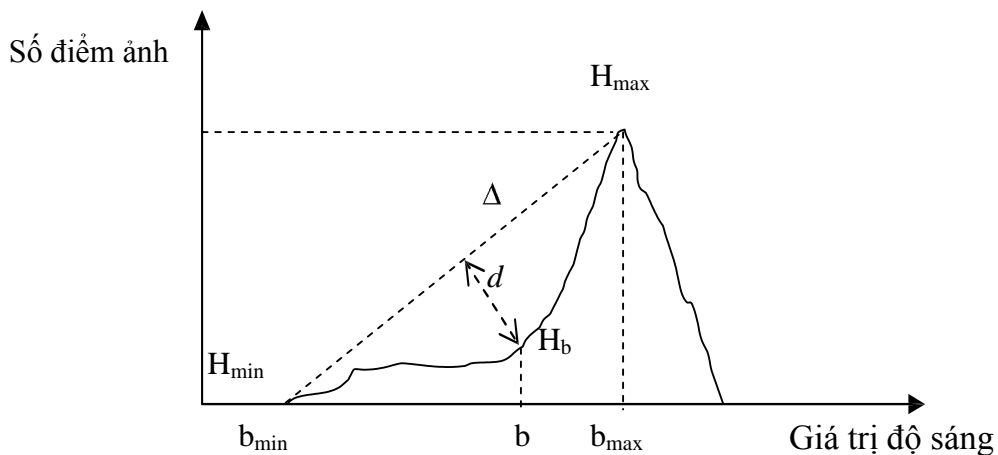
Kỹ thuật này dễ dàng điều chỉnh được cho phù hợp với tình huống ảnh có các đối tượng sáng trên một nền tối.

2.3.3 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 họa thuật toán tam giác bởi hình vẽ như sau:



Hình 3. Minh họa thuật toán tam giác

2.3.4 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:

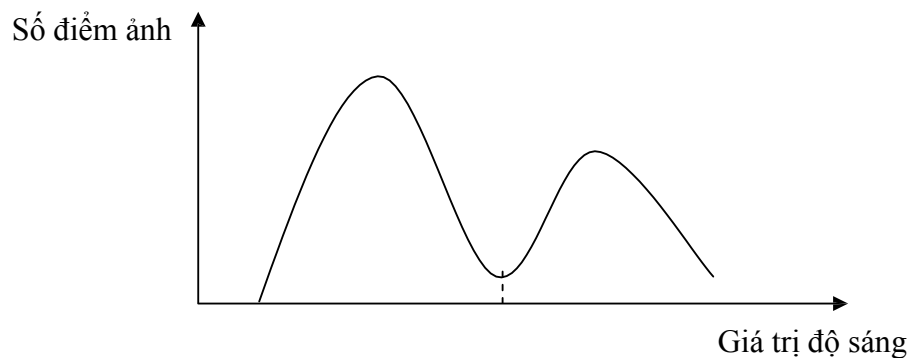
a/ Các đối tượng sáng:

$$\text{TopHat}(A, B) = A - (A \circ B) = A - \max_B(\min_A(A)) \quad (2.3)$$

b/ Các đối tượng tối:

$$\text{TopHat}(A, B) = A - (A \circ B) = A - \min_B(\max_A(A)) \quad (2.4)$$

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.

2.4 Phân ngưỡng tối ưu dựa trên sự không ổn định của lớp và tính đồng nhất của vùng

2.4.1 Giới thiệu

Phần này trình bày một phương pháp được nghiên cứu gần đây bởi Punam K.Shaha, Jayaram K.Udupa, đó là cách phân đoạn bằng cách chọn ngưỡng tối ưu thông qua sự không ổn định của lớp và tính đồng nhất của vùng.

Đặc điểm chung của các phương pháp dùng Histogram là không tận dụng được các thuộc tính trung gian của ảnh như cường độ sáng và không gian của ảnh. Trong thực tế rất khó để chọn ra được ngưỡng tối ưu chỉ từ histogram mà không xem ảnh. Trong khi đó phương pháp dựa vào không gian lại cho ta hình dạng các đối tượng rất rõ ràng. Do đó người ta đã phát triển những phương pháp, trong đó sử dụng đến hình thái của ảnh. Trong phần này ta trình bày một phương pháp như vậy.

Biên của đối tượng trong bất kì ảnh nào cũng rất mờ do quá trình thu nhận ảnh gây ra những hiện tượng nhòe hay không đồng đều về những hiệu suất thu trên các vùng của thiết bị thu nhận ảnh. Hầu hết các phương pháp phân ngưỡng đều tận dụng một phần của tính không chắc chắn khi một điểm thuộc vào một lớp hay tiêu chuẩn entropy để chọn ngưỡng tối ưu. Do đó có thể thấy rằng nếu tiêu chuẩn tối ưu được đề ra, sau đó tìm ngưỡng tối ưu thì các phần tử ảnh trong vùng lân cận biên đối tượng sẽ có giá trị của độ không ổn định cao.

Điều này rất khó thực hiện được nếu không có những hiểu biết về đối tượng hoặc các thông tin phụ về đối tượng từ ảnh. Một số tiêu chuẩn đơn giản về sự đồng nhất của cường độ có thể được dùng để thu thập những thông tin sơ bộ về đối tượng. Một tập các tiêu chuẩn tối ưu trong đó sử dụng những tiêu chuẩn về độ không ổn định và tiêu chuẩn về tính thuần nhất vùng sẽ cho ta những dấu hiệu của ngưỡng mà tại đó các phần tử ảnh trong lân cận biên sẽ có giá trị hàm tối ưu đạt giá trị cao tại ngưỡng tối ưu.

2.4.2 Cơ sở lý thuyết và thuật toán

2.4.2.1 Ký hiệu và định nghĩa

Chia không gian Euclidean R^n thành các hình hộp bởi n họ mặt phẳng trực giao, mỗi họ chứa các mặt phẳng song song cách đều. Mặc dù phương pháp này có thể mở rộng cho các ảnh lưới không đẳng hướng, nhưng để đơn giản ta giả sử các ảnh được áp dụng có độ phân giải đẳng hướng. Ta gọi các siêu hộp là các Spels (viết tắt của spatial elements) và giả sử rằng các mặt phẳng liên tiếp trong cùng một họ mặt phẳng cách đều nhau cùng một đơn vị khoảng cách. Ta xây dựng một hệ tọa độ sao cho tâm của mỗi Spels có tọa độ là (c_1, c_2, \dots, c_n) với các thành phần c_i là các số nguyên. Từ đó ta có một tương ứng 1-1 giữa các Spels với không gian Z^n như sau: một n -bộ c trong không gian Z^n được sơ đồ hóa bởi một Spels sao cho c_j ($1 \leq j \leq n$) là tọa độ thứ j của tâm Spels tương ứng với c . Từ giờ ta có thể coi Z^n như một tập các Spels trong R^n và sẽ sử dụng các khái niệm Spels và tọa độ tâm của chúng thay cho nhau.

Quy ước với mỗi tập con mờ (một quan hệ mờ) của một tham chiếu X có một hàm μ_A là hàm thành phần của A trong X (tương ứng trong $X \times X$). Giá trị hàm μ_A nằm trong khoảng $[0, 1]$. Một quan hệ mờ α trong Z^n được gọi là một quan hệ liên kề mờ giữa các spel nếu nó có tính phản xạ và đối xứng. Có thể thấy rằng α thỏa mãn: giá trị hàm thành phần của nó $\mu_\alpha(c, d)$ ($\forall c, d \in Z^n$) là một hàm không tăng của khoảng cách $\|c - d\|$ giữa c và d với $\|\cdot\|$ là chuẩn trong R^n . Để đơn giản trong tính toán, ta giả thiết μ_α được tính như sau:

- Với mỗi spel c , $\mu_\alpha(c, c) = 1$
- $\forall c, d \in Z^n$, $\mu(c, d) = 1$ nếu c, d khác nhau đúng 1, các trường hợp còn lại $\mu(c, d) = 0$

Ta gọi cặp (Z^n, α) là một không gian số mờ. Không gian số mờ là một khái niệm mô tả hệ thống lưới số cơ bản độc lập của bất kỳ quan niệm nào liên quan tới ảnh.

Một cảnh trong một không gian số mờ (Z^n, α) là một cặp $C = (C, f)$ với $C = \{c \mid -b_j \leq c \leq b_j, \text{ với } b \in Z_+^n\}$, f là một hàm mà có miền xác định là C , gọi là miền ảnh và có miền giá trị là tập hợp các số nguyên thuộc đoạn $\Gamma = [\text{MIN}, \text{MAX}]$. Để tránh các điều kiện không xác định, ta sẽ tìm ngưỡng tối ưu T bên trong khoảng Γ trừ hai điểm tận cùng của đoạn. Tương ứng với một ảnh từ bốn bức xám trở lên, ta chọn: $\Gamma = [\text{MIN} + 2, \text{MAX} - 1]$

Phương pháp mô tả ở đây hoàn toàn có thể được áp dụng cho ảnh vector, có nghĩa là khi f là một hàm có giá trị vector. Nhưng để cho đơn giản ta sẽ coi f như một hàm vô hướng. Một ảnh trên không gian (Z^n, α) sẽ được coi $\Gamma = [0, 1]$.

2.4.2.2. Độ không ổn định của lớp dựa vào cường độ

Nguyên lý về độ không ổn định của lớp dựa trên cường độ đã được rất nhiều nhà nghiên cứu sử dụng trong các phương pháp chọn ngưỡng tối ưu trước đây. Trong phương pháp mới này, người ta đã đưa ra các công thức chung.

Ý tưởng đằng sau khái niệm độ không ổn định của lớp dựa trên cường độ là để xác định độ không ổn định của việc phân lớp một spel vào một lớp đối tượng nào đó dựa trên các lớp khác nhau. Dù vấn đề có thể khái quát thành bài toán xác định đa ngưỡng để phân thành nhiều lớp, song ta sẽ chỉ xem xét đến bài toán phân lớp hai đối tượng, tức là phân một ảnh thành đối tượng và nền. Việc mở rộng cho bài toán nhiều đối tượng sẽ được tìm hiểu tiếp theo.

Để xác định độ không ổn định của lớp, ta giả sử đã biết các xác suất tiên nghiệm của hàm phân bố cường độ của đối tượng và nền và xác suất để một spel thuộc vào đối tượng. Gọi F_0 và F_b lần lượt là lớp đối tượng và nền, θ là xác suất để một spel thuộc vào lớp đối tượng còn $1 - \theta$ là xác suất để một spel thuộc vào lớp nền. Gọi $p_0(g)$ là xác suất để một spel c của đối tượng có cường độ g . Tức là ta có:

$$p_0(g) = P(f(c)=g \mid c \in F_0) \quad (2.5)$$

Gọi $p_b(g)$ là xác suất để một spel của nền có cường độ g . Tức là ta có:

$$p_b(g) = P(f(c)=g | c \in F_b) \quad (2.6)$$

Như vậy nếu gọi $p(g)$ là xác suất để một spel có cường độ g thì:

$$p(g) = \theta p_0(g) + (1-\theta)p_b(g) \quad (2.7)$$

Suy ra theo quy tắc Bayes ta có xác suất để một spel có giá trị xám g thuộc vào lớp đối tượng là:

$$P(c \in F_0 | f(c) = g) = \frac{\theta p_0(g)}{p(g)} \quad (2.8)$$

Còn xác suất để một spel có giá trị xám g thuộc vào lớp nền là:

$$P(c \in F_b | f(c) = g) = \frac{(1-\theta) p_b(g)}{p(g)}$$

Theo Shannon và Weaver, sau khi biết giá trị cường độ tại spel c là g thì độ không chắc chắn của việc phân lớp c vào đối tượng hay nền là entropy của hai xác suất hậu nghiệm (hai công thức trên).

Ta gọi tắt độ không ổn định của lớp dựa trên cường độ là độ không ổn định lớp. Độ không ổn định lớp $H(g)$ tại cường độ g được cho bởi:

$$H(g) = - \frac{\theta p_0(g)}{p(g)} \log \frac{\theta p_0(g)}{p(g)} - \frac{(1-\theta) p_b(g)}{p(g)} \log \frac{(1-\theta) p_b(g)}{p(g)} \quad (2.9)$$

Từ (2.8) và (2.9), nếu ta biết θ , p_0 và p_b thì độ không chắc chắn có thể tính được tại bất kỳ cường độ nào. Với mỗi ảnh $C = (C, f)$ trên không gian (Z^n, α) vì khoảng cách Γ của f là một tập hữu hạn các số nguyên nên xác suất p_0, p_b xác định từ công thức (2.5) và (2.6) cũng là các hàm mật độ của biến ngẫu nhiên g . Trong các nghiên cứu gần đây, các hàm này giả sử được phân bố bởi chuẩn Gauss G_{m_0, σ_0} và G_{m_b, σ_b} với kỳ vọng và phương sai tương ứng là $m_0, m_b, \sigma_0, \sigma_b$ và được ước lượng từ ảnh đã cho như một hàm của ngưỡng như sau:

Giả sử $t \in \Gamma$ là một ngưỡng cường độ bất kỳ. Đặt $F_{0,t}$ và $F_{b,t}$ tương ứng là tập các spel thuộc lớp đối tượng và lớp nền được cho bởi ngưỡng t . Tức là:

$$F_{0,t} = \{ c | c \in C \text{ và } f(c) \geq t \} \quad (2.10)$$

$$F_{b,t} = \{ c \mid c \in C \text{ và } f(c) < t \} \quad (2.11)$$

Đặt $m_0(t)$ và $m_b(t)$ tương ứng là giá trị trung bình của cường độ các spel trong tập $F_{0,t}$ và $F_{b,t}$, $\sigma_o(t)$ và $\sigma_b(t)$ là độ lệch trung bình. Đặt $\theta(t) = \frac{|F_{0,t}|}{|C|}$ trong đó $|X|$ là lực lượng của tập X . Chú ý rằng $F_{b,t} = \Phi$ khi $t = \text{Min}$, $\sigma_b(t) = 0$ khi $t = \text{Min}+1$ và $\sigma_o(t) = 0$ khi $t = \text{Max} - 1$. Đó là lý do tại sao ta cần $t \in \Gamma$. Các hàm mật độ với chỉ số t biểu thị sự phụ thuộc của chúng vào ngưỡng được tính như sau:

$$p_{0,t}(g) = \frac{1}{\sqrt{2\pi\sigma_o}} e^{-\frac{(g-m_0)^2}{2\sigma_o^2}} \quad \text{với } g \in \Gamma \quad (2.12)$$

$$p_{b,t}(g) = \frac{1}{\sqrt{2\pi\sigma_b}} e^{-\frac{(g-m_b)^2}{2\sigma_b^2}} \quad \text{với } g \in \Gamma \quad (2.13)$$

$$p_t(g) = \theta(t) p_{0,t}(g) + (1 - \theta(t)) p_{b,t}(g) \quad \text{với } g \in \Gamma \quad (2.14)$$

Từ (2.9) độ không ổn định của lớp được tính hàm của ngưỡng t đối với $\forall g \in \Gamma$ như sau:

$$H_t(g) = - \frac{\theta(t) p_{0,t}(g)}{p_t(g)} \log \frac{\theta(t) p_{0,t}(g)}{p_t(g)} - \frac{(1-\theta(t)) p_{b,t}(g)}{p_t(g)} \log \frac{(1-\theta(t)) p_{b,t}(g)}{p_t(g)} \quad (2.15)$$

Từ (2.10) và (2.15) cho ta một thuật toán để tính trên máy tính một thuật toán không ổn định $H_t(g)$ đối với mỗi cường độ g trong Γ , tại bất kỳ ngưỡng t nào trong Γ .

2.4.2.3. Tính thuần nhất vùng

Chúng ta nghĩ về khái niệm độ thuần nhất trên vùng (hay gọi tắt là thuần nhất) như một đặc tính của mỗi spel trong ảnh đã cho $C = (C, f)$ trên không gian (Z^n, α) . Có nghĩa là tính thuần nhất μ_τ là một hàm $\mu_\tau: C \rightarrow [0,1]$. Nó phụ thuộc vào quan hệ liên kề giữa các spel α và một quan hệ mờ khác trên C là ψ được gọi là độ đồng nhất mờ. Khi mối liên kết $\mu_\psi(c,d)$ càng lớn thì cường độ ảnh trong vùng lân cận của c và d trong C càng giống nhau. Ta sử dụng phương pháp tính

toán dựa trên phạm vi đồng nhất để xác định Ψ sẽ mô tả sau. Dạng hàm của μ_τ mà ta sử dụng trong chương trình này như sau:

$$\mu_\tau(c) = \frac{\sum_{d \in C} \mu_\alpha(c, d) \mu_\psi(c, d)}{\sum_{d \in C} \mu_\alpha(c, d)} \quad (2.16)$$

Nói cách khác độ đồng nhất tại spel c là trung bình trọng số của các lực hấp dẫn đối với c của các spels trong C . Do tính liên kết giữa các spel ở xa nhau là bằng 0 nên các spel thực sự cần quan tâm trong công thức (2.16) là các spel trong một lân cận gần của c .

Nguồn gốc của ý tưởng lập công thức dựa vào phạm vi đồng nhất giữa hai spel c và d là để tận dụng khái niệm kích thước của một cấu trúc địa phương hay phạm vi tại c và d . Phạm vi trong một ảnh C tại mỗi spel c được định nghĩa là bán kính $r(c)$ của hình cầu lớn nhất có tâm tại c nằm hoàn toàn trong cùng một vùng đối tượng được xác định dựa trên tiêu chuẩn xấp xỉ độ đồng nhất cường độ. Bằng trực giác ta thấy để phân đoạn ảnh, trước hết cần xác định phạm vi của đối tượng. Người ta đưa ra một thuật toán đơn giản nhưng rất hiệu quả, trong đó ước lượng $r(c)$ tại mỗi spel $c \in C$ trong một ảnh bất kỳ không cần phân đoạn trực tiếp ảnh đó mà dựa trên sự liên tục của tính đồng nhất về cường độ.

Xác định phạm vi:

Một hình cầu $B_k(c)$ tâm tại c , bán kính k được định nghĩa như sau:

$$B_k(c) = \{e \in C \mid \|c - e\| \leq k\} \quad (2.17)$$

Với mỗi hình cầu $B_k(c)$ ta xác định một phân số $FO_k(c)$, biểu diễn tỉ lệ giữa tập các spel trên biên hình cầu có cường độ gần như đồng nhất với c và tập các điểm trên biên:

$$FO_k(c) = \frac{\sum_{e \in B_k \setminus B_{k-1}} W_\psi(e) \mu_\alpha(c, e)}{|B_k \setminus B_{k-1}|} \quad (2.18)$$

Trong đó W_ψ là hàm thành phần tương ứng với mệnh đề mờ: “ x is small”. Có nhiều cách chọn hàm này có thể áp dụng được, nhưng trong bài này ta chọn hàm

phân bố chuẩn Gauss với kỳ vọng bằng 0 và độ lệch chuẩn σ_ψ . Sau đây là thuật toán ước lượng phạm vi của đối tượng:

Thuật toán OSE

Input: $C, c \in C, W_\psi$, một ngưỡng cố định t_s

Output: $r(c)$

Begin

Set $k = 1$;

While $FO_k(c) \geq t_s$ do

Set k to $k + 1$

Endwhile;

Set $r(c)$ to $k - 1$;

Output $r(c)$;

End.

Thuật toán này lặp tăng bán kính cầu k lên 1, bắt đầu từ 1, và kiểm tra $FO_k(c)$ xem đối tượng chứa c có nằm trong biên hình cầu không. Ngay lần đầu tiên khi phân số này nhỏ hơn t_s , ta xem như hình cầu rơi ra ngoài đối tượng đang xét và vào vùng của đối tượng khác. Ta thường dùng $t_s = 0.85$.

Tham số σ_ψ được xác định như sau: Trên toàn miền C , các hiệu cường độ địa phương $|f(c) - f(d)|$ được tính cho tất cả các cặp (c, d) với $c \neq d$ và $\mu_\alpha(c, d) > 0$. Trên 10% các giá trị này bị loại đi do phân biên chung giữa các đối tượng. Còn lại dưới 90% các giá trị này dùng để tính kỳ vọng M_h và độ lệch trung bình σ_h . Khi đó ta tính $\sigma_\psi = M_h + 3 \sigma_h$.

Lý do căn bản của sự lựa chọn này là trong phân phối chuẩn, ba lần độ lệch chuẩn về cả hai phía của giá trị trung bình chiếm tới 99,7% tổng số.

2.4.2.4. Xác định tính tương đồng

Để xác định sự giống nhau giữa hai spel $c, d \in C$, ta xét hai hình cầu: một có tâm tại c , một có tâm tại d , ký hiệu lần lượt là $B_{cd}(c)$ và $B_{cd}(d)$, cùng có bán kính bằng $\min[r(c), r(d)]$:

$$B_{cd}(c) = \{ e \in C \mid \|c - e\| \leq \min[r(c), r(d)] \} \quad (2.19)$$

$$B_{cd}(d) = \{ e \in C \mid \|d - e\| \leq \min[r(c), r(d)] \} \quad (2.20)$$

Hai hình cầu này được sử dụng như là các lân cận của c, d và có kích thước được xác định bởi các bán kính $r(c), r(d)$.

Để xác định μ_ψ , chỉ xét những điểm $c, d: \mu_\alpha(c, d) > 0$. Xét các điểm $e \in B_{cd}(c)$ và $e' \in B_{cd}(d)$ (đó là các điểm tương ứng với nhau trên $B_{cd}(c)$ và $B_{cd}(d)$) sao cho $c - e = d - e'$. Ta sẽ xác định hai tổng trọng số $D^+(c, d)$ và $D^-(c, d)$ của sự khác nhau về cường độ giữa hai hình cầu. Ta tính các hàm sau:

Nếu $f(e) - f(e') > 0$

$$\delta_{cd}^+(e, e') = \begin{cases} f(e) - f(e') \\ 0 \end{cases} \quad (2.21)$$

Nếu khác

Nếu $f(e') - f(e) < 0$

$$\delta_{cd}^-(e, e') = \begin{cases} f(e') - f(e) \\ 0 \end{cases} \quad (2.22)$$

Nếu khác

Ta tính:

$$D^+(c, d) = \sum_{\substack{e \in B_{cd}(c) \\ e' \in B_{cd}(d) \\ c - e = d - e'}} 1 \cdot W_\psi \delta_{cd}^+(e, e') \quad (2.23)$$

$$D^-(c, d) = \sum_{\substack{e \in B_{cd}(c) \\ e' \in B_{cd}(d) \\ c - e = d - e'}} 1 \cdot W_\psi \delta_{cd}^-(e, e') \quad (2.24)$$

Với ω_{cd} là một hàm thành phần ứng với mệnh đề mờ: " x is small ". Tương tự như W_{ψ} , hàm Gauss với giá trị trung bình 0 và độ lệch tiêu chuẩn $\sigma_{\omega_{cd}} = \text{Min} \{r(c), r(d)\}$ được dùng để tính ω_{cd} .

Kết hợp các đẳng thức trên với quan hệ tương đồng ψ như sau: Có hai loại cường độ quanh c, d là biến nội và biến liên kết. Thành phần biến nội nói chung là ngẫu nhiên, do đó thường gần như toàn bộ bằng 0. Trái lại thành phần liên kết thì có hướng. Nó tăng hay giảm theo hướng được cho bởi c - d và độ lớn hơn biến nội. Do đó rất hợp lý nếu giả sử rằng số nhỏ hơn trong hai số $D^+(c, d)$ và $D^-(c, d)$ sẽ biểu diễn cho biến nội và số còn lại biểu thị cho tác động liên kết của hai thành phần.

Chú ý rằng: khi δ_{cd}^+ nhỏ thì $D^+(c, d)$ nhỏ, δ_{cd}^- nhỏ thì $D^-(c, d)$ cũng nhỏ.

Nếu có một vài thành phần của nền nằm trong lân cận đang xét thì thành phần đó không gây ra những ảnh hưởng đáng kể so với thành phần liên kết. Điều này dẫn ta đến dạng hàm cho μ_{ψ} .

$$\mu_{\psi}(c, d) = 1 - \frac{|D^+(c, d) - D^-(c, d)|}{\sum_{e \in B_{cd}} \omega_{cd} \|c - e\|} \quad (2.25)$$

Chú ý rằng: $|D^+(c, d) - D^-(c, d)|$ biểu thị cho độ không đồng nhất của vùng chứa c và d. Giá trị này thấp khi cả c và d nằm trong vùng của đối tượng, giá trị này cao khi cả c và d nằm ở lân cận hoặc dọc trên biên.

Từ (2.16), (2.19) và (2.25) định nghĩa về độ đồng nhất $\mu_r(c)$ tại mỗi spel c trong một ảnh cho trước đã hoàn chỉnh. Cần nhấn mạnh lại rằng định nghĩa này chỉ phụ thuộc duy nhất vào tham số σ_{ψ} .

2.4.2.5. Chọn ngưỡng tối ưu

Trong phần này ta mô tả phương thức chọn ngưỡng dự kiến thỏa mãn cả độ không chắc ổn định và tính thuần nhất vùng. Định hướng và cơ sở cho phương pháp này là mệnh đề sau:

Mệnh đề A: " Tại mỗi cạnh (ảnh - scene) với biên mờ, tại mỗi đoạn chia tối ưu của lớp đối tượng, điểm với độ không ổn định cao xuất hiện trong lân cận của biên đối tượng ".

Quả thực, khó mà chứng minh hay phủ nhận mệnh đề trên bằng toán học chỉ dựa trên những phát biểu của mệnh đề đó. Tuy nhiên, bằng trực giác ta có thể biện luận như sau:

Khi p_0 và p_b là các hàm phân bố Gauss, ngưỡng tối ưu t_0 là tại điểm giao của hai hàm phân bố, nghĩa là t_0 thỏa mãn $p_0(t_0) = p_b(t_0)$

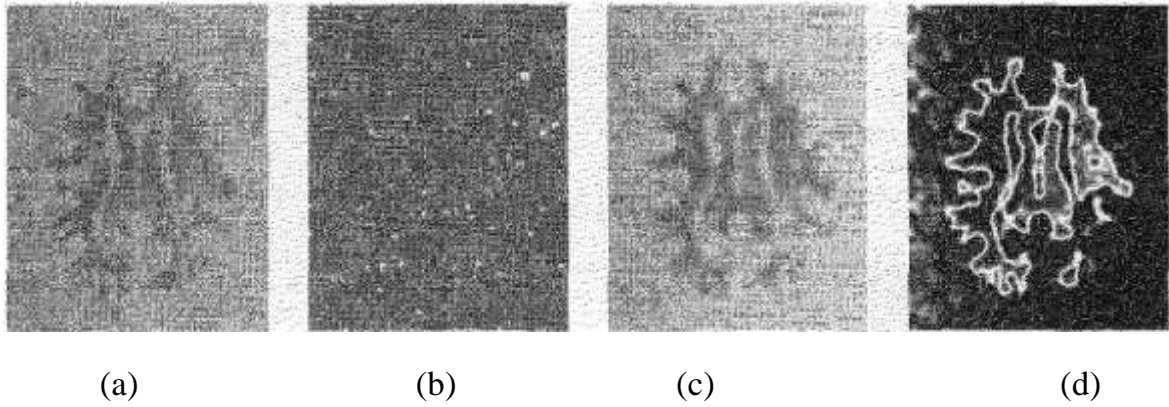
Điều này dựa trên lý thuyết nhận dạng đối tượng đối với sự phân lớp các nét đặc trưng dựa trên cực tiểu tỷ lệ lỗi.

Rõ ràng từ (2.15), cường độ ảnh trong lân cận t_0 tạo ra sự không ổn định cao từ ngưỡng t_0 .

Trước hết chúng ta hãy xem xét một ảnh với biên không mờ, trình bày trong hình (a) dưới đây:

Cảnh trong hình này được tạo ra bằng cách thêm vào nhiễu Gauss có kỳ vọng 0 cho một ảnh được tạo ra bằng cách phân đoạn thô (bằng tay) vùng ảnh trắng trên lát cắt MRI của đầu bệnh nhân. Trước khi thêm nhiễu vào, giá trị xám trong hai vùng được đặt tương ứng với giá trị xám trung bình trong vùng WM và vùng xám (gray matter). Chú ý rằng không có hiện tượng nhòe được thêm vào đây, vì đây là ảnh thật nên ngưỡng tối ưu giả định có thể xác định bằng cách xem xét một cách toàn diện để tìm những phần chồng lên nhau cực đại giữa vùng ảnh trắng và vùng ảnh kết quả từ phân ngưỡng.

Hình (b) mô tả độ không ổn định của ảnh trong hình (a) như là một ngưỡng tối ưu, với các vùng sáng nhất biểu thị tính không ổn định cao.



Hình : Minh họa các nguyên lý của mệnh đề A: ảnh thu được bằng cách phân đoạn ảnh lát cắt MRI.

Như đã chỉ ra trên hình, vùng có độ không ổn định cao được phân rải rác một cách ngẫu nhiên trong toàn ảnh và nó không xảy ra trong lân cận biên của đối tượng.

Nhưng ví dụ này không mâu thuẫn với mệnh đề trên vì giả thiết của mệnh đề này đòi hỏi một ảnh có biên mờ, có nghĩa là cường độ thay đổi trơn qua giao diện giữa hai vùng đối tượng.

Bây giờ, xét một ảnh có biên mờ minh họa trên hình trên. Cảnh trong hình (c) được tạo ra cùng cách với hình (a) trừ việc làm mờ, ngoài ra cường độ nền được thêm vào trước khi thêm nhiễu. Bây giờ không khó khăn gì thấy rằng các pixel ở lân cận biên có cường độ nằm trong khoảng của trung bình cường độ đối tượng và trung bình cường độ nền. Do đó ta có một lớp có độ không ổn định cao tại mỗi ngưỡng thích hợp.

Hình (d), ảnh của lớp không ổn định tại ngưỡng tối ưu, minh họa một thực tế rằng, các pixel với độ không ổn định cao (cường độ cao) xuất hiện ở lân cận biên.

Dù mệnh đề phát biểu cho các ảnh có biên mờ hoặc bị nhòe nhưng nó cũng áp dụng được cho các ảnh thu được từ hầu hết các thiết bị ảnh vì các ảnh đó đều có biên mờ do độ phân giải hạn chế, hiệu ứng không toàn bộ và tài liệu không đồng nhất

Từ mệnh đề A ta thấy, để xác định một ngưỡng tối ưu càng gần với ngưỡng tối ưu lý thuyết càng tốt, ta cần một dạng thông tin về tính thuần nhất vùng. Ta có thể tạo ra tiêu chuẩn thỏa mãn giả thiết rằng nó sẽ có giá trị cao khi cả độ không ổn định lớp và tính thuần nhất cùng cao hay thấp.

Mục đích của chúng ta là tạo ra một hàm ngưỡng năng lượng tiêu chuẩn E mà cực tiểu của hàm này sẽ cho ta ngưỡng tối ưu.

Hàm E có thể bao gồm cả nhân tố về độ không ổn định và nhân tố về độ thuần nhất. Trong phần tiếp theo ta sẽ mô tả công thức của E theo những định hướng trên.

Thay vì tính trực tiếp $\mu_\tau(c)$, chúng ta sẽ dùng khoảng thông thường $\mu_\varphi(c)$ của những giá trị này trong ảnh để biểu thị độ đồng nhất tại c.

$$\mu_\varphi(c) = \frac{LC(\mu_\tau)}{LC(\mathbb{1})} \quad (2.26)$$

$$\text{Với } LC(x) = \sum_{y \in Y, y \leq x} L(y) \quad (2.27)$$

trong đó: $L(y)$ là số các spel $c \in C$ trong ảnh với $\mu_\tau(c) = y$ và Y là tập tất cả các $z \in [0, 1]$ sao cho tồn tại $c \in C$ thỏa mãn $\mu_\tau(c) = z$. Chú ý rằng Y là một tập rời rạc với lực lượng hữu hạn. Sự chuẩn hóa dựa trên khoảng độ đồng nhất cho bởi $\mu_\varphi(c)$ làm cho nó đỡ nhạy cảm hơn so với cách chọn thực tế của tham số σ_ψ .

Cuối cùng ta định nghĩa ngưỡng năng lượng cho mỗi $t \in \Gamma$ và mỗi lớp $C = (C, f)$ trong không gian (Z^n, α) , như sau:

$$E(t) = \sum_{c \in C} H_t(\mu_\varphi) + \mathbb{1} - H_t(\mu_\varphi) \quad (2.28)$$

Khi cả hai đại lượng là tính không ổn định và tính thuần nhất vùng đều cao thì $E(t)$ cũng cao và thành phần thứ nhất trong công thức (2.28) đóng vai trò chủ đạo. Còn khi cả hai thành phần đều thấp thì $E(t)$ vẫn cao nhưng thành phần thứ hai trong công thức (2.28) chiếm ưu thế. Tương tự khi một trong hai thành phần thấp còn thành phần kia cao thì $E(t)$ thấp. Bởi vậy, tại mỗi ngưỡng t , $E(t)$ biểu diễn tổng

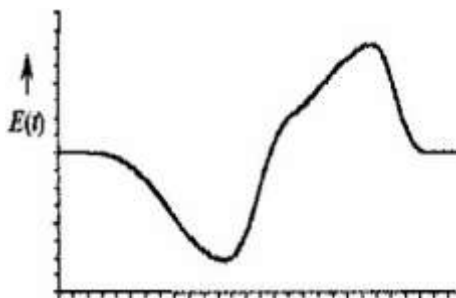
sự khác nhau của giá trị spel trong toàn vùng với việc chú trọng tới hàm phân bố cường độ và độ thuần nhất vùng. Điều này cũng giống như các nguyên lý trong mệnh đề A đã cho thấy rằng biên của đối tượng có quan hệ với tập hỗn độn cao. Mục đích của ta là đi tìm một ngưỡng t tại đó $E(t)$ đạt cực tiểu. Tức là tìm:

$$t_{MHUE} = \arg \min\{E(t) \mid t \in \Gamma\} \quad (2.29)$$

Vì $E(t)$ biểu thị năng lượng dựa trên cả tính thuần nhất vùng và tính không ổn định nên chúng ta đề cập tới phương pháp này để xác định ngưỡng tối ưu như một phương thức cực tiểu hóa năng lượng dựa trên tính thuần nhất và tính không ổn định.

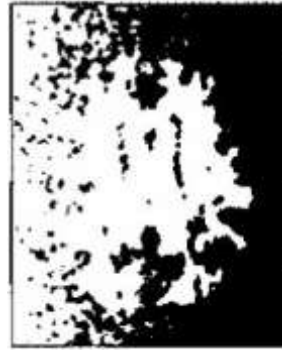
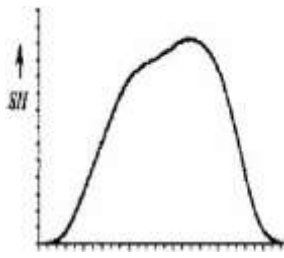
2.4.2.6. Kết luận

Một phương pháp phân đoạn ảnh mới đã được đưa ra. Trong phương pháp này xác định ngưỡng tối ưu bằng cách xem xét đến hình thái của ảnh. Phương pháp này đã kết hợp một cách hiệu quả những thông tin về độ đồng nhất và thông tin về tính không ổn định khi thuộc vào một lớp của một điểm dựa trên histogram. Nguyên lý cơ bản để thực hiện công việc này là dựa trên mệnh đề phát biểu rằng trong hầu hết những ảnh trong thực tế, tại ngưỡng tối ưu, các phần tử có độ không ổn định cao nhất thì đều xuất hiện trong lân cận của biên đối tượng. Được dẫn đường bởi mệnh đề này, chúng ta đã chứng minh rằng thông tin về độ đồng nhất có thể được sử dụng như thế nào trong việc kết hợp với độ không ổn định để nâng cao chất lượng của ước cho ngưỡng tối ưu.



(a)





(b)

Hình : a) $E(t)$ là một hàm của ngưỡng t đối với một ảnh nhị phân của ảnh hình (a) ở títt trên.

b) ảnh nhị phân đã được phân ngưỡng tối ưu tìm được bằng phương pháp MHUE (Minimization of Homogeneity and Uncertainty based Energy).

CHƯƠNG 3. PHÂN ĐOẠN THEO MIỀN ĐỒNG NHẤT

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 như: mức xám, màu sắc (đối với ảnh màu), kết cấu sợi vv...

Giả sử ảnh X phải được phân thành N vùng khác nhau R_1, R_2, \dots, R_N và nguyên tắc phân đoạn là một vị từ của công thức $P(R)$. Việc phân đoạn ảnh chia tập X thành các tập con $R_i, i = 1 \dots N$ phải thỏa mãn:

- Các vùng $R_i, i=1..N$ phải lấp kín hoàn toàn ảnh:

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

- Hai vùng khác nhau phải là những tập hợp rời nhau:

$$R_i \cap R_j = 0 \text{ với } i \neq j$$

- Mỗi vùng R_i phải có tính đồng nhất:

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

- Nếu R_i, R_j là hai vùng rời nhau thì $(R_i \cup R_j)$ phải là một vùng ảnh không đồng nhất:

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

Kết quả của việc phân vùng ảnh phụ thuộc vào dạng của vị từ P và các đặc trưng được biểu diễn bởi vectơ đặc trưng. Thường thì vị từ P có dạng $P(R, X, t)$, trong đó X là vectơ đặc trưng gắn với một điểm ảnh và t là một tập hợp các tham số (thường là các ngưỡng). Trong trường hợp đơn giản nhất, vectơ đặc trưng X chỉ chứa giá trị mức xám của ảnh $I(k,l)$ và vectơ ngưỡng chỉ gồm một ngưỡng T . Một nguyên tắc phân đoạn đơn giản có công thức:

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

Trong trường hợp các ảnh màu, vectơ đặc trưng X có thể là ba thành phần ảnh RGB $[f_R(k,l), f_G(k,l), f_B(k,l)]^T$. Lúc đó luật phân ngưỡng có dạng:

$$P(R,x,t): ((f_R(k,l) < T_R) \&\& (f_G(k,l) < T_R) \&\& (f_B(k,l) < T_R))$$

3.2 Phương pháp tách cây tứ phân

Phương pháp tách cây tứ phân dựa trên nguyên tắc kiểm tra tính hợp thức của tiêu chuẩn đồng nhất một cách tổng thể trên miền lớn. Nếu tiêu chuẩn được thoả việc phân đoạn coi như kết thúc. Trong trường hợp ngược lại, chia miền đang xét thành 4 miền nhỏ hơn, áp dụng đệ quy bằng phương pháp trên cho mỗi miền nhỏ hơn cho đến khi tất cả các miền đều thoả mãn tiêu chuẩn đồng nhất.

Thuật toán được mô tả như sau:

Procedure

PhanDoan(Mien) Begin

If miền đang xét không thoả Then

Begin

Chia miền đang xét thành 4 miền: Z1, Z2, Z3, Z4

For i=1 to 4 Do

PhanDoan(Zi) End

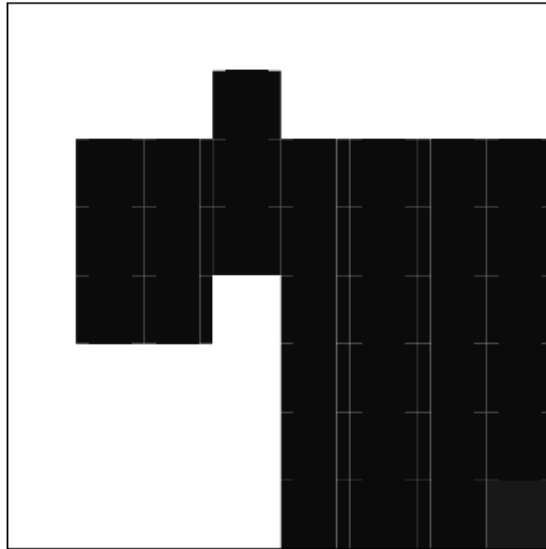
Else Exit

End;

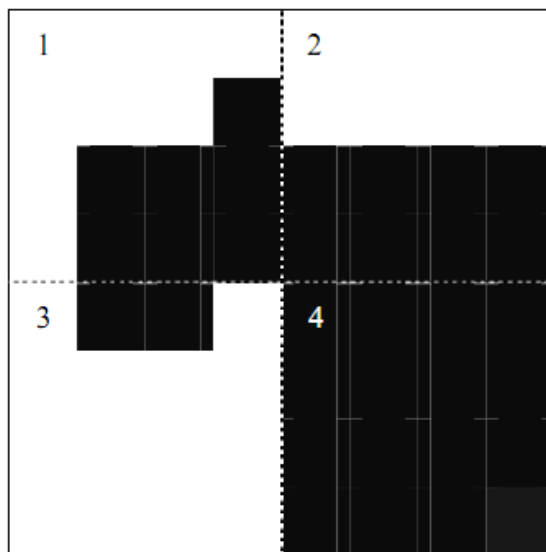
Thuật toán này tạo nên một cây mà mỗi nút cha có 4 nút con ở mọi mức, trừ mức ngoài cùng. Vì thế cây này có tên là cây tứ phân. Gốc của cây là ảnh ban đầu, một vùng thoả tiêu chuẩn tạo nên một nút lá, nếu không sẽ tạo nên một nút nhánh.

Giả sử chọn tiêu chuẩn phân vùng là màu sắc và quy ước 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.

Hình 4.1 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



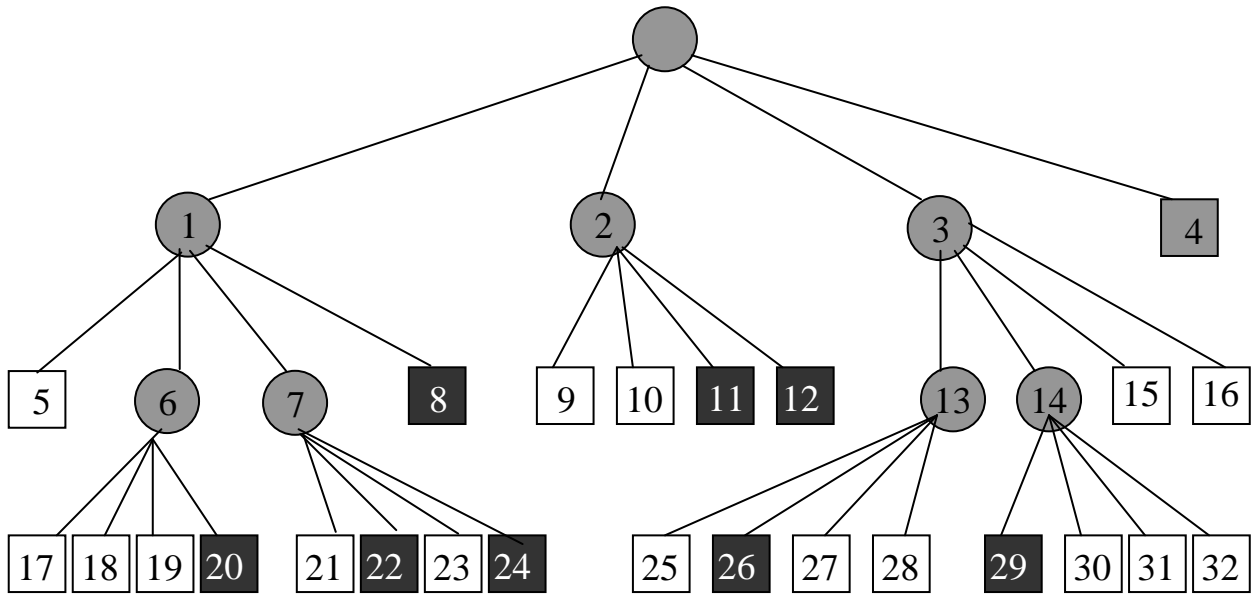
b. phân mức 1

5	6	9	10
7	8	11	12
13	14	4	
15	16		

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



Hình 4. Tách cây tứ phân

3.3 Phương pháp phân vùng hợp

Phương pháp phân vùng bởi hợp thao tác ngược lại với phương pháp tách cây tứ phân, nghĩa là xuất phát từ các miền nhỏ nhất – các điểm ảnh rời hợp chúng lại nếu thỏa mãn tiêu chuẩn đề ra để được miền đồng nhất lớn hơn. Tiếp tục với các miền thu được cho đến khi ta không thể hợp nhất chúng với nhau nữa, lúc này số miền còn lại chính là các phân vùng của ảnh. Việc hợp nhất hai miền phải thỏa mãn hai nguyên tắc sau:

- Hai vùng phải kề cận.
- Hai vùng phải đáp ứng tiêu chuẩn, như cùng màu, cùng mức xám hay cùng kết cấu vv ...

Giả sử vùng R_i có n điểm, lúc đó giá trị trung bình m_i và độ lệch tiêu chuẩn σ_i được tính theo công thức:

$$m_i = \frac{1}{n} \sum_{(k,l) \in R_i} I(k,l) \quad (3.1)$$

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{(k,l) \in R_i} (I(k,l) - m_i)^2} \quad (3.2)$$

Hai vùng R_1 và R_2 có thể hợp thành một vùng nếu $|m_1 - m_2| < T$ và điểm $I(k,l)$ sẽ được hợp với vùng R_i nếu $|I(k,l) - m_i| < T$, với T là một ngưỡng.

Đầu tiên chúng ta cố gắng hợp điểm (k, l) với một trong các vùng lân cận R_i . Nếu việc hợp không thành công thì ta hợp với các vùng khác đã có. Nếu vẫn không thành công hoặc không có vùng lân cận tồn tại thì điểm này được coi là một vùng mới.

Sau khi hợp nhất (k,l) vào vùng R thì ta phải cập nhật lại giá trị trung bình và độ lệch tiêu chuẩn:

$$m_i = \frac{1}{n+1}(I(k,l) + n * m_i) \quad (3.3)$$

$$\sigma_i^2 = \sqrt{\frac{1}{n+1} \left(n\sigma_i^2 + \frac{n}{n+1} (I(k,l) - m_i)^2 \right)} \quad (3.4)$$

Nếu có nhiều hơn một vùng lân cận thỏa mãn thì hợp điểm (k,l) với vùng R_i sao cho sự khác biệt $|I(k,l) - m_i|$ nhỏ nhất.

Cũng trong phương pháp phân vùng bởi hợp, có một cách tiếp cận khác với kỹ thuật trên, đó là phương pháp phân vùng dựa vào đồ thị. Phân vùng dựa trên đồ thị tìm cách hợp nhất hai miền R_i và R_j theo tính chất so sánh giữa hai cặp miền. Thuật toán này được chúng tôi trình bày chi tiết ở chương 3.

3.4 Phương pháp tách hợp (Split- Meger)

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 giảm số vùng liên thông xuống mức tối thiểu nhưng cấu trúc hàng ngang dàn trải, không cho ta thấy mối liên hệ giữa các vùng. Chính vì nhược điểm này ta nghĩ đến việc phối hợp cả 2 phương pháp. Trước tiên dùng phương pháp tách để tạo nên cây tứ phân, phân đoạn theo hướng từ gốc lên 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

1.1. 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 thì chuyển sang bước 1.2

1.2. 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.

3.5 Nhận xét

Các kĩ thuật phân đoạn dựa trên miền đồng nhất sử dụng cả hai yếu tố về không gian và màu sắc nhằm ước lượng tính đồng nhất. Các bước xử lý chung của thuật toán là, chọn ra một vùng hạt nhân (nhỏ), thực hiện thao tác tăng hoặc phân chia vùng từ điểm hạt nhân này, sau đó trộn các vùng đồng nhất và dừng lại khi đã thỏa mãn các tiêu chuẩn về tăng hoặc phân chia vùng. Theo bản chất, quá trình xử lý tuân theo trình tự trên khi mỗi pixel và các điểm ảnh lân cận nó được ước lượng, tuy nhiên, trình tự này sẽ bị thay đổi nếu các điểm có cùng giá trị đồng nhất. Một vấn đề khác nữa là làm sao có thể lựa chọn được vùng hạt nhân phù hợp, do vùng hạt nhân được chọn ban đầu này sẽ chi phối toàn bộ quá trình tăng vùng hay phân chia vùng.

CHƯƠNG 4. PHÂN ĐOẠN DỰA VÀO ĐỒ THỊ

Phân đoạn ảnh dựa vào đồ thị là một phương pháp tiếp cận khá hiện đại dựa trên thuộc tính non-local của ảnh đầu vào. Phương pháp này phát hiện ra biên giữa hai vùng của ảnh bằng cách so sánh sự khác nhau giữa nội vùng (inter-component) với sự khác nhau với các vùng khác. Thuật toán phân đoạn dựa vào đồ thị tuân theo 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.

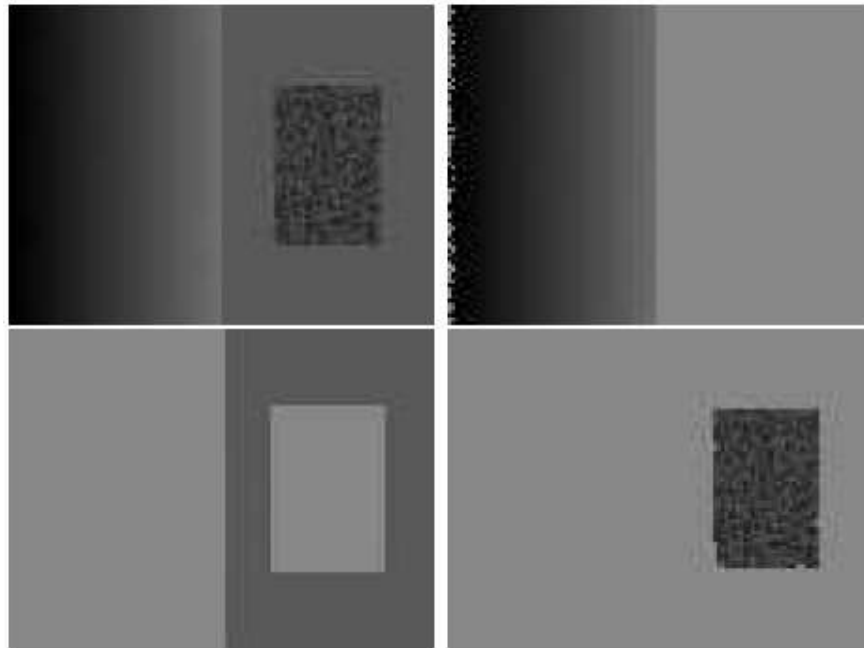
4.1 Giới thiệu

Các phương pháp phân đoạn ảnh cổ điển đều có chung một nhược điểm là chạy rất chậm trong các ứng dụng XLA và hầu như không nắm bắt được các thuộc tính non-local quan trọng của ảnh. Vì vậy, hầu hết các nghiên cứu của những năm gần đây đều có xu hướng tìm kiếm một kỹ thuật phân đoạn có khả năng xử lý trong cơ sở dữ liệu ảnh lớn một cách nhanh chóng, chính xác và hiệu quả. Kỹ thuật phân đoạn dựa vào đồ thị được mô tả ở đây không những vừa nắm bắt được các đặc tính non-local mà độ phức tạp tính toán chỉ là $O(n \log n)$ đối với bức ảnh có n điểm ảnh (pixel).

Giống như các phương pháp phân cụm cổ điển, phương pháp này cũng dựa trên việc chọn các cạnh từ một đồ thị. Đồ thị này được xây dựng bằng cách coi mỗi điểm ảnh là một đỉnh, hai điểm ảnh kề nhau thì được nối bởi một cạnh vô hướng, trọng số trên một cạnh thể hiện sự khác nhau giữa hai điểm ảnh. Tuy nhiên, phương pháp này thực hiện việc điều chỉnh sự phân đoạn dựa vào mức độ thay đổi giữa các miền lân cận của ảnh.

Lấy một ví dụ đơn giản thể hiện việc nắm bắt được các đặc tính non-local của phương pháp này. Hãy để ý vào ảnh phía trên bên trái của hình ; hầu hết ta đều nói rằng bức ảnh này có ba miền phân biệt: một hình chữ nhật ở nửa bên trái, một hình chữ nhật đặc ở giữa nửa bên phải và phần bao quanh hình chữ nhật đặc này. Vì sao ta khẳng định được như thế? Chắc chắn đó là thuộc tính quan trọng của sự

tri giác (perceptually) và chúng tôi tin rằng các đặc trưng này cũng sẽ được nắm bắt bởi thuật toán phân đoạn.



Hình 5. Ví dụ về nhận dạng các vùng ảnh

Phương pháp phân đoạn dựa vào đồ thị sẽ tìm dấu hiệu đường biên giữa hai vùng bằng cách so sánh hai đại lượng: một là dựa vào cường độ khác nhau dọc theo đường biên và hai là dựa vào cường độ khác nhau giữa các điểm ảnh với mỗi vùng.

4.2 Phân đoạn dựa vào đồ thị

Cho $G = (V, E)$ là một đồ thị vô hướng với các đỉnh $vi \in V$, là tập hợp các phần tử cần được phân đoạn và các cạnh $(vi, vj) \in E$, tương ứng với các cặp đỉnh lân cận nhau. Mỗi cạnh $(vi, vj) \in E$ có một trọng số tương ứng, trọng số là một số không âm đo sự khác nhau giữa hai phần tử lân cận vi và vj , ký hiệu $w(vi, vj)$. Ở đây trọng số của các cạnh đo sự khác nhau giữa hai điểm nối bởi cạnh đó (có nhiều mức độ khác nhau: màu sắc, vị trí, sự vận động hoặc các thuộc tính khác).

Như vậy phân đoạn một bức ảnh là việc phân chia V thành các thành phần, mà mỗi thành phần (hoặc miền) $C \in V$ tương đương với một thành phần liên thông trong đồ thị $G' = \langle V, E' \rangle$, $E' \subseteq E$.

4.3 Tính chất của so sánh cặp miền

Để có thể dễ dàng định lượng dấu hiệu của một đường biên giữa hai vùng trong ảnh, chúng ta định nghĩa một tính chất D. Tính chất này dựa vào độ đo sự khác nhau giữa các phần tử dọc theo một đường biên của hai thành phần liên quan nhằm đo sự khác nhau giữa các phần tử lân cận trong mỗi thành phần. Kết quả là so sánh sự khác nhau giữa nội vùng (inter-component) với sự khác nhau với các vùng khác.

Trước hết, ta định nghĩa độ khác nội vùng (internal difference) và độ khác giữa hai vùng (difference between two components).

Độ khác nội vùng (internal difference) của một thành phần $C \subseteq V$ là trọng số lớn nhất trong cây tủa nhánh tối thiểu của thành phần đó, kí hiệu $\text{Int}(C)$. Khi đó:

$$\text{Int}(C) = \max_{e \in \text{MST}(C,E)} w(e) \quad (4.1)$$

Độ khác giữa hai vùng (difference between two components) $C_1, C_2 \subseteq V$, là trọng số nhỏ nhất giữa hai vùng, kí hiệu là $\text{Dif}(C_1, C_2)$. Khi đó:

$$\text{Dif}(C_1, C_2) = \min_{\substack{V_i \in C_1, V_j \in C_2, (V_i, V_j) \in E}} w(V_i, V_j) \quad (4.2)$$

Nếu không có cạnh nối nào giữa C_1 và C_2 thì đặt $\text{Dif}(C_1, C_2) = \infty$. Độ đo sự khác nhau này là về nguyên lý thì vẫn có vẻ mơ hồ, vì nó chỉ phản ánh được cạnh có trọng số nhỏ nhất nối giữa hai thành phần.

Một khái niệm có liên quan trong định nghĩa về tính chất D là giá trị khác nội vùng nhỏ nhất, kí hiệu MInt . Giá trị MInt được định nghĩa như sau:

$$\text{MInt}(C_1, C_2) = \min(\text{Int}(C_1) + \tau(C_1), \text{Int}(C_2) + \tau(C_2)) \quad (4.3)$$

Hàm ngưỡng τ điều khiển mức độ khác nhau giữa hai thành phần, sao cho giá trị này phải lớn hơn các giá trị khác nội vùng của các thành phần để nhằm mục đích nhận ra đường biên giữa chúng. Đối với các thành phần nhỏ $\text{Int}(C)$ là không đủ tốt để ước lượng các đặc tính của dữ liệu. Trong một số trường hợp khi $|C| = 1$ thì $\text{Int}(C) = 0$

với $|C|$ là kích thước của thành phần C . Khi đó chúng ta sử dụng một hàm ngưỡng dựa trên kích thước của thành phần:

$$\tau(C) = k / |C| \quad (4.4)$$

Với k là một tham số hằng. Trong thực tế thì k được chọn không nhỏ hơn kích thước của thành phần nhỏ nhất.

Lúc này tính chất so sánh giữa hai cặp miền C_1 và C_2 , kí hiệu $D(C_1, C_2)$ được định nghĩa như sau:

$$D(C_1, C_2) = \begin{cases} \text{true if } D(C_1, C_2) > \text{MInt}(C_1, C_2) \\ \text{false otherwise} \end{cases} \quad (4.5)$$

4.4 Thuật toán và các tính chất

Trong mục này chúng tôi đưa ra một thuật toán phân đoạn sử dụng tiêu chuẩn quyết định D đã mô tả ở trên. Ta sẽ chỉ ra rằng phân đoạn bằng thuật toán này sẽ tuân theo các thuộc tính không quá thô (too coarse) và cũng không quá mịn (too fine), theo các định nghĩa sau đây.

4.4.1 Định nghĩa 1

Một phân đoạn được xem là quá mịn nếu tồn tại một số cặp miền $C_1, C_2 \in S$ mà giữa hai miền này không có dấu hiệu của đường biên.

Để định nghĩa được những khái niệm bổ sung cho phân đoạn quá thô, chúng ta đưa ra khái niệm tinh chỉnh (refinement) của một phân đoạn.

Cho hai phân đoạn S và T của cùng một tập cơ sở, ta nói rằng T là một tinh chỉnh (refinement) của S khi mỗi thành phần của T được chứa trong (hoặc bằng) một số thành phần của S . Và ta cũng nói rằng T là một tinh chỉnh đúng (proper refinement) của S khi $T \neq S$.

Chú ý rằng nếu T là tinh chỉnh đúng của S thì T có thể được chứa bởi một hoặc một số các miền trong S và S được gọi là thô hơn T .

4.4.2 Định nghĩa 2

Một phân đoạn được xem là quá thô khi tồn tại một tinh chỉnh đúng của S mà phân đoạn đó vẫn chưa là quá mịn.

Vấn đề đặt ra là liệu có phải luôn luôn tồn tại phân đoạn không quá thô cũng không quá mịn hay không? Và nếu tồn tại thì phân đoạn đó có là duy nhất không?

Thực tế cho thấy là nói chung luôn có thể có nhiều hơn một phân đoạn không quá thô cũng không quá mịn, do đó phân đoạn này là không duy nhất. Đây là một tính chất đặc biệt của phân đoạn ảnh dựa trên đồ thị và được chứng minh chi tiết ở mục 4.4.3.

4.4.3 Tính chất

Với một đồ thị hữu hạn $G = (V, E)$ bất kỳ luôn tồn tại một số phân đoạn S không quá thô mà cũng không quá mịn.

Chứng minh: Chúng ta dễ dàng nhận thấy là tính chất này đúng. Thật vậy, nếu phân đoạn mà tất cả các phần tử đều nằm trong một thành phần, thì phân đoạn này là không quá mịn, vì nó chỉ có đúng một thành phần (định nghĩa 1). Nếu mà phân đoạn này cũng không quá thô thì coi như xong. Ngược lại, theo định nghĩa 2, thì sẽ có một tinh chỉnh đúng mà ko quá mịn. Lấy một trong số các tinh chỉnh đó và lặp lại thủ tục này cho đến khi chúng ta sẽ thu được một phân đoạn không quá thô.

Trở lại với thuật toán phân đoạn dựa trên đồ thị, thuật toán này gắn với thuật toán Kruskal xây dựng cây tủa nhánh tối thiểu của một đồ thị. Độ phức tạp của thuật toán là $O(m \log m)$, trong đó m là số cạnh của đồ thị.

4.4.4 Thuật toán

Thuật toán phân đoạn

Input: Đồ thị $G = (V, E)$ gồm n đỉnh và m cạnh

Output: Một phân đoạn của V thành các thành phần $S = (C_1, C_2, \dots)$.

Thuật toán

- Bước 0: Sắp xếp các cạnh của G theo thứ tự không giảm của trọng số.

$$\pi = (o_1, o_2, \dots, o_m)$$

- Bước 1: Bắt đầu với phân đoạn S^0 , lúc này mỗi đỉnh nằm trong một thành phần.

- Bước 2: Lặp lại bước 3 với $q = 1, \dots, m$

- Bước 3: Xây dựng S^q từ S^{q-1} như sau: Cho v_i và v_j là hai đỉnh nối với nhau bởi cạnh thứ q , tức là $o_q = (v_i, v_j)$. Nếu v_i và v_j nằm trong hai thành phần tách rời nhau của S^{q-1} và $\omega(o_q)$ nhỏ hơn sự khác nhau nội vùng của cả hai thành phần thì trộn hai thành phần này với nhau, ngược lại không làm gì cả. Cụ thể hơn gọi C_i^{q-1} là thành phần của S^{q-1} chứa v_i và C_j^{q-1} là thành phần của S^{q-1} chứa v_j . Nếu $C_i^{q-1} \neq C_j^{q-1}$ và $\omega(o_q) \leq \text{Mint}(C_i^{q-1}, C_j^{q-1})$ thì S^q thu được từ S^{q-1} bằng cách trộn C_i^{q-1} với C_j^{q-1} . Ngược lại $S^q = S^{q-1}$.

- Bước 4: Trả về kết quả $S = S^m$.

Chúng ta sẽ chứng minh rằng phân đoạn S được xây dựng trong thuật toán trên là tuân theo các thuộc tính toàn cục khi sử dụng tính chất so sánh cặp miền đã định nghĩa trong phần trước. Nghĩa là mặc dù thuật toán chỉ dựa vào các quyết định tham lam nhưng phân đoạn được xây dựng vẫn thỏa mãn các thuộc tính toàn cục.

Để chứng minh điều này chúng ta xem xét các bổ đề và các định lý sau đây:

4.4.5 Bổ đề 1

Giả sử C_i^{q-1} và C_j^{q-1} biểu diễn hai thành phần được nối với nhau bằng cạnh $O_q = (v_i, v_j)$ thì $C_i = C_i^{q-1}$ hoặc $C_j = C_j^{q-1}$ mà C_i là thành phần chứa v_i và C_j là thành phần chứa v_j trong phân đoạn S cuối cùng.

Chứng minh: Khi hai thành phần không được trộn với nhau thì có hai trường hợp có thể xảy ra.

Hoặc là $\omega(o_q) > \text{Int}(C_i^{q-1}) + \tau(C_i^{q-1})$, hoặc là $\omega(o_q) > \text{Int}(C_j^{q-1}) + \tau(C_j^{q-1})$.

Vì các cạnh được sắp xếp theo chiều không giảm của trọng số nên $\omega(o_k) > \omega(o_q)$ với $\forall k \geq q+1$. Do đó không có phép trộn nào xảy ra nữa.

4.4.6 Định lý 1

Phân đoạn S sử dụng tính chất so sánh miền D và thuật toán 1 là không quá mịn theo định nghĩa 1.

Chứng minh: Theo định nghĩa, để S là quá mịn thì phải có một số cặp thành phần nào đó mà D không nắm bắt được. Thế thì phải tồn tại ít nhất một cạnh giữa hai thành phần cùng cặp vì theo bước 3 của thuật toán thì chúng không được trộn thành một miền. Chẳng hạn cho $O_q = (v_i, v_j)$ là cạnh có thứ tự đầu tiên. Trong trường hợp này thì theo thuật toán trên C_i^{q-1} và C_j^{q-1} không được trộn vào nhau, nghĩa là $\omega(o_q) > \text{Mint}(C_i^{q-1}, C_j^{q-1})$. Theo bổ đề 1 chúng ta biết rằng $C_i = C_i^{q-1}$ hoặc $C_j = C_j^{q-1}$, một trong hai điều này xảy ra nghĩa là $\omega(o_q) > \text{Mint}(C_i, C_j)$, điều này chứng tỏ D nắm bắt được cả C_i và C_j . Đây là điều mâu thuẫn, vậy định lý được chứng minh.

4.4.7 Định lý 2

Phân đoạn S sử dụng tính chất so sánh miền D và thuật toán 1 là không quá thô theo định nghĩa 2.

Chứng minh: Để S là không quá thô thì phải có một số phép tinh chỉnh hợp lý T, sao cho nó vẫn chưa là quá mịn. Xem xét cạnh có trọng số nhỏ nhất e nằm trong thành phần $C \in S$ nhưng khác với $A, B \in T$. Theo định nghĩa về phép tinh chỉnh thì $A \subset C$ và $B \subset C$. Vì T là không quá mịn nên hoặc là $\omega(e) > \text{Int}(A) + \tau(A)$, hoặc là $\omega(e) > \text{Int}(B) + \tau(B)$. Không mất tính tổng quát, giả sử biểu thức đầu đúng khi đó bằng cách xây dựng một kết nối từ A tới một thành phần con của C, trọng số của cạnh này cùng lắm bằng $\omega(e)$ vì $\omega(e) > \text{Int}(A)$. Mà theo thuật toán thì

khi xem xét đến các cạnh trong tập đã sắp xếp theo chiều không giảm của trọng số, phải xem xét tất cả các cạnh trong cây khung MST(A,E) trước khi xem xét các cạnh từ A đến một thành phần khác của C. Do đó thuật toán phải xây dựng A trước C, và trong bước xây dựng C thì phải trộn A với một thành phần con của C. Trọng số của cạnh nối giữa A và thành phần này lớn nhất là $w(e)$. Tuy nhiên thuật toán đã không trộn A vì $\omega(e) > \text{Int}(A) + \tau(A)$, đây là một mâu thuẫn. Vậy định lý đã được chứng minh.

4.4.8 Định lý 3

Phân đoạn theo thuật toán 1 không phụ thuộc vào việc sắp xếp các cạnh theo thứ tự không giảm của trọng số.

Chứng minh: Bất kì một thứ tự sắp xếp nào cũng được thay đổi chỉ bằng cách đảo vị trí của các phần tử liên kề. Điều này chỉ ra rằng bằng cách đổi chỗ thứ tự của hai cạnh liên kề cùng trọng số thì sự sắp xếp không giảm của các cạnh vẫn không thay đổi, và kết quả phân đoạn được sinh ra theo thuật toán 1 cũng không thay đổi.

Cho e_1 và e_2 là hai cạnh có cùng trọng số liên kề nhau trong dãy cạnh sau khi sắp xếp. Rõ ràng là khi thuật toán xem xét cạnh đầu tiên trong hai cạnh thì chúng kết nối giữa hai thành phần tách rời hay nói chính xác hơn là hai cặp của các thành phần, nên thứ tự của hai cạnh này là không thành vấn đề. Chỉ còn trường hợp chúng ta cần thiết kiểm tra đó là kiểm tra là khi e_1 kết nối giữa thành phần A và thành phần B và e_2 là kết nối giữa một trong hai thành phần A hoặc B với một thành phần C.

Bây giờ chúng ta chỉ ra rằng e_1 là căn nguyên của phép trộn khi xem xét sau e_2 . Điều này có nghĩa rằng $\omega(e_1) \leq \text{MInt}(A, B)$. Nếu e_2 thay vì được xem xét trước e_1 , thì cả e_2 và e_1 vẫn là căn nguyên của phép trộn, hoặc là e_2 cũng là căn nguyên của phép trộn trong trường hợp thành phần mới của $B \cup C$ có $\text{Int}(B \cup C) = \omega(e_2) = \omega(e_1)$. Do đó chúng ta biết rằng $\omega(e_1) \leq \text{MInt}(A, B \cup C)$ vẫn ngụ ý rằng

e_1 vẫn là căn nguyên của phép trộn, Mặt khác, giả sử e_1 không phải là căn nguyên của phép trộn nếu nó được xem xét sau e_2 . Tức là, $\omega(e_1) > \text{MInt}(A, B)$. Do đó hoặc là (2.8) $\omega(e_1) > \text{Int}(A) + \tau_A$ hoặc là (2.9) $\omega(e_1) > \text{Int}(B) + \tau_B$. Trong trường hợp (2.8) vẫn đúng nếu e_2 được xem xét trước. Trong trường hợp (2.9) nếu e_2 được xem xét trước thì nó có thể không phải là căn nguyên của một phép trộn vì $\omega(e_1) = \omega(e_2)$, do đó $\omega(e_2) > \text{MInt}(B, C)$. Vậy khi xem xét e_1 sau e_2 chúng ta vẫn có $\omega(e_1) > \text{MInt}(A, B)$ và e_1 không là căn nguyên của phép trộn.

4.4.9. Độ phức tạp tính toán

Thời gian thực hiện của thuật toán này chia làm hai phần:

Một là thời gian cần thiết để sắp xếp dãy trọng số theo chiều không giảm (bước 0). Đối với dãy số nguyên thì điều này có thể thực hiện trong thời gian tuyến tính, Có rất nhiều phương pháp sắp xếp có thể thực hiện trong thời gian $O(m \log m)$ với m là số lượng cạnh.

Hai là thời gian thực hiện bước 1-3. Để kiểm tra được hai đỉnh có cùng chung trong một thành phần hay không chúng tôi sử dụng biến set-find trên mỗi đỉnh nhằm lưu lại số hiệu thành phần mà đỉnh đó đang phụ thuộc vào. Để trộn hai thành phần lại với nhau chúng tôi chỉ việc hiệu chỉnh lại các biến set-find của một trong hai tập đỉnh. Mint được tính trong một hằng thời gian nếu biết được Int và kích thước của mỗi thành phần. Int cũng được tính trong một hằng thời gian cho mỗi phép trộn, vì cạnh có trọng số lớn nhất trong cây khung nhỏ nhất của một thành phần là căn nguyên của một phép trộn. Có được điều này vì bổ đề 1 nói rằng căn nguyên của phép trộn chính là cạnh có trọng số nhỏ nhất giữa hai thành phần được trộn. Kích thước của thành phần sau khi trộn bằng tổng kích thước của hai thành phần trước khi trộn. Vậy độ phức tạp tính toán từ bước 1 đến bước 3 của thuật toán là $O(m\alpha(m))$ trong đó α là hàm Ackerman nghịch đảo, m là số cạnh của đồ thị.

4.5 Nhận xét

Phân đoạn dựa trên đồ thị được coi là một trong những kỹ thuật phân đoạn hiệu quả nhất cả về không gian và thời gian trong các ứng dụng thời gian thực. Mặc dù không tránh khỏi những hạn chế như: độ khác nội vùng được xác định một cách chủ quan, không biểu diễn chính xác được các thành phần, hay rất khó để chọn được giá trị phù hợp của tham số k mà hàm ngưỡng yêu cầu xác định nhằm điều khiển kích cỡ của vùng đã phân đoạn. Tuy nhiên, phương pháp này có những ưu điểm hơn hẳn các phương pháp trước đó: Nó sử dụng một sửa đổi đơn giản nhưng hiệu quả của thuật toán Kruskal, phân đoạn một ảnh thành các vùng bằng cách định nghĩa một tính chất xác định dấu hiệu đường biên giữa hai vùng sử dụng biểu diễn đồ thị của ảnh. Một đặc tính quan trọng của phương pháp này là khả năng lưu giữ chi tiết nhỏ trong các vùng ảnh ít biến đổi, trong khi nó bỏ đi những khía cạnh này trong các vùng ảnh có sự biến đổi cao.

Mặt khác, mặc dù thuật toán này đưa ra các quyết định tham lam, nhưng kết quả phân đoạn thỏa mãn được các thuộc tính toàn cục. Thuật toán phân đoạn dựa trên đồ thị chạy trong thời gian gần như tuyến tính với số cạnh của đồ thị và đạt hiệu quả cao trong thực nghiệm.

CHƯƠNG 5. CÀI ĐẶT THỬ NGHIỆM

Với mục tiêu thử nghiệm để đánh giá kết quả của các phương pháp em xin cài đặt thử nghiệm một số thuật toán điển hình của phân đoạn ảnh. Chương trình được cài đặt trên ngôn ngữ Visual C++ 6.0.

Với các ảnh màu thuật toán được thực hiện 3 lần, một lần đối với thành phần màu đỏ (red), một lần với thành phần màu xanh lá cây (green) và một lần với màu xanh da trời (blue). Em cũng đã thử chạy một lần với cả 3 màu trên, khi đó trọng số của các cạnh được lấy bằng sự khác nhau về màu sắc của các điểm ảnh trong không gian màu. Tuy nhiên, thực hành cho thấy kết quả tốt nhất thu được khi chạy với từng thành phần màu riêng của ảnh.

Tham số k khi thực hiện thuật toán được chọn nhằm tính ra giá trị của hàm ngưỡng τ . Như trong chương 4 đã nói, hàm ngưỡng của thành phần C được tính bằng $\tau(C) = k/|C|$. Trong đó $|C|$ là số lượng phần tử của C . Em sử dụng hai tham số khác nhau cho từng bức ảnh dựa vào độ phân giải khác nhau của bức ảnh. Chẳng hạn với ảnh có độ phân giải 128×128 điểm ảnh thì giá trị của k được lựa chọn là $k = 150$. Còn đối với ảnh 320×240 thì sử dụng k với giá trị $k = 300$. Ảnh có độ phân giải càng cao thì giá trị của k được chọn càng lớn.

Các thuật toán được cài đặt thử nghiệm trong chương trình :

- Dùng ngưỡng cố định
- Sử dụng Entropy
- Thuật toán đẳng liêu
- Thuật toán tam giác
- Thuật toán GraphBased

5.1 Thuật toán Đẳng liêu :

```
void CDibView::OnMnDanglieu()  
{
```

```

// TODO: Add your command handler code here

//Bat dau chung

LPSTR  lpDIBHdr;      // Pointer to BITMAPINFOHEADER
LPSTR  lpDIBBits;    // Pointer to DIB bits
BYTE *lpDIBBytes;

BOOL   bSuccess=FALSE; // Success/fail flag
HPALETTE hPal=NULL;   // Our DIB's palette
HPALETTE hOldPal=NULL; // Previous palette
HDIB hDIB;

DWORD height;

DWORD width;

DWORD size;

CDibDoc *pDoc = CDibView::GetDocument();

hDIB = pDoc->GetHDIB ();

/* Check for valid DIB handle */
if (hDIB == NULL)

    return;

/* Lock down the DIB, and get a pointer to the beginning of the bit
* buffer
*/

lpDIBHdr = (LPSTR) ::GlobalLock((HGLOBAL) hDIB);
lpDIBBits = ::FindDIBBits(lpDIBHdr);
lpDIBBytes = (BYTE *)lpDIBBits;
height = DIBHeight(lpDIBHdr);
width = DIBWidth(lpDIBHdr);
size = height*width;

DWORD nBytes = (DWORD)(size/8);

```

```

ToGrey(lpDIBBytes,height,width); //MyTools

DWORD i,j;

//Ket thuc chung

DWORD hist[256];

DWORD t;

double tt, tOld;

double m0, m1, mI;

DWORD np0, np1;

//Tinh histogram

for (i=0; i<256; i++) hist[i]=0;

for(i=0;i<height;i++)

for(j=0;j<width;j++)

{

    hist[getpixel(lpDIBBytes,height,width,i,j)] += 1;

}

//Tinh xong hist

/* Tinh toan tim nguong toi uu */

mI = 0.0;

for (i=0; i<256; i++) mI += (i*hist[i]/size);

//Xuat phat t0 cu

tt = 128.0;

t = 128;

//Xuat phat t0 cai tien

/*tt = mI;

t = (DWORD)tt;*/

do

{

```

```

np0 = 0;

np1 = 0;

for (i=0; i <= t; i++) np0 += hist[i];

for (i=t+1; i < 256; i++) np1 += hist[i];

m0 = 0.0; m1 = 0.0;

for (i=0; i <= t; i++) m0 += (i*hist[i]/np0);

for (i=t+1; i < 256; i++) m1 += (i*hist[i]/np1);

tOld = t;

tt = (m0+m1)/2;

t = (DWORD)tt;

}

while (abs(tt - tOld) < 0.5);

CString strNguong;

strNguong.FormatMessage(_T("The threshold: %1!d!"),t);

MessageBox((LPCTSTR)strNguong);

/* Phan nguong theo t */

for(i=0;i<height;i++)

for(j=0;j<width;j++)

{

    if (getpixel(lpDIBBytes,height,width,i,j) < t)

        putpixel(lpDIBBytes,height,width,i,j,0);

    else putpixel(lpDIBBytes,height,width,i,j,255);

}

//Bat dau chung

CDibView::Invalidate ();

return;

//Ket thuc chung

```

```
}
```

5.2 Thuật toán Tam giác :

```
void CDibView::OnMnTamgiac()
```

```
{
```

```
    // TODO: Add your command handler code here
```

```
//Bat dau chung
```

```
    LPSTR  lpDIBHdr;        // Pointer to BITMAPINFOHEADER
```

```
    LPSTR  lpDIBBits;      // Pointer to DIB bits
```

```
    BYTE *lpDIBBytes;
```

```
    BOOL  bSuccess=FALSE;  // Success/fail flag
```

```
    HPALETTE hPal=NULL;    // Our DIB's palette
```

```
    HPALETTE hOldPal=NULL; // Previous palette
```

```
    HDIB hDIB;
```

```
    DWORD height;
```

```
    DWORD width;
```

```
    DWORD size;
```

```
    CDibDoc *pDoc = CDibView::GetDocument();
```

```
    hDIB = pDoc->GetHDIB ();
```

```
    /* Check for valid DIB handle */
```

```
    if (hDIB == NULL)
```

```
        return;
```

```
    /* Lock down the DIB, and get a pointer to the beginning of the bit
```

```
    * buffer
```

```
    */
```

```
    lpDIBHdr = (LPSTR) ::GlobalLock((HGLOBAL) hDIB);
```

```
    lpDIBBits = ::FindDIBBits(lpDIBHdr);
```

```
    lpDIBBytes = (BYTE *)lpDIBBits;
```

```

height = DIBHeight(lpDIBHdr);
width = DIBWidth(lpDIBHdr);
size = height*width;
DWORD nBytes = (DWORD)(size/8);
ToGrey(lpDIBBytes,height,width); //MyTools
DWORD i,j;
//Ket thuc chung
double hist[256], d[256];
int t;
int start, end;
int bMax, bMax0, bMax1;
double A, B, C, D;
//Tinh histogram, d
for (i=0; i<256; i++) { hist[i]=0.0; d[i] = 0.0;}
for(i=0;i<height;i++)
for(j=0;j<width;j++)
{
    hist[getpixel(lpDIBBytes,height,width,i,j)] += 1.0;
}
for (i=0; i<256; i++) hist[i] /= (double)size;
//Tinh xong hist
start = 0;
while (hist[start] = 0.0) start += 1;
end = 255;
while (hist[end] = 0.0) end -= 1;
bMax = start;
for (i = start+1; i <= end; i++)

```

```

        if (hist[i]>hist[bMax]) bMax = i;

A = (double)(hist[bMax] - hist[start]);
B = (double)(start - bMax);
C = (double)(-hist[start]*B - start*A);
D = (double)(sqrt(A*A+B*B));
for (i=start+1; i<bMax; i++)
    d[i] = abs(A*i + B*hist[i] + C)/D;
A = (double)(hist[end] - hist[bMax]);
B = (double)(bMax - end);
C = (double)(-hist[bMax]*B - bMax*A);
D = (double)(sqrt(A*A+B*B));
for (i=bMax+1; i<end; i++)
    d[i] = abs(A*i + B*hist[i] + C)/D;

/* Tinh toan tim nguong toi uu */
t = start;
for (i=start+1; i < end; i++)
    if ((d[i] > d[t])&&(i != bMax)) t = i;

CString strNguong;
strNguong.FormatMessage(_T("bMax la: %1!d!"),bMax);
MessageBox((LPCTSTR)strNguong);
strNguong.FormatMessage(_T("The threshold: %1!d!"),t);
MessageBox((LPCTSTR)strNguong);

/* Phan nguong theo t */
for(i=0;i<height;i++)
for(j=0;j<width;j++)
{
    if (getpixel(lpDIBBytes,height,width,i,j) < t)

```



```

        putpixel(lpDIBBytes,height,width,i,j,0);
    else putpixel(lpDIBBytes,height,width,i,j,255);
    }
//Bat dau chung
    CDibView::Invalidate ();
    return;
//Ket thuc chung
}

```

5.3 Thuật toán GraphBased :

```

void CDibView::OnGraphBased()
{
    // TODO: Add your command handler code here
    //Bat dau chung
    LPSTR lpDIBHdr; // Pointer to BITMAPINFOHEADER
    LPSTR lpDIBBits; // Pointer to DIB bits
    BYTE *lpDIBBytes;
    BOOL bSuccess=FALSE; // Success/fail flag
    HPALETTE hPal=NULL; // Our DIB's palette
    HPALETTE hOldPal=NULL; // Previous palette
    HDIB hDIB;
    DWORD height;
    DWORD width;
    DWORD size;
    CDibDoc *pDoc = CDibView::GetDocument();
    hDIB = pDoc->GetHDIB ();
    /* Check for valid DIB handle */
    if (hDIB == NULL)

```

```

        return;

    /* Lock down the DIB, and get a pointer to the beginning of the bit
    * buffer
    */

    lpDIBHdr = (LPSTR) ::GlobalLock((HGLOBAL) hDIB);
    lpDIBBits = ::FindDIBBits(lpDIBHdr);
    lpDIBBytes = (BYTE *)lpDIBBits;
    height = DIBHeight(lpDIBHdr);
    width = DIBWidth(lpDIBHdr);
    size = height*width;

    DWORD nBytes = (DWORD)(size/8);

// de dua ve cung cap xam

    //ToGrey(lpDIBBytes,height,width); //MyTools

    DWORD x, y, pos;

    ////////////

    //lay du lieu tu lpDIBBytes dua vao imPtr.

    image<rgb> *im = new image<rgb>(height, width);

    for (x = 0; x < height ; x++)

        for (y = 0; y < width; y++) {

            if (Position(x,y,width,height,pos)) {

                imRef(im, x, y).r = (unsigned char)lpDIBBytes[pos*3];
                imRef(im, x, y).g = (unsigned char)lpDIBBytes[pos*3 + 1];
                imRef(im, x, y).b = (unsigned char)lpDIBBytes[pos*3 + 2];

            }

        }

}

```

```

CDOption frmOption;

    frmOption.DoModal();

    int num_ccs;

    // goi thu tuc de phan doan theo graph_based
    image<rgb> *seg = segment_image(im, sigma, K, min_size, &num_ccs);

    // dua du lieu tu *seg sang lpDIBBytes
    for (x = 0; x < height; x++) {
        for (y = 0; y < width; y++) {
            Position(x,y,width, height,pos);

            lpDIBBytes[pos*3] = (BYTE)imRef(seg, x, y).r;
            lpDIBBytes[pos*3+1] = (BYTE)imRef(seg, x, y).g;
            lpDIBBytes[pos*3+2] = (BYTE)imRef(seg, x, y).b;
        }
    }

//Bat dau chung
    CDibView::Invalidate ();

    return;

//Ket thuc chung
}

```

5.4 Kết quả đạt được

*Với thuật toán Đẳng liệu :



Ảnh gốc a



Ảnh b

Hình a là ảnh gốc, hình b là kết quả của chương trình với ngưỡng bằng 71.

*Với Thuật toán Tam giác :



Ảnh gốc a



Ảnh b

Hình a là ảnh gốc, hình b là kết quả chạy chương trình với ngưỡng bằng 103.

* Với thuật toán GraphBased :



Ảnh gốc a



Ảnh b

Kết quả chạy chương trình với $\text{Sigma}=0.8$, $K = 500$, $\text{Minsize}=50$

KẾT LUẬN

1. Các kết quả đạt được

Có thể nói, cùng với sự phát triển như vũ bão của khoa học kỹ thuật trong một vài thập kỷ gần đây, xử lý ảnh ngày càng thể hiện vai trò quan trọng trong các lĩnh vực nghiên cứu, ứng dụng thực tế và thu hút sự quan tâm đặc biệt từ các nhà khoa học trên thế giới.

Trong một số lượng lớn các ứng dụng về xử lý ảnh và hiển thị máy tính, phân đoạn đóng vai trò chính yếu như là bước đầu tiên trước khi áp dụng các thao tác xử lý ảnh mức cao hơn như: nhận dạng, giải thích ngữ nghĩa, và biểu diễn ảnh.

Chính vì thế, để đạt được các kết quả phân đoạn ảnh tốt nhất, các nhà khoa học đang nỗ lực nghiên cứu và phát triển nhiều phương pháp mới, đạt hiệu quả cao trong các ứng dụng thời gian thực.

Xuất phát từ sự quan tâm và hứng thú đối với các vấn đề liên quan đến xử lý ảnh số và các ứng dụng của nó. Đồ án đã nghiên cứu và trình bày được một số nội dung chính sau :

Tìm hiểu được một cách tổng quan các vấn đề về xử lý ảnh, phân đoạn ảnh và tác động của nó đến cuộc sống con người.

Qua đó, có một cách nhìn có hệ thống về các phương pháp phân đoạn ảnh và các thuật toán sử dụng trong mỗi phương pháp. Đồng thời biết được ưu, nhược điểm của từng phương pháp để có thể đưa ra cách lựa chọn phù hợp với từng loại ảnh.

Tìm hiểu và cài đặt phương pháp phân đoạn dựa vào đồ thị - một trong những phương pháp phân đoạn tối ưu nhất hiện nay. Phương pháp này đạt hiệu quả phân đoạn nhanh và hiệu quả cả về không gian và thời gian. Ngoài ra, nó có khả năng nắm bắt cả những thuộc tính cục bộ cũng như toàn cục của bức ảnh để tăng cường khả năng phân đoạn chính xác.

Ngoài ra, trong quá trình nghiên cứu em cũng tự tích lũy thêm cho mình các kiến thức về toán học, về kỹ thuật lập trình,... Và quan trọng là rèn luyện kỹ năng để thực hiện một nghiên cứu khoa học. Tuy mới chỉ là bước đầu, nhưng những kết quả này sẽ giúp ích cho em trong những nghiên cứu sau này để thu được kết quả tốt hơn

Tuy nhiên, do thời gian và kinh nghiệm thực tế còn thiếu nên những nội dung mà đề án thực hiện được chỉ ở mức độ đơn giản và chưa hoàn thiện.

Cụ thể trong từng chương trình đã có kết quả sau:

Chương 1 là giới thiệu tổng quan về xử lý ảnh, các khái niệm cơ bản trong xử lý ảnh.

Chương 2 là trình bày về phân đoạn ảnh dựa vào ngưỡng. Chương này giới thiệu về phương pháp phân đoạn dựa vào ngưỡng và trình bày một số phương pháp chọn ngưỡng tự động : dựa vào Histogram (có thuật toán đẳng liệu, thuật toán đối xứng nên, thuật toán tam giác, chọn ngưỡng với Bimodal Histogram), phương pháp sử dụng entropy,... và phương pháp mới được nghiên cứu gần đây đó là : thông qua sự không ổn định của lớp và tính thuần nhất vùng.

Phương pháp phân đoạn theo miền đồng nhất được trình bày trong chương 3. Chương này giới thiệu một số phương pháp phân đoạn theo miền đồng nhất : phương pháp tách cây tứ phân, phân vùng hợp, phương pháp tách-hợp.

Chương 4 trình bày về ứng dụng đồ thị trong phân đoạn ảnh.

Chương 5 cài đặt thử nghiệm một số thuật toán và kết quả chạy chương trình.

2. Những hạn chế của đề án

Bên cạnh những kết quả đạt được bản đề án vẫn còn một số hạn chế :

Chưa đưa ra được một phương pháp phân đoạn mới hoàn toàn. Đề án mới chỉ trình bày lại các kiến thức tìm hiểu được chứ chưa đề xuất được một phương pháp hoàn toàn mới.

Do thời gian có hạn, nên việc trình bày các thuật toán phân đoạn cũng chưa được hệ thống và khoa học. Có nhiều thuật toán được trình bày sơ lược.

Đề án cũng chưa chỉ ra được các ứng dụng thực tế của mỗi thuật toán phân đoạn.

TÀI LIỆU THAM KHẢO

1. Nhập môn xử lý ảnh số - Tác giả: Lương Mạnh Bà & Nguyễn Thanh Thủy
- Nhà xuất bản khoa học và kỹ thuật.
2. Kỹ thuật xử lý ảnh và video - Tác giả Nguyễn Kim Sách - Nhà xuất bản
khoa học và kỹ thuật
3. Giáo trình xử lý ảnh - Tác giả Võ Đức Khánh - Nhà xuất bản Thống Kê
4. Internet