



**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH
VIỄN THÔNG**

BÀI GIẢNG MÔN HỌC

XỬ LÝ ẢNH

Người soạn : PGS. TS. ĐỖ NĂNG TOÀN

Chỉnh sửa lần 1, Tháng 11 năm 2013

LỜI NÓI ĐẦU

Khoảng hơn mười năm trở lại đây, phần cứng máy tính và các thiết bị liên quan đã có sự tiến bộ vượt bậc về tốc độ tính toán, dung lượng chứa, khả năng xử lý v.v.. và giá cả đã giảm đến mức máy tính và các thiết bị liên quan đến xử lý ảnh đã không còn là thiết bị chuyên dụng nữa. Khái niệm ảnh số đã trở nên thông dụng với hầu hết mọi người trong xã hội và việc thu nhận ảnh số bằng các thiết bị cá nhân hay chuyên dụng cùng với việc đưa vào máy tính xử lý đã trở nên đơn giản.

Trong hoàn cảnh đó, xử lý ảnh là một lĩnh vực đang được quan tâm và đã trở thành môn học chuyên ngành của sinh viên ngành công nghệ thông tin trong nhiều trường đại học trên cả nước. Tuy nhiên, tài liệu giáo trình còn là một điều khó khăn. Hiện tại chỉ có một số ít tài liệu bằng tiếng Anh hoặc tiếng Pháp, tài liệu bằng tiếng Việt thì rất hiếm. Với mong muốn đóng góp vào sự nghiệp đào tạo và nghiên cứu trong lĩnh vực này, chúng tôi biên soạn cuốn giáo trình *Xử lý ảnh* dựa trên đề cương môn học đã được duyệt. Cuốn sách tập trung vào các vấn đề cơ bản của xử lý ảnh nhằm cung cấp một nền tảng kiến thức đầy đủ và chọn lọc nhằm giúp người đọc có thể tự tìm hiểu và xây dựng các chương trình ứng dụng liên quan đến xử lý ảnh.

Giáo trình được chia làm 7 chương: Chương 1, trình bày Tổng quan về xử lý ảnh, các khái niệm cơ bản, sơ đồ tổng quát của một hệ thống xử lý ảnh và các vấn đề cơ bản trong xử lý ảnh. Chương 2, trình bày các vấn đề liên quan đến thu nhận ảnh, bao gồm các thiết bị thu nhận ảnh cơ bản, quá trình lấy mẫu và lượng tử hóa, một số phương pháp biểu diễn ảnh, cũng như một số định dạng ảnh phổ dụng. Chương 3, trình bày các kỹ thuật nâng cao chất lượng ảnh dựa vào các thao tác với điểm ảnh, nâng cao chất lượng ảnh thông qua việc xử lý các điểm ảnh trong lân cận điểm ảnh đang xét. Chương này cũng trình bày các kỹ thuật nâng cao chất lượng ảnh nhờ vào việc lọc nhiễu. Chương 4, trình bày các kỹ thuật cơ bản trong việc phát hiện biên của các đối tượng ảnh theo cả hai khuynh hướng: Phát hiện biên trực tiếp và phát hiện biên gián tiếp. Chương 5 thể hiện các kỹ thuật phân vùng ảnh, đây là khâu quan trọng hỗ trợ cho việc trích chọn các thuộc tính của ảnh để tiến tới hiểu ảnh. Chương 6, trình bày các vấn đề và kỹ thuật liên quan đến nhận dạng ảnh, theo tiếp cận không gian, tiếp cận cấu trúc và tiếp cận mạng nơron. Và cuối cùng là Chương 7 với các kỹ thuật nén ảnh, đây là vấn đề luôn được quan tâm trong xử lý ảnh.

Giáo trình được biên soạn dựa trên kinh nghiệm giảng dạy của tác giả trong nhiều năm tại các khóa đại học và cao học của ĐH Công nghệ - ĐHQG Hà Nội, ĐH Khoa học tự nhiên - ĐHQG Hà Nội, Khoa Công nghệ thông tin - ĐH Thái Nguyên, Khoa Công nghệ thông tin - Học viện Công nghệ BCVT v.v.. Cuốn sách có thể làm tài liệu tham khảo cho sinh viên các hệ kỹ sư, cử nhân và các bạn quan tâm đến vấn đề nhận dạng và xử lý ảnh.

Các tác giả bày tỏ lòng biết ơn chân thành tới các bạn đồng nghiệp trong Viện Công nghệ thông tin, Khoa Công nghệ thông tin - Học viện Công nghệ BCVT, Khoa Công nghệ thông tin - ĐH Thái Nguyên, Khoa Công nghệ thông tin - ĐH Công nghệ - ĐHQG Hà Nội, Khoa Toán - Cơ - Tin, ĐH Khoa học tự nhiên, ĐHQG Hà Nội đã đồng viên, góp ý và giúp đỡ để hoàn chỉnh nội dung cuốn sách này. Xin cảm ơn Lãnh đạo Khoa Công nghệ thông

tin – Học viện Công nghệ BCVT, Ban Giám đốc Học viện Công nghệ BCVT đã hỗ trợ và tạo điều kiện để cho ra đời giáo trình này.

Mặc dù rất cố gắng nhưng tài liệu này chắc chắn không tránh khỏi những sai sót. Chúng tôi xin trân trọng tiếp thu tất cả những ý kiến đóng góp của bạn đọc cũng như các bạn đồng nghiệp để có chỉnh lý kịp thời.

Thư góp ý xin gửi về: Đỗ Năng Toàn,
Viện Công nghệ thông tin.
18 Hoàng Quốc Việt, Cầu Giấy, Hà Nội
Điện thoại: 04.3.7567935 Email: dntoan@ioit.ac.vn

Hà Nội, ngày 28 tháng 06 năm 2010

TÁC GIẢ

PTIT

MỤC LỤC

LỜI NÓI ĐẦU	2
Chương 1: NHẬP MÔN XỬ LÝ ẢNH	7
1.1. TỔNG QUAN VỀ HỆ THỐNG XỬ LÝ ẢNH	7
1.2. NHỮNG VẤN ĐỀ CƠ BẢN TRONG XỬ LÝ ẢNH	8
1.2.1. Một số khái niệm cơ bản	8
1.2.2. Nắn chỉnh biến dạng	8
1.2.3. Khử nhiễu	8
1.2.4. Chỉnh số mức xám	9
1.2.5. Phân tích ảnh	9
1.2.6. Nhận dạng và phân loại ảnh	9
1.2.7. Nén ảnh	10
Chương 2: THU NHẬN ẢNH	11
2.1. CÁC THIẾT BỊ THU NHẬN ẢNH	11
2.2. LẤY MẪU VÀ LƯỢNG TỬ HOÁ	11
2.2.1. Giai đoạn lấy mẫu	11
2.2.2. Lượng tử hóa	12
2.3. MỘT SỐ PHƯƠNG PHÁP BIỂU DIỄN ẢNH	12
2.3.1. Mô hình Raster	12
2.3.2. Mô hình Vector	13
2.4. CÁC ĐỊNH DẠNG ẢNH CƠ BẢN	13
2.4.1. Định dạng ảnh IMG	13
2.4.2. Định dạng ảnh PCX	14
2.4.3. Định dạng ảnh TIFF	15
2.4.4. Định dạng file ảnh BITMAP	16
2.5. KỸ THUẬT TÁI NHẬN ẢNH	17
2.5.1. Kỹ thuật chụp ảnh	17
2.5.2. Kỹ thuật in ảnh	18
2.6. KHÁI NIỆM ẢNH ĐEN TRẮNG, ẢNH MÀU	20
2.6.1. Phân loại ảnh số	20
2.6.2. Màu sắc	21
2.7. CÂU HỎI ÔN TẬP CHƯƠNG	28
Chương 3: XỬ LÝ NÂNG CAO CHẤT LƯỢNG ẢNH	29
3.1. CẢI THIÊN ẢNH SỬ DỤNG CÁC TOÁN TỬ ĐIỀM	29
3.1.1. Giới thiệu	29
3.1.2. Tăng giảm độ sáng	29
3.1.3. Tách ngưỡng	29
3.1.4. Cân bằng histogram	30
3.1.5. Kỹ thuật tìm tách ngưỡng tự động	31

3.1.6. Biến đổi cấp xám tổng thể	32
3.2. CẢI THIẾN ẢNH SỬ DỤNG CÁC TOÁN TỬ KHÔNG GIAN	32
3.2.1. Phép cửa sổ di chuyển.....	32
3.2.2. Phép nhân chập và mẫu	33
3.2.3. Lọc trung vị.....	35
3.2.4. Lọc trung bình	37
3.2.5. Lọc trung bình theo k giá trị gần nhất	38
3.3. KHÔI PHỤC ẢNH.....	39
3.3.1. Nhiễu và mô hình nhiễu.....	39
3.3.2. Các loại nhiễu	39
3.3.3. Các kỹ thuật lọc nhiễu	40
3.4. CÂU HỎI ÔN TẬP CHƯƠNG.....	42
Chương 4: CÁC PHƯƠNG PHÁP PHÁT HIỆN BIÊN	43
4.1. KHÁI QUÁT VỀ BIÊN VÀ PHÂN LOẠI CÁC KỸ THUẬT DÒ BIÊN CƠ BẢN.....	43
4.1.1. Giới thiệu	43
4.1.2. Kỹ thuật phát hiện biên Gradient.....	43
4.1.3. Kỹ thuật phát hiện biên Laplace	47
4.1.4. Kỹ thuật Canny.....	47
4.2. PHƯƠNG PHÁP PHÁT HIỆN BIÊN CỤC BỘ.....	49
4.2.1. Biên và độ biến đổi về mức xám.....	49
4.2.2. Phát hiện biên dựa vào trung bình cục bộ.....	50
4.3. DÒ BIÊN THEO QUY HOẠCH ĐỘNG.....	53
4.3.1. Một số khái niệm cơ bản.....	53
4.3.2. Chu tuyến của một đối tượng ảnh.....	54
4.3.3. Thuật toán dò biên tổng quát.....	55
4.4. CÁC PHƯƠNG PHÁP KHÁC.....	58
4.4.1. Các phép toán hình thái cơ bản	58
4.4.2. Một số tính chất của phép toán hình thái.....	59
4.4.3. Xấp xỉ trên và xấp xỉ dưới đối tượng ảnh	62
4.4.4. Thuật toán phát hiện biên dựa vào phép toán hình thái	64
4.5. CÂU HỎI ÔN TẬP CHƯƠNG.....	65
Chương 5: PHÂN VÙNG ẢNH.....	66
5.1. GIỚI THIỆU	66
5.2. PHÂN VÙNG ẢNH THEO NGUỒN BIÊN ĐỘ	67
5.2.1. Thuật toán đẳng liệu.....	68
5.2.2. Thuật toán đối xứng nền.....	69
5.2.3. Thuật toán tam giác	70
5.2.3. Chọn ngưỡng đối với Bimodal Histogram.....	70
5.3. PHÂN VÙNG THEO MIỀN ĐỒNG NHẤT	71
5.3.1. Giới thiệu	71
5.3.2. Phương pháp tách cây tứ phân (quad tree).....	72

5.3.3. Các phương pháp phân vùng bởi hợp	75
5.3.4. Phương pháp tách hợp (split-merge)	76
5.4. PHÂN VÙNG DỰA THEO ĐƯỜNG BIÊN	77
5.4.1. Phát hiện biên	77
5.4.2. Làm mảnh biên	78
5.4.3. Nhị phân hóa đường biên	78
5.4.4. Miêu tả đường biên	78
5.5. PHÂN VÙNG THEO KẾT CẤU BỀ MẶT	79
5.5.1. Phương pháp thống kê	79
5.5.2. Phương pháp cấu trúc	81
5.5.3. Tiếp cận theo tính kết cấu	81
5.6. CÂU HỎI ÔN TẬP CHƯƠNG	82
Chương 6: NHẬN DẠNG ẢNH	83
6.1. GIỚI THIỆU	83
6.2. NHẬN DẠNG DỰA THEO MIỀN KHÔNG GIAN	84
6.2.1. Phân hoạch không gian	84
6.2.2. Hàm phân lớp hay hàm ra quyết định	84
6.2.3. Nhận dạng thống kê	85
6.2.4. Một số thuật toán nhận dạng tiêu biểu trong tự học	86
6.3. NHẬN DẠNG DỰA THEO CẤU TRÚC	88
6.3.1. Biểu diễn định tính	88
6.3.2. Phương pháp ra quyết định dựa vào cấu trúc	88
6.4. NHẬN DẠNG DỰA THEO MẠNG NƠON	90
6.4.1. Mạng Hopfield	90
6.4.2. Mạng Kohonen	92
Chương 7: NÉN DỮ LIỆU ẢNH	95
7.1. GIỚI THIỆU	95
7.2. CÁC PHƯƠNG PHÁP NÉN THỂ HỆ THỨ NHẤT	96
7.2.1. Phương pháp mã hóa loạt dài	96
7.2.2. Phương pháp mã hóa Huffman	96
7.2.3. Phương pháp LZW	98
7.2.4. Phương pháp mã hóa khối	102
7.2.5. Phương pháp thích nghi	103
7.2.6. Biến đổi Cosin và chuẩn nén JPEG	104
7.3. CÁC PHƯƠNG PHÁP NÉN THỂ HỆ THỨ HAI	109
7.3.1. Phương pháp Kim tự tháp Laplace (Pyramide Laplace)	109
7.3.2. Phương pháp mã hóa dựa vào biểu diễn ảnh	110
7.4. CÂU HỎI ÔN TẬP CHƯƠNG	112
TÀI LIỆU THAM KHẢO	113

Chương 1:

NHẬP MÔN XỬ LÝ ẢNH

1.1. TỔNG QUAN VỀ HỆ THỐNG XỬ LÝ ẢNH

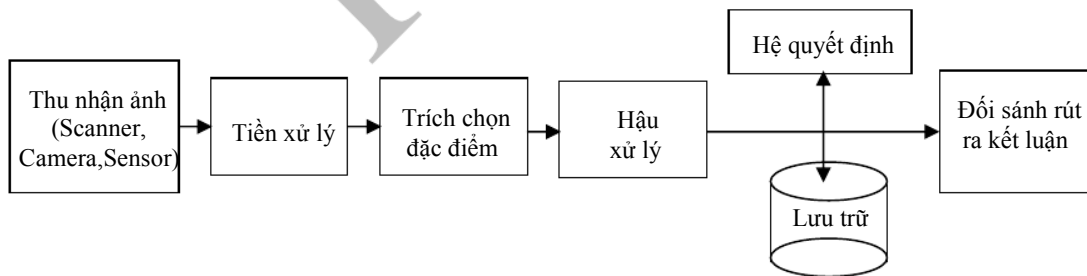
Con người thu nhận thông tin qua các giác quan, trong đó thị giác đóng vai trò quan trọng nhất. Những năm trở lại đây với sự phát triển của phần cứng máy tính, xử lý ảnh và đồ họa đã phát triển một cách mạnh mẽ và có nhiều ứng dụng trong cuộc sống. Xử lý ảnh và đồ họa đóng một vai trò quan trọng trong tương tác người máy.

Quá trình xử lý ảnh được xem như là quá trình thao tác ảnh đầu vào nhằm cho ra kết quả mong muốn. Kết quả đầu ra của một quá trình xử lý ảnh có thể là một ảnh “tốt hơn” hoặc một kết luận.



Hình 1.1. Quá trình xử lý ảnh

Ảnh có thể xem là tập hợp các điểm ảnh và mỗi điểm ảnh được xem như là đặc trưng cường độ sáng hay một dấu hiệu nào đó tại một vị trí nào đó của đối tượng trong không gian và nó có thể xem như một hàm n biến $P(c_1, c_2, \dots, c_n)$. Do đó, ảnh trong xử lý ảnh có thể xem như ảnh n chiều.



Hình 1.2. Các bước cơ bản trong một hệ thống xử lý ảnh

Hình 1.2 biểu diễn sơ đồ tổng quát của một hệ thống xử lý ảnh:

- **Khối thu nhận ảnh:** có nhiệm vụ tiếp nhận ảnh đầu vào.
- **Khối tiền xử lý:** có nhiệm vụ xử lý nâng cao chất lượng ảnh như giảm nhiễu, phân vùng, tìm biên v.v..
- **Khối trích chọn đặc điểm:** có nhiệm vụ trích chọn các đặc trưng quan trọng của các bức ảnh đã được tiền xử lý để sử dụng trong hệ quyết định

- **Khối hậu xử lý:** có nhiệm vụ xử lý các đặc điểm đã trích chọn, có thể lược bỏ hoặc biến đổi các đặc điểm này để phù hợp với các kỹ thuật cụ thể sử dụng trong hệ quyết định
- **Khối hệ quyết định và lưu trữ:** có nhiệm vụ đưa ra quyết định (phân loại) dựa trên dữ liệu đã học lưu trong khối lưu trữ
- **Khối kết luận:** đưa ra kết luận dựa vào quyết định của khối quyết định

1.2. NHỮNG VẤN ĐỀ CƠ BẢN TRONG XỬ LÝ ẢNH

1.2.1. Một số khái niệm cơ bản

* Ảnh và điểm ảnh:

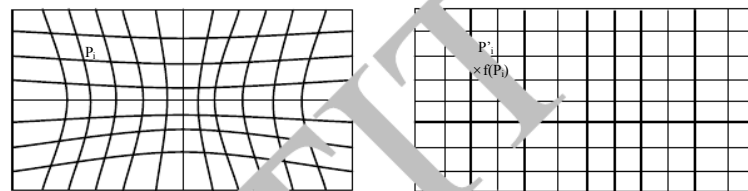
Điểm ảnh được xem như là dấu hiệu hay cường độ sáng tại một vị trí nào đó của đối tượng trong không gian và ảnh được xem như là một tập hợp các điểm ảnh.

* Mức xám, màu

Là số các giá trị có thể có của các điểm ảnh của ảnh

1.2.2. Nắn chỉnh biến dạng

Ảnh thu nhận thường bị biến dạng do các thiết bị quang học và điện tử.



Ảnh thu nhận

Ảnh mong muốn

Hình 1.3. Ảnh thu nhận và ảnh mong muốn

Để khắc phục người ta sử dụng các phép chiếu được xây dựng trên tập các điểm điều khiển. Giả sử có hai ảnh I và I' tương ứng với ảnh thu nhận được và ảnh mong muốn. P_i là một điểm thuộc I tương ứng với một điểm P'_i trên I' , ta có n các cặp điểm điều khiển như vậy.

Nắn chỉnh biến dạng là tìm hàm $f: P_i \rightarrow f(P_i)$ sao cho

$$\sum_{i=1}^n |f(P_i) - P'_i|^2 \rightarrow \min$$

Thông thường các điểm điều khiển được lấy bằng cách chụp những ảnh có cấu trúc định sẵn như là ảnh lưới giống như ảnh mong muốn trên Hình 1.3.

1.2.3. Khử nhiễu

Có 2 loại nhiễu cơ bản trong quá trình thu nhận ảnh mà chúng ta cần loại bỏ:

- **Nhiều hệ thống:** là nhiễu có quy luật có thể khử bằng các phép biến đổi
- **Nhiều ngẫu nhiên:** vết bản không rõ nguyên nhân có thể khắc phục bằng các phép lọc

1.2.4. Chính số mức xám

Chính số mức xám là nhằm khắc phục tính không đồng đều của hệ thống xử lý ảnh, thông thường có 2 hướng tiếp cận:

- **Giảm số mức xám:** Thực hiện bằng cách nhóm các mức xám gần nhau thành một bó. Trường hợp giảm xuống 2 mức xám thì chính là chuyển về ảnh đen trắng.
- **Tăng số mức xám:** Thực hiện nội suy ra các mức xám trung gian bằng kỹ thuật nội suy. Kỹ thuật này nhằm tăng cường độ mịn cho ảnh.

1.2.5. Phân tích ảnh

Là khâu quan trọng trong quá trình xử lý ảnh để tiến tới hiểu ảnh. Trong phân tích ảnh việc trích chọn đặc điểm là một bước quan trọng. Các đặc điểm của đối tượng được trích chọn tùy theo mục đích nhận dạng trong quá trình xử lý ảnh. Có thể nêu ra một số đặc điểm của ảnh sau đây:

- **Đặc điểm không gian:** Phân bố mức xám, phân bố xác suất, biên độ, điểm uốn v.v..
- **Đặc điểm biến đổi:** Các đặc điểm loại này được trích chọn bằng việc thực hiện lọc vùng (zonal filtering). Các bộ vùng được gọi là “mặt nạ đặc điểm” (feature mask) thường là các khe hẹp với hình dạng khác nhau (chữ nhật, tam giác, cung tròn v.v..)
- **Đặc điểm biên và đường biên:** Đặc trưng cho đường biên của đối tượng và do vậy rất hữu ích trong việc trích chọn các thuộc tính bất biến được dùng khi nhận dạng đối tượng (Ví dụ đặc điểm góc, cạnh v.v..). Các đặc điểm này có thể được trích chọn thông qua ảnh biên. Để thu được ảnh biên ta có thể sử dụng toán tử gradient, toán tử Laplace v.v..

Việc trích chọn hiệu quả các đặc điểm giúp cho việc nhận dạng các đối tượng ảnh chính xác, với tốc độ tính toán cao và giảm thiểu dung lượng lưu trữ.

1.2.6. Nhận dạng và phân loại ảnh

Nhận dạng tự động (automatic recognition), mô tả đối tượng, phân loại và phân nhóm các mẫu là những vấn đề quan trọng trong thị giác máy, được ứng dụng trong nhiều ngành khoa học khác nhau. Tuy nhiên, một câu hỏi đặt ra là: mẫu (pattern) là gì? Watanabe, một trong những người đi đầu trong lĩnh vực này đã định nghĩa: “Ngược lại với hỗn loạn (chaos), mẫu là một thực thể (entity), được xác định một cách ang áng (vaguely defined) và có thể gán cho nó một tên gọi nào đó”. Ví dụ mẫu có thể là ảnh của vân tay, ảnh của một vật nào đó được chụp, một chữ viết, khuôn mặt người hoặc một ký đồ tín hiệu tiếng nói. Khi biết một mẫu nào đó, để nhận dạng hoặc phân loại mẫu đó có thể sử dụng hai cách chính:

- **Phân loại có mẫu (supervised classification):** ví dụ phân tích phân biệt (discriminant analysis), trong đó mẫu đầu vào được định danh thành một phần của một lớp đã xác định.
- **Phân loại không có mẫu (unsupervised classification hay clustering):** Các mẫu được gán vào các lớp khác nhau dựa trên một tiêu chuẩn đồng dạng nào đó. Các lớp này cho đến thời điểm phân loại vẫn chưa biết hay chưa được định danh.

Có nhiều cách tiếp cận khác nhau trong lý thuyết nhận dạng trong đó có: Đối sánh mẫu dựa trên các đặc trưng được trích chọn, phân loại thống kê, đối sánh cấu trúc, phân loại dựa trên mạng nơ-ron nhân tạo...

Hệ thống nhận dạng tự động bao gồm ba khâu tương ứng với ba giai đoạn chủ yếu sau đây:

1. Thu nhận dữ liệu và tiền xử lý.
2. Biểu diễn dữ liệu.
3. Nhận dạng, ra quyết định.

Trong ứng dụng thực tiễn, không thể chỉ dùng có một cách tiếp cận đơn lẻ để phân loại “tối ưu” vì vậy các phương thức phân loại tổ hợp thường được sử dụng khi nhận dạng. Cho đến nay các hệ thống lai (hybrid system) sử dụng nhiều phương pháp và cách tiếp cận khác nhau đã cho những kết quả nhiều triển vọng.

1.2.7. Nén ảnh

Nén ảnh là kỹ thuật nhằm giảm thiểu không gian lưu trữ. Có hai hướng tiếp cận chính là nén có bảo toàn và không bảo toàn thông tin. Nén không bảo toàn thì thường có khả năng nén cao hơn nhưng không phục hồi được ảnh gốc, ngược lại nén có bảo toàn cho phép khôi phục hoàn toàn ảnh gốc. Nén ảnh nói chung có 4 cách tiếp cận cơ bản:

- **Nén ảnh thống kê:** Kỹ thuật nén này dựa vào việc thống kê tần suất xuất hiện của giá trị các điểm ảnh, trên cơ sở đó mà có chiến lược mã hóa thích hợp. Một ví dụ điển hình cho kỹ thuật mã hóa này là *.TIF
- **Nén ảnh không gian:** Kỹ thuật này dựa vào vị trí không gian của các điểm ảnh để tiến hành mã hóa. Kỹ thuật lợi dụng sự giống nhau của các điểm ảnh trong các vùng gần nhau. Ví dụ cho kỹ thuật này là mã nén *.PCX
- **Nén ảnh sử dụng phép biến đổi:** Đây là kỹ thuật tiếp cận theo hướng nén không bảo toàn và do vậy tỉ lệ nén tương đối cao. *.JPG chính là tiếp cận theo kỹ thuật nén này.
- **Nén ảnh Fractal:** Sử dụng tính chất Fractal của các đối tượng ảnh. Tính chất Fractal của ảnh thể hiện sự lặp lại của các chi tiết tại nhiều vị trí khác nhau với kích thước và hướng khác nhau. Kỹ thuật nén sẽ tính toán để chỉ cần lưu trữ phần gốc ảnh và quy luật sinh ra ảnh theo nguyên lý Fractal.

THU NHẬN ẢNH

2.1. CÁC THIẾT BỊ THU NHẬN ẢNH

Các thiết bị thu nhận ảnh có hai loại chính ứng với hai loại ảnh thông dụng Raster, Vector và có thể cho ảnh đen trắng hoặc ảnh màu.

Các thiết bị thu nhận ảnh Raster thông thường là camera, scanner. Còn các thiết bị thu nhận ảnh Vector thông thường là sensor hoặc bàn số hoá digitalizer hoặc được chuyển đổi từ ảnh Raster.

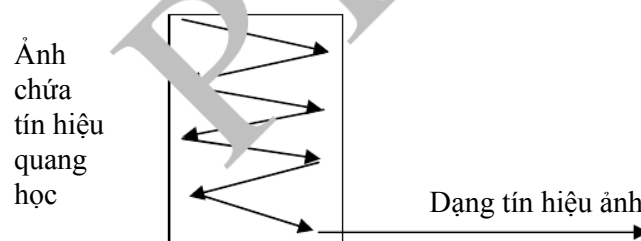
Các hệ thống thu nhận ảnh sử dụng chung một quá trình thu nhận:

- Cảm biến: biến đổi năng lượng quang học thành năng lượng điện (giai đoạn lấy mẫu)
- Tổng hợp năng lượng điện thành ảnh (giai đoạn lượng tử hóa)

2.2. LẤY MẪU VÀ LƯỢNG TỬ HOÁ

2.2.1. Giai đoạn lấy mẫu

Người ta sử dụng bộ cảm biến hoặc máy quét để biến tín hiệu quang của ảnh thành tín hiệu điện liên tục. Phương pháp sử dụng máy quét phổ biến hơn. Máy quét sẽ quét theo chiều ngang để tạo ra tín hiệu điện của ảnh, kết quả cho ra một tín hiệu điện hai chiều $I(x,y)$ liên tục.



Xét ảnh liên tục được biểu diễn bởi hàm $I(x, y)$, gọi Δx là khoảng cách giữa hai điểm được giữ lại theo trục x , gọi Δy là khoảng cách giữa hai điểm được giữ lại theo trục y . Δx , Δy được gọi là chu kỳ lấy mẫu theo trục x và y .

Giai đoạn lấy mẫu sẽ biến hàm liên tục $I(x,y)$ thành hàm rời rạc $I(n \Delta x, m \Delta y)$. Với m,n là nguyên.

Theo SHANON để đảm bảo không xảy ra hiện tượng chồng phổ, cho phép tái tạo lại ảnh gốc từ ảnh đã số hóa thì tần số lấy mẫu phải ít nhất phải lớn hơn hoặc bằng hai tần số cao nhất của tín hiệu ảnh. Gọi $f_{x\max}$, $f_{y\max}$ là tần số cao nhất của tín hiệu theo trục x , y thì:

- $f_x \geq 2f_{x\max}$ với $f_x \gg \frac{1}{\Delta x}$ là tần số lấy mẫu theo trục x .
- $f_y \geq 2f_{y\max}$ với $f_y \gg \frac{1}{\Delta y}$ là tần số lấy mẫu theo trục y .

2.2.2. Lượng tử hóa

Ảnh sau khi lấy mẫu sẽ có dạng $I(m,n)$ với m, n là nguyên nhưng giá trị $I(m, n)$ vẫn là giá trị vật lý liên tục. Quá trình biến đổi giá trị $I(m,n)$ thành một số nguyên thích hợp để lưu trữ gọi là lượng tử hoá. Đây là quá trình ánh xạ một biến liên tục u vào biến rời rạc u^* thuộc tập hữu hạn $[u_1, u_2, \dots, u_L]$ xác định trước, L là mức lượng tử hoá được tạo ra.

Ví dụ:

- Với ảnh đa cấp xám thì $L=256$, $f(m,n) \in [0, 255]$
- Với ảnh 24 bit thì $L=2^{24}$, $f(m, n) \in [0, 2^{24} - 1]$

2.3. MỘT SỐ PHƯƠNG PHÁP BIỂU DIỄN ẢNH

Ảnh trên máy tính là kết quả thu nhận theo các phương pháp số hoá được nhưng trong các thiết bị kỹ thuật khác nhau. Cấu trúc dữ liệu để lưu trữ thông tin ảnh trong bộ nhớ có ảnh hưởng rất lớn đến việc hiển thị, in ấn và xử lý. Quá trình lưu trữ ảnh nhằm biểu diễn ảnh dưới dạng lưu trữ được với hai tiêu chí chính:

- Tiết kiệm bộ nhớ
- Giảm thời gian xử lý

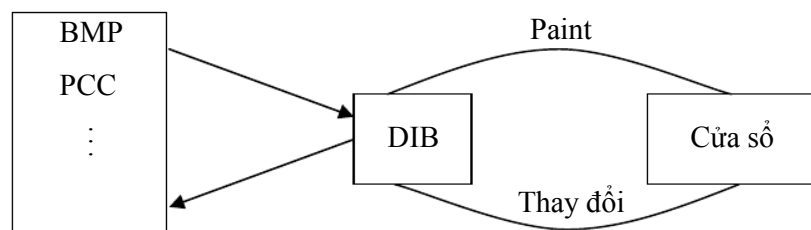
Ảnh được biểu diễn theo hai mô hình cơ bản, ảnh Raster và ảnh Vector.

2.3.1. Mô hình Raster

Đây là cách biểu diễn ảnh thông dụng nhất hiện nay, ảnh được biểu diễn dưới dạng ma trận các điểm (điểm ảnh). Thường thu nhận qua các thiết bị như camera, scanner. Tuy theo yêu cầu thực tế mà mỗi điểm ảnh được biểu diễn qua 1 hay nhiều bit. Mật độ điểm ảnh trên một đơn vị kích thước vật lý được gọi là độ phân giải. Ảnh có độ phân giải cao thì càng đẹp, càng mịn và càng thể hiện rõ chi tiết. Việc lựa chọn độ phân giải thích hợp tùy thuộc vào nhu cầu sử dụng và đặc trưng của mỗi ảnh cụ thể.

Mô hình Raster thuận lợi cho hiển thị và in ấn. Ngày nay công nghệ phần cứng cung cấp những thiết bị thu nhận ảnh Raster phù hợp với tốc độ nhanh và chất lượng cao cho cả đầu vào và đầu ra. Một thuận lợi cho việc hiển thị trong môi trường Windows là Microsoft đưa ra khuôn dạng ảnh DIB (Device Independent Bitmap) làm trung gian. Hình 2.1 thể hiện quy trình chung để hiển thị ảnh Raster thông qua DIB.

Một trong những hướng nghiên cứu cơ bản trên mô hình biểu diễn này là kỹ thuật nén ảnh bao gồm cả nén bảo toàn và không bảo toàn. Hiện nay trên thế giới có trên 50 khuôn dạng ảnh thông dụng (BMP, TIF, GIF, PCX v.v..) bao gồm cả trong đó các kỹ thuật nén có khả năng phục hồi dữ liệu 100% và nén có khả năng phục hồi với độ sai số nhất định.



Hình 2.1. Quá trình hiển thị và chỉnh sửa, lưu trữ ảnh thông qua DIB

2.3.2. Mô hình Vector

Biểu diễn ảnh ngoài tiêu chí tiết kiệm không gian lưu trữ, dễ dàng cho hiển thị và in ấn còn cần đảm bảo dễ dàng trong lựa chọn, sao chép, di chuyển và tìm kiếm v.v.. Theo những yêu cầu này kỹ thuật biểu diễn Vector tỏ ra ưu việt hơn.

Trong mô hình vector người ta sử dụng hướng giữa các vector của điểm ảnh lân cận để mã hoá và tái tạo hình ảnh ban đầu ảnh vector được thu nhận trực tiếp từ các thiết bị số hoá như Digital hoặc được chuyển đổi từ ảnh Raster thông qua các chương trình số hoá

Công nghệ phần cứng cung cấp những thiết bị xử lý với tốc độ nhanh và chất lượng cho cả đầu vào và ra nhưng lại chỉ hỗ trợ cho ảnh Raster. Do vậy, những nghiên cứu về biểu diễn Vector đều tập trung vào việc chuyển đổi ảnh từ mô hình Raster.



Hình 2.2. Sự chuyển đổi giữa các mô hình biểu diễn ảnh

2.4. CÁC ĐỊNH DẠNG ẢNH CƠ BẢN

Hiện nay trên thế giới có trên 50 khuôn dạng ảnh thông dụng. Sau đây là một số định dạng ảnh hay dùng trong quá trình xử lý ảnh hiện nay.

2.4.1. Định dạng ảnh IMG

Ảnh IMG là ảnh đen trắng, phần đầu của ảnh IMG có 16 byte chứa các thông tin:

- 6 byte đầu: dùng để đánh dấu định dạng ảnh. Giá trị của 6 byte này viết dưới dạng Hexa: 0x0001 0x0008 0x0001
- 2 byte tiếp theo: chứa độ dài mẫu tín. Đó là độ dài của dãy các byte kề liền nhau mà dãy này sẽ được lặp lại một số lần nào đó. Số lần lặp này sẽ được lưu trong byte đếm. Nhiều dãy giống nhau được lưu trong một byte.
- 4 byte tiếp: mô tả kích cỡ pixel.
- 2 byte tiếp: số pixel trên một dòng ảnh.
- 2 byte cuối: số dòng ảnh trong ảnh.

Ảnh IMG được nén theo từng dòng, mỗi dòng bao gồm các gói (pack). Các dòng giống nhau cũng được nén thành một gói. Có 4 loại gói sau:

- **Loại 1: Gói các dòng giống nhau.** Quy cách gói tin này như sau: 0x00 0x00 0xFF Count. Ba byte đầu tiên cho biết số các dãy giống nhau, byte cuối cho biết số các dòng giống nhau.
- **Loại 2: Gói các dãy giống nhau.** Quy cách gói tin này như sau: 0x00 Count. Byte thứ hai cho biết số các dãy giống nhau được nén trong gói. Độ dài của dãy ghi ở đầu tệp.
- **Loại 3: Dãy các Pixel không giống nhau, không lặp lại và không nén được.** Quy cách gói tin này như sau: 0x80 Count. Byte thứ hai cho biết độ dài dãy các pixel không giống nhau không nén được.

- **Loại 4: Dây các Pixel giống nhau.** Tuỳ theo các bit cao của byte đầu tiên được bật hay tắt. Nếu bit cao được bật (giá trị 1) thì đây là gói nén các byte chỉ gồm bit 0, số các byte được nén được tính bởi 7 bit thấp còn lại. Nếu bit cao tắt (giá trị 0) thì đây là gói nén các byte gồm toàn bit 1. Số các byte được nén được tính bởi 7 bit còn lại.

Các gói tin của file IMG rất đa dạng do ảnh IMG là ảnh đen trắng, do vậy chỉ cần 1 bit cho 1 pixel thay vì 4 hoặc 8 như đã nói ở trên. Toàn bộ ảnh chỉ có những điểm sáng và tối tương ứng với giá trị 1 hoặc 0. Tỷ lệ nén của kiểu định dạng này là khá cao.

2.4.2. Định dạng ảnh PCX

Định dạng ảnh PCX là một trong những định dạng ảnh cổ điển. Nó sử dụng phương pháp mã hoá loạt dài RLE (Run – Length – Encoded) để nén dữ liệu ảnh. Quá trình nén và giải nén được thực hiện trên từng dạng ảnh. Thực tế, phương pháp giải nén PCX kém hiệu quả hơn so với kiểu IMG. Tập PCX gồm 3 phần: đầu tệp (header), dữ liệu ảnh (Image data) và bảng màu mở rộng.

Header của tệp PCX có kích thước cố định gồm 128 byte và được phân bố như sau:

- 1 byte: chỉ ra kiểu định dạng. Nếu là PCX/PCC thì nó luôn có giá trị là 0Ah.
- 1 byte: chỉ ra version sử dụng để nén ảnh, có thể có các giá trị sau:
 - + 0: version 2.5.
 - + 2: version 2.8 với bảng màu.
 - + 3: version 2.8 hay 3.0 không có bảng màu.
 - + 5: version 3.0 có bảng màu.
- 1 byte: chỉ ra phương pháp mã hoá. Nếu là 0 thì mã hoá theo phương pháp BYTE PACKED, ngược lại là phương pháp RLE.
- 1 byte: Số bit cho một điểm ảnh phẳng (plane).
- 1 word: toạ độ góc trái của ảnh. Với kiểu PCX nó có giá trị là (0,0), cũn PCC thì khác (0,0).
- 1 word: toạ độ góc phải dưới.
- 1 word: kích thước bề rộng và bề cao của ảnh.
- 1 word: số điểm ảnh.
- 1 word: độ phân giải màn hình.
- 1 word.
- 48 byte: chia nó thành 16 nhóm, mỗi nhóm 3 byte. Mỗi nhóm này chứa thông tin về một thanh ghi màu. Như vậy ta có 16 thanh ghi màu.
- 1 byte: không dùng đến và luôn đặt là 0.
- 1 byte: số bit plane mà ảnh sử dụng. Với ảnh 16 màu, giá trị này là 4, với ảnh 256 màu (1pixel/8bits) thì số bit plane lại là 1.
- 1 byte: số bytes cho một dòng quét ảnh.
- 1 word: kiểu bảng màu.
- 58 byte: không dùng.

Định dạng ảnh PCX thường được dùng để lưu trữ ảnh và thao tác đơn giản, cho phép nén và giải nén nhanh. Tuy nhiên, vì cấu trúc của nó cố định, nên trong một số trường hợp làm tăng kích thước lưu trữ. Cũng vì nhược điểm này mà một số ứng dụng sử dụng một kiểu định dạng khác mềm dẻo hơn: định dạng TIFF (Targed Image File Format) sẽ mô tả dưới đây.

2.4.3. Định dạng ảnh TIFF

Kiểu định dạng TIFF được thiết kế để làm nhẹ bớt các vấn đề liên quan đến việc mở rộng file ảnh cố định. Về cấu trúc, nó cũng gồm 3 phần chính:

- **Phần Header (IFH):** có trong tất cả các tệp TIFF và gồm 8 byte:
 - + 1 word: chỉ ra kiểu tạo tệp trên máy tính PC hay máy Macintosh. Hai loại này khác nhau rất lớn ở thứ tự các byte lưu trữ trong các số dài 2 hay 4 byte. Nếu trường này có giá trị là 4D4Dh thì đó là ảnh cho máy Macintosh, nếu là 4949h là của máy PC.
 - + 1 word: version. từ này luôn có giá trị là 42. đây là đặc trưng của file TIFF và không thay đổi.
 - + 2 word: giá trị Offset theo byte tính từ đầu tới cấu trúc IFD là cấu trúc thứ hai của file. Thứ tự các byte này phụ thuộc vào dấu hiệu trường đầu tiên.
- **Phần thứ 2 (IFD):** Không ở ngay sau cấu trúc IFH mà vị trí được xác định bởi trường Offset trong đầu tệp. Có thể có một hay nhiều IFD cùng tồn tại trong một file. Một IFD bao gồm:
 - + 2 byte: chứa các DE (Directory Entry).
 - + 12 byte là các DE xếp liên tiếp, mỗi DE chiếm 12 byte.
 - + 4 byte: chứa Offset trở tới IFD tiếp theo. Nếu đây là IFD cuối cùng thì trường này có giá trị 0.
- **Phần thứ 3:** các DE: các DE có độ dài cố định gồm 12 byte và chia làm 4 phần:
 - + 2 byte: chỉ ra dấu hiệu mà file ảnh đó được xây dựng.
 - + 2 byte: kiểu dữ liệu của tham số ảnh. Có 5 kiểu tham số cơ bản:
 - 1: BYTE (1 byte)
 - 2: ASCII (1 byte)
 - 3: SHORT (2 byte).
 - 4: LONG (4 byte)
 - 5: RATIONAL (8 byte)
 - + 4 byte: trường độ dài chưa số lượng chỉ mục của kiểu dữ liệu đó chỉ ra. Nó không phải là tổng số byte cần thiết để lưu trữ. Để có số liệu này ta cần nhân số chỉ mục với kiểu dữ liệu đã dùng.
 - + 4 byte: đó là Offset tới điểm bắt đầu dữ liệu liên quan tới dấu hiệu, tức là liên quan với DE không phải lưu trữ vật lý cùng với nó nằm ở một vị trí nào đó trong file.

Dữ liệu chứa trong tệp thường được tổ chức thành các nhóm dòng (cột) quét của dữ liệu ảnh. Cách tổ chức này làm giảm bộ nhớ cần thiết cho việc đọc tệp. Việc giải nén được thực hiện theo 4 kiểu khác nhau được lưu trữ trong byte dấu hiệu nén.

2.4.4. Định dạng file ảnh BITMAP

Mỗi file BITMAP gồm đầu file chứa các thông tin chung về file, đầu thông tin chứa các thông tin về ảnh, một bảng màu và một mảng dữ liệu ảnh. Khuôn dạng được cho như sau:

```

BITMAPFILEHEADER    bmfh;
BITMAPINFOHEADER    bmih;
RGBQUAD             aColors[];
BYTE                 aBitmapBits[];

```

Trong đó, các cấu trúc được định nghĩa như sau:

```

typedef struct tagBITMAPFILEHEADER { /* bmfh */
    UINT        bfType;
    DWORD       bfSize;
    UINT        bfReserved1;
    UINT        bfReserved2;
    DWORD       bfOffBits;
} BITMAPFILEHEADER;

typedef struct tagBITMAPINFOHEADER { /* bmih */
    DWORD       biSize;
    LONG        biWidth;
    LONG        biHeight;
    WORD        biPlanes;
    WORD        biBitCount;
    DWORD       biCompression;
    DWORD       biSizeImage;
    LONG        biXPelsPerMeter;
    LONG        biYPelsPerMeter;
    DWORD       biClrUsed;
    DWORD       biClrImportant;
} BITMAPINFOHEADER, *LPBITMAPINFOHEADER;

```


Với biến chi tiết như sau

biSize	kích thước của BITMAPINFOHEADER
biWidth	Chiều rộng của ảnh, tính bằng số điểm ảnh
biHeight	Chiều cao của ảnh, tính bằng số điểm ảnh
biPlanes	Số plane của thiết bị, phải bằng 1
biBitCount	Số bit cho một điểm ảnh
biCompression	Kiểu nén
biSizeImage	Kích thước của ảnh tính bằng byte
biXPelsPerMeter	độ phân giải ngang của thiết bị, tính bằng điểm ảnh trên met
biYPelsPerMeter	độ phân giải dọc của thiết bị, tính bằng điểm ảnh trên met
biClrUsed	Số lượng các màu thực sự được sử dụng
biClrImportant	Số lượng các màu cần thiết cho việc hiển thị, bằng 0 nếu tất cả các màu đều cần để hiển thị

Nếu `bmi.biBitCount > 8` thì mảng màu `rgbq[]` trống, ngược lại thì mảng màu có $2^{\text{bmi.biBitCount}}$ phần tử.

```
typedef struct tagRGBQUAD { /* rgbq */
    BYTE    rgbBlue;
    BYTE    rgbGreen;
    BYTE    rgbRed;
    BYTE    rgbReserved;
} RGBQUAD;

typedef struct tagBITMAPINFO {
    BITMAPINFOHEADER    bmiHeader;
    RGBQUAD              bmiColors[1];
} BITMAPINFO, *PBITMAPINFO;
```

2.5. KỸ THUẬT TÁI NHẬN ẢNH

2.5.1 Kỹ thuật chụp ảnh

Phương pháp sao chụp ảnh là phương pháp đơn giản, giá thành thấp, chất lượng cao. Sau bước chụp là kỹ thuật phòng tối nhằm tăng cường ảnh như mong muốn. Ví dụ kỹ thuật phòng tối như: phóng đại ảnh, thu nhỏ ảnh..., tùy theo ứng dụng. Kỹ thuật chụp ảnh màn hình màu khá đơn giản. Nó bao gồm các bước sau:

- Đặt camera trong phòng tối, cách màn hình khoảng 10 feet (1 feet = 0,3048m)
- Mở ống kính để phẳng mặt cong màn hình, do vậy ảnh sẽ dàn đều hơn

- Tắt phím sáng tối (Brightness) và phím tương phản (Contrast) của màn hình để tạo độ rõ cho ảnh. Các màu chói, cường độ cao trên ảnh sẽ giảm đi.
- Đặt tốc độ ống kính từ 1/8 đến 1/2 giây.

2.5.2 Kỹ thuật in ảnh

Người ta dùng kỹ thuật nửa cường độ để thể hiện ảnh trên sách báo, tạp chí. Theo kỹ thuật này, một ảnh tạo nên bởi một chuỗi các điểm in trên giấy. Thực chất, mỗi điểm ảnh có thể được coi là một hình vuông trắng bao quanh một chấm đen. Do vậy, nếu chấm đen càng lớn điểm ảnh sẽ càng xẫm màu. Màu xám có thể coi như chấm đen chiếm nửa vùng trắng. Vùng trắng là vùng gồm một chùm các điểm ảnh có rất ít hoặc không có chấm đen.

Từ đặc điểm cảm nhận của mắt người, sự thay đổi cường độ chấm đen trong các phần tử ảnh trắng tạo nên mô phỏng của một ảnh liên tục. Như vậy, mắt người cảm nhận từ một ảnh mà màu biến đổi từ đen qua xám rồi đến trắng. Tổng số cường độ duy nhất hiện diện sẽ xác định các kích thước khác nhau của chấm đen. Thông thường, báo ảnh tạo ảnh nửa cường độ với độ phân giải từ 60 đến 80 DPI(dot per inch: số điểm ảnh trên một inch), sách có thể in đến 150 DPI.

Tuy nhiên, các máy in ghép nối với máy tính không có khả năng sắp xếp các chấm đen có kích thước khác nhau của ảnh, vì vậy chúng ta cần các kỹ thuật nửa cường độ kỹ thuật số (Digital Halftoning): phân ngưỡng, chọn mẫu (Patterning), Dithering hoặc khuếch tán lỗi (Error Diffusion).

a) Phân ngưỡng

Kỹ thuật này đặt ngưỡng để hiển thị các tông màu liên tục. Các điểm trong ảnh được so sánh với ngưỡng định trước. Giá trị của ngưỡng sẽ quyết định điểm có được hiển thị hay không. Do vậy ảnh kết quả sẽ mất đi một số chi tiết. Có nhiều kỹ thuật chọn ngưỡng áp dụng cho các đối tượng khác nhau:

Hiển thị 2 màu: chỉ dùng ảnh đen trắng có 256 mức xám. Bản chất của phương pháp này là chọn ngưỡng dựa trên lược đồ mức xám của ảnh. Để đơn giản có thể lấy ngưỡng với giá trị là 127. Như vậy nghĩa là các điểm có giá trị điểm ảnh lớn hơn 127 sẽ là 1, ngược lại là 0.

Hiển thị 4 màu: sử dụng 4 màu để khắc phục nhược điểm của kỹ thuật hiển thị 2 màu. Một ví dụ của Bảng mã 4 màu được cho ở Bảng 2.1.

Mã màu	Màn hình monochrome (đơn sắc)	Màn hình màu
0	Đen	Đen
1	Xám đậm	Đỏ
2	Xám nhạt	Xanh
3	Trắng	Vàng

Bảng 2.1 Bảng mã 4 màu

b) Kỹ thuật chọn theo mẫu (Patterning)

Kỹ thuật này sử dụng một nhóm các phần tử trên thiết bị ra (máy in chẳng hạn) để biểu diễn một pixel trên ảnh nguồn. Các phần tử của nhóm quyết định độ sáng tối của cả nhóm. Các phần tử này mô phỏng các chấm đen trong kỹ thuật nửa cường độ. Nhóm thường được chọn có dạng ma trận vuông. Nhóm $n \times n$ phần tử sẽ tạo nên $n^2 + 1$ mức sáng. Ma trận mẫu thường được chọn là ma trận Rylander. Ma trận Rylander cấp 4 có dạng như Bảng 2.2.

0	8	2	10
4	12	6	14
3	11	1	9
7	15	5	13

Bảng 2.2. Ma trận Rylander cấp 4

Việc chọn kích thước của nhóm như vậy sẽ làm giảm độ mịn của ảnh. Vì vậy kỹ thuật này chỉ áp dụng trong trường hợp mà độ phân giải của thiết bị ra lớn hơn độ phân giải của ảnh nguồn.

Ví dụ: thiết bị ra có độ phân giải 640×480 khi sử dụng nhóm có kích thước 4×4 sẽ chỉ còn 160×120 .

c) Kỹ thuật Dithering

Dithering là việc biến đổi một ảnh đa cấp xám (nhiều mức sáng tối) sang ảnh nhị phân (hai mức sáng tối). Kỹ thuật Dithering được áp dụng để tạo ra ảnh đa cấp sáng khi độ phân giải nguồn và đích là như nhau. Kỹ thuật này sử dụng một ma trận mẫu gọi là ma trận Dither. Ma trận này gần giống như ma trận Rylander.

Để tạo ảnh, mỗi phần tử của ảnh gốc sẽ được so sánh với phần tử tương ứng của ma trận Dither. Nếu lớn hơn, phần tử ở đầu ra sẽ sáng và ngược lại.

d) Kỹ thuật khuếch tán lỗi (Error diffusion)

Khuếch tán lỗi cho phép giảm thiểu mức độ mất chi tiết của ảnh khi tách ngưỡng bằng cách phân tán lỗi do lượng tử hóa ra các điểm ảnh xung quanh. Bằng cách này tổng giá trị điểm ảnh của một vùng nhỏ được giữ tương đối gần với giá trị trên ảnh gốc.

Khuếch tán lỗi có hai dạng, khuếch tán lỗi một chiều và khuếch tán lỗi hai chiều. Khuếch tán lỗi một chiều chỉ phân tán lỗi sang điểm ảnh lân cận trên cùng một dòng trong khi đó khuếch tán lỗi hai chiều phân tán lỗi sang các điểm ảnh lân cận theo một tỷ lệ xác định bởi một ma trận khuếch tán lỗi.

- **Kỹ thuật khuếch tán lỗi một chiều:**

- Ảnh được duyệt từ trái qua phải, từ trên xuống dưới.
- Tại mỗi điểm ảnh, giá trị điểm ảnh được tách theo ngưỡng có sẵn.
- Phần dư do lượng tử hóa được chuyển sang điểm ảnh tiếp theo trên cùng dòng.

- Các bước được lặp lại cho đến hết dòng, phần dư của điểm ảnh cuối cùng sẽ được loại bỏ.
- Kỹ thuật khuếch tán lỗi hai chiều:
 - Các bước được thực hiện như khuếch tán lỗi một chiều, tuy nhiên lỗi do lượng tử hóa sẽ được phân tán ra các điểm xung quanh theo tỷ lệ xác định bởi ma trận khuếch tán
 - Có hai ma trận phổ biến (Hình 2.3):
 - Ma trận của Floyd và Steinberg.
 - Ma trận của J F Jarvis, C N Judice, và W H Ninke tại phòng thí nghiệm Bell.

$$\frac{1}{16} \begin{bmatrix} - & \# & 7 \\ 3 & 5 & 1 \end{bmatrix} \qquad \frac{1}{48} \begin{bmatrix} - & - & \# & 7 & 5 \\ 3 & 5 & 7 & 5 & 3 \\ 1 & 3 & 5 & 3 & 1 \end{bmatrix}$$

Floyd và Steinberg Phòng thí nghiệm Bell

Hình 2.3. Ma trận khuếch tán lỗi,

"#" chỉ điểm đang xét và "-" chỉ các điểm đã xét, các con số chỉ tỷ lệ lỗi được phân tán từ điểm hiện tại

2.6. KHÁI NIỆM ẢNH ĐEN TRẮNG, ẢNH MÀU

Ảnh trong thực tế là một ảnh liên tục về không gian và giá trị cường độ sáng. Để có thể xử lý ảnh bằng máy tính người ta cần thiết phải số hóa ảnh. Trong quá trình số hóa, ta biến đổi tín hiệu liên tục sang tín hiệu rời rạc thông qua quá trình lấy mẫu (rời rạc hóa về không gian) và lượng hóa thành phần giá trị mà về nguyên tắc bằng mắt thường không phân biệt được hai điểm kề nhau. Một ảnh sẽ là một tập hợp các phần tử ảnh (Picture element) hay còn được gọi là Pixel.

Như vậy, ảnh số là một tập hợp các điểm ảnh. Khi được số hóa, nó thường được biểu diễn bởi một mảng hai chiều $I(n,p)$ với n là số dòng và p là số cột. Ta nói ảnh gồm $n \times p$ pixels. Người ta thường ký hiệu $I(x,y)$ để chỉ một pixel I chính là giá trị cường độ sáng của ảnh tại pixel đó. Ảnh có thể được biểu diễn với các độ phân giải khác nhau. Một pixel có thể được lưu trữ trên 1, 4, 8 hay 24 bit tùy thuộc vào các mức cường độ sáng cần phân biệt của ảnh và số lượng thành phần màu cơ bản chứa trong ảnh đó.

2.6.1. Phân loại ảnh số

Người ta thường chia ảnh số ra làm ba loại chính là:

- Ảnh đen trắng: Mỗi phần tử ảnh nhận một trong hai giá trị tương ứng với hai mức sáng đen và trắng (còn gọi là ảnh nhị phân)
- Ảnh đa mức xám (Grayscale): Các phần tử ảnh chứa thông số về cường độ sáng đã được mã hoá thành N mức (8, 256 hoặc nhiều hơn) tương ứng với 3 bit, 8 bit hoặc hơn nữa, ảnh xám có một đặc trưng là lược đồ xám (histogram).
- Ảnh màu: Mỗi phần tử ảnh được lưu trữ trong ảnh dưới dạng một cấu trúc có 3 trường chứa thông tin về 3 màu cơ bản là đỏ, xanh, lơ (red, blue, green). Màu của

2.3.2. Mô hình Vector

Biểu diễn ảnh ngoài tiêu chí tiết kiệm không gian lưu trữ, dễ dàng cho hiển thị và in ấn còn cần đảm bảo dễ dàng trong lựa chọn, sao chép, di chuyển và tìm kiếm v.v.. Theo những yêu cầu này kỹ thuật biểu diễn Vector tỏ ra ưu việt hơn.

Trong mô hình vector người ta sử dụng hướng giữa các vector của điểm ảnh lân cận để mã hoá và tái tạo hình ảnh ban đầu ảnh vector được thu nhận trực tiếp từ các thiết bị số hoá như Digital hoặc được chuyển đổi từ ảnh Raster thông qua các chương trình số hoá

Công nghệ phần cứng cung cấp những thiết bị xử lý với tốc độ nhanh và chất lượng cho cả đầu vào và ra nhưng lại chỉ hỗ trợ cho ảnh Raster. Do vậy, những nghiên cứu về biểu diễn Vector đều tập trung vào việc chuyển đổi ảnh từ mô hình Raster.



2.4. CÁC ĐỊNH DẠNG ẢNH CƠ BẢN

Hiện nay trên thế giới có trên 50 khuôn dạng ảnh thông dụng. Sau đây là một số định dạng ảnh hay dùng trong quá trình xử lý ảnh hiện nay.

2.4.1. Định dạng ảnh IMG

Ảnh IMG là ảnh đen trắng, phần đầu của ảnh IMG có 16 byte chứa các thông tin:

- 6 byte đầu: dùng để đánh dấu định dạng ảnh. Giá trị của 6 byte này viết dưới dạng Hexa: 0x0001 0x0008 0x0001
- 2 byte tiếp theo: chứa độ dài mẫu tin. Đó là độ dài của dãy các byte kề liền nhau mà dãy này sẽ được lặp lại một số lần nào đó. Số lần lặp này sẽ được lưu trong byte đếm. Nhiều dãy giống nhau được lưu trong một byte.
- 4 byte tiếp: mô tả kích cỡ pixel.
- 2 byte tiếp: số pixel trên một dòng ảnh.
- 2 byte cuối: số dòng ảnh trong ảnh.

Ảnh IMG được nén theo từng dòng, mỗi dòng bao gồm các gói (pack). Các dòng giống nhau cũng được nén thành một gói. Có 4 loại gói sau:

- **Loại 1: Gói các dòng giống nhau.** Quy cách gói tin này như sau: 0x00 0x00 0xFF Count. Ba byte đầu tiên cho biết số các dãy giống nhau, byte cuối cho biết số các dòng giống nhau.
- **Loại 2: Gói các dãy giống nhau.** Quy cách gói tin này như sau: 0x00 Count. Byte thứ hai cho biết số các dãy giống nhau được nén trong gói. Độ dài của dãy ghi ở đầu tệp.
- **Loại 3: Dãy các Pixel không giống nhau, không lặp lại và không nén được.** Quy cách gói tin này như sau: 0x80 Count. Byte thứ hai cho biết độ dài dãy các pixel không giống nhau không nén được.

- **Loại 4: Dãy các Pixel giống nhau.** Tùy theo các bit cao của byte đầu tiên được bật hay tắt. Nếu bit cao được bật (giá trị 1) thì đây là gói nén các byte chỉ gồm bit 0, số các byte được nén được tính bởi 7 bit thấp còn lại. Nếu bit cao tắt (giá trị 0) thì đây là gói nén các byte gồm toàn bit 1. Số các byte được nén được tính bởi 7 bit còn lại.

Các gói tin của file IMG rất đa dạng do ảnh IMG là ảnh đen trắng, do vậy chỉ cần 1 bit cho 1 pixel thay vì 4 hoặc 8 như đã nói ở trên. Toàn bộ ảnh chỉ có những điểm sáng và tối tương ứng với giá trị 1 hoặc 0. Tỷ lệ nén của kiểu định dạng này là khá cao.

2.4.2. Định dạng ảnh PCX

Định dạng ảnh PCX là một trong những định dạng ảnh cổ điển. Nó sử dụng phương pháp mã hoá loạt dài RLE (Run – Length – Encoded) để nén dữ liệu ảnh. Quá trình nén và giải nén được thực hiện trên từng dạng ảnh. Thực tế, phương pháp giải nén PCX kém hiệu quả hơn so với kiểu IMG. Tập PCX gồm 3 phần: đầu tệp (header), dữ liệu ảnh (Image data) và bảng màu mở rộng.

Header của tệp PCX có kích thước cố định gồm 128 byte và được phân bố như sau:

- 1 byte: chỉ ra kiểu định dạng. Nếu là PCX/PCC thì nó luôn có giá trị là 0Ah.
- 1 byte: chỉ ra version sử dụng để nén ảnh, có thể có các giá trị sau:
 - + 0: version 2.5.
 - + 2: version 2.8 với bảng màu.
 - + 3: version 2.8 hay 3.0 không có bảng màu.
 - + 5: version 3.0 có bảng màu.
- 1 byte: chỉ ra phương pháp mã hoá. Nếu là 0 thì mã hoá theo phương pháp BYTE PACKED, ngược lại là RLE.
- 1 byte: Số bit cho một điểm ảnh phẳng (plane).

- 1 word: toạ độ góc trái của ảnh. Với kiểu PCX nó có giá trị là (0,0), cũn PCC thì khác (0,0).
- 1 word: toạ độ góc phải dưới.
- 1 word: kích thước bề rộng và bề cao của ảnh.
- 1 word: số điểm ảnh.
- 1 word: độ phân giải màn hình.
- 1 word.
- 48 byte: chia nó thành 16 nhóm, mỗi nhóm 3 byte. Mỗi nhóm này chứa thông tin về một thanh ghi màu. Như vậy ta có 16 thanh ghi màu.
- 1 byte: không dùng đến và luôn đặt là 0.
- 1 byte: số bít plane mà ảnh sử dụng. Với ảnh 16 màu, giá trị này là 4, với ảnh 256 màu (1pixel/8bits) thì số bít plane lại là 1.
- 1 byte: số bytes cho một dòng quét ảnh.
- 1 word: kiểu bảng màu.
- 58 byte: không dùng.

Định dạng ảnh PCX thường được dùng để lưu trữ ảnh và thao tác đơn giản, cho phép nén và giải nén nhanh. Tuy nhiên, vì cấu trúc của nó cố định, nên trong một số trường hợp làm tăng kích thước lưu trữ. Cũng vì nhược điểm này mà một số ứng dụng sử dụng một kiểu định dạng khác mềm dẻo hơn: định dạng TIFF (Targed Image File Format) sẽ mô tả dưới đây.

2.4.3. Định dạng ảnh TIFF

Kiểu định dạng TIFF được thiết kế để làm nhẹ bớt các vấn đề liên quan đến việc mở rộng file ảnh cố định. Về cấu trúc, nó cũng gồm 3 phần chính:

- **Phần Header (IFH):** có trong tất cả các tệp TIFF và gồm 8 byte:
 - + 1 word: chỉ ra kiểu tạo tệp trên máy tính PC hay máy Macintosh. Hai loại này khác nhau rất lớn ở thứ tự các byte lưu trữ trong các số dài 2 hay 4 byte. Nếu trường này có giá trị là 4D4Dh thì đó là ảnh cho máy Macintosh, nếu là 4949h là của máy PC.
 - + 1 word: version. từ này luôn có giá trị là 42. đây là đặc trưng của file TIFF và không thay đổi.
 - + 2 word: giá trị Offset theo byte tính từ đầu tới cấu trúc IFD là cấu trúc thứ hai của file. Thứ tự các byte này phụ thuộc vào dấu hiệu trường đầu tiên.
- **Phần thứ 2 (IFD):** Không ở ngay sau cấu trúc IFH mà vị trí được xác định bởi trường Offset trong đầu tệp. Có thể có một hay nhiều IFD cùng tồn tại trong một file. Một IFD bao gồm:
 - + 2 byte: chứa các DE (Directory Entry).
 - + 12 byte là các DE xếp liên tiếp, mỗi DE chiếm 12 byte.
 - + 4 byte: chứa Offset trở tới IFD tiếp theo. Nếu đây là IFD cuối cùng thì trường này có giá trị 0.
- **Phần thứ 3:** các DE: các DE có độ dài cố định gồm 12 byte và chia làm 4 phần:
 - + 2 byte: chỉ ra dấu hiệu mà file ảnh đó được xây dựng.
 - + 2 byte: kiểu dữ liệu của tham số ảnh. Có 5 kiểu tham số cơ bản:
 - 1: BYTE (1 byte)
 - 2: ASCII (1 byte)
 - 3: SHORT (2 byte).
 - 4: LONG (4 byte)
 - 5: RATIONAL (8 byte)
 - + 4 byte: trường độ dài chứa số lượng chỉ mục của kiểu dữ liệu đó chỉ ra. Nó không phải là trường bắt buộc cần thiết để lưu trữ. Để có số liệu này ta cần nhân số chỉ mục với kích thước của kiểu dữ liệu đã dùng.
 - + 4 byte: đó là Offset tới điểm bắt đầu dữ liệu liên quan tới dấu hiệu, tức là

liên quan với DE không phải lưu trữ vật lý cùng với nó nằm ở một vị trí nào đó trong file.

Chương 2: Thu nhận ảnh

Dữ liệu chứa trong tệp thường được tổ chức thành các nhóm dòng (cột) quét của dữ liệu ảnh. Cách tổ chức này làm giảm bộ nhớ cần thiết cho việc đọc tệp. Việc giải nén được thực hiện theo 4 kiểu khác nhau được lưu trữ trong byte dấu hiệu nén.

2.4.4. Định dạng file ảnh BI

 0 |  0

Mỗi file BITMAP gồm đầu file chứa các thông tin chung về file, đầu thông tin chứa

các thông tin về ảnh, một bảng màu và một mảng dữ liệu ảnh. Khuôn dạng được cho như sau:

```
BITMAPFILEHEADER    bmfh;  
BITMAPINFOHEADER    bmih;  
RGBQUAD             aColors[];  
BYTE                aBitmapBits[];
```

Trong đó, các cấu trúc được định nghĩa như sau:

```
typedef struct tagBITMAPFILEHEADER { /* bmfh */
```

```
    UINT        bfType;  
    DWORD       bfSize;  
    UINT        bfReserved1;  
    UINT        bfReserved2;  
    DWORD       bfOffBits;
```

```
} BITMAPFILEHEADER;
```

```
typedef struct tagBITMAPINFOHEADER { /* bmih */
```

```
    DWORD       biSize;  
    LONG        biWidth;  
    LONG        biHeight;  
    WORD        biPlanes;  
    WORD        biBitCount;  
    DWORD       biCompression;  
    DWORD       biSizeImage;  
    LONG        biXPelsPerMeter;  
    LONG        biYPelsPerMeter;  
    DWORD       biClrUsed;  
    DWORD       biClrImportant;
```

```
} BITMAPINFOHEADER, *LPBITMAPINFOHEADER;
```

Chương 2: Thu nhận ảnh

Với biến chi tiết như sau

biSize	kích thước của BITMAPINFOHEADER
biWidth	Chiều rộng của ảnh, tính bằng số điểm ảnh
biHeight	Chiều cao của ảnh, tính bằng số điểm ảnh
biPlanes	Số plane của thiết bị, phải bằng 1
biBitCount	Số bit cho một điểm ảnh
biCompression	Kiểu nén
biSizeImage	Kích thước của ảnh tính bằng byte
biXPelsPerMeter	độ phân giải ngang của thiết bị, tính bằng điểm ảnh trên met
biYPelsPerMeter	độ phân giải dọc của thiết bị, tính bằng điểm ảnh trên met
biClrUsed	Số lượng các màu thực sự được sử dụng
biClrImportant	Số lượng các màu cần thiết cho việc hiển thị, bằng 0 nếu tất cả các màu đều có thể hiển thị

Nếu `bmih.biBitCount > 8` thì mảng màu `rgbq[]` trống, ngược lại thì mảng màu có 2^{biBitCount}

```
typedef struct tagRGBQUAD { /* rgbq */
    BYTE    rgbBlue;
    BYTE    rgbGreen;
    BYTE    rgbRed;
    BYTE    rgbReserved;
} RGBQUAD;

typedef struct tagBITMAPINFO {
    BITMAPINFOHEADER    bmiHeader;
    RGBQUAD              bmiColors[1];
} BITMAPINFO, *PBITMAPINFO;
```

2.5. KỸ THUẬT TÁI NHẬN ẢNH

2.5.1 Kỹ thuật chụp ảnh

Phương pháp sao chụp ảnh là phương pháp đơn giản, giá thành thấp, chất lượng cao. Sau bước chụp là kỹ thuật phòng tối nhằm tăng cường ảnh như mong muốn. Ví dụ kỹ thuật phòng tối như: phóng đại ảnh, thu nhỏ ảnh..., tùy theo ứng dụng. Kỹ thuật chụp ảnh màn hình màu khá đơn giản. Nó bao gồm các bước sau:

- Đặt camera trong phòng tối, cách màn hình khoảng 10 feet (1 feet=0,3048m)
- Mở ống kính để phẳng mặt cong màn hình, do vậy ảnh sẽ dàn đều hơn

- Tắt phím sáng tối (Brightness) và phím tương phản (Contrast) của màn hình để tạo độ rõ cho ảnh. Các màu chói, cường độ cao trên ảnh sẽ giảm đi.
- Đặt tốc độ ống kính từ 1/8 đến 1/2 giây.

2.5.2 Kỹ thuật in ảnh

Người ta dùng kỹ thuật nửa cường độ để thể hiện ảnh trên sách báo, tạp chí. Theo kỹ thuật này, một ảnh tạo nên bởi một chuỗi các điểm in trên giấy. Thực chất, mỗi điểm ảnh có thể được coi là một hình vuông trắng bao quanh một chấm đen. Do vậy, nếu chấm đen càng lớn điểm ảnh sẽ càng xẫm màu. Màu xám có thể coi như chấm đen chiếm nửa vùng trắng. Vùng trắng là vùng gồm một chùm các điểm ảnh có rất ít hoặc không có chấm đen.



Từ đặc điểm cảm nhận của mắt người, sự thay đổi cường độ chấm đen trong các phần tử ảnh trắng tạo nên mô phỏng của một ảnh liên tục. Như vậy, mắt người cảm nhận từ một ảnh mà màu biến đổi từ đen qua xám rồi đến trắng. Tổng số cường độ duy nhất hiện diện sẽ xác định các kích thước khác nhau của chấm đen. Thông thường, báo ảnh tạo ảnh nửa cường độ với độ phân giải từ 60 đến 80 DPI(dot per inchs: số điểm ảnh trên một inch), sách có thể in đến 150 DPI.

Tuy nhiên, các máy in ghép nối với máy tính không có khả năng sắp xếp các chấm đen có kích thước khác nhau của ảnh, vì vậy chúng ta cần các kỹ thuật nửa cường độ kỹ thuật số (Digital Halftoning): phân ngưỡng, chọn mẫu (Patterning), Dithering hoặc khuếch tán lỗi (Error Diffusion).

a) Phân ngưỡng

Kỹ thuật này đặt ngưỡng để hiển thị các tông màu liên tục. Các điểm trong ảnh được so sánh với ngưỡng định trước. Giá trị của ngưỡng sẽ quyết định điểm có được hiển thị hay không. Do vậy ảnh kết quả sẽ mất đi một số chi tiết. Có nhiều kỹ thuật chọn ngưỡng áp dụng cho các đối tượng khác nhau:

Hiển thị 2 màu: chỉ dùng ảnh đen trắng có 256 mức xám. Bản chất của phương pháp này là chọn ngưỡng dựa trên lược đồ mức xám của ảnh. Để đơn giản có thể lấy ngưỡng với giá trị là 127. Như vậy nghĩa là các điểm có giá trị điểm ảnh lớn hơn 127 sẽ là 1, ngược lại là 0.

Hiển thị 4 màu: sử dụng  0 |  0 khắc phục nhược điểm của kỹ thuật hiển thị 2 màu. Một ví dụ của Bảng mã 4 màu được cho ở Bảng 2.1.

Mã màu	Màn hình monochrome (đơn sắc)	Màn hình màu
0	Đen	Đen
1	Xám đậm	Đỏ
2	Xám nhạt	Xanh
3	Trắng	Vàng

Bảng 2.1 Bảng mã 4 màu

b) Kỹ thuật chọn theo mẫu (Patterning)

Kỹ thuật này sử dụng một nhóm các phần tử trên thiết bị ra (máy in chẳng hạn) để biểu diễn một pixel trên ảnh nguồn. Các phần tử của nhóm quyết định độ sáng tối của cả nhóm. Các phần tử này mô phỏng các chấm đen trong kỹ thuật nửa cường độ. Nhóm thường được chọn có dạng ma trận vuông. Nhóm $n \times n$ phần tử sẽ tạo nên $n^2 + 1$ mức sáng. Ma trận mẫu thường được chọn là ma trận Rylander. Ma trận Rylander cấp 4 có dạng như Bảng 2.2.

0	8	2	10
4	12	6	14
3	11	1	9
7	15	5	13

Bảng 2.2. Ma trận Rylander cấp 4

Việc chọn kích thước của nhóm như vậy sẽ làm giảm độ mịn của ảnh. Vì vậy kỹ thuật này chỉ áp dụng trong trường hợp mà độ phân giải của thiết bị ra lớn hơn độ phân giải của ảnh nguồn.

Ví dụ: thiết bị ra có độ phân giải 640×480 khi sử dụng nhóm có kích thước 4×4 sẽ chỉ còn 160×120 .

c) Kỹ thuật Dithering

Dithering là việc biến đổi một ảnh đa cấp xám (nhiều mức sáng tối) sang ảnh nhị phân (hai mức sáng tối). Kỹ thuật Dithering được áp dụng để tạo ra ảnh đa cấp sáng khi độ phân giải nguồn và đích là như nhau. Kỹ thuật này sử dụng một ma trận mẫu gọi là ma trận Dither. Ma trận này gần giống như ma trận Rylander.

Để tạo ảnh, mỗi phần tử của ảnh gốc sẽ được so sánh với phần tử tương ứng của ma trận Dither. Nếu lớn hơn, phần tử ở đầu ra sẽ sáng và ngược lại.

d) Kỹ thuật khuếch tán lỗi (Error diffusion)

Khuếch tán lỗi cho phép giảm thiểu mức độ mất chi tiết của ảnh khi tách ngưỡng bằng cách phân tán lỗi do lượng tử hóa ra các điểm ảnh xung quanh. Bằng cách này tổng giá trị điểm ảnh của một vùng nhỏ được giữ tương đối gần với giá trị trên ảnh gốc.

Khuếch tán lỗi có hai dạng, khuếch tán lỗi một chiều và khuếch tán lỗi hai chiều. Khuếch tán lỗi một chiều chỉ phân tán lỗi sang điểm ảnh lân cận trên cùng một dòng trong khi đó khuếch tán lỗi hai chiều phân tán lỗi sang các điểm ảnh lân cận theo một tỷ lệ xác định bởi một ma trận khuếch tán lỗi.

- Kỹ thuật khuếch tán lỗi một chiều:**

- Ảnh được duyệt từ trái qua phải, từ trên xuống dưới.
- Tại mỗi điểm ảnh, giá trị điểm ảnh được tách theo ngưỡng có sẵn.
- Phần dư do lượng tử hóa được chuyển sang điểm ảnh tiếp theo trên cùng dòng.

Chương 2: Thu nhận ảnh

- Các bước được lặp lại cho đến hết dòng, phần dư của điểm ảnh cuối cùng sẽ được loại bỏ.
- Kỹ thuật khuếch tán lỗi hai chiều:
 - Các bước được thực hiện như khuếch tán lỗi một chiều, tuy nhiên lỗi do lượng tử hóa sẽ được phân tán ra các điểm xung quanh theo tỷ lệ xác định bởi ma trận khuếch tán
 - Có hai ma trận phổ biến (Hình 2.3):
 - Ma trận của Floyd và Steinberg.
 - Ma trận của J. P. Borge, C N Judice, và W H Ninke tại phòng thí nghiệm Bell.

$$\frac{1}{16} \begin{bmatrix} - & \# & 7 \\ 3 & 5 & 1 \end{bmatrix}$$

Floyd và Steinberg

$$\frac{1}{48} \begin{bmatrix} - & - & \# & 7 & 5 \\ 3 & 5 & 7 & 5 & 3 \\ 1 & 3 & 5 & 3 & 1 \end{bmatrix}$$

Phòng thí nghiệm Bell

Hình 2.3. Ma trận khuếch tán lỗi,

"#" chỉ điểm đang xét và "-" chỉ các điểm đã xét, các con số chỉ tỷ lệ lỗi được phân tán từ điểm hiện tại

2.6. KHÁI NIỆM ẢNH ĐEN TRẮNG, ẢNH MÀU

Ảnh trong thực tế là một ảnh liên tục về không gian và giá trị cường độ sáng. Để có thể xử lý ảnh bằng máy tính người ta cần thiết phải số hóa ảnh. Trong quá trình số hóa, ta biến đổi tín hiệu liên tục sang tín hiệu rời rạc thông qua quá trình lấy mẫu (rời rạc hóa về không gian) và lượng hóa thành phần giá trị mà về nguyên tắc bằng mắt thường không phân biệt được hai điểm kề nhau. Một ảnh sẽ là một tập hợp các phần tử ảnh (Picture element) hay còn được gọi là Pixel.

Như vậy, ảnh số là một tập hợp các điểm ảnh. Khi được số hóa, nó thường được biểu diễn bởi một mảng hai chiều $I(n,p)$ với n là số dòng và p là số cột. Ta nói ảnh gồm $n \times p$ pixels. Người ta thường ký hiệu $I(x,y)$ để chỉ một pixel I chính là giá trị cường độ sáng của ảnh tại pixel đó. Ảnh có thể được biểu diễn với các độ phân giải khác nhau. Một pixel có thể được lưu trữ trên 1, 4, 8 hay 24 bit tùy thuộc vào các mức cường độ sáng cần phân biệt của ảnh và số lượng thành phần màu cơ bản chứa trong ảnh đó.

2.6.1. Phân loại ảnh số

Người ta thường chia ảnh số ra làm ba loại chính là:

- Ảnh đen trắng: Mỗi phần tử ảnh nhận một trong hai giá trị tương ứng với hai mức sáng đen và trắng (còn gọi là ảnh nhị phân)
- Ảnh đa mức xám (Grayscale): Các phần tử ảnh chứa thông số về cường độ sáng đã được mã hoá thành N mức (8, 256 hoặc nhiều hơn) tương ứng với 3 bit, 8 bit hoặc hơn nữa, ảnh xám có một đặc trưng là lược đồ xám (histogram).
- Ảnh màu: Mỗi phần tử ảnh được lưu trữ trong ảnh dưới dạng một cấu trúc có 3 trường chứa thông tin về 3 màu cơ bản là đỏ, xanh, lơ (red, blue, green). Màu của

Chương 2: Thu nhận ảnh

ảnh sẽ là tổng hợp của 3 giá trị trên. Mỗi trường biểu diễn giá trị màu có thể dùng 8 bit, 16 bit hoặc 24 bit để mã hoá.

Như vậy ảnh màu là ảnh mang thông tin về đối tượng đầy đủ nhất so với ảnh đen trắng và ảnh đa cấp xám. Tuy nhiên, tất cả cách biểu diễn này đều chỉ là sự mô phỏng hệ màu trong tự nhiên. Thực tế thì một màu được phân biệt qua 3 thuộc tính là độ chói (Intensity), sắc thái màu (Hue) và độ bão hoà (Saturation). Tuy nhiên với khả năng cảm nhận của mắt người thì cách biểu diễn ảnh dưới dạng cấu trúc 3 màu RGB chất lượng ảnh thu được là có thể chấp nhận được.

2.6.2. Màu sắc

Màu sắc được tạo ra bởi các ánh sáng với các bước sóng khác nhau, mắt người bao gồm ba loại tế bào cảm nhận màu có thể nhìn được bảy triệu màu nhưng thực chất chúng ta chỉ có thể cảm nhận sự khác biệt vài ngàn màu. Một màu có thể được biểu diễn bởi ba thuộc tính: Sắc thái màu (Hue), độ bão hoà (Saturation), và độ chói (Intensity).

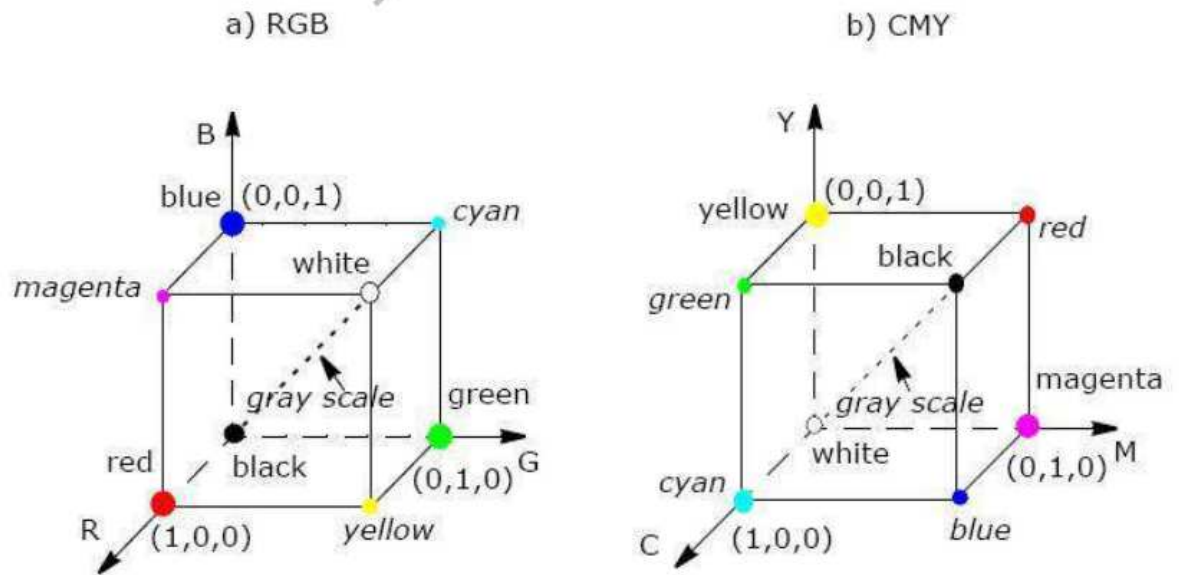
Trong xử lý ảnh và đồ họa, mô hình màu là một chỉ số kỹ thuật của một hệ tọa độ màu 3 chiều có thể dùng để biểu diễn tất cả các màu. Ví dụ như mô hình màu RGB (Red, Green, Blue): là một đơn vị tập các màu thành phần sắp xếp theo hình lập phương của hệ trục tọa độ Đề các.

Mục đích của mô hình màu là cho phép biểu diễn một phần các màu nhìn thấy được bằng các chỉ số kỹ thuật quy ước. Sau đây, ta xem xét một số mô hình hay được sử dụng nhất.

2.6.2.1. Mô hình màu RGB (Red Green Blue)

Màu đỏ, lục – xanh lá, xanh da trời (RGB) được sử dụng phổ biến nhất trong hiển thị. Các màu gốc này được tổ hợp với nhau theo một tỷ lệ để tái tạo màu sắc, hệ

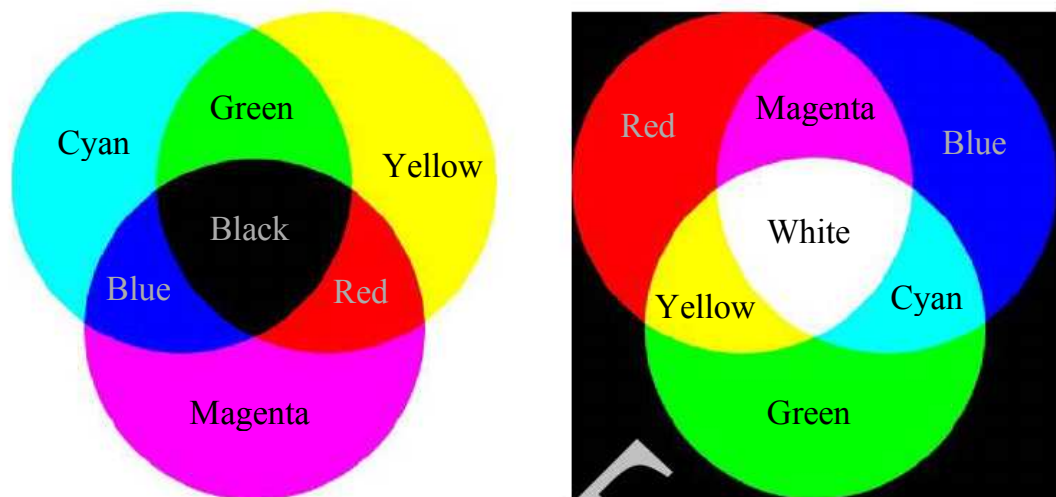
màu này sử dụng phối màu cộng. Tập hợp các màu được sắp xếp theo khối lập phương đơn vị. Đường chéo chính của khối lập phương thể hiện các màu được phối bởi ba màu gốc với tỷ lệ tương đương nhau ứng với các mức độ xám từ đen là $(0,0,0)$ tới trắng $(1,1,1)$.



Hình 2.4. Mô hình màu RGB và CMY

2.6.2.2. Mô hình màu CMY (Cyan, Magenta, Yellow)

Xanh nhạt (Cyan), vàng (Yellow) và tím (Magenta) là các phần bù tương ứng cho các màu đỏ, lục và lam. Chúng có thể được sử dụng để lọc loại trừ các màu này từ ánh sáng trắng. Vì vậy CMY còn được gọi là các phần bù loại trừ của màu gốc. Các màu trong hệ màu CMY cũng được biểu diễn trong hệ tọa độ Đề-các nhưng phương pháp phối màu sử dụng là phối màu trừ. Các màu được tạo thành bằng cách loại bỏ hoặc được bù từ ánh sáng trắng thay vì là được thêm vào vùng tối.



Hình 2.5. Trộn màu cộng và trộn màu trừ

Khi bề mặt được bao phủ bởi lớp mực màu xanh tím, sẽ không có tia màu đỏ phản chiếu từ bề mặt đó. Màu xanh tím đã loại bỏ phần màu đỏ phản xạ khi có tia sáng trắng, mà bản chất là tổng của 3 màu đỏ, lục, lam. Vì thế ta có thể coi màu Cyan là màu trắng trừ đi màu đỏ và đó cũng là màu lam cộng màu lục. Tương tự như vậy ta có màu đỏ thẫm (magenta) hấp thụ màu lục, vì thế nó tương đương với màu đỏ cộng màu lam. Và cuối cùng màu vàng (yellow) hấp thụ màu lam, nó sẽ bằng màu đỏ cộng với lục.

Khi bề mặt của thực thể được bao phủ bởi xanh tím và vàng, chúng sẽ hấp thụ hết các phần màu đỏ và xanh lam của bề mặt. Khi đó chỉ tồn tại duy nhất màu lục bị phản xạ từ sự chiếu sáng của ánh sáng trắng. Trong trường hợp khi bề mặt được bao phủ bởi cả 3 màu xanh tím, vàng, đỏ thẫm, hiện tượng hấp thụ xảy ra trên cả 3 màu đỏ, lục và lam. Do đó, màu đen sẽ màu của bề mặt. Những mối liên hệ này có thể được miêu tả bởi:

$$\begin{bmatrix} C \\ M \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

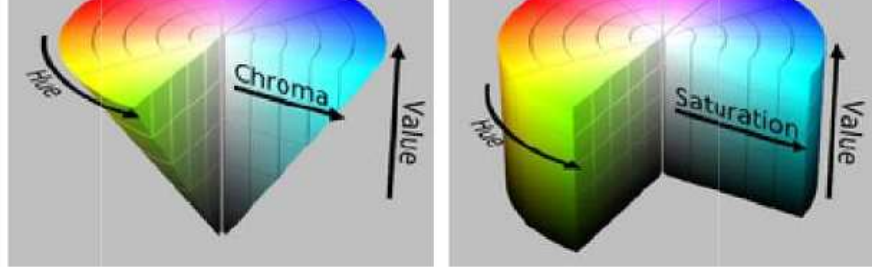
2.6.2.3. Mô hình màu HSV (Hue, Saturation, Value)

Các mô hình màu RGB, CMY có thể hiển thị được tất cả các màu cần thiết, tuy vậy rất khó khăn để con người có thể phối màu trên hai hệ màu này. Để phối màu dễ dàng hơn Smith định nghĩa mô hình màu HSV của hay HSB với B là Brightness (độ sáng) hướng người sử dụng dựa trên cơ sở nền tảng trực giác về tông màu, sắc độ và sắc thái mỹ thuật.

Chương 2: Thu nhận ảnh

Hệ thống tọa độ có dạng hình trụ và tập màu thành phần của không gian bên trong mô hình màu được xác định bởi các trục màu trong hình 2.7.





Hình 2.7. Mô hình màu HSV

Sắc màu (hue) hoặc H được đo bởi góc quanh trục đứng với màu đỏ là 0° , màu lục là 120° , màu lam là 240° . Các màu bù nằm ở vị trí đối diện với những màu gốc. S lấy giá trị từ 0 trên đường trục tâm (trục V) đến 1 trên các mặt bên tại cửa hình chóp sáu cạnh. Sự bão hòa được hiểu là mức độ tươi của màu, các màu xám từ đen tới trắng có S là 0. Giá trị V được hiểu là độ sáng của màu, V là 0 thì là màu đen, V là 1 thì là màu có độ sáng tối đa.

Giải mã chuyển đổi từ RGB sang HSV

Hàm RGB_HSV_Conversion

H: Sắc độ màu [0-360] với màu đỏ tại điểm 0

S: Độ bão hòa [0-1]

V: Giá trị cường độ sáng [0-1]

Max: Hàm lấy giá trị cực đại

Min: Hàm lấy giá trị nhỏ nhất

```
{
    //Xác định giá trị cường độ sáng
    V= Max(R,G,B)
    //Xác định độ bão hòa
    Temp= Min(R,G,B)
    If V=0 then
        S= 0
    Else
        S= (V-Temp)/V
    End
    //Xác định sắc màu
    IF s=0 THEN
        H= Undefined
    Else
        Cr= (V-R)/(V-Temp);
        Cg= (V-G)/(V-Temp);
        Cb= (V-B)/(V-Temp);
        // Màu nằm trong khoảng giữa vàng (Yellow) và đỏ tía (Magenta)
        If R=V then
            H= Cb-Cg
```

Chương 2: Thu nhận ảnh

```
// Màu nằm trong khoảng giữa xanh tím (cyan) và vàng (yellow)
If G= V then
    H= 2+Cr-Cb
// Màu nằm trong khoảng giữa đỏ tươi (magenta) và xanh (cyan)
If B=V then
    H= 4+ Cg - Cr
H= 60*H // Chuyển sang độ
//Loại các giá trị âm
If H < 0 then
    H= H+360
}
```

Giải mã chuyển đổi từ HSV sang RGB

Hàm HSV_RGB_Conversion()

H: Sắc độ màu [0-360]  0 |  0 độ tại điểm 0

S: Độ bão hòa [0-1]

V: Giá trị màu [0-1]


```

V: Giá trị cường độ sáng [0-1]
{
    //Kiểm tra trường hợp ánh sáng không màu
    If S=0 then
        If H=Undefined then
            R= V
            G= V
            B= V
        Endif
    Else
        If H=360 then
            H= 0
        Else
            H= H/60
        endif
    I= Floor(H)
    F= H-I
    M= V*(1-S)
    N= V*(1-S*F)
    K= V*(1-S*(1-F))
    //(R,G,B)=(V,K,M)  $\Leftrightarrow$  R= V; C= K; B= M
    If I=0 then
        (R,G,B)=(V,K,M);
    If I=1 then
        (R,G,B)=(N,V,M);
    If I=2 then
        (R,G,B)=(M,V,K);

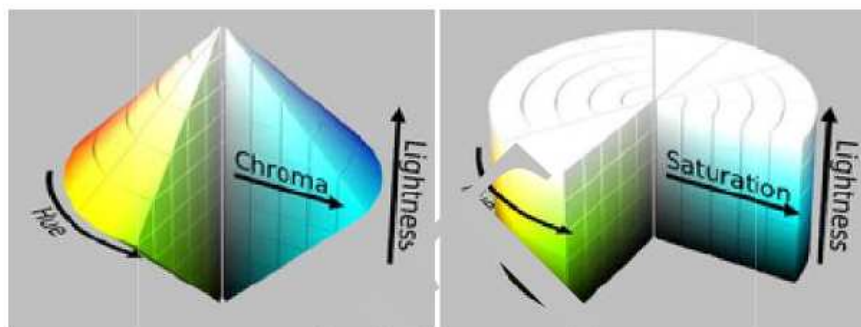
```

Chương 2: Thu nhận ảnh

```
If I=3 then
    (R, G, B) = (M, N, V) ;
If I=4 then
    (R, G, B) = (K, M, V) ;
If I=5 then
    (R, G, B) = (V, M, N) ;
}
```

2.6.2.4. Mô hình màu HSL

Mô hình màu HSL tương tự như mô hình HSV nhưng thay vì biểu diễn tập hợp màu hình nón sáu như mô hình HSV thì HSL biểu diễn các tập hợp màu hình chóp đôi dưới dạng không gian hình trụ. Sắc màu H (hue) vẫn là góc quanh trục đứng với màu đỏ tại góc 0° . Chúng ta có thể xem hệ màu HSL như một sự biến dạng của hệ HSV.



Hình 2.8. Mô hình màu HSL

Giải mã chuyển đổi từ RGB sang HSL

Hàm RGB_HLS_Conversion()

H: Sắc độ màu [0-360] với màu đỏ tại điểm 0

S: Độ bão hòa [0-1]

V: Giá trị cường độ

Max: Hàm lấy giá trị cực đại

Min: Hàm lấy giá trị nhỏ nhất

```
{  
    //Xác định độ sáng  
    M1= Max (R,G,B)  
    M2= Min (R,G,B)  
    L= (M1+M2)  
    //Xác định độ bão hòa  
    If M1=M2 //Trường hợp không màu  
        S= 0  
        H= Undefined  
    Else  
        If L <= 0.5 then //Trường hợp màu  
            S= (M1-M2) / (M1+M2)  
        Else  
            S= (M1-M2) / (2-M1-M2)
```

```
Endif
//Xác định sắc độ
Cr= (M1-R) / (M1-M2)
Cg= (M1-G) / (M1-M2)
Cb= (M1-B) / (M1-M2)
if R=M1 then
    H= Cb-Cg
If G=M1 then
    H= 2+Cr-Cb
If B=M1 then
    H= 4+Cg-Cr
H= H*60
if H<0 then
    H= H+360
endif
}
```

Giải mã chuyển đổi từ HSL sang RGB

```
Hàm HLS_RGB_Conversion()
H: Sắc độ màu [0-360] với màu đỏ tại điểm 0
S: Độ bão hòa [0-1]
V: Giá trị cường độ sáng [0-1]
{
    If L <= 0.5 then
        M2= L*(1+S)
    Else
        M2= L+S-L*S
    Endif
    M1= 2*L-M2
    //Kiểm tra độ bão hòa = 0
    If S=0 then
        If H=Undefined
            R=L
            G=L
            B=L
        Else //Error: Dữ liệu nhập sai
            Endif
    Else //Xác định giá trị của RGB
        RGB(H+120, M1,M2,Value)
        R= Value
        RGB(H, M1,M2,V
        G= Value
        RGB(H-120, M1,M2,Value)
```

```
RGB(H-120, M1, M2, Value),  
B= Value
```

Chương 2: Thu nhận ảnh

```
Endif  
}  
  
//Hàm điều chỉnh giá trị của H cho phù hợp khoảng xác định  
Hàm RGB(H, M1, M2, Value)  
{  
    If H < 0 then  
        H= H+360  
    If H < 60 then
```

```
    Value= M1+(M2-M1)*H/60
  If H >=60 and H < 180 then
    Value= M2
  If H>= 180 and H < 240 then
    Value = M1+(M2-M1)*(240-H)/60
  If H > 240 and H <= 360 then
    Value= M1
  Return
}
```

2.7. CÂU HỎI ÔN TẬP CHƯƠNG

Câu 1: Thế nào là mô hình Raster và Vector? Trình bày sự giống và khác nhau của hai mô hình này.

Câu 2: Nêu các định dạng ảnh phổ biến. Trình bày cách nén ảnh của định dạng ảnh BITMAP.

Câu 3: Kỹ thuật khuếch tán lỗi (Error diffusion) là gì? Thực hiện khuếch tán lỗi một chiều với ảnh sau, được biết ngưỡng là 127.

$$I = \begin{bmatrix} 1 & 23 & 156 & 22 & 45 \\ 133 & 13 & 12 & 12 & 212 \\ 12 & 232 & 127 & 32 & 21 \end{bmatrix}$$

Câu 4: Kỹ thuật khuếch tán lỗi hai chiều khác gì với một chiều. Thực hiện khuếch tán lỗi hai chiều với ảnh ở câu 3 sử dụng ma trận khuếch tán lỗi Floyd và Steinberg, được biết ngưỡng là 127.

Câu 5: Hệ màu HSL và HSV được phát triển vì mục đích gì? Dựa theo giả mã đã trình bày ở trên, hãy viết một phần mềm nhỏ có thể chuyển ảnh giữa các hệ màu RGB, CMY, HSV và HSV.

PL

Chương 3:

XỬ LÝ NÂNG CAO CHẤT LƯỢNG ẢNH

3.1. CẢI THIỆN ẢNH SỬ DỤNG CÁC TOÁN TỬ ĐIỂM

3.1.1. Giới thiệu

Các phép toán không phụ thuộc không gian là các phép toán không phụ thuộc vị trí của điểm ảnh. Ví dụ: Phép tăng giảm độ sáng, phép thống kê tần suất, biến đổi tần suất v.v..

Một trong những khái niệm quan trọng trong xử lý ảnh là biểu đồ tần suất (Histogram): Biểu đồ tần suất của mức xám g của ảnh I là số điểm ảnh có giá trị g của ảnh I . Ký hiệu là $h(g)$

Ví dụ:

$$I = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 1 & 0 & 0 & 7 \\ 2 & 2 & 1 & 0 \\ 4 & 1 & 2 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix}$$

g	0	1	2	4	5	6	7
$h(g)$	5	7	5	2	0	0	1

3.1.2. Tăng giảm độ sáng

Giả sử ta có ảnh I ~ kích thước $m \times n$ và số nguyên c

Khi đó, kỹ thuật tăng, giảm độ sáng được thể hiện

$$I[i, j] \leftarrow I[i, j] + c; \forall (i, j)$$

Nếu $c > 0$ thì ảnh sáng lên và $c < 0$ thì ảnh tối đi

3.1.3. Tách ngưỡng

Giả sử ta có ảnh I ~ kích thước $m \times n$, hai số Min, Max và ngưỡng θ

Khi đó, kỹ thuật tách ngưỡng được thể hiện

$$I[i, j] \geq I[i, j] \times \theta \text{? Max: Min; } \forall (i, j)$$

Nếu Min " 0, Max " 1 thì ảnh thu được sau tách ngưỡng là ảnh đen trắng.

Ảnh I được gọi là cân bằng "lý tưởng" nếu với mọi mức xám g, g' ta có $h(g) \gg h(g')$

Giả sử, ta có:

$I \sim$ kích thước $m \times n$

$new_level \sim$ số mức xám mới của ảnh sau cân bằng, thông thường giá trị này có thể bằng đúng số mức xám của ảnh gốc

$TB = \frac{m \times n}{new_level} \sim$ số điểm ảnh trung bình của mỗi mức xám của ảnh đã cân bằng

$t(g) = \sum_{i=0}^g h(i) \sim$ số điểm ảnh có mức xám $\leq g$ trên ảnh gốc

Cần xác định hàm $f: g \mapsto f(g)$ sao cho:

$$f(g) = \max \left\{ 0, \text{round} \left(\frac{t(g)}{TB} \right) - 1 \right\}$$

Ví dụ: Cân bằng ảnh sau với $new_level = 4$

$$I = \begin{pmatrix} 1 & 2 & 4 & 6 & 7 \\ 2 & 1 & 3 & 4 & 5 \\ 7 & 2 & 6 & 9 & 1 \\ 4 & 1 & 2 & 1 & 2 \end{pmatrix}$$

g	$h(g)$	$t(g)$	$f(g)$
1	5	5	0
2	5	10	1
3	1	11	1
4	3	14	2
5	1	15	2
6	2	17	2
7	2	19	3
9	1	20	3

$$I_{kq} = \begin{pmatrix} 0 & 1 & 2 & 2 & 3 \\ 1 & 0 & 1 & 2 & 2 \\ 3 & 1 & 2 & 3 & 0 \\ 2 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Ảnh sau khi thực hiện cân bằng chưa chắc đã là cân bằng "lý tưởng", tuy nhiên chúng ta có thể thấy biểu đồ tần suất của ảnh sau cân bằng đồng đều hơn rất nhiều so với ảnh gốc.

3.1.5. Kỹ thuật tìm tách ngưỡng tự động

Ngưỡng θ trong kỹ thuật tách ngưỡng thường được cung cấp bởi người sử dụng. Kỹ thuật tìm tách ngưỡng tự động xác định ngưỡng θ một cách tự động dựa vào histogram theo một nguyên lý trong vật lý là vật thể có thể tách làm 2 phần nếu tổng độ lệch trong từng phần là tối thiểu.

Giả sử, ta có ảnh

$I \sim$ kích thước $m \times n$

$G \sim$ là số mức xám của ảnh kể cả khuyết thiếu

$t(g) \sim$ số điểm ảnh có mức xám $\leq g$

$$m(g) = \frac{1}{t(g)} \sum_{i=0}^g i \cdot h(i)$$

mô-men quán tính TB có mức xám $\leq g$

Hàm $f: g \mapsto f(g)$ được định nghĩa là:

$$f(g) = \frac{t(g)}{mxn - t(g)} [m(g) - m(G-1)]^2$$

Ngưỡng θ được xác định sao cho:

$$f(\theta) = \max_{0 \leq g < G-1} \{f(g)\}$$

Ví dụ: Tìm ngưỡng tự động của ảnh sau

$$I = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

g	h(g)	t(g)	g.h(g)	$\sum_{i=0}^g ih(i)$	m(g)	f(g)
0	15	15	0	0	0	1.35
1	5	20	5	5	0,25	<u>1.66</u>
2	4	24	8	13	0,54	1.54
3	3	27	9	22	0,81	1.10
4	2	29	8	30	1,03	0.49
5	1	30	5	35	1,16	0

Ngưỡng tìm được $\theta = 1$ ứng với $f(\theta) = 1.66$.

3.1.6. Biến đổi cấp xám tổng thể

Nếu biết ảnh và hàm biến đổi thì ta có thể tính được ảnh kết quả và do đó ta sẽ có được histogram của ảnh biến đổi. Nhưng thực tế nhiều khi ta chỉ biết histogram của ảnh gốc và hàm biến đổi, câu hỏi đặt ra là liệu ta có thể có được histogram của ảnh biến đổi. Nếu có như vậy ta có thể hiệu chỉnh hàm biến đổi để thu được ảnh kết quả có phân bố histogram như mong muốn.

Bài toán đặt ra là biết histogram của ảnh, biết hàm biến đổi hãy vẽ histogram của ảnh mới. Giả sử ta có ảnh gốc I với các mức xám g và biểu đồ tần suất $h(g)$. Phép biến đổi $f(g)$ biến ảnh I thành I' với các mức xám q như vậy $q = f(g)$. Biểu đồ tần suất $h(q)$ có thể được tính bằng $h(q) = \sum_{i \in f^{-1}(q)} h(i)$.

Ví dụ:

g	1	2	3	4
$h(g)$	4	2	1	2

$$f(g) = \begin{cases} g + 1 & \text{nếu } g \leq 2 \\ g & \text{nếu } g = 3 \\ g - 1 & \text{nếu } g > 3 \end{cases}$$

g	1	2	3	4
$q = f(g)$	2	3	3	3

q	1	2	3	4
$h(q)$	0	4	5	0

3.2. CẢI THIỆN ẢNH SỬ DỤNG CÁC TOÁN TỬ KHÔNG GIAN

3.2.1. Phép cửa sổ di chuyển (Moving Window)

Hầu hết các phương pháp xử lý ảnh sử dụng toán tử không gian là phương pháp cửa sổ di chuyển. Về cơ bản phương pháp cửa sổ di chuyển thực hiện biến đổi trên một điểm dựa vào giá trị điểm ảnh nằm trong một cửa sổ bao trùm các điểm ảnh lân cận.

Gọi I và J là ảnh sao cho $J = T[I]$. $T[X]$ là một phép biến đổi sao cho

$$J(i, j) = T[I](i, j) \\ = f(\{I(x, y) | x \in \{i, \dots, i + s\}, y \in \{j, \dots, j + d\}\})$$

Đây là phép biến đổi cửa sổ di chuyển với kích thước $(s+1) \times (d+1)$ và giá trị được biến đổi nằm ở góc trên bên trái cửa sổ.

3.2.2. Phép nhân chập và mẫu

Nhân chập là một phép cửa sổ di chuyển, phép biến đổi được định nghĩa thông qua một ma trận (mẫu). Giả sử ta có ảnh I kích thước $M \times N$, mẫu T có kích thước $m \times n$ khi đó, ảnh I nhân chập theo mẫu T được xác định bởi công thức.

$$I \otimes T(x, y) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} I(x+i, y+j) * T(i, j) \quad (3.1)$$

$$\text{Hoặc } I \otimes T(x, y) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} I(x-i, y-j) * T(i, j) \quad (3.2)$$

Theo công thức (3.1) thì giá trị được biến đổi nằm ở góc trên bên trái của sổ, trong khi đó theo công thức (3.2) thì giá trị được biến đổi nằm ở góc dưới bên phải của sổ.

Ví dụ:

$$I = \begin{pmatrix} 1 & 2 & 4 & 5 & 8 & 7 \\ 2 & 1 & 1 & 4 & 2 & 2 \\ 4 & 5 & 5 & 8 & 8 & 2 \\ 1 & 2 & 1 & 1 & 4 & 4 \\ 7 & 2 & 2 & 1 & 5 & 2 \end{pmatrix}$$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Tính theo công thức (3.1)

$$\begin{aligned} I \otimes T(x, y) &= \sum_{i=0}^1 \sum_{j=0}^1 I(x+i, y+j) * T(i, j) = I(x, y) * T(0, 0) + I(x+1, y+1) * T(1, 1) \\ &= I(x, y) + I(x+1, y+1) \end{aligned}$$

$$I \otimes T = \begin{pmatrix} 2 & 3 & 8 & 7 & 10 & * \\ 7 & 6 & 9 & 12 & 4 & * \\ 6 & 6 & 6 & 12 & 12 & * \\ 3 & 4 & 2 & 6 & 6 & * \\ * & * & * & * & * & * \end{pmatrix}$$

Chương 3: Xử lý nâng cao chất lượng ảnh

Tính theo công thức (3.2)

$I \otimes T''$

*	*	*	*	*	*
*	2	3	8	7	10
*	7	6	0	0	4
*	6	6	0	12	12
*	3	4	2	6	6

Trong quá trình thực hiện phép nhân chập có một số thao tác ra ngoài ảnh, ảnh không được xác định tại những vị trí đó dẫn đến ảnh thu được có kích thước nhỏ hơn.

Ảnh thực hiện theo công thức 3.1 và 3.2 chỉ sai khác nhau 1 phép dịch chuyển để đơn giản ta sẽ hiểu phép nhân chập là theo công thức 3.1

Một số mẫu nhân chập thông dụng

$$T_1'' = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Dùng để khử nhiễu \Rightarrow Các điểm có tần số cao

Ví dụ 1:

$$I'' = \begin{pmatrix} 1 & 2 & 4 & 5 & 8 & 7 \\ 2 & 31 & 1 & 4 & 2 & 2 \\ 4 & 5 & 5 & 8 & 8 & 2 \\ 1 & 2 & 1 & 1 & 4 & 4 \\ 7 & 2 & 2 & 1 & 5 & 2 \end{pmatrix}$$

$$I \otimes T_1'' = \begin{pmatrix} 55 & 65 & 45 & 46 & * & * \\ 52 & 58 & 34 & 35 & * & * \\ 29 & 27 & 35 & 35 & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix}$$

Áp dụng kỹ thuật cộng hằng số với $c'' - 27$, ta có:

$$I_{kq}'' = \begin{pmatrix} 28 & 38 & 18 & 19 & * & * \\ 25 & 31 & 7 & 8 & * & * \\ 2 & 0 & 8 & 8 & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix}$$

Chương 3: Xử lý nâng cao chất lượng ảnh

$$T_2'' = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

Dùng để phát hiện các điểm có tần số cao

Ví dụ 2:

$$I \otimes T_2'' = \begin{pmatrix} 114 & -40 & 0 & -14 & * & * \\ -22 & 5 & 14 & 16 & * & * \\ -1 & -6 & -10 & -2 & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix}$$

3.2.3. Lọc trung vị

*** Định nghĩa 3.1 (Trung vị)**

Cho dãy $x_1; x_2; \dots; x_n$ (giảm). Khi đó trung vị của dãy ký hiệu là $\text{Med}(\{x_n\})$, được định nghĩa:

+ Nếu n lẻ $x^{\left\lceil \frac{n}{2} + 1 \right\rceil}$

+ Nếu n chẵn: $x^{\left\lceil \frac{n}{2} \right\rceil}$ hoặc $x^{\left\lceil \frac{n}{2} + 1 \right\rceil}$

*** Mệnh đề 3.1**

$$\sum_{i=1}^n |x - x_i| \rightarrow \min \text{ tại } Med(\{x_n\})$$

*** Chứng minh mệnh đề 3.1**

Xét trường hợp n chẵn

$$\text{Đặt } M = \frac{n}{2}$$

Ta có:

$$\begin{aligned} \sum_{i=1}^n |x - x_i| &= \sum_{i=1}^M |x - x_i| + \sum_{i=1}^M |x - x_{M+i}| \\ &= \sum_{i=1}^M (|x - x_i| + |x_{M+i} - x|) \geq \sum_{i=1}^M |x_{M+i} - x_i| \\ &= \sum_{i=1}^M (|x_{M+1} - x_M| + |x_M - x_i|) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{i=1}^M |x_{M+i} - \text{Med}(\{x_i\})| + \sum_{i=1}^M |x_i - \text{Med}(\{x_i\})| \\
 &= \sum_{i=1}^n |x_i - \text{Med}(\{x_i\})|
 \end{aligned}$$

Nếu n lẻ:

Bổ sung thêm phần tử $\text{Med}(\{x_i\})$ vào dãy. Theo trường hợp n chẵn ta có:

$$\begin{aligned}
 \sum_{i=1}^n |x - x_i| + |\text{Med}(\{x_i\}) - \text{Med}(\{x_i\})| &\rightarrow \min \text{ tại } \text{Med}(\{x_n\}) \\
 \sum_{i=1}^n |x - x_i| &\rightarrow \min \text{ tại } \text{Med}(\{x_n\})
 \end{aligned}$$

* Kỹ thuật lọc trung vị

Giả sử ta có ảnh I ngưỡng θ của sổ W(P) và điểm ảnh P

Khi đó kỹ thuật lọc trung vị phụ thuộc không gian bao gồm các bước cơ bản sau:

+ **Bước 1:** Tìm trung vị

$$\{I(q) | q \in W(P)\} \rightarrow \text{Med}(P)$$

+ **Bước 2:** Gán giá trị

$$I(P) = \begin{cases} I(P) & |I(P) - \text{Med}(P)| \leq \theta \\ \text{Med}(P) & |I(P) - \text{Med}(P)| > \theta \end{cases}$$

Ví dụ:

$$I'' = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 4 & 16 & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix}$$

với cửa sổ W(3 × 3) và

$$\begin{pmatrix} 1 & 2 & 3 & 2 \\ \end{pmatrix}$$

$$I_{kq}'' \begin{bmatrix} 4 & \textcircled{2} & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{bmatrix}$$

Giá trị 16, sau phép lọc có giá trị 2, các giá trị còn lại không thay đổi.

3.2.4. Lọc trung bình

* Định nghĩa 3.2 (Trung bình)

Cho dãy x_1, x_2, \dots, x_n khi đó trung bình của dãy ký hiệu $AV(\{x_n\})$ được định nghĩa:

$$AV(\{x_n\}) = \text{round} \left(\frac{1}{n} \sum_{i=1}^n x_i \right)$$

* Mệnh đề 3.2

$$\sum_{i=1}^n (x - x_i)^2 \rightarrow \min \text{ tại } AV(\{x_n\})$$

* Chứng minh mệnh đề 3.2

$$\text{Đặt: } \phi(x) = \sum_{i=1}^n (x - x_i)^2$$

Ta có:

$$\phi(x) = 2 \sum_{i=1}^n (x - x_i)$$

$$\phi'(x) = 0$$

$$\Leftrightarrow \sum_{i=1}^n (x - x_i) = 0$$

$$\Leftrightarrow x = \frac{1}{n} \sum_{i=1}^n x_i = AV(\{x_i\})$$

$$\text{Mặt khác, } \phi''(x) = 2n > 0$$

$$\Rightarrow \phi \rightarrow \min \text{ tại } x = AV(\{x_i\})$$

Kỹ thuật lọc trung bình

Giả sử ta có ảnh I , điểm ảnh P , cửa sổ $W(P)$ và ngưỡng θ . Khi đó kỹ thuật lọc trung bình phụ thuộc không gian bao gồm các bước cơ bản sau:

+ **Bước 1:** Tìm trung bình

$$\{I(q) | q \in W(P)\} \rightarrow AV(P)$$

+ **Bước 2:** Gán giá trị

$$I(P) = \begin{cases} I(P) & |I(P) - AV(P)| \leq \theta \\ AV(P) & |I(P) - AV(P)| > \theta \end{cases}$$

Chương 3: Xử lý nâng cao chất lượng ảnh

Ví dụ:

$I'' = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 4 & 16 & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix}$

với cửa sổ $W(3 \times 3)$ và $\theta = 2$

$I_{kq}'' = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 4 & 4 & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix}$

 0

 0

Giá trị 16 sau phép lọc trung bình có giá trị 4, các giá trị còn lại giữ nguyên sau phép lọc.

3.2.5. Lọc trung bình theo k giá trị gần nhất

Giả sử ta có ảnh I, điểm ảnh P, cửa sổ W(P), ngưỡng θ và số k. Khi đó, lọc trung bình theo k giá trị gần nhất bao gồm các bước sau:

+ **Bước 1:** Tìm K giá trị gần nhất

$$\{I(q) \mid q \in W(p)\} \rightarrow \{k \sim \text{giá trị gần } I(P) \text{ nhất}\}$$

+ **Bước 2:** Tính trung bình

$$\{k \sim \text{giá trị gần } I(P) \text{ nhất}\} \rightarrow AV_k(P)$$

+ **Bước 3:** Gán giá trị

$$I(P) = \begin{cases} I(P) & |I(P) - AV_k(P)| \leq \theta \\ AV_k(P) & |I(P) - AV_k(P)| > \theta \end{cases}$$

Ví dụ:

$$I'' = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 4 & 16 & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix}$$

với cửa sổ W(3 × 3), $\theta'' = 2$ và k'' = 3

$$I_{kq}'' = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 4 & \textcircled{8} & 2 & 1 \\ 4 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix}$$

Nếu k lớn hơn kích thước cửa sổ thì kỹ thuật chính là kỹ thuật lọc trung bình, nếu k là 1 thì ảnh kết quả không thay đổi. Vậy nên chất lượng của kỹ thuật phụ thuộc vào số phân tử lựa chọn k.

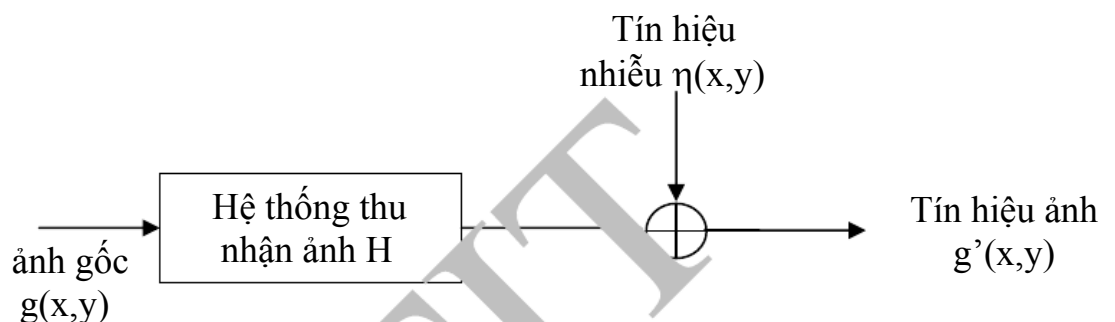
3.3. KHÔI PHỤC ẢNH

3.3.1. Nhiều và mô hình nhiều

Ảnh được coi là một miền đồng nhất về mức xám, tức là các điểm ảnh lân cận có sự biến đổi liên tục về mức xám. Như vậy sau quá trình số hoá thì trong mỗi cửa sổ đang xét các điểm ảnh đều có giá trị gần bằng nhau. Thực tế quan sát có những điểm ảnh có giá trị khác hơn nhiều so với các điểm ảnh xung quanh. Đó chính là nhiễu. Như vậy, nhiễu trong ảnh số được xem như là sự dịch chuyển đột ngột của tín hiệu ảnh trên một khoảng cách nhỏ.

Mô hình liên tục

Hệ thống thu nhận ảnh chuyển các hình ảnh thực của môi trường xung quanh $g(x,y)$ thành dạng tín hiệu ảnh $g'(x,y)$. Tuy nhiên trong quá trình chuyển đổi có nhiều yếu tố tác động tạo thành nhiễu. Do đó tín hiệu $g'(x,y)$ có thể chứa các thành phần nhiễu trong đó. Quá trình thu nhận ảnh có nhiễu có thể mô tả một cách trực quan như sau:



Hình 3.1. Mô hình nhiễu

Mô hình rời rạc:

Trên cơ sở mô hình nhiễu liên tục, ta có thể xây dựng một mô hình nhiễu rời rạc tương ứng với ảnh số. Khi đó $g(x,y)$ sẽ chuyển thành ảnh rời rạc $g[m,n]$, ảnh liên tục

tương ứng với ảnh số. Khi đó $g(x,y)$ sẽ chuyển thành ảnh rời rạc $g[m,n]$, ảnh rời rạc $g'(x,y)$ sẽ chuyển thành ma trận điểm ảnh $g'[m,n]$ và nhiễu cũng phân bố rời rạc tại các điểm ảnh $\eta[m,n]$. Giả sử H là hàm tuyến tính bất biến trong phạm vi $M \times N$ (kích thước ảnh) thì ta có:

$$g'[m,n] = h[m,n] * g[m,n] + \eta[m,n] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h[m-k, n-l] \cdot g[k,l] + \eta[m,n] \quad (3.3)$$

Đối với xử lý số ảnh thì ta chỉ sử dụng mô hình nhiễu rời rạc.

3.3.2. Các loại nhiễu

Các tín hiệu nhiễu thường được chia thành các loại chính như sau:

- **Nhiễu do thiết bị thu nhận ảnh:** là loại nhiễu gây ra do giới hạn nhiễu xạ và quang sai của thấu kính, nhiễu do bộ phận cảm quang, ảnh mờ nhòe do ống kính, nhiễu do rung động thiết bị trong quá trình thu nhận.
- **Nhiễu ngẫu nhiên độc lập:** là các loại nhiễu gây ra do ảnh hưởng của môi trường xung quanh, do ảnh hưởng của khí quyển.
- **Nhiễu do vật quan sát:** là nhiễu gây ra do bề mặt của bản thân vật có độ nhám gồ ghề. Chính nhiễu này gây hiện tượng tán xạ của các tia đơn sắc và sinh ra hiện tượng nhiễu lốm đốm.

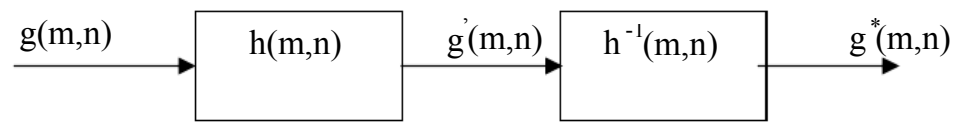
Thường người ta xấp xỉ các loại nhiễu bằng các quá trình tuyến tính bất biến vì có nhiều công cụ tuyến tính có thể giải quyết vấn đề khôi phục ảnh hơn là các công cụ phi tuyến. Việc xử lý nhiễu bằng cách xấp xỉ tuyến tính cũng giúp cho công việc dễ dàng hơn trong trường hợp dùng cách biến đổi phi tuyến.

3.3.3. Các kỹ thuật lọc nhiễu

Trong phần này ta sẽ xét hai kỹ thuật lọc tuyến tính hay sử dụng là lọc đảo và lọc giả đảo.

Lọc đảo (inverse filter):

Nguyên lý của lọc đảo là sử dụng hàm ngược của đáp ứng xung $h[m,n]$ để khôi phục lại một ảnh xấp xỉ ảnh nguyên gốc $g[m,n]$ từ ảnh $g'[m,n]$ đã biết nguyên lý này được biểu diễn mô tả theo sơ đồ sau:



Hình 3.2. Lọc đảo khôi phục ảnh nguyên

Như vậy ảnh khôi phục $g^*[m,n]$ sẽ được tính theo công thức:

$$g^*[m,n] = g'[m,n] * h^{-1}[m,n] \quad (3.4)$$

Trong đó $h^{-1}[m,n]$ chính là hàm của bộ lọc đảo. Vì $H^{-1}(H(x)) = x$ nên ta có giá trị đầu ra là $g^*[m,n]$ cũng bằng giá trị đầu vào $g[m,n]$. Như vậy ta đã khôi phục được ảnh $g[m,n]$ nhờ dùng hàm ngược của đáp ứng xung $h[m,n]$. Nếu dùng biến đổi Fourier ta có:

$$G[u,v] = G'[u,v] \cdot H^T[u,v] = \frac{G'[u,v]}{H[u,v]} \quad (3.5)$$

Qua đó ta thấy, đáp ứng tần số của bộ lọc đảo là nghịch đảo của đáp ứng tần số của hệ thu nhận ảnh. Nếu đánh giá được mức nhiễu ta có thể xấp xỉ gần hơn với ảnh nguyên gốc:

$$G[u,v] = \frac{G'[u,v]}{H[u,v]} + \frac{M[u,v]}{H[u,v]} \quad \text{với } M[u,v] \text{ là nhiễu ước lượng} \quad (3.6)$$

Nếu $H[u,v]$ bằng 0 hoặc khá nhỏ thì hệ thống khôi phục sẽ không ổn định (hàm $H^T[u,v]$ không xác định). Đây chính là nhược điểm của phương pháp lọc đảo. Tuy bộ lọc đảo có khả năng ngăn nhiễu do hệ thống nhận ảnh gây ra khá tốt nhưng việc thiết kế bộ lọc này lại là khá phức tạp.

Lọc giả đảo (Pseudoinverse filter):

Kỹ thuật lọc này khắc phục được nhược điểm của kỹ thuật lọc đảo là làm cho hàm

$H^T[u,v]$ luôn xác định:

Chương 3: Xử lý nâng cao chất lượng ảnh

$$H^T[u,v] = \begin{cases} \frac{1}{H[u,v]} & \text{khi } |H[u,v]| \geq \varepsilon \\ 0 & \text{khi } |H[u,v]| < \varepsilon \end{cases}$$

(3.7)

Giá trị ε là một giá trị cho trước

Với ϵ là một giá trị cho trước.

Trong trường hợp ảnh nguyên gốc $g[m,n]$ chuyển động tịnh tiến theo phương x và y sinh ra hiện tượng nhoè ảnh thì $H^T[u,v]$ sẽ được xác định theo các thông số chuyển động, và lọc giả đảo có khả năng khôi phục được ảnh nhoè này.

Lọc nhiễu lốm đốm:

Ta đã biết nhiễu lốm đốm gây ra do tính chất gồ ghề của bề mặt vật thể gây ra hiện tượng tán xạ các tia đơn sắc. Phần này ta sẽ nghiên cứu một kỹ thuật lọc nhiễu lốm đốm là kỹ thuật trung bình thống kê sử dụng bộ lọc đồng cầu.

Kỹ thuật lọc nhiễu trung bình thống kê tiến hành thống kê các cường độ sáng của đối tượng ảnh bị nhiễu lốm đốm bằng N lần thu nhận độc lập và lấy trung bình các cường độ đó.

Ta giả thiết hệ thống chỉ có nhiễu lốm đốm và nhiễu này có thể coi như là tổng vô số hạn các pha độc lập và đồng nhất. Ta có thể biểu diễn nhiễu lốm đốm như sau:

$$a[m,n] + a_R[m,n] + j.a_L[m,n]$$

Với a_R và a_L là các biến ngẫu nhiên độc lập phân bố theo Gaussian, ứng với mỗi toạ độ $[m,n]$ có trung bình bằng 0 và phương sai σ^2 . Ta có cường độ S :

$$S[m,n] = |a[m,n]|^2 = a_R^2 + a_L^2 \quad (3.8)$$

Như vậy ảnh thu được trong lần nhận thứ i ($1, \dots, N$) sẽ là:

$$g'_i[m,n] = g[m,n].s_i[m,n] \quad (3.9)$$

Trị trung bình thống kê của N lần thu nhận ảnh sẽ là:

$$g'_N[m,n] = \frac{1}{N} \sum_{i=1}^N g'_i[m,n] = g[m,n].\left(\frac{1}{N} \sum_{i=1}^N s'_i[m,n]\right) = g[m,n].s_N[m,n] \quad (3.10)$$

Trong đó $s_N[m,n]$ là trị trung bình thống kê của trường cường độ nhiễu lốm đốm

Lọc đồng cầu (Homomorphic filter):

Nếu lấy logarit 2 vế của biểu thức 3.10 ta thu được:

$$\log g'_N[m,n] = \log g[m,n] + \log s_N[m,n] \quad (3.11)$$

Đặt $w_N[m,n] = \log g'_N[m,n]$, $z[m,n] = \log g[m,n]$, $\eta_N[m,n] = \log s_N[m,n]$

Ta có mô hình quan sát có nhiễu lốm đốm như sau:

$$w_N[m,n] = z[m,n] + \eta_N[m,n] \quad (3.12)$$

Từ công thức 3.12 ta thấy có thể tìm ra $z[m,n]$ từ $w_N[m,n]$ và $\eta_N[m,n]$. Như vậy là có thể tách nhiễu lốm đốm và khôi phục lại ảnh ban đầu.

Trong công thức 3.12 thì $\eta_N[m,n]$ là nhiễu trắng dừng. Với $N \geq 2$ thì $\eta_N[m,n]$ có thể mô tả gần với nhiễu ngẫu nhiên Gaussian với mật độ phổ được định nghĩa như sau:

$$S_{\eta_N}(\xi_1, \xi_2) = \sigma^2 = \begin{cases} \pi^2 / 6 & N = 1 \\ 1/N & N > 1 \end{cases} \quad (3.13)$$

3.4. CÂU HỎI ÔN TẬP CHƯƠNG

Câu 1: Biểu đồ tần suất là gì? Hãy tìm biểu đồ tần suất $h(g)$ cho ảnh I sau, được biết các điểm ảnh có giá trị từ 0 đến 9:

$$I = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 & 1 \\ 4 & 4 & 3 & 2 & 4 & 0 \\ 6 & 9 & 2 & 3 & 2 & 1 \\ 6 & 2 & 0 & 5 & 3 & 0 \\ 3 & 4 & 0 & 5 & 1 & 5 \\ 5 & 6 & 8 & 9 & 3 & 6 \end{bmatrix}$$

Câu 2: Thực hiện các phép biến đổi sau với ảnh I ở câu 1.

- Tăng sáng với $c = 4$.
- Tách ngưỡng với $\theta = 3$.
- Cân bằng tần suất với ảnh kết quả có cùng số mức xám với ảnh gốc.
- Tìm ngưỡng tự động và tách ngưỡng.

Lưu ý: Các điểm ảnh có giá trị từ 0 đến 9.

Câu 3: Từ biểu đồ tần suất $h(g)$ tính được ở câu 1 hãy tính biểu đồ tần suất $h'(g)$ của

ảnh I' sau khi được biến đổi từ ảnh I sử dụng biểu thức:

$$f(g) = |2g - 9|$$

Câu 4: Thực hiện nhân chập ảnh I với ma trận nhân chập H:

$$I = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 15 & 15 & 15 & 2 & 2 \\ 2 & 2 & 13 & 17 & 16 & 2 & 2 \\ 2 & 2 & 15 & 15 & 15 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix} H = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Câu 5: Thực hiện lọc với cửa sổ 3x3 và $\theta = 3$ sử dụng các phương pháp sau với ảnh I ở câu 1:

- Lọc trung vị với.
- Lọc trung bình.
- Lọc trung bình k giá trị gần nhất với $k = 4$.

Chương 4:

CÁC PHƯƠNG PHÁP PHÁT HIỆN BIÊN

4.1. KHÁI QUÁT VỀ BIÊN VÀ PHÂN LOẠI CÁC KỸ THUẬT DÒ BIÊN CƠ BẢN

4.1.1. Giới thiệu

Biên là vấn đề quan trọng trong trích chọn đặc điểm nhằm tiến tới hiểu ảnh. Cho đến nay chưa có định nghĩa chính xác về biên, trong mỗi ứng dụng người ta đưa ra các độ đo khác nhau về biên, một trong các độ đo đó là độ đo về sự thay đổi đột ngột về cấp xám. Ví dụ: Đối với ảnh đen trắng, một điểm được gọi là điểm biên nếu nó là điểm đen có ít nhất một điểm trắng bên cạnh. Tập hợp các điểm biên tạo nên biên hay đường bao của đối tượng. Xuất phát từ cơ sở này người ta thường sử dụng hai phương pháp phát hiện biên cơ bản:

- **Phát hiện biên trực tiếp:** Phương pháp này làm nổi biên dựa vào sự biến thiên mức xám của ảnh. Kỹ thuật chủ yếu dùng để phát hiện biên ở đây là kỹ thuật lấy đạo hàm. Nếu lấy đạo hàm bậc nhất của ảnh ta có các kỹ thuật Gradient, nếu lấy đạo hàm bậc hai của ảnh ta có kỹ thuật Laplace. Ngoài ra còn có một số các tiếp cận khác
- **Phát hiện biên gián tiếp:** Nếu bằng cách nào đó ta phân được ảnh thành các vùng thì ranh giới giữa các vùng đó gọi là biên. Kỹ thuật dò biên và phân vùng ảnh là hai bài toán đối ngẫu nhau vì dò biên để thực hiện phân lớp đối tượng mà khi đã phân lớp xong nghĩa là đã phân vùng được ảnh và ngược lại, khi đã phân vùng ảnh đã được phân lớp thành các đối tượng, do đó có thể phát hiện được biên.

Phương pháp phát hiện biên trực tiếp tỏ ra khá hiệu quả và ít chịu ảnh hưởng của nhiễu, song nếu sự biến thiên độ sáng không đột ngột, phương pháp tỏ ra kém hiệu quả, phương pháp phát hiện biên gián tiếp có ưu điểm là không cần cài đặt, song lại áp dụng khá tốt trong trường hợp này. Sự khác biệt cơ bản giữa hai phương pháp này là: Phương pháp phát hiện biên trực tiếp cho ta kết quả là ảnh biên, còn phương pháp phát hiện biên gián tiếp cho ta kết

quả là đường biên.

4.1.2. Kỹ thuật phát hiện biên Gradient

Theo định nghĩa, gradient là một vectơ có các thành phần biểu thị tốc độ thay đổi giá trị của điểm ảnh, ta có:

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

Trong đó, dx, dy là khoảng cách (tính bằng số điểm) theo hướng x và y.

Tuy ta nói là lấy đạo hàm nhưng thực chất chỉ là mô phỏng và xấp xỉ đạo hàm bằng các kỹ thuật nhân chập vì ảnh số là tín hiệu rời rạc nên đạo hàm không tồn tại.

Giả sử với dx và dy nhỏ nhất là 1, ta có:

$$\begin{cases} \frac{\partial f}{\partial x} \approx f(x+1, y) - f(x, y) \\ \frac{\partial f}{\partial y} \approx f(x, y+1) - f(x, y) \end{cases}$$

Do đó, mặt nạ nhân chập theo hướng x là $A'' = \begin{pmatrix} 1 & 1 \end{pmatrix}$

và hướng y là $B'' = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

Ví dụ:

$$I'' = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & 3 \\ 0 & 3 & 3 & 3 \\ 0 & 3 & 3 & 3 \end{pmatrix}$$

Ta có,

$$I \otimes A'' = \begin{pmatrix} 0 & 0 & 0 & * \\ 3 & 0 & 0 & * \\ 3 & 0 & 0 & * \\ 3 & 0 & 0 & * \end{pmatrix}; I \otimes B'' = \begin{pmatrix} 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ * & * & * & * \end{pmatrix}$$

$$I \otimes A + I \otimes B'' = \begin{pmatrix} 0 & 0 & 0 & * \\ 3 & 0 & 0 & * \\ 3 & 0 & 0 & * \\ * & * & * & * \end{pmatrix}$$

4.1.2.1. Kỹ thuật Prewitt

Kỹ thuật sử dụng 2 mặt nạ nhân chập xấp xỉ đạo hàm theo 2 hướng x và y là:

$$H_x'' = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

$$H_y'' = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Chương 4: Các phương pháp phát hiện biên

- Các bước tính toán của kỹ thuật Prewitt
- + **Bước 1:** Tính $I \otimes H_x$ và $I \otimes H_y$
 - + **Bước 2:** Tính $I \otimes H_x + I \otimes H_y$
 - + **Bước 3:** Phân ngưỡng theo θ để có ảnh biên

Ví dụ:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 5 & 0 & 0 & 0 & 0 \\ 5 & 5 & 5 & 5 & 0 & 0 \\ 5 & 5 & 5 & 5 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned}
 I'' &= \begin{pmatrix} 5 & 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
 I \otimes H_x'' &= \begin{pmatrix} 0 & 0 & -10 & -10 & * & * \\ 0 & 0 & -15 & -15 & * & * \\ 0 & 0 & -10 & -10 & * & * \\ 0 & 0 & -5 & -5 & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix} \\
 I \otimes H_y'' &= \begin{pmatrix} 15 & 15 & 10 & 5 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ -15 & -15 & -10 & -5 & * & * \\ -15 & -15 & -10 & -5 & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix} \\
 I \otimes H_x + I \otimes H_y'' &= \begin{pmatrix} 15 & 15 & 0 & -5 & * & * \\ 0 & 0 & -15 & -15 & * & * \\ -15 & -15 & -20 & -15 & * & * \\ -15 & -15 & -15 & -10 & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix}
 \end{aligned}$$

Chương 4: Các phương pháp phát hiện biên

4.1.2.2. Kỹ thuật Sobel

Tương tự như kỹ thuật Prewitt kỹ thuật Sobel sử dụng 2 mặt nạ nhân chập theo 2 hướng x, y là:

$$H_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$
$$H_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Các bước tính toán tương tự Prewitt

+ **Bước 1:** Tính $I \otimes H_x$ và $I \otimes H_y$

+ **Bước 2:** Tính $I \otimes H_x + I \otimes H_y$

+ **Bước 3:** Phân ngưỡng theo θ để có ảnh biên

4.1.2.3. Kỹ thuật la bàn

Kỹ thuật sử dụng 8 mặt nạ nhân chập theo 8 hướng $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$

$$H_1 = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad H_2 = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

$$\begin{aligned}
 H_3'' &= \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} & H_4'' &= \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \\
 H_5'' &= \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} & H_6'' &= \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \\
 H_7'' &= \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} & H_8'' &= \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}
 \end{aligned}$$

Các bước tính toán thuật toán La bàn

+ **Bước 1:** Tính $I \otimes H_i; i \in \overline{1,8}$

+ **Bước 2:** $\sum_{i=1}^8 I \otimes H_i$

+ **Bước 3:** Phân ngưỡng theo θ để có ảnh biên

4.1.3. Kỹ thuật phát hiện biên Laplace

Các phương pháp đánh giá gradient ở trên làm việc khá tốt khi mà độ sáng thay đổi rõ nét. Khi mức xám thay đổi chậm, miền chuyển tiếp trải rộng, phương pháp cho hiệu quả hơn đó là phương pháp sử dụng đạo hàm bậc hai Laplace.

Toán tử Laplace được định nghĩa như sau:

Ta có:

$$\begin{aligned}\nabla^2 f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \\ \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) \approx \frac{\partial}{\partial x} [f(x+1, y) - f(x, y)] \\ &\approx [f(x+1, y) - f(x, y)] - [f(x, y) - f(x-1, y)] \\ &\approx f(x+1, y) - 2f(x, y) + f(x-1, y)\end{aligned}$$

Tương tự,

$$\begin{aligned}\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) \approx \frac{\partial}{\partial y} [f(x, y+1) - f(x, y)] \\ &\approx [f(x, y+1) - f(x, y)] - [f(x, y) - f(x, y-1)] \\ &\approx f(x, y+1) - 2f(x, y) + f(x, y-1)\end{aligned}$$

Vậy: $\nabla^2 f = f(x+1, y) + f(x, y+1) - 4f(x, y) + f(x-1, y) + f(x, y-1)$

Dẫn tới:

$$H = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Trong thực tế, người ta thường dùng nhiều kiểu mặt nạ khác nhau để xấp xỉ rồi rạc đạo hàm bậc hai Laplace. Dưới đây là ba kiểu mặt nạ thường dùng:

$$H_1 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad H_2 = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad H_3 = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

4.1.4. Kỹ thuật Canny

Đây là một thuật toán được phát triển khá sớm nhưng cho đến nay vẫn là một trong những kỹ thuật được sử dụng rộng rãi, cho các kết quả tương đối tốt, có khả năng đưa ra đường biên mảnh, phân biệt được điểm biên với điểm nhiễu.





Hình 4.1. Kết quả tìm biên của Canny

Thuật toán Canny gồm năm bước:

+ **Bước 1:** Làm trơn ảnh. Chúng ta sử dụng

Tính $I \otimes H$, với:

$$H = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix}$$

Gọi G là kết quả lọc nhiễu: $G = I \otimes H$

+ **Bước 2:** Tính gradient của ảnh bằng mặt nạ PreWitt (hoặc các ma trận phát hiện biên khác như Roberts, Sobel...), kết quả đặt vào G_x, G_y .

$$G_x = G \otimes H_x, G_y = G \otimes H_y$$

+ **Bước 3:** Tính độ lớn và hướng của gradient tại mỗi điểm (i, j) của ảnh. Độ lớn $G(i, j)$ tại điểm (i, j) được tính dựa vào định lý Pitago:

$$G(i, j) = \sqrt{G_x(i, j)^2 + G_y(i, j)^2}$$

Ngoài ra hướng của véc tơ gradient được tính với công thức:

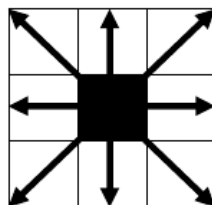
$$\theta(i, j) = \text{atan2}(G_x(i, j), G_y(i, j))$$

với atan2 là hàm \arctan với hai biến nhằm nhận thêm thông tin hướng và xác định góc trong khoảng 360° , hàm atan2 có thể được định nghĩa như sau:

$$\text{atan2}(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & y \geq 0, x < 0 \\ \arctan\left(\frac{y}{x}\right) - \pi & y < 0, x < 0 \\ +\frac{\pi}{2} & y > 0, x = 0 \\ -\frac{\pi}{2} & y < 0, x = 0 \\ \text{undefined} & y = 0, x = 0 \end{cases}$$

Chương 4: Các phương pháp phát hiện biên

Sau đó hướng gradient được làm tròn thành bốn góc 0° , 45° , 90° và 135° đại diện cho chiều dọc, ngang và hai chiều chéo.



+ **Bước 4:** Bước này loại bỏ những điểm không phải là cực đại địa phương để xóa bỏ những điểm không thực sự là biên, bước này sẽ giúp biên mỏng hơn. Xét (i,j) , $\theta(i,j)$ là hướng gradient tại (i,j) . I_1 , I_2 là hai điểm lân cận của (i,j) theo hướng θ . Theo định nghĩa điểm biên cực bộ thì (i,j) là biên nếu $I(i,j)$ cực đại địa phương theo hướng gradient nghĩa là nếu $I(i,j) > I_1$ và $I(i,j) > I_2$ thì mới giữ lại $I(i,j)$ để xét tiếp trong bước sau.

+ **Bước 5:** Tại bước này, chúng ta sử dụng hai ngưỡng cao và thấp. Đầu tiên, Canny lọc các điểm được giữ lại sử dụng ngưỡng cao, chỉ những điểm có độ lớn gradient cao hơn ngưỡng này mới được chọn. Từ những điểm được chọn Canny dò theo biên sử dụng hướng

của gradient tại các điểm. Khi thực hiện việc dò theo biên, Canny sử dụng ngưỡng thấp để xác định điểm dừng của biên (nếu giá trị điểm tiếp theo thấp hơn ngưỡng này, việc dò theo biên này sẽ kết thúc)

Các kỹ thuật thông thường để có được ảnh biên thường sử dụng một ngưỡng, điểm ảnh có độ lớn gradient lớn hơn ngưỡng này sẽ được chọn làm biên. Tuy nhiên các kỹ thuật sử dụng một ngưỡng không thể vừa chọn biên chính xác và mảnh như Canny, và thông thường sẽ có quá nhiều biên giả hoặc xác định thiếu biên. Đây chính là một trong những điều khiến Canny vẫn là kỹ thuật dò biên tốt nhất hiện nay.

4.2. PHƯƠNG PHÁP PHÁT HIỆN BIÊN CỤC BỘ

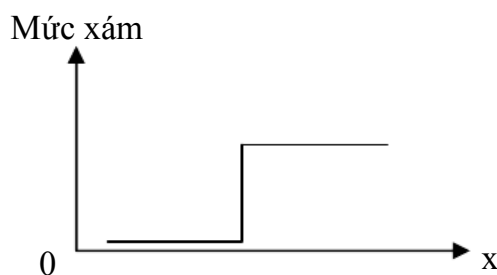
4.2.1. Biên và độ biến đổi về mức xám

Như đã trình bày ở trên, trong thực tế người ta thường dùng hai phương pháp phát hiện biên cơ bản là: Phát hiện biên trực tiếp và gián tiếp. Phần này đề cập đến kỹ thuật dựa vào trung bình cục bộ trên cơ sở đánh giá độ chênh lệch về giá trị mức xám của điểm ảnh so với các điểm lân cận do đó kết hợp được ưu điểm của cả hai khuynh hướng trực tiếp và gián tiếp.

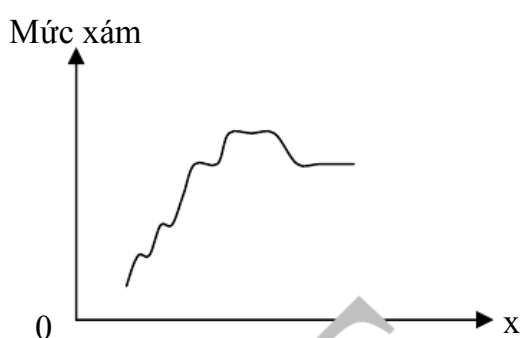
Đối với các ảnh màu theo mô hình nào đó đều có thể chuyển sang mô hình gồm 3 thành phần màu R, G, B. Sau đó dễ dàng chuyển các ảnh màu sang dạng ảnh đa cấp xám. Giá trị xám của điểm ảnh có thể được tính là trung bình cộng của ba thành phần màu R, G và B. Việc xử lý, thao tác trên các ảnh xám có một ưu điểm là dễ xử lý hơn các ảnh màu mà vẫn giữ được các đặc tính của ảnh. Các ảnh trắng đen tuy dễ xử lý nhất nhưng sẽ bị mất nhiều chi tiết sau khi chuyển đổi.

Chương 4: Các phương pháp phát hiện biên

Một cách lý tưởng đồ thị biến thiên mức xám của điểm ảnh khi qua biên phải có dạng:



Trong thực tế dạng đồ thị này chỉ gặp trong các ảnh trắng đen (ảnh xám có hai màu), còn với các ảnh thực thì đồ thị của nó có dạng:



Khó khăn cho việc phân tích các ảnh trên thực tế là ở chỗ sự biến thiên về mức xám của điểm ảnh không phải chỉ được thể hiện theo một hướng duy nhất mà phải xét theo cả tám hướng của các điểm ảnh láng giềng, tại các vùng biên và lân cận biên sự biến thiên mức xám của các điểm ảnh thường không đột ngột mà trải qua một khoảng biến thiên không đều nhưng có tốc độ biến thiên nhanh.

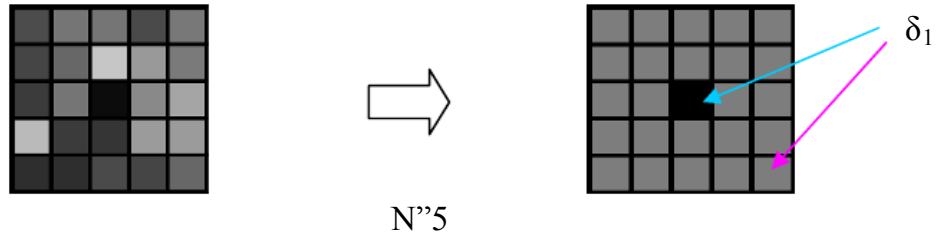
Chúng ta có thể xác định được các đường biên như thế này bằng kỹ thuật Laplace nhưng như ở trên đã nói kỹ thuật này rất nhạy cảm với nhiễu mà nhiễu hầu như lại là vấn đề mà ở trong bức ảnh nào cũng có. Ngoài ra, trong thực tế khi dò biên cho các ảnh xám tùy theo mục đích xử lý sau này ta có thể muốn lấy biên của tất cả các đối tượng trong ảnh hoặc chỉ một số đối tượng trong ảnh. Các kỹ thuật đạo hàm do sử dụng các mặt nạ là các ma trận nhân chập nên khó điều chỉnh độ chi tiết của ảnh biên thu được.

Muốn làm được điều này lại phải tính toán lại các giá trị của các phần tử trong ma trận theo các công thức nhất định, rất phức tạp và tốn kém. Không những thế ảnh thu được sau khi lọc không làm mất đi được tất cả các điểm không thuộc đường biên mà chỉ làm nổi lên các điểm nằm trên biên và muốn nhận dạng được các đối tượng thì ta còn phải xử lý thêm một vài bước nữa thì mới thu được ảnh biên thực sự. Có thể nhận thấy là các thuật toán dò biên truyền thống mà chúng ta hay dùng vẫn chưa đạt được sự hoàn thiện như mong muốn.

4.2.2. Phát hiện biên dựa vào trung bình cục bộ

Ý tưởng chính của thuật toán được đề xuất là: Xác định tất cả các điểm nằm trên biên không theo hướng tìm kiếm và sử dụng các ma trận lọc, thông qua việc so sánh độ chênh lệch về mức xám của nó so với mức xám chung của các điểm ảnh lân cận (mức xám nền). Trước hết giá trị xám trung bình của các điểm ảnh nằm trong phạm vi của ma trận 3×3 hoặc 5×5 có tâm là điểm ảnh đang xét sẽ được tính toán. Nếu như độ chênh lệch mức xám giữa điểm đang xét với giá trị xám trung bình thỏa mãn lớn hơn một mức tối thiểu δ_1 nào

đó ($PTB + \delta_1 < P$) thì chúng ta sẽ coi nó là điểm biên và ghi nhận lại, còn các điểm không thỏa mãn điều kiện trên sẽ được coi là điểm nền.



a) Ma trận điểm ảnh trước khi lọc b) Ma trận điểm ảnh sau khi lọc

Hình 4.2. Ma trận điểm ảnh trước và sau lọc

Thuật toán có thể được mô tả như sau:

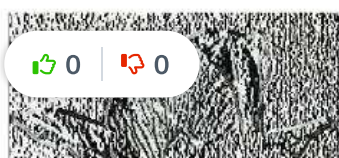
```
for (i=0; i< biHeight; i++){
    for (j=0; j< biWidth; j++){
        tt_GrayScale=0;
        for (ii=i-1; ii<=i+1; ii++){
            for (jj=j-1; jj<=j+1; jj++){
                tt_GrayScale+=GetPoint(pOrgImg,ii,jj);
            }
        }
        if (tt_GrayScale>9*GetPoint(pOrgImg,i,j)+ $\delta_1$ ) {
            SetPoint(pBdImg,i,j,BLACK);
        }
    }
}
```

Trong đó:

- biWidth, biHeight: là chiều rộng và chiều cao của ảnh tính theo đơn vị Pixel.
- pOrgImg, pBdImg: lần lượt là các con trỏ trỏ đến các vùng dữ liệu của ảnh gốc và ảnh biên.
- tt_GrayScale: là tổng giá trị độ xám của các điểm ảnh thuộc ma trận 3×3 có tâm là điểm ảnh đang xét.
- δ_1 : là độ chênh lệch mức xám của điểm ảnh đang xét so với giá trị xám trung bình của ma trận.
- SetPoint() và GetPoint(): là các hàm đọc, ghi giá trị điểm ảnh.

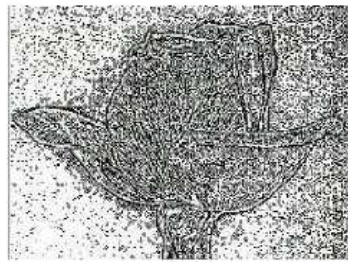
Chương 4: Các phương pháp phát hiện biên

ảnh biên thu được khi sử dụng thuật toán phát hiện biên đề xuất dựa vào trung bình cục bộ với giá trị δ_1 khác nhau. Hình 4.3a là ảnh biên thu được với $\delta_1'' = 25$, Hình 4.3b là ảnh biên thu được với $\delta_1'' = 250$.

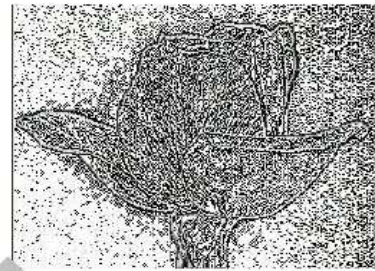




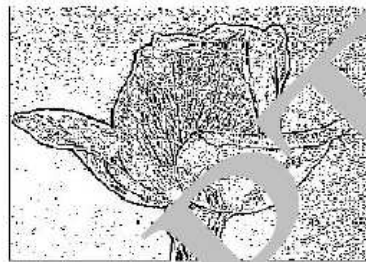
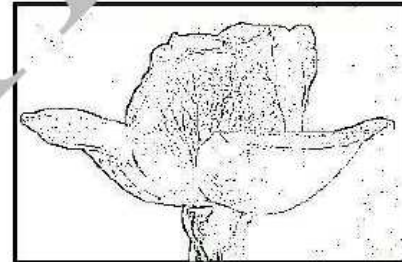
a) Ảnh gốc

b) Ảnh qua lọc Sobel H_x c) Ảnh qua lọc Sobel H_y 

d) Ảnh qua lọc Kirsh



e) Ảnh qua lọc Laplace

Hình 4.3. Các ảnh biên theo các thuật toán phát hiện biên truyền thốnga) Ảnh biên thu được với $\delta_1 = 25$ b) Ảnh biên thu được với $\delta_1 = 250$ **Hình 4.4. Các ảnh biên kết quả thu được theo thuật toán đề xuất**

Chúng ta có nhận xét là ảnh gốc sử dụng trong chương trình có màu nền khá tối và có rất nhiều nhiễu. Các bộ lọc sử dụng trong minh họa trên đều mắc phải vấn đề này. Thuật toán dò biên sử dụng trong chương trình tuy đã hạn chế được nhiều nhiễu so với việc sử dụng các bộ lọc và làm nổi rõ các đường biên nhưng vẫn không loại bỏ được hầu hết các nhiễu. Khi áp dụng thuật toán trên chúng ta vẫn có thể làm giảm bớt nhiễu đi nhiều hơn nữa bằng cách tăng giá trị của hệ số δ_1 lên. Nhưng khi đó các đường biên thu được cũng bị đứt đoạn và mờ đi nhiều.

Thuật toán có độ phức tạp tỷ lệ với kích thước ảnh và kích thước cửa sổ. Độ phức tạp của thuật toán là $O(n^2)$ nên nó thực hiện việc tìm biên khá nhanh, ảnh biên thu được chỉ gồm các điểm ảnh và điểm biên nên dễ xử lý, bản thân thuật toán này cũng ít chịu ảnh hưởng của nhiễu hơn là kỹ thuật Sobel mặc dù nó có khả năng phát hiện khá tốt các vùng biên nhiễu. Nhưng cũng giống các phương pháp phát hiện biên trực tiếp khác là nó cho kết quả đường biên có độ dày không đều.

4.3. DÒ BIÊN THEO QUY HOẠCH ĐỘNG

4.3.1. Một số khái niệm cơ bản

*Ảnh và điểm ảnh

Ảnh số là một mảng số thực 2 chiều (I_{ij}) có kích thước ($M \times N$), trong đó mỗi phần tử I_{ij} ($i = 1, \dots, M; j = 1, \dots, N$) biểu thị mức xám của ảnh tại (i, j) tương ứng.

Ảnh được gọi là ảnh nhị phân nếu các giá trị I_{ij} chỉ nhận giá trị 0 hoặc 1.

Ở đây ta chỉ xét tới ảnh nhị phân vì ảnh bất kỳ có thể đưa về dạng nhị phân bằng kỹ thuật phân ngưỡng. Ta ký hiệu S là tập các điểm vùng (điểm đen) và \bar{S} là tập các điểm nền (điểm trắng).

*Các điểm 4 và 8-láng giềng

Giả sử (i, j) là một điểm ảnh, các điểm 4-láng giềng là các điểm kề trên, dưới, trái, phải của (i, j) :

$$N_4(i, j) = \{(i', j') : |i - i'| + |j - j'| = 1\},$$

và những điểm 8-láng giềng gồm:

$$N_8(i, j) = \{(i', j') : \max(|i - i'|, |j - j'|) = 1\}.$$

Trong Hình 3.1 biểu diễn ma trận 8 láng giềng kề nhau, các điểm P_0, P_2, P_4, P_6 là các 4-láng giềng của điểm P , còn P_1, P_3, P_5, P_7 là các 8-láng giềng của P .

P_3	P_2	P_1
P_4	P	P_0
P_5	P_6	P_7

Hình 4.5. Ma trận 8-láng giềng kề nhau

***Đối tượng ảnh**

Hai điểm $P_s, P_e \in E$, $E \subseteq \mathfrak{S}$ hoặc $\overline{\mathfrak{S}}$ được gọi là 8-liên thông (hoặc 4-liên thông) trong E nếu tồn tại tập các điểm được gọi là **đường đi** $(i_0, j_0) \dots (i_n, j_n)$ sao cho $(i_0, j_0) \sim P_s$, $(i_n, j_n) \sim P_e$, $(i_r, j_r) \in E$ và (i_r, j_r) là 8-láng giềng (hoặc 4-láng giềng tương ứng) của (i_{r-1}, j_{r-1}) với $r = 1, 2, \dots, n$.

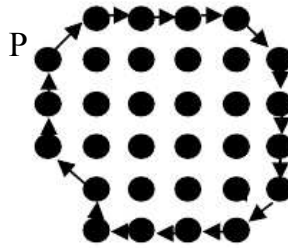
Quan hệ ***k*-liên thông trong E** ($k \in \{4, 8\}$) là một quan hệ phản xạ, đối xứng và bắc cầu. Bởi vậy đó là một quan hệ tương đương. Mỗi lớp tương đương được gọi là một thành phần *k*-liên thông của ảnh. Về sau ta sẽ gọi mỗi thành phần *k*-liên thông của ảnh là một đối tượng ảnh.

4.3.2. Chu tuyến của một đối tượng ảnh

Định nghĩa 4.1: [Chu tuyến]

Chu tuyến của một đối tượng ảnh là dãy các điểm của đối tượng ảnh P_1, \dots, P_n sao cho P_i và P_{i+1} là các 8-láng giềng của nhau ($i=1, \dots, n-1$) và P_1 là 8-láng giềng của P_n , $\forall i \exists Q$ không thuộc đối tượng ảnh và Q là 4-láng giềng của P_i . Kí hiệu " $P_1 P_2 \dots P_n$ ".

Tổng các khoảng cách giữa hai điểm kế tiếp của chu tuyến là độ dài của chu tuyến và kí hiệu $Len(C)$ và hướng $P_i P_{i+1}$ là hướng chẵn nếu P_i và P_{i+1} là các 4-láng giềng (trường hợp còn lại thì $P_i P_{i+1}$ là hướng lẻ). Hình 4.6 dưới đây biểu diễn chu tuyến của ảnh, trong đó, P là điểm khởi đầu chu tuyến.



Hình 4.6. Ví dụ về chu tuyến của đối tượng ảnh

Định nghĩa 4.2 [Chu tuyến đối ngẫu]

Hai chu tuyến $C = \langle P_1 P_2 \dots P_n \rangle$ và $C^\perp = \langle Q_1 Q_2 \dots Q_m \rangle$ được gọi là đối ngẫu của nhau nếu và chỉ nếu $\forall i \exists j$ sao cho:

- (i) P_i và Q_j là 4-láng giềng của nhau.
- (ii) Các điểm P_i là vùng thì Q_j là nền và ngược lại.

Định nghĩa 4.3 [Chu tuyến ngoài]

Chu tuyến C được gọi là chu tuyến ngoài (Hình 4.7a) nếu và chỉ nếu

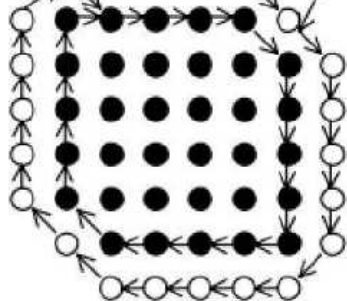
- (i) Chu tuyến đối ngẫu C^\perp là chu tuyến của các điểm nền
- (ii) Độ dài của C nhỏ hơn độ dài C^\perp

Định nghĩa 4.4 [Chu tuyến trong]

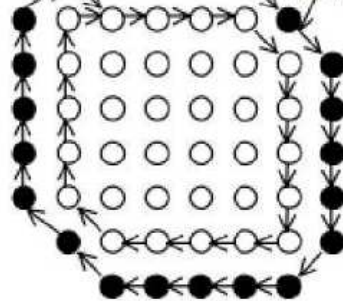
Chu tuyến C được gọi là chu tuyến trong (Hình 4.7b) nếu và chỉ nếu:

- (i) Chu tuyến đối ngẫu C^\perp là chu tuyến của các điểm nền
- (ii) Độ dài của C lớn hơn độ dài C^\perp





a) Chu tuyến ngoài



b) Chu tuyến trong

Hình 4.7. Chu tuyến trong, chu tuyến ngoài

Định nghĩa 4.5 [Điểm trong và điểm ngoài chu tuyến]

Giả sử C'' " $P_1P_2..P_n$ " là chu tuyến của một đối tượng ảnh và P là một điểm ảnh. Khi đó:

- (i) Nếu nửa đường thẳng xuất phát từ P sẽ cắt chu tuyến C tại số lẻ lần, thì P được gọi là điểm trong chu tuyến C và kí hiệu $in(P,C)$
- (ii) Nếu $P \notin C$ và P không phải là điểm trong của C , thì P được gọi là điểm ngoài chu tuyến C và kí hiệu $out(P,C)$.

Bổ đề 4.1 [Chu tuyến đối ngẫu]

Giả sử $E \subseteq \mathfrak{S}$ là một đối tượng ảnh và C'' " $P_1P_2..P_n$ " là chu tuyến của E , C^{\perp} " $Q_1Q_2..Q_m$ " là chu tuyến đối ngẫu tương ứng. Khi đó:

- (i) Nếu C là chu tuyến trong thì $in(Q_i,C) \forall i (i=1,...,m)$
- (ii) Nếu C là chu tuyến ngoài thì $in(P_i,C^{\perp}) \forall i (i=1,...,n)$

Bổ đề 4.2 [Phần trong/ngoài của chu tuyến]

Giả sử $E \subseteq \mathfrak{S}$ là một đối tượng ảnh và C là chu tuyến của E . Khi đó:

- (i) Nếu C là chu tuyến ngoài thì $\forall x \in E$ sao cho $x \notin C$, ta có $in(x,C)$
- (ii) Nếu C là chu tuyến trong thì $\forall x \in E$ sao cho $x \notin C$, ta có $out(x,C)$

Định lý 4.1 [Tính duy nhất của chu tuyến ngoài]

Giả sử $E \subseteq \mathfrak{S}$ là một đối tượng ảnh và C_E là chu tuyến ngoài của E . Khi đó C_E là duy nhất.

4.3.3. Thuật toán dò biên tổng quát

Biểu diễn đối tượng ảnh theo chu tuyến thường dựa trên các kỹ thuật dò biên. Có hai kỹ thuật dò biên cơ bản. Kỹ thuật thứ nhất xét ảnh biên thu được từ ảnh vùng sau một lần duyệt như một đồ thị, sau đó áp dụng các thuật toán duyệt cạnh đồ thị. Kỹ thuật thứ hai dựa trên ảnh vùng, kết hợp đồng thời quá trình dò biên và tách biên. Ở đây ta quan tâm cách tiếp cận thứ hai.

Trước hết, giả sử ảnh được xét chỉ bao gồm một vùng ảnh 8-liên thông \mathfrak{S} , được bao bọc bởi một vành đai các điểm nền. Dễ thấy \mathfrak{S} là một vùng 4-liên thông chỉ là một trường riêng của trường hợp trên.

Về cơ bản, các thuật toán dò biên trên một vùng đều bao gồm các bước sau:

- Xác định điểm biên xuất phát
- Dự báo và xác định điểm biên tiếp theo
- Lặp bước 2 cho đến khi gặp điểm xuất phát

Do xuất phát từ những tiêu chuẩn và định nghĩa khác nhau về điểm biên, và quan hệ liên thông, các thuật toán dò biên cho ta các đường biên mang các sắc thái rất khác nhau.

Kết quả tác động của toán tử dò biên lên một điểm biên r_i là điểm biên r_{i+1} (8-láng giềng của r_i). Thông thường các toán tử này được xây dựng như một hàm đại số Boolean trên các 8-láng giềng của r_i . Mỗi cách xây dựng các toán tử đều phụ thuộc vào định nghĩa quan hệ liên thông và điểm biên. Do đó sẽ gây khó khăn cho việc khảo sát các tính chất của đường biên. Ngoài ra, vì mỗi bước dò biên đều phải kiểm tra tất cả các 8-láng giềng

Chương 4: Các phương pháp phát hiện biên



của mỗi điểm nên thuật toán thường kém hiệu quả. Để khắc phục các hạn chế trên, thay vì sử dụng một điểm biên ta sử dụng cặp điểm biên (một thuộc \mathcal{S} , một thuộc $\overline{\mathcal{S}}$), các cặp điểm này tạo nên tập nền vùng, kí hiệu là NV và phân tích toán tử dò biên thành 2 bước:

- Xác định cặp điểm nền vùng tiếp theo.
- Lựa chọn điểm biên

Trong đó bước thứ nhất thực hiện chức năng của một ánh xạ trên tập NV lên NV và bước thứ hai thực hiện chức năng chọn điểm biên.

Thuật toán dò biên tổng quát

Bước 1: Xác định cặp nền-vùng xuất phát

Bước 2: Xác định cặp  0 |  0 theo

Bước 3: Lựa chọn điểm biên vùng

Bước 4: Nếu gặp lại cặp xuất phát thì dừng, nếu không quay lại bước 2

Việc xác định cặp nền-vùng xuất phát được thực hiện bằng cách duyệt ảnh lần lượt từ trên xuống dưới và từ trái qua phải rồi kiểm tra điều kiện lựa chọn cặp nền-vùng. Do việc chọn điểm biên chỉ mang tính chất quy ước, nên ta gọi ánh xạ xác định cặp nền-vùng tiếp theo là toán tử dò biên.

Định nghĩa 4.6 [Toán tử dò biên]

Giả sử T là một ánh xạ như sau: $T: NV \rightarrow NV$
 $(b,r) \mapsto (b',r')$

Gọi T là một toán tử dò biên cơ sở nếu nó thỏa mãn điều kiện: b',r' là các 8-láng giềng của r .

Giả sử $(b,r) \in NV$; gọi $K(b,r)$ là hàm chọn điểm biên. Biên của một dạng \mathfrak{I} có thể định nghĩa theo một trong ba cách:

- Tập những điểm thuộc \mathfrak{I} có mặt trên NV , tức là $K(b,r)''r$
- Tập những điểm thuộc $\overline{\mathfrak{I}}$ có trên NV , tức là $K(b,r)''b$
- Tập những điểm ảo nằm giữa cặp nền-vùng, tức là $K(b,r)$ là những điểm nằm giữa hai điểm b và r .

Cách định nghĩa thứ ba tương ứng mỗi cặp nền-vùng với một điểm biên. Còn đối với cách định nghĩa thứ nhất và thứ hai một số cặp nền-vùng có thể có chung một điểm biên. Bởi vậy, quá trình chọn điểm biên được thực hiện như sau:

```
i := 1; (bi, ri) := (bo, ro);  
While K(bi, ri) <> K(bn, rn) and i ≤ 8 do  
Begin  
    (bi+1, ri+1) = T(bi, ri); i := i+1;  
End;
```

Điều kiện dừng: Cặp nền-vùng thứ n trùng với cặp nền vùng xuất phát: $(b_n, r_n)''(b_o, r_o)$

Chương 4: Các phương pháp phát hiện biên

*** Xác định cặp nền – vùng xuất phát**

Cặp nền vùng xuất phát được xác định bằng cách duyệt ảnh lần lượt từ trên xuống dưới và từ trái sang phải điểm đem đầu tiên gặp được cùng với điểm trắng trước đó (theo hướng 4) để tạo nên cặp nền vùng xuất phát.

*** Xác định cặp nền vùng tiếp theo**

Đầu vào: pt, dir

Ví dụ: (3, 2) 4

```
Point orient [] = {(1,0); (1,-1); (0,-1); (-1,-1); (-1,0); (-1,1); (0,1); (1,1)};
//Hàm tìm hướng có điểm đen gần nhất
BYTE GextNextDir(POINT pt, BYTE dir)
{
    BYTE pdir= (dir + 7)%8;
    do{
        if(getpixel(pt. x+orient [pdir]. x,pt.y+orient [pdir]. y)==BLACK)
            return pdir;
        pdir = (pdir + 7) %8;
    }while(pdir != dir);
    return. ERR; //Điểm cô lập
}
//Gán giá trị cho bước tiếp theo
pdir = GetNextDir(pt, dir);
if(pdir==ERR) //Kiểm tra có là điểm cô lập không?
    return. ERR; //Điểm cô lập
pt. x = pt. x + orient [pdir]. x;
pt. y = pt. y + orient [pdir]. y;
```

Để tính giá trị cho hướng tiếp theo ta lập bảng dựa trên giá trị pdir đã tính được trước đó theo các khả năng có thể xảy ra:

<i>pdir</i>	<i>Điểm trắng trước đó</i>	<i>Trắng so với đen mới</i>
0	1	2
1	2	4
2	3	4
3	4	6
4	5	6
5	6	0
6	7	0
7	0	2

⇒ Do đó công thức để tính hướng tiếp theo sẽ là:

dir'' ((pdir+3)/ 2 * 2)%8;

4.4. CÁC PHƯƠNG PHÁP KHÁC

4.4.1. Các phép toán hình thái cơ bản

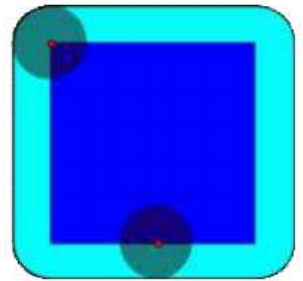
Hình thái là thuật ngữ chỉ sự nghiên cứu về cấu trúc hay hình học topo của đối tượng trong ảnh. Phần lớn các phép toán của "Hình thái" được định nghĩa từ hai phép toán cơ bản là phép "giãn nở" (Dilation) và phép "co" (Erosion).

Các phép toán này được định nghĩa như sau: Giả sử ta có đối tượng X và phần tử cấu trúc (mẫu) B trong không gian Euclide hai chiều. Kí hiệu B_x là dịch chuyển của B tới vị trí x.

Định nghĩa 4.7 (DILATION)

Phép "giãn nở" của X theo mẫu B là hợp của tất cả các B_x với x thuộc X. Ta có:

$$X \oplus B = \bigcup_{x \in X} B_x$$



Định nghĩa 4.7 (EROSION)

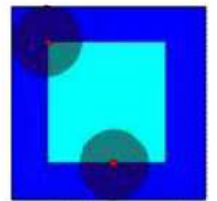
Phép "co" của X theo B là tập hợp tất cả các điểm x sao cho B_x nằm trong X. Ta có:

$$X \ominus B = \{x: B_x \subseteq X\}$$

Ví dụ: Ta có tập X như sau: $X = \begin{pmatrix} 0 & x & 0 & x & x \\ x & 0 & x & x & 0 \\ 0 & x & x & 0 & 0 \\ 0 & x & 0 & x & 0 \\ 0 & x & x & x & 0 \end{pmatrix}$

$B = \begin{pmatrix} \otimes & x \end{pmatrix}$

$$X \oplus B = \begin{pmatrix} 0 & x & x & x & x & x \\ x & x & x & x & x & 0 \\ 0 & x & x & x & 0 & 0 \\ 0 & x & x & x & x & 0 \\ 0 & x & x & x & x & 0 \end{pmatrix} \text{ và } X \ominus B = \begin{pmatrix} 0 & 0 & 0 & x & 0 \\ 0 & 0 & x & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & x & x & 0 & 0 \end{pmatrix}$$



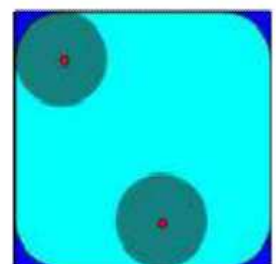
Định nghĩa 4.9 (OPEN)

Phép toán mở (OPEN) của X theo cấu trúc B là tập hợp các điểm của ảnh X sau khi đã co và giãn nở liên tiếp theo B. Ta có:

$$\text{OPEN}(X, B) = (X \ominus B) \oplus B$$

Ví dụ: Với tập X và B trong ví dụ trên ta có

$$\text{OPEN}(X, B) = (X \ominus B) \oplus B = \begin{pmatrix} 0 & 0 & 0 & x & x \\ 0 & 0 & x & x & 0 \\ 0 & x & x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$



Chương 4: Các phương pháp phát hiện biên

Định nghĩa 4.10 (CLOSE)

Phép toán đóng (CLOSE) của X theo cấu trúc B là tập hợp các điểm của ảnh X sau khi đã giãn nở và co liên tiếp theo B. Ta có:

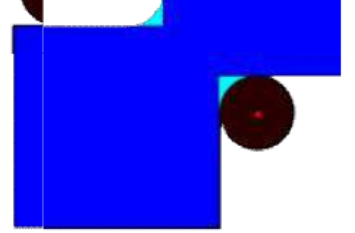
$$\text{CLOSE}(X, B) = (X \oplus B) \ominus B$$

Theo ví dụ trên ta có:

👍 0 | 👎 0



$$\text{CLOSE}(X, B) \text{'' } (X \oplus B) \ominus B \text{'' } \begin{pmatrix} 0 & x & x & x & x \\ x & x & x & x & 0 \\ 0 & x & x & 0 & 0 \\ 0 & x & x & x & 0 \\ 0 & x & x & x & 0 \end{pmatrix}$$



4.4.2. Một số tính chất của phép toán hình thái

* **Mệnh đề 4.1 [Tính gia tăng]:**

$$\begin{aligned} \text{(i) } X \subseteq X' &\Rightarrow \begin{cases} X \ominus B \subseteq X' \ominus B & \forall B \\ X \oplus B \subseteq X' \oplus B & \forall B \end{cases} \\ \text{(ii) } B \subseteq B' &\Rightarrow \begin{cases} X \ominus B \supseteq X \ominus B' & \forall X \\ X \oplus B \subseteq X \oplus B' & \forall X \end{cases} \end{aligned}$$

Chứng minh:

$$\text{(i) } X \oplus B \text{'' } \bigcup_{x \in X} B_x \subseteq \bigcup_{x \in X'} B_x = X' \oplus B$$

$$X \ominus B \text{'' } \{x / B_x \subseteq X\} \subseteq \{x / B_x \subseteq X'\} \text{'' } X' \ominus B$$

$$\text{(ii) } X \oplus B \text{'' } \bigcup_{x \in X} B_x \subseteq \bigcup_{x \in X} B'_x = X \oplus B'$$

Theo định nghĩa:

$$X \ominus B' \text{'' } \{x / B'_x \subseteq X\} \subseteq \{x / B_x \subseteq X\} \text{'' } X \ominus B.$$

* **Mệnh đề 4.2 [Tính phân phối với phép \cup]:**

$$\text{(i) } X \oplus (B \cup B') \text{'' } (X \oplus B) \cup (X \oplus B')$$

$$\text{(ii) } X \ominus (B \cup B') \text{'' } (X \ominus B) \cap (X \ominus B')$$

Chương 4: Các phương pháp phát hiện biên

Chứng minh:

$$(i) \quad X \oplus (B \cup B') \supseteq (X \oplus B) \cup (X \oplus B')$$

Ta có: $B \cup B' \supseteq B$

$$X \oplus (B \cup B') \supseteq X \oplus B \quad (\text{tính gia tăng})$$

Tương tự:

$$X \oplus (B \cup B') \supseteq X \oplus B'$$

$$X \oplus (B \cup B') \supseteq (X \oplus B) \cup (X \oplus B') \quad (4.1)$$

Mặt khác,

$$\forall y \in X \oplus (B \cup B') \Rightarrow \exists x \in X \text{ sao cho } y \in (B \cup B')_x$$

$$\Rightarrow \begin{cases} y \in B_x \\ y \in B'_x \end{cases} \Rightarrow \begin{cases} y \in X \oplus B \\ y \in X \oplus B' \end{cases}$$

$$\Rightarrow y \in (X \oplus B) \cup (X \oplus B')$$

$$\Rightarrow X \oplus (B \cup B') \subseteq (X \oplus B) \cup (X \oplus B') \quad (4.2)$$

$$\text{Từ (4.1) và (4.2) } X \oplus (B \cup B') = (X \oplus B) \cup (X \oplus B')$$

$$(ii) \quad X \ominus (B \cup B') \supseteq (X \ominus B) \cap (X \ominus B')$$

$$\text{Ta có: } B \cup B' \supseteq B$$

Ta có: $B \cup B' \supseteq B$

$$\Rightarrow X \ominus (B \cup B') \subseteq X \ominus B \quad (\text{tính gia tăng})$$

Tương tự: $X \ominus (B \cup B') \subseteq X \ominus B'$

$$\Rightarrow X \ominus (B \cup B') \subseteq (X \ominus B) \cap (X \ominus B') \quad (4.3)$$

Mặt khác,

$$\forall x \in (X \ominus B) \cap (X \ominus B')$$

$$\text{Suy ra, } \begin{cases} x \in X \ominus B \\ x \in X \ominus B' \end{cases} \Rightarrow \begin{cases} B_x \subseteq X \\ B'_x \subseteq X \end{cases}$$

$$\Rightarrow (B \cup B')_x \subseteq X$$

$$\Rightarrow x \in X \ominus (B \cup B')$$

$$\Rightarrow X \ominus (B \cup B') \supseteq (X \ominus B) \cap (X \ominus B') \quad (4.4)$$

Từ (4.3) và (4.4) ta có: $X \ominus (B \cup B') = (X \ominus B) \cap (X \ominus B')$.

*** Ý nghĩa:**

Ta có thể phân tích các mẫu phức tạp trở thành các mẫu đơn giản thuận tiện cho việc cài đặt.

*** Mệnh đề 4.3 [Tính phân phối với phép \cap]:**

$$(X \cap Y) \ominus B = (X \ominus B) \cap (Y \ominus B)$$

Chứng minh:

Ta có, $X \cap Y \subseteq X$

$$\Rightarrow (X \cap Y) \ominus B \subseteq X \ominus B$$

Tương tự: $(X \cap Y) \ominus B \subseteq Y \ominus B$

Chương 4: Các phương pháp phát hiện biên

$$\Rightarrow (X \cap Y) \ominus B \subseteq (X \ominus B) \cap (Y \ominus B) \quad (4.5)$$

Mặt khác,

$$\begin{aligned} & \forall x \in (X \ominus B) \cap (Y \ominus B) \\ \text{Suy ra } & \begin{cases} x \in X \ominus B \\ x \in Y \ominus B \end{cases} \Rightarrow \begin{cases} B_x \subseteq X \\ B_x \subseteq Y \end{cases} \\ & \Rightarrow B_x \subseteq X \cap Y \\ & \Rightarrow x \in (X \cap Y) \ominus B \\ & \Rightarrow (X \cap Y) \ominus B \supseteq (X \ominus B) \cap (Y \ominus B) \end{aligned} \quad (4.6)$$

Từ (4.5) và (4.6) ta có: $(X \cap Y) \ominus B = (X \ominus B) \cap (Y \ominus B)$.

*** Mệnh đề 4.4 [Tính kết hợp]**

- (i) $(X \oplus B) \oplus B' = X \oplus (B \oplus B')$
- (ii) $(X \ominus B) \ominus B' = X \ominus (B \oplus B')$

Chứng minh:

$$(i) (X \oplus B) \oplus B' = X \oplus (B' \oplus B)$$

$$\begin{aligned} \text{Ta có, } (X \oplus B) \oplus B' &= \left(\bigcup_{x \in X} B_x \right) \oplus B' \\ &= \bigcup_{x \in X} (B_x \oplus B') = \bigcup_{x \in X} (B' \oplus B)_x \\ &= X \oplus (B' \oplus B) \end{aligned}$$

$$(i) (X \ominus B) \ominus B' = X \ominus (B \oplus B')$$

Trước hết ta đi chứng minh: $B'_x \subseteq X \ominus B \Leftrightarrow (B' \oplus B)_x \subseteq X$

Thật vậy, do $B'_x \subseteq X \ominus B$ nên $\forall y \in B'_x \Rightarrow y \in X \ominus B$

$$\Rightarrow B_y \subseteq X$$

$$\Rightarrow \bigcup_{y \in B'_x} B_y \subseteq X$$

$$\Rightarrow (B' \oplus B)_x \subseteq X$$

Mặt khác, $(B' \oplus B)_x \subseteq X \Leftrightarrow (B'_x \oplus B) \subseteq X$

$$\Leftrightarrow \bigcup_{y \in B'_x} B_y \subseteq X$$

$$\Rightarrow \forall y \in B'_x \text{ ta có } B_y \subseteq X$$

$$\Rightarrow \text{hay } \forall y \in B'_x \text{ ta có } y \in X \ominus B$$

Do đó, $B'_x \subseteq X \ominus B$

Ta có, $(X \ominus B) \ominus B' \quad \text{''} \quad \{x / B_x \subseteq X\} \ominus B'$

$$\text{''} \quad \{x / B'_x \subseteq X \ominus B\}$$

$$\text{''} \quad \{x / (B' \oplus B)_x \subseteq X\} \text{ (do chứng minh ở trên)}$$

$$X \ominus (B \oplus B').$$

*** Định lý 4.2 [X bị chặn bởi các cận OPEN và CLOSE]**

Giả sử, X là một đối tượng ảnh, B là mẫu, khi đó, X sẽ bị chặn trên bởi tập CLOSE của X theo B và bị chặn dưới bởi tập OPEN của X theo B.

Tức là:

$$(X \oplus B) \ominus B \supseteq X \supseteq (X \ominus B) \oplus B$$

Chứng minh:

$$\text{Ta có: } \forall x \in X \Rightarrow B_x \subseteq X \oplus B \quad (\text{Vì } X \oplus B \supseteq \bigcup_{x \in X} B_x)$$

$$\Rightarrow x \in (X \oplus B) \ominus B \quad (\text{theo định nghĩa phép co})$$

$$\Rightarrow (X \oplus B) \ominus B \supseteq X \quad (4.7)$$

Mặt khác,

$$\forall y \in (X \ominus B) \oplus B, \text{ suy ra:}$$

$$\exists x \in X \ominus B \text{ sao cho } y \in B_x \quad (\text{Vì } (X \ominus B) \oplus B \supseteq \bigcup_{x \in X \ominus B} B_x)$$

$$\Rightarrow B_x \subseteq X \Rightarrow y \in X$$

$$\text{Suy ra: } X \supseteq (X \ominus B) \oplus B \quad (4.8)$$

Từ (4.7) và (4.8) Ta có: $(X \oplus B) \ominus B \supseteq X \supseteq (X \ominus B) \oplus B$.

***Hệ quả 4.1 [Tính bất biến]:**

$$(i) \quad ((X \oplus B) \ominus B) \oplus B \supseteq X \oplus B$$

$$(ii) \quad ((X \ominus B) \oplus B) \ominus B \supseteq X \ominus B$$

Chứng minh:

$$(i) \text{ Thật vậy, từ định lý 4.2 ta có } X \subseteq (X \oplus B) \ominus B$$

$$\Rightarrow X \oplus B \subseteq ((X \oplus B) \ominus B) \oplus B \quad (\text{do tính chất gia tăng}) \quad (4.9)$$

Mặt khác, cũng từ định lý 4.2 ta có $(X \ominus B) \oplus B \subseteq X \quad \forall X$

$$\text{Do đó, thay } X \text{ bởi } X \oplus B \text{ ta có, } ((X \oplus B) \ominus B) \oplus B \subseteq X \oplus B \quad (4.10)$$

$$\text{Từ (4.9) và (4.10) Ta có: } ((X \oplus B) \ominus B) \oplus B \supseteq X \oplus B$$

$$(ii) \text{ Thật vậy, từ định lý 4.2 ta có } (X \ominus B) \oplus B \subseteq X$$

$$\Rightarrow ((X \ominus B) \oplus B) \ominus B \subseteq X \ominus B \quad (\text{do tính chất gia tăng}) \quad (4.11)$$

Mặt khác, cũng từ định lý 4.2 ta có $X \subseteq (X \oplus B) \ominus B \quad \forall X$

$$\text{Do đó, thay } X \text{ bởi } X \ominus B \text{ ta có, } X \ominus B \subseteq ((X \ominus B) \oplus B) \ominus B \quad (4.12)$$

Từ (4.11) và (4.12) Ta có: $((X \ominus B) \oplus B) \ominus B \supseteq X \ominus B$ (đpcm).

4.4.3. Xấp xỉ trên và xấp xỉ dưới đối tượng ảnh

Biên là vấn đề quan trọng trong xử lý ảnh và nhận dạng, vì các đặc điểm trích chọn trong quá trình nhận dạng chủ yếu dựa vào biên. Trong thực tế người ta thường dùng hai phương pháp phát hiện biên là phát hiện biên trực tiếp và gián tiếp. Phần này đề

Chương 4: Các phương pháp phát hiện biên

cập đến một tiếp cận mới trong phát hiện biên dựa vào các phép toán hình thái thông qua các kỹ thuật xấp xỉ trên và xấp xỉ dưới đối tượng.

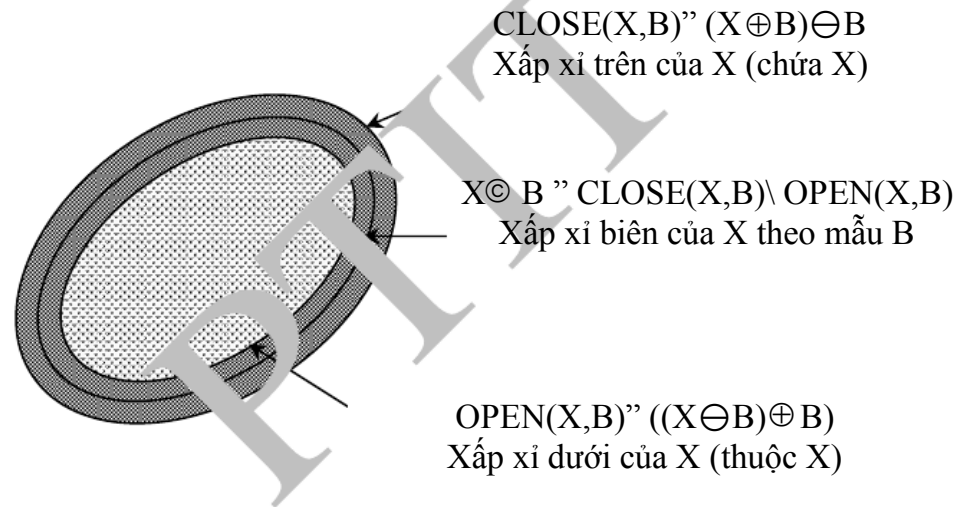
Cũng đã có một vài tác giả đề cập đến kỹ thuật phát hiện biên dựa vào phép toán hình thái. Nhưng các kỹ thuật phát hiện biên trực tiếp, gián tiếp và dựa vào các phép toán hình thái kể trên đều xuất phát từ quan điểm biên của đối tượng là một tập hợp con của đối tượng. Trong thực tế chúng ta thường hiểu đường biên là khu vực ranh giới bao gồm cả hai phần thuộc đối tượng và không thuộc đối tượng. Ở phần dưới đây, chúng tôi đề xuất một kỹ thuật phát hiện biên dựa vào phép toán hình thái theo quan niệm này, xuất phát từ cơ sở định lý 4.2 đã được chứng minh.

Biên (hay đường biên) có thể hiểu đơn giản là các đường bao của các đối tượng trong ảnh chính là ranh giới giữa đối tượng và nền. Việc xem ranh giới là phần được tạo lập bởi

ảnh chính là ranh giới giữa đối tượng và nền. Việc xem ranh giới là phần được tạo lập bởi các điểm thuộc đối tượng và thuộc nền cho phép ta xác định biên dựa trên các phép toán hình thái.

Theo định lý 4.2 ta có: $(X \oplus B) \ominus B \supseteq X \forall B$

Như vậy, tập $CLOSE(X, B) \supseteq (X \oplus B) \ominus B$ có thể được xem như là xấp xỉ trên của tập X theo mẫu B (Hình 4.8).



Hình 4.8. Xấp xỉ trên và dưới theo mẫu B của X

Cũng theo định lý 4.2 ta có, $(X \ominus B) \oplus B \subseteq X \forall B$

Do vậy, tập $OPEN(X, B) \supseteq (X \ominus B) \oplus B$ có thể được xem như là xấp xỉ dưới của tập X theo mẫu B .

Từ đó, tập $CLOSE(X, B) \setminus OPEN(X, B)$ có thể được xem như là xấp xỉ biên của tập X theo mẫu và quá trình xấp xỉ biên của X theo mẫu B kí hiệu là $X \odot B$.

Để tăng độ chính xác, người ta thường xem B là dãy các phần tử cấu trúc.

$B \supseteq \{B_i, 1 \leq i \leq n\}$

Và xấp xỉ biên của X theo tập cấu trúc B được xác định:

$$X \odot B = \bigcup_{i=1}^n (X \odot B_i)$$

4.4.4. Thuật toán phát hiện biên dựa vào phép toán hình thái

Vào : Ảnh X và dãy mẫu B'' $\{B_i, 1 \leq i \leq n\}$;

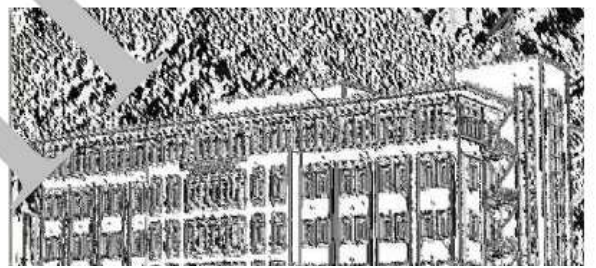
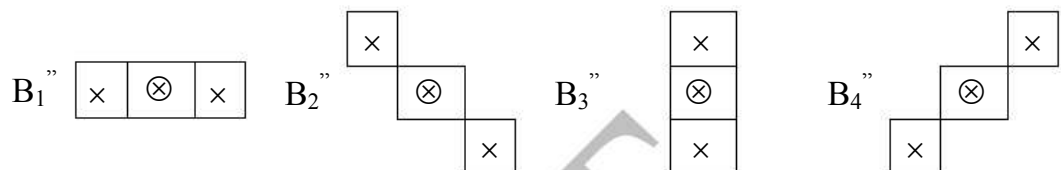
Ra : Biên của đối tượng theo mẫu B

Phương pháp:

Bước 1: Tính $X \odot B_i \forall i=1, n$

Bước 2: Tính
$$\bigcup_{i=1}^n (X \odot B_i)$$

Trong Hình 4.9a dưới đây là ảnh gốc với 256 mức xám, Hình 4.9b là ảnh biên thu được qua phát hiện biên bằng Sobel, Hình 4.9c là ảnh biên thu được qua phát hiện biên bằng Laplace. Hình 4.9d là ảnh biên kết quả thực hiện bởi thuật toán phát hiện biên bằng các phép toán hình thái với ngưỡng tách $\theta = 128$ và các mẫu tách biên B_i là:

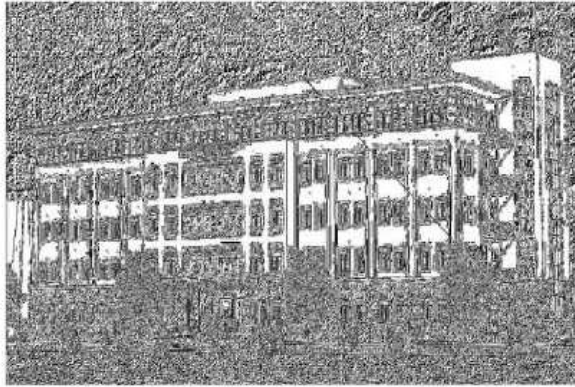




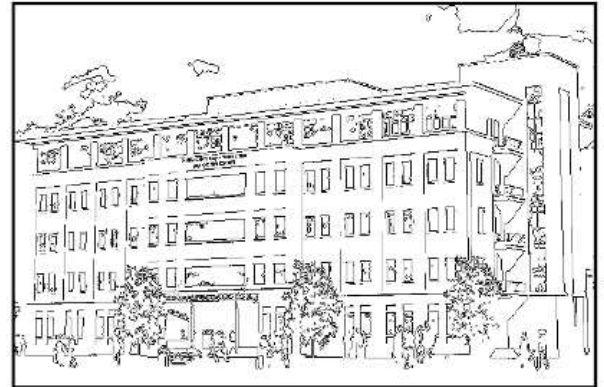
a) Ảnh gốc đa cấp xám



b) Ảnh biên thu được qua Sobel



c) Ảnh biên thu được qua Laplace



d) Ảnh biên kết quả dựa vào phép toán hình thái

Hình 4.9. Phát hiện biên bởi thuật toán dựa vào phép toán hình thái

4.5. CÂU HỎI ÔN TẬP CHƯƠNG

Câu 1: Điểm biên được định nghĩa như thế nào? Kỹ thuật phổ biến nhất để phát hiện biên là gì?

Câu 2: Thực hiện cài đặt phương pháp tách biên Canny. Đầu vào của hệ thống là một ma trận ảnh đa cấp xám (0-255) và đầu ra là ảnh với chỉ hai loại điểm ảnh, điểm nền có giá trị là 0 và điểm biên có giá trị là 255.

Câu 3: Thực hiện phép co và giãn hình X với một phần tử cấu trúc B sau để được hai hình X_1 và X_2 . Thực hiện phép trừ để tìm ảnh biên. Được biết tâm của B là ở chính giữa.

$$X = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Câu 4: Thực hiện phép đóng và mở hình X với một phần tử cấu trúc B ở câu 3 để được hai hình X_1 và X_2 . Thực hiện phép trừ để tìm ảnh biên. Được biết tâm của B là ở chính giữa.

5.1. GIỚI THIỆU

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

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

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

$$X = \bigcup_{i=1}^N R_i \quad (5.1)$$

$$R_i \cap R_j = \emptyset \text{ với } i \neq j \quad (5.2)$$

$$P(R_i) = \text{TRUE} \text{ với } i = 1..N \quad (5.3)$$

$$P(R_i \cup R_j) = \text{FALSE} \text{ với } i \neq j \quad (5.4)$$

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