

**Đề tài: Tìm hiểu phương pháp phát hiện biên cho ảnh đa cấp xám và ảnh màu.**

**Giáo viên hướng dẫn: PGS. TS. Ngô Quốc Tạo**

**Sinh viên thực hiện: Đặng Thị Thương.**

**LỜI CẢM ƠN**

Đầu tiên em xin gửi lời cảm ơn chân thành và trân trọng nhất tới PQS TS Ngô Quốc Tạo, người đã trực tiếp hướng dẫn, giúp đỡ em trong việc hình thành, phát triển và hoàn thành đồ án tốt nghiệp này.

Em cũng xin gửi lời cảm ơn trân trọng tới các thầy cô giáo trong khoa Công nghệ thông tin – Đại học Dân lập Hải Phòng, những người đã tận tình dạy dỗ và dìu dắt em trong suốt bốn năm học vừa qua để chúng em có những kinh nghiệm, kiến thức vững chắc để hoàn thành đồ án tốt nghiệp này.

Em xin chân thành cảm ơn gia đình, bạn bè đã luôn động viên, giúp đỡ em về mọi mặt trong quá trình học tập và nghiên cứu, cũng như góp ý cho đồ án tốt nghiệp.

Em rất mong nhận được sự khích lệ, quan tâm, giúp đỡ của các quý thầy cô và các bạn trong quá trình học tập và công tác sau này, em rất mong muốn được mang một công sức, kiến thức của mình để xây dựng xã hội, xứng đáng với sự giúp đỡ và tình cảm mà mọi người đã dành cho em.

Em xin chân thành cảm ơn!

Hải Phòng, ... tháng ... năm 2011

Sinh viên

Đặng Thị Thương

---



---

**MỤC LỤC**

<b>MỤC LỤC</b>	<b>3</b>
<b>DANH SÁCH CÁC HÌNH ẢNH.....</b>	<b>5</b>
<b>MỞ ĐẦU</b>	<b>6</b>
<b>CHƯƠNG I: TỔNG QUAN VỀ XỬ LÝ ẢNH VÀ BIÊN.....</b>	<b>8</b>
1.1. Tổng quan về xử lý ảnh .....	8
1.1.1 Xử lý ảnh là gì ?.....	8
1.1.2 Các quá trình cơ bản trong hệ thống xử lý ảnh.....	8
1.1.3 Một số vấn đề cơ bản trong xử lý ảnh .....	11
1.2 Nâng cao chất lượng ảnh và toán tử không gian .....	15
1.2.1 Làm trơn nhiễu bằng lọc tuyến tính.....	16
1.2.2 Làm trơn nhiễu bằng lọc phi tuyến.....	18
1.2.3 Lọc thông thấp, thông cao và lọc giải thông.....	19
1.3 Tổng quan về biên .....	20
1.3.1 Biên và các kiểu biên cơ bản .....	20
1.3.2 Vai trò của biên trong nhận dạng.....	21
<b>CHƯƠNG II: PHƯƠNG PHÁP PHÁT HIỆN BIÊN CHO ẢNH ĐA MỨC XÁM</b>	<b>23</b>
2.1 Giới thiệu .....	23
2.2 Quy trình phát hiện biên .....	23
2.3 Phương pháp phát hiện biên cơ bản .....	24
2.3.1 Phương pháp phát hiện biên Gradient .....	24
2.3.2 Phương pháp phát hiện biên Laplace.....	26
2.4 Phương pháp phát hiện biên nâng cao.....	29
2.4.1 Phương pháp phát hiện biên Canny .....	29
2.4.2 Phương pháp phát hiện biên Wavelet.....	32
<b>CHƯƠNG III: PHƯƠNG PHÁP PHÁT HIỆN BIÊN CHO ẢNH MÀU ....</b>	<b>34</b>
3.1 Cơ sở phát hiện biên ảnh màu .....	34
3.2 Mô hình ba màu.....	34
3.3 Các phương pháp phát hiện biên ảnh màu .....	37
3.3.1 Toán tử Vector Gradient.....	37

---



---

3.3.2 Phương pháp Vector Field .....	39
3.3.3 Bộ dò biên Vector Order-Statistic .....	40
<b>CHƯƠNG IV: PHƯƠNG PHÁP PHÁT HIỆN BIÊN THEO WAVELET..</b>	<b>43</b>
4.1 Giới thiệu .....	43
4.2 Biến đổi Wavelet liên tục (The Continuous Wavelet Transform – CWT) .....	44
4.3 MRA – Multi-resolution Analysis.....	45
4.4 Xây dựng một Wavelet.....	46
4.5 Phân hoạch và tái thiết wavelet .....	46
4.3 Biến đổi Wavelet rời rạc (DWT – Discrete Wavelet Transform ) .....	47
4.4 Phương pháp phát hiện biên DWT .....	49
<b>CHƯƠNG V: CÀI ĐẶT THỬ NGHIỆM VÀ NHẬN XÉT ĐÁNH GIÁ CÁC PHƯƠNG PHÁP .....</b>	<b>51</b>
4.1 So sánh các phương pháp phát hiện biên (Gradient, Laplace, Canny) .....	51
4.2 Cài đặt thử nghiệm chương trình Wavelet Transform .....	52
<b>KẾT LUẬN .....</b>	<b>60</b>
<b>TÀI LIỆU THAM KHẢO:.....</b>	<b>61</b>

## **DANH SÁCH CÁC HÌNH ẢNH**

Hình 1.1: Quá trình xử lý ảnh

Hình 1.2: Các bước cơ bản trong một quá trình xử lý ảnh

Hình 1.3 Một ví dụ về điểm ảnh

Hình 1.4 Biểu diễn ảnh với độ phân giải

Hình 1.5: lân cận các điểm ảnh của  $P(x,y)$

Hình 1.6: Đường bao của ảnh

Hình 2.1: Quy trình phát hiện biên

Hình 2.2: Hình mô tả các điểm biên lân cận của P

Hình 3.1: Các màu quang phổ theo tiêu chuẩn CIE

Hình 3.2: Biểu diễn mô hình màu RGB trên trục tọa độ

Hình 3.3: Biểu diễn mô hình màu CMY trên trục tọa độ

Hình 4.1: Sơ đồ kim tự tháp Laplace phát triển bởi Burt và Adelson

Hình 4.2: DWT của hai chiều tín hiệu

Hình 5.1: Biên ảnh theo Gradient, Laplace, Canny.

Hình 5.2: Biến đổi ảnh với tỉ lệ 1 và bộ lọc daub1.

Hình 5.3: DWT hai chiều tín hiệu.

Hình 5.4: Loại bỏ thành phần tần số thấp.

Hình 5.5: Ảnh biến đổi sau khi xây dựng lại

Hình 5.6: Ảnh sau khi được làm nổi biên.

Hình 5.7: Biên ảnh sau khi được khuếch đại.

## MỞ ĐẦU

Trong hai thập kỷ qua, chúng ta đã chứng kiến một sự tăng trưởng bùng nổ trong đa dạng cả về kỹ thuật và phạm vi của các ứng dụng xử lý ảnh. Xử lý ảnh là một trong những chuyên ngành quan trọng và lâu đời của Công nghệ thông tin. Xử lý ảnh được áp dụng trong nhiều lĩnh vực khác nhau như y học, vật lý, hoá học, tìm kiếm tội phạm, trong quân sự và trong một số lĩnh vực khác...

Phần lớn con người thu nhận thông tin bằng thị giác, cụ thể đó là các hình ảnh. Vì vậy xử lý ảnh là vấn đề không thể thiếu và hết sức quan trọng để thu được hình ảnh tốt hơn, đẹp hơn, nhằm đáp ứng yêu cầu thông tin khác nhau của người nhận.

Tuy nhiên biên cạnh việc xử lý ảnh xám, phạm vi xử lý ảnh màu thỉnh thoảng vẫn còn được che phủ, mặc dù đã trở thành phổ biến, người tiêu dùng lựa chọn các tiện ích trong ảnh màu hơn ảnh xám truyền thống. Với những tiến bộ trong bộ cảm biến hình ảnh, truyền hình kỹ thuật số, cơ sở dữ liệu hình ảnh, video, đa phương tiện và hệ thống, và với sự gia tăng của máy in màu, hiển thị hình ảnh màu sắc, các thiết bị DVD, và đặc biệt là máy ảnh kỹ thuật số và hình ảnh kích hoạt điện tử tiêu dùng, xử lý ảnh màu dường như đã trở thành trọng tâm chính của hội nghiên cứu xử lý ảnh.

Trong xử lý ảnh, việc nhận dạng và phân lớp đối tượng cần trải qua các quá trình và các thao tác khác nhau. Phát hiện biên là một giai đoạn rất quan trọng vì các kỹ thuật phân đoạn ảnh chủ yếu dựa vào giai đoạn này. Mục đích của việc dò biên sẽ đánh dấu những điểm trong một ảnh số mà có sự thay đổi đột ngột về độ xám, tập hợp nhiều điểm biên tạo nên một đường bao quanh ảnh (đường biên). Nhờ có đường biên mà chúng ta có thể phân biệt giữa đối tượng và nền, phân biệt giữa các vùng khác nhau và định vị được đối tượng từ đó mà nhận dạng đối tượng. Đây là cơ sở quan trọng trong việc ứng dụng phương pháp này vào thực tiễn của cuộc sống, đặc biệt là trong điều kiện đất nước ta đang từng bước phát triển và đi lên nên việc nghiên cứu các ứng dụng vấn đề này cần được quan tâm và phát triển.

Từ đó, em đã chọn đề tài **“Tìm hiểu phương pháp phát hiện biên cho ảnh đa cấp xám và ảnh màu”** Mục đích chính của đề tài là hệ thống hóa kiến thức về các phương pháp phát hiện biên đối với ảnh đa cấp xám và ảnh màu, từ các kỹ thuật dò biên cài đặt chương trình để đưa ra các nhận xét, so sánh, đánh giá về các phương pháp phát hiện biên. Qua đó có cái nhìn tổng quát về các phương pháp phát hiện biên.

Sau một thời gian tìm hiểu về đề tài, ngoài phần mở đầu và kết luận, kết cấu tài liệu báo cáo bao gồm 5 chương với các nội dung cụ thể như sau:

**Chương I: Tổng quan về xử lý ảnh và biên**

Trong chương này trình bày sơ lược về xử lý ảnh, giới thiệu các bước xử lý trong một hệ thống xử lý ảnh. Một số thành phần cốt tử trong xử lý ảnh, như điểm ảnh, mức xám, biên,... được trình bày như là các khái niệm.

**Chương II: Các phương pháp phát hiện biên cho ảnh đa cấp xám**

Nội dung của chương này sẽ đề cập đến một số phương pháp phát hiện biên cơ bản trong phương pháp đạo hàm bậc nhất và phương pháp đạo hàm bậc hai và vài phương pháp nâng cao như Canny, Wavelet.

**Chương III: Các phương pháp phát hiện biên cho ảnh màu**

Trong chương này đề cập đến phương pháp phát hiện biên vector Gradient, vector Field,...

**Chương IV: Phương pháp phát hiện biên theo Wavelet**

Tìm hiểu về biến đổi Wavelet, từ đó tìm hiểu phương pháp phát hiện biên DWT...

**Chương V: Cài đặt thử nghiệm và nhận xét đánh giá các phương pháp phát hiện biên**

Qua việc cài đặt thử nghiệm các phương pháp phát hiện biên đã trình bày trong các chương trước, từ các kết quả mô phỏng thực nghiệm khi chạy chương trình, trong chương này đưa ra các nhận xét đánh giá, so sánh các phương pháp phát hiện biên. Chỉ ra phương pháp phát hiện biên phù hợp với loại ảnh cần xử lý.

Sinh viên

Đặng Thị Thương

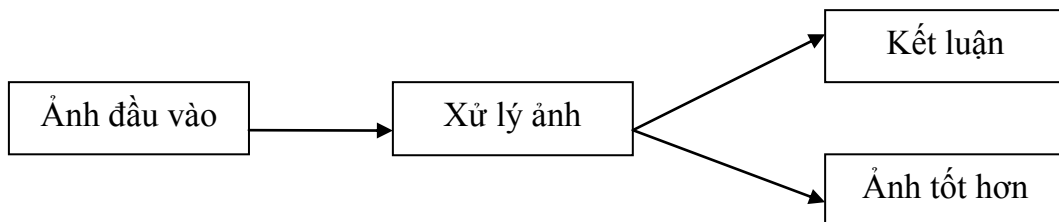
## CHƯƠNG I: TỔNG QUAN VỀ XỬ LÝ ẢNH VÀ BIÊN

### 1.1. Tổng quan về xử lý ảnh

#### 1.1.1 Xử lý ảnh là gì ?

Con người thu nhận thông tin qua các giác quan trong đó thị giác đóng vai trò quan trọng nhất. Cùng với sự phát triển nhanh của phần cứng máy tính, xử lý ảnh và đồ họa đã phát triển mạnh mẽ và ngày càng có nhiều ứng dụng trong cuộc sống. Nhờ thế mà xử lý ảnh từng bước đóng một vai trò quan trọng trong tương tác người máy.

Quá trình xử lý nhận dạng ảnh là một quá trình thao tác nhằm biến đổi một ảnh đầu vào để cho ra một kết quả mong muốn. Xử lý ảnh thông thường gồm 3 bước:



Hình 1.1: Quá trình xử lý ảnh.

– *Bước 1:* Nhập một hình ảnh với một máy quét quang học hoặc trực tiếp thông qua nghệ thuật chụp ảnh số.

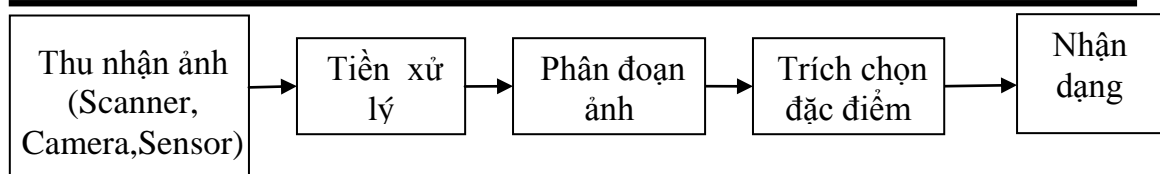
– *Bước 2:* Thao tác hoặc phân tích các hình ảnh bằng một cách nào đó. Giai đoạn này có thể bao gồm kỹ thuật nâng cao chất lượng hình ảnh và nén dữ liệu, hoặc hình ảnh có thể được phân tích để tìm ra các hình dáng mà mắt người không thể thấy được. Ví dụ: các nhà khí tượng học sử dụng xử lý ảnh để phân tích các ảnh vệ tinh.

– *Bước 3:* Kết quả đầu ra - hình ảnh có thể bị thay đổi bằng cách này hay cách khác, nó có thể là một ảnh “tốt hơn” (Ví dụ: ảnh mờ được xử lý để nhìn rõ hơn) hoặc một kết luận (Ví dụ: phân tích ảnh để trích chọn các đặc trưng vân tay hay ảnh một tai nạn giao thông phác họa hiện trường tại nạn).

#### 1.1.2 Các quá trình cơ bản trong hệ thống xử lý ảnh

Một ảnh đầu vào cần thông qua rất nhiều bước khác nhau để có được một ảnh đầu ra mong muốn. Các quá trình cơ bản của một hệ thống xử lý ảnh được thể hiện thông qua sơ đồ dưới đây:





Hình 1.2: Các bước cơ bản trong một quá trình xử lý

### 1.1.2.1 Thu nhận ảnh

Để thực hiện được quá trình đầu tiên trong hệ thống xử lý ảnh, ta cần sử dụng các thiết bị thu nhận ảnh để chuyển các thông tin dưới dạng hình ảnh thành các cấu trúc lưu trữ được trong máy tính và được hiển thị ra màn hình, máy in,... Ảnh có thể thu nhận từ vệ tinh qua các bộ cảm ứng, qua camera, máy chụp ảnh đơn sắc (màu), hay các tranh, ảnh được quét trên máy quét ảnh. Nếu ảnh thu nhận được chưa phải là dạng số hóa ta phải chuyển đổi hay số hóa ảnh trước khi chuyển sang giai đoạn tiếp theo.

### 1.1.2.2 Tiền xử lý

Sau khi được thu nhận bởi các thiết bị thu nhận ảnh, ảnh sẽ được cải thiện về độ tương phản, nhiễu,... bởi các kỹ thuật xử lý ảnh để làm ảnh tốt hơn theo mục đích sử dụng nhằm phục vụ cho quá trình xử lý tiếp theo.

Một số tiền trình trong quá trình tiền xử lý là:

- *Điều chỉnh độ chiếu sáng*: khắc phục hậu quả của sự chiếu sáng không đồng đều.

- *Khử nhiễu*: Nhiễu được chia làm 2 loại cơ bản là nhiễu hệ thống và nhiễu ngẫu nhiên. Trong đó, nhiễu hệ thống là nhiễu có quy luật có thể khử bằng các phép biến đổi Fourier và loại bỏ các đỉnh điểm. Đối với nhiễu ngẫu nhiên – vết bẩn không rõ nguyên nhân thì được khắc phục bằng các phép lọc (lọc trung bình, lọc trung vị,...).

- *Hiệu chỉnh mức xám*: có thể tăng hay giảm số mức xám nhằm khắc phục tính không đồng bộ gây nên từ hiệu ứng của thiết bị thu nhận hình ảnh hoặc độ tương phản giữa các vùng ảnh.

- *Chuẩn hóa độ lớn, hình dạng và màu sắc*.

- *Nắn chỉnh hình học*: ảnh thu nhận thường bị biến dạng do các thiết bị quang học và điện tử, để khắc phục điều này người ta sử dụng các phép chiếu được xây dựng trên tập các điểm điều khiển.

### 1.1.2.3 Phân đoạn ảnh

Phân đoạn ảnh là một quá trình thao tác ở mức thấp trong toàn bộ hệ thống xử lý ảnh. Quá trình này thực hiện việc phân vùng ảnh thành các vùng rời rạc và đồng nhất với nhau hay nói cách khác là xác định các biên của các vùng ảnh đó. Phân đoạn ảnh là chia ảnh thành các vùng không trùng lắp, mỗi vùng gồm 1 nhóm pixel liên thông và đồng nhất theo 1 tiêu chí nào đó. Ví dụ: đồng

nhất về màu sắc, mức xám, kết cấu, độ sâu của các layer,... Sau khi phân đoạn mỗi pixel chỉ thuộc về một dùng duy nhất.

Để đánh giá chất lượng của quá trình phân đoạn là rất khó, vì vậy cần phải xác định rõ mục tiêu của quá trình phân đoạn là gì? Ví dụ: để nhận dạng chữ (mã vạch) trên phong bì thư cho mục đích phân loại bưu phẩm, cần chia các câu chữ về địa chỉ hoặc tên người gửi thành các từ, các chữ, các số (các vạch) riêng biệt để nhận dạng.

Kết quả của quá trình phân đoạn ảnh thường được cho dưới dạng dữ liệu điểm ảnh thô, trong đó hàm chứa biên của một vùng ảnh hoặc tập hợp tất cả các điểm ảnh thuộc về chính vùng ảnh đó. Trong cả hai trường hợp, sự chuyển đổi dữ liệu thô này thành một dạng thích hợp hơn cho việc xử lý trong máy tính là hết sức cần thiết, nghĩa là nên biểu diễn một vùng ảnh dưới dạng biên hay dưới dạng một vùng hoàn chỉnh gồm tất cả những điểm ảnh thuộc về nó.

– *Biểu diễn dạng biên* cho một vùng phù hợp với những ứng dụng chỉ quan tâm đến các đặc trưng hình dạng bên ngoài của đối tượng, ví dụ như các góc cạnh và điểm uốn trên biên.

– *Biểu diễn dạng vùng* lại thích hợp cho những ứng dụng khai thác các tính chất bên trong của đối tượng. Ví dụ như vân ảnh hoặc cấu trúc xương của nó. Và trong một số ứng dụng thì cả hai cách biểu diễn trên đều cần thiết.

#### 1.1.2.4 Trích chọn đặc điểm

Dựa trên các thông tin thu nhận được qua quá trình phân đoạn, kết hợp với các kỹ thuật xử lý, thủ tục phân tích dữ liệu để đưa ra các đặc điểm đặc trưng, đối tượng ảnh cũng như các thông tin cần thiết trong quá trình xử lý. Nhờ đó việc nhận dạng các đối tượng ảnh chính xác hơn, tốc độ tính toán cao và dung lượng nhớ lưu trữ giảm xuống.

Các đặc điểm của đối tượng được trích chọn tùy theo mục đích nhận dạng trong hệ thống xử lý ảnh. Sau đây là một vài đặc điểm của ảnh:

– *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. Các đặc điểm này có thể được trích chọn nhờ toán tử gradient, toán tử la bàn, toán tử Laplace, toán tử “chéo không” (zero crossing) v.v..

### 1.1.2.5 Nhận dạng

Nhận dạng ảnh là quá trình cuối cùng của hệ thống xử lý ảnh - quá trình liên quan đến các mô tả đối tượng mà người ta muốn đặc tả nó. Quá trình này thường đi sau quá trình trích chọn đặc điểm trong hệ thống xử lý ảnh.

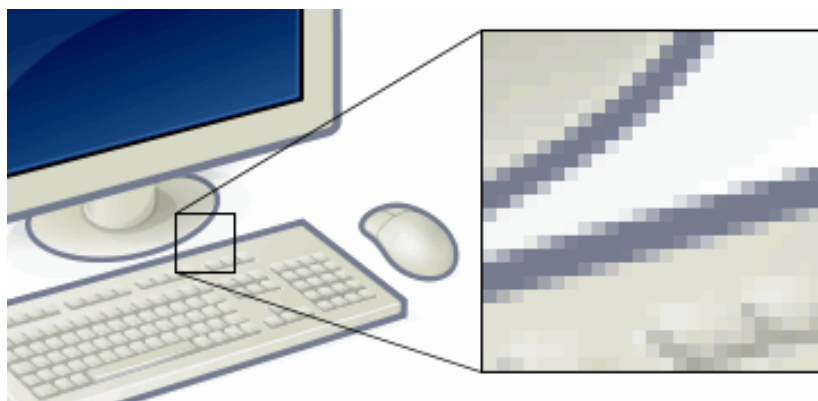
Có 2 kiểu nhận dạng ảnh cơ bản:

- Nhận dạng theo tham số (mô tả tham số).
- Nhận dạng theo cấu trúc (mô tả theo cấu trúc).

Hiện nay, người ta đã áp dụng kỹ thuật nhận dạng khá thành công với nhiều đối tượng khác nhau như: nhận dạng ảnh vân tay, khuôn mặt, nhận dạng chữ (chữ cái, chữ số, chữ có dấu), nhận dạng chữ in (đánh máy) phục vụ cho việc tự động hóa quá trình đọc tài liệu, tăng tốc độ và chất lượng nhận thông tin từ máy tính. Ngoài ra kỹ thuật nhận dạng dựa vào kỹ thuật mạng nơ ron đang được áp dụng và cho kết quả khả quan.

### 1.1.3 Một số vấn đề cơ bản trong xử lý ảnh

#### 1.1.3.1 Ảnh và điểm ảnh



Hình 1.3 Biểu diễn ảnh với điểm ảnh

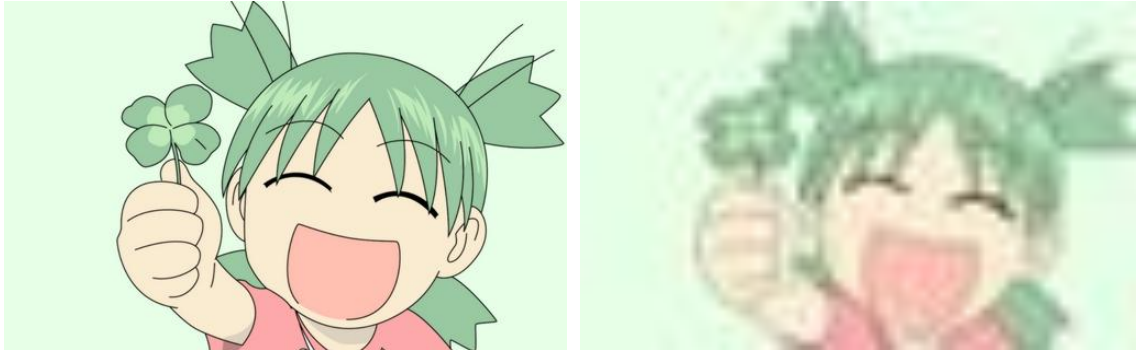
(Ví dụ này cho thấy một hình ảnh với một phần mở rộng rất nhiều, trong đó các điểm ảnh riêng lẻ được kết xuất như hình vuông nhỏ và có thể dễ dàng được nhìn thấy.)

Ảnh trong thực tế là ảnh liên tục về không gian và giá trị độ sáng. Để xử lý ảnh người ta phải tiến hành số hóa, quá trình số hóa là quá trình 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 (rời rạc hóa biên độ giá trị) mà về nguyên tắc mắt người không phân biệt được 2 điểm kề nhau. Trong quá trình này người ta sử dụng khái niệm PEL: *Picture Element* mà ta quen gọi (viết tắt) là Pixel – phần tử ảnh (điểm ảnh). Mỗi pixel gồm một cặp tọa độ (x, y) (trong khuôn khổ ảnh 2 chiều) với độ xám hoặc màu nhất định

Từ khái niệm điểm ảnh, ta có thể định nghĩa ảnh là tập hợp hữu hạn các điểm ảnh, thường được biểu diễn bằng một mảng hai chiều  $I(n,m)$  với n là số

hàng,  $m$  là số cột. Ta ký hiệu  $P(x,y)$  – 1 phần tử trong ma trận là một điểm ảnh tại vị trí  $(x,y)$ . Số lượng điểm ảnh trên mỗi hàng hoặc các hàng xác định độ phân giải của ảnh.

### 1.1.3.2 Độ phân giải của ảnh



a) Ảnh với độ phân giải 512 x 320

b) Ảnh với độ phân giải 64 x 40

Hình 1.4 Biểu diễn ảnh với độ phân giải

Độ phân giải (*Resolution*) của ảnh là mật độ điểm ảnh được ấn định trên một ảnh số được hiển thị.

Khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ phân bố, đó chính là độ phân giải và được phân bố theo trục  $x$  và  $y$  (512 và 320 ở hình a) trong không gian hai chiều.

### 1.1.3.3 Mức xám của ảnh và phân loại ảnh

Mỗi Pixel (điểm ảnh) có hai đặc trưng cơ bản là vị trí  $(x, y)$  của điểm ảnh và mức xám của nó. Mức xám (Gray level) là kết quả của sự mã hóa thương ứng một cường độ sáng của mỗi điểm ảnh với một giá trị số (giá trị nguyên dương) - kết quả của quá trình lượng hóa.

Các thang giá trị mức xám thường dùng là 16, 32, 64, 128 hay 256 mức. Mã hóa 256 mức là phổ dụng nhất do lý do kỹ thuật. Vì  $2^8 = 256$  (tức là từ 0  $\rightarrow$  255) nên với 256 mức, mỗi pixel sẽ được mã hóa bởi 1 byte (8 bit).

Từ định nghĩa mức xám ta phân loại ảnh theo giá trị mức xám của nó:

\* **Ảnh nhị phân:** Giá trị mức xám của tất cả điểm ảnh chỉ nhận giá trị 1 hoặc 0:

- 1: biểu diễn đối tượng ảnh và được gọi là điểm đen.
- 0: biểu diễn ảnh nền (phông ảnh), được gọi là điểm trắng

Như vậy mỗi điểm ảnh trong ảnh nhị phân được biểu diễn bởi 1 bit.

\* **Ảnh xám:** Giá trị nằm trong khoảng từ 0  $\rightarrow$  255, như vậy mỗi điểm ảnh trong ảnh xám được biểu diễn bởi 1 byte. Ảnh có nhiều mức xám được gọi là ảnh đa cấp xám. Ta có thể chuyển đổi từ ảnh đa mức xám về ảnh nhị phân theo công thức:

$$Y(m, n) = \begin{cases} 1 & \text{nếu } X(m, n) \geq \theta \\ 0 & \text{nếu } X(m, n) < \theta \end{cases} \quad \text{với } \theta \text{ là ngưỡng tự chọn}$$

\* **Ảnh màu:** Ảnh màu theo lý thuyết của Thomas là ảnh tổ hợp từ 3 màu cơ bản được thu nhận trên các dải băng tần khác nhau:

+ Đỏ – RED (R)

+ Lục – GREEN (G)

+ Lam – BLUE (B)

Với ảnh màu, cách biểu diễn cũng tương tự như với ảnh đen trắng, chỉ khác là tại mỗi phần tử của ma trận biểu diễn cho ba màu riêng rẽ gồm: đỏ, lục và lam.

Để lưu trữ ảnh màu người ta có thể lưu trữ từng màu riêng biệt, mỗi màu lưu trữ như một ảnh đa cấp xám, mỗi mức xám của ảnh sẽ được biểu diễn bởi 3 thành phần: R, G, B (mỗi thành phần được biểu diễn bởi 1 byte). Do đó, không gian nhớ dành cho một ảnh màu lớn gấp 3 lần một ảnh đa cấp xám cùng kích cỡ.

Ta có thể chuyển đổi ảnh màu về ảnh đa mức xám theo công thức:

$$G = 0.299R + 0.587G + 0.114B$$

(ảnh hưởng của màu đến mức xám giảm dần từ R đến B)

hoặc:

$$G = 0.333R + 0.333G + 0.333B$$

(Coi ảnh hưởng của R, G, B là như nhau)

Như vậy để xử lý ảnh màu ta phải xử lý trên 3 ảnh xám R, G, B của nó và kết hợp lại sẽ có kết quả như mong muốn.

#### 1.1.3.4 Quan hệ giữa các điểm ảnh

##### a, Các lân cận của điểm ảnh (Image Neighbors)

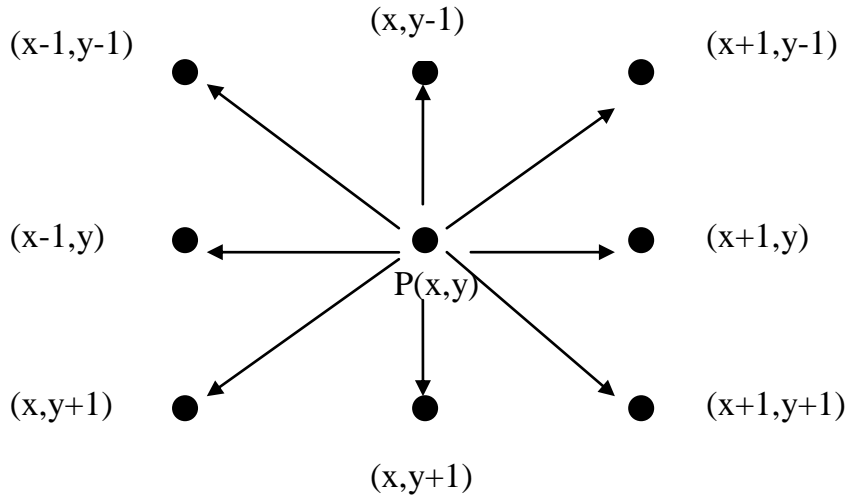
Giả sử ta có điểm ảnh P tại tọa độ (x,y), khi đó P(x,y) có 2 điểm lân cận đứng và 2 điểm lân cận ngang: (x+1,y), (x-1,y), (x,y+1), (x,y-1). Một tập các điểm như vậy được gọi là tập 4 điểm lân cận của P và được ký hiệu là  $N_4(P)$

$$\{(x+1,y); (x-1,y); (x,y+1); (x,y-1)\} = N_4(P)$$

Trong đó 1 là giá trị logic.

Ngoài ra điểm ảnh P còn có các lân cận chéo  $N_D(P)$

$$\{(x+1, y+1); (x+1, y-1); (x-1, y+1); (x-1, y-1)\} = N_D(P)$$

Hình 1.5: Lân cận các điểm ảnh của  $P(x,y)$ 

→ ta có tập kết hợp  $N_8(P) = N_4(P) + N_D(P)$  là tập hợp 8 lân cận của điểm ảnh  $P$ .

Chú ý: Nếu  $P(x,y)$  nằm ở biên (mép) ảnh, một số điểm lân cận sẽ nằm ở ngoài ảnh.

### b, Các mối liên kết điểm ảnh

Các mối liên kết được sử dụng để xác định giới hạn (Boundaries) của đối tượng vật thể hoặc xác định vùng trong một ảnh. Một liên kết được đặc trưng bởi tính liên kết giữa các điểm và mức xám của chúng.

Giả sử  $V$  là một tập các giá trị mức xám. Một ảnh có các giá trị cường độ sáng từ thang mức xám từ 32 đến 64 được mô tả như sau :

$$V = \{32, 33, \dots, 63, 64\}.$$

Có 3 loại liên kết:

– Liên kết 4: Hai điểm ảnh  $p$  và  $q$  được nói là liên kết 4 với các giá trị cường độ sáng  $V$  nếu  $q$  nằm trong một các lân cận của  $p$ , tức  $q$  thuộc  $N_4(p)$ .

– Liên kết 8: Hai điểm ảnh  $p$  và  $q$  nằm trong một các lân cận 8 của  $p$ , tức  $q$  thuộc  $N_8(p)$ .

– Liên kết  $m$  (liên kết hỗn hợp): Hai điểm ảnh  $p$  và  $q$  với các giá trị cường độ sáng  $V$  được nói là liên kết  $m$  khi và chỉ khi:

+  $q$  thuộc  $N_4(p)$  hoặc

+  $q$  thuộc  $N_D(p)$  và  $\{N_4(p) \cap N_4(q)\} = \emptyset$

Ví dụ: cho tập  $V = \{1, 2\}$

0	1	1
0	2	0
0	0	1

Liên kết 4

0	1	1
0	2	0
0	0	1

Liên kết 8

0	1	1
0	2	0
0	0	1

Liên kết m

### c, Đo khoảng cách giữa các điểm ảnh

Cho các điểm ảnh  $p$ ,  $q$  và  $z$  với tọa độ  $(x, y)$ ,  $(s, t)$ ,  $(u, v)$  tương ứng, gọi  $D$  là hàm khoảng cách giữa hai điểm ảnh  $p$  và  $q$ . Khi đó hàm khoảng cách  $D$  (Distance) có tính chất sau:

1.  $D(p, q) \geq 0$  (Với  $D(p, q) = 0$  khi và chỉ khi  $p = q$ )
2.  $D(p, q) = D(q, p)$
3.  $D(p, z) \leq D(p, q) + D(q, z)$

Ngoài ra còn có các biện pháp đo khoảng cách giữa các điểm ảnh khác:

\* Khoảng cách O-clit (Euclidean): khoảng cách O-clit giữa 2 điểm ảnh  $p$  và  $q$  được định nghĩa như sau:

$$D_e(p, q) = [(x - s)^2 + (y - t)^2]^{1/2}$$

\* Khoảng cách khối: khoảng cách  $D_4(p, q)$  được gọi là khoảng cách khối đồ thị và được xác định như sau:

$$D_4(p, q) = |x - s| + |y - t|$$

\* Khoảng cách bàn cờ  $D_8(p, q)$ : là khoảng cách được xác định như sau:

$$D_8(p, q) = \max(|x - s|, |y - t|)$$

### 1.2 Nâng cao chất lượng ảnh và toán tử không gian

Thông thường ảnh thu nhận có nhiều cần phải loại bỏ nhiễu hay ảnh không sắc nét bị mờ hoặc cần làm rõ các chi tiết như các đường biên ảnh. Các toán tử không gian dùng trong kỹ thuật tăng cường ảnh được phân nhóm theo công dụng: làm trơn nhiễu, nổi biên. Để làm trơn nhiễu hay tách nhiễu, người ta sử dụng các bộ lọc tuyến tính (lọc trung bình, thông thấp) hay lọc phi tuyến (trung vị, giả trung vị, lọc đồng hình). Từ bản chất của nhiễu (thường tương ứng với tần số cao) và từ cơ sở lý thuyết lọc là: bộ lọc chỉ cho tín hiệu có tần số nào đó thông qua do đó, để lọc nhiễu người ta thường dùng lọc thông thấp (theo quan điểm tần số không gian) hay lấy tổ hợp tuyến tính để san bằng (lọc trung bình). Để làm nổi biên (ứng với tần số cao), người ta dùng các bộ lọc thông cao, lọc Laplace..

Để hiểu rõ hơn các kỹ thuật áp dụng, cần phải phân biệt các loại nhiễu can thiệp trong quá trình xử lý ảnh. Trên thực tế tồn tại khá nhiều loại nhiễu như sự thay đổi độ nhạy của cảm biến, sự biến đổi của môi trường, sai số của quá trình lượng tử hóa, sai số của kênh truyền...; tuy nhiên người ta thường xem xét 3 loại nhiễu chính và phổ biến là: nhiễu cộng, nhiễu nhân và nhiễu xung:

– *Nhiễu cộng* (Additive noise): thường phân bố khắp ảnh và được biểu diễn bởi:

$$Y = X + n$$

– *Nhiễu nhân*: cũng thường phân bố khắp ảnh và được biểu diễn bởi:

$$Y = X * n$$

Chú ý: với Y: ảnh quan sát, X: ảnh gốc và n là nhiễu.

– *Nhiễu xung* (Impulse noise): là một loại nhiễu khá đặc biệt có thể sinh ra bởi nhiều lý do khác nhau chẳng hạn: lỗi truyền tín hiệu, lỗi bộ nhớ, hay lỗi định thời trong quá trình lượng tử hóa. Nhiễu này thường gây đột biến tại một số điểm ảnh.

### 1.2.1 Làm trơn nhiễu bằng lọc tuyến tính

Do có nhiều loại nhiễu can thiệp vào quá trình xử lý ảnh nên cần có nhiều bộ lọc thích hợp. Với nhiễu cộng và nhiễu nhân ta dùng các bộ lọc thông thấp, trung bình và lọc đồng hình (Homomorphie), với nhiễu xung ta dùng lọc trung vị, giả trung vị, lọc ngoại (Outlier).

#### 1.2.1.1 Lọc trung bình không gian

Với lọc trung bình, mỗi điểm ảnh được thay thế bằng trung bình trọng số các điểm lân cận và được định nghĩa như sau:

$$v(m,n) = \sum_{(k,l) \in w} a(k,l) y(m-k, n-l)$$

Nếu trong kỹ thuật lọc trên, ta chọn các trọng số bằng nhau, phương trình trên sẽ có dạng:

$$v(m,n) = \frac{1}{N} \sum_{(k,l) \in w} a(k,l) y(m-k, n-l)$$

Với:  $y(m,n)$ : ảnh đầu vào

$v(m,n)$ : ảnh đầu ra

$a(k,l)$ : là trọng số lọc

$a_{k,l} = \frac{1}{N}$  và  $N_w$  là số điểm ảnh trong cửa sổ lọc W

Lọc trung bình có trọng số chính là thực hiện chập ảnh đầu vào với nhân chập H. Nhân chập H trong trường hợp này có dạng:



$$H = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Trong lọc trung bình, thường người ta ưu tiên cho các hướng để bảo vệ biên của ảnh khỏi bị mờ khi làm trơn ảnh. Các kiểu mặt nạ được sử dụng tùy theo các trường hợp khác nhau. Các bộ lọc trên là bộ lọc tuyến tính theo nghĩa là điểm ảnh ở tâm cửa sổ sẽ được thay bởi tổ hợp các điểm lân cận chập với mặt nạ.

Giả sử ảnh đầu vào biểu diễn bởi ma trận:

$$I = \begin{bmatrix} 4 & 7 & 3 & 7 & 1 \\ 5 & 7 & 1 & 7 & 1 \\ 6 & 6 & 1 & 8 & 3 \\ 5 & 7 & 5 & 7 & 1 \\ 5 & 7 & 6 & 1 & 2 \end{bmatrix}$$

Ảnh số thu được bởi lọc trung bình  $Y = H \otimes I$  có dạng:

$$Y = \frac{1}{9} \begin{bmatrix} 23 & 26 & 31 & 19 & 16 \\ 35 & 39 & 46 & 31 & 27 \\ 36 & 43 & 49 & 34 & 27 \\ 36 & 48 & 48 & 34 & 22 \\ 24 & 35 & 33 & 22 & 11 \end{bmatrix}$$

Lọc trung bình trọng số là một trường hợp riêng của lọc thông thấp.

### 1.2.1.2 Lọc thông thấp

Lọc thông thấp thường được sử dụng để làm trơn nhiều. Trong kỹ thuật này ta sử dụng một số nhân chập có dạng sau:

$$H_{tl} = \frac{1}{8} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$H_b = \frac{1}{(b+2)^2} * \begin{bmatrix} 1 & b & 1 \\ b & b^2 & b \\ 1 & b & 1 \end{bmatrix}$$

Ta dễ dàng nhận thấy khi  $b = 1$ ,  $H_b$  chính là  $H_{tl}$  (lọc trung bình). Để hiểu rõ hơn bản chất khử nhiễu cộng của các bộ lọc này, ta viết phương trình thu nhận ảnh dưới dạng:

$$X_{qs} [m,n] = X_{gốc} [m,n] + \eta[m,n]$$

trong đó  $\eta[m,n]$  là nhiễu cộng có phương sai  $\sigma_n^2$ . Như vậy, theo cách tính lọc trung bình ta có:

$$Y[m,n] = \frac{1}{N_w} \sum_{k,l \in w} X_{qs}(m-k, n-l) + \eta[m,n]$$

$$\text{hay } Y[m,n] = \frac{1}{N_w} \sum_{k,l \in w} X_{qs}(m-k, n-l) + \frac{\sigma_n^2}{N_w}$$

Như vậy, nhiễu cộng trong ảnh đã giảm đi  $N_w$  lần.

### 1.2.1.3 Lọc đồng hình (Homomorphie Filter)

Kỹ thuật lọc này hiệu quả với ảnh có nhiễu nhân. Thực tế, ảnh quan sát được gồm ảnh gốc nhân với hệ số nhiễu. Gọi  $\bar{X}(m,n)$  là ảnh thu được,  $X(m,n)$  là ảnh gốc và  $\eta(m,n)$  là nhiễu, ta có:

$$X(m,n) = \bar{X}(m,n) * \eta(m,n)$$

Lọc đồng hình thực hiện lấy Logarit của ảnh quan sát. Do vậy ta có kết quả sau:

$$\text{Log}(X(m, n)) = \log(\bar{X}(m,n)) + \log(\eta(m,n))$$

Rõ ràng, nhiễu nhân có trong ảnh sẽ bị giảm. Sau quá trình lọc tuyến tính, ta chuyển về ảnh cũ bằng phép biến đổi hàm e mũ.

### 1.2.2 Làm trơn nhiễu bằng lọc phi tuyến

Các bộ lọc phi tuyến cũng hay được dùng trong kỹ thuật tăng cường ảnh. Trong kỹ thuật này, người ta dùng bộ lọc trung vị, giả trung vị, lọc ngoài. Với lọc trung vị, điểm ảnh đầu vào sẽ được thay thế bởi trung vị các điểm ảnh, còn lọc giả trung vị sẽ dùng trung bình cộng của 2 giá trị “trung vị” (trung bình cộng của max và min).

#### 1.2.2.1 Lọc trung vị

Khái niệm trung vị được viết bởi công thức:

$$v(m,n) = \text{Trungvi}(y(m-k, n-l)) \quad \text{với } (k, l) \text{ thuộc } W$$

Kỹ thuật này đòi hỏi giá trị các điểm ảnh trong cửa sổ phải xếp theo thứ tự tăng hay giảm dần so với giá trị trung vị. Kích thước cửa sổ thường được chọn sao cho số điểm ảnh trong cửa sổ là lẻ. Các cửa sổ hay dùng là cửa sổ có kích thước 3x3, hay 5x5 hay 7x7.

*Ví dụ:* Nếu  $y(m) = \{2, 3, 8, 4, 2\}$  và cửa sổ  $W = (-1, 0, 1)$  thì ảnh kết quả thu được sau lọc trung vị là  $v(m) = \{2, 3, 4, 4, 2\}$ .

do đó:

$$\begin{aligned} v[0] &= 2 <\text{giá trị biên}> & v[3] &= \text{Trungvi}(8, 4, 2) = 4 \\ v[1] &= \text{Trungvi}(2, 3, 8) = 3 & v[4] &= 2 <\text{giá trị biên}> \\ v[2] &= \text{Trungvi}(3, 8, 4) = 4 \end{aligned}$$

\* *Tính chất của lọc trung vị:*

– Lọc trung vị là loại lọc phi tuyến. Điều này được thể hiện:

$$\text{Trungvi}(x(m) + y(m)) \neq \text{Trungvi}(x(m)) + \text{Trungvi}(y(m)).$$

– Có lợi cho việc loại bỏ các điểm ảnh hay các hàng mà vẫn bảo toàn bộ phân giải.

– Hiệu quả giảm khi số điểm trong cửa sổ lớn hay bằng một nửa số điểm trong cửa sổ. Điều này dễ giải thích vì trung vị là  $(N_w+1)/2$  giá trị lớn nhất nếu  $N_w$  lẻ. Lọc trung vị cho trường hợp 2 chiều coi như lọc trung vị tách được theo từng chiều.

### 1.2.2.2 Lọc ngoài (Outlier Filter)

Giả thiết có ngưỡng nào đó cho các mức nhiễu (có thể dựa vào lược đồ xám). Tiến hành so sánh giá trị độ xám của một điểm ảnh với trung bình số học 8 lân cận của nó. Nếu sai lệch lớn hơn ngưỡng, điểm ảnh này được coi như nhiễu. Trong trường hợp đó, thay thế giá trị của điểm ảnh bằng giá trị trung bình 8 lân cận vừa tính được. Bộ lọc ngoài có thể diễn tả bằng công thức sau:

$$Y(m,n) = \begin{cases} \alpha(w) & \text{khi } |u(m,n) - \alpha(w)| > \delta \\ u(m,n) & \text{không} \end{cases}$$

với  $\alpha(w)$  là trung bình cộng các điểm trong lân cận  $w$ ;  $\delta$  là ngưỡng ngoài.

Các cửa sổ tính toán thường chọn là  $3 \times 3$ . Tuy nhiên, cửa sổ có thể mở rộng đến  $5 \times 5$  hay  $7 \times 7$  để đảm bảo tính tương quan giữa các điểm ảnh. Vấn đề quan trọng là xác định ngưỡng để loại nhiễu mà vẫn không làm mất thông tin của ảnh.

### 1.2.3 Lọc thông thấp, thông cao và lọc giải thông

Toán tử trung bình không gian là lọc thông thấp. Nếu  $H_{LP}(m,n)$  biểu diễn bộ lọc thông thấp thì bộ lọc thông cao  $H_{HP}(m,n)$  có thể được định nghĩa:

$$H_{HP}(m,n) = \delta(m,n) - H_{LP}(m,n)$$

Và bộ lọc giải thông được định nghĩa:

$$H_{HP}(m,n) = H_{L1}(m,n) - H_{L2}(m,n)$$

với  $H_{L1}$  và  $H_{L2}$  là các bộ lọc thông thấp.

Bộ lọc thông thấp thường dùng làm trơn nhiễu và nội suy ảnh. Bộ lọc thông cao dùng nhiễu trong trích chọn biên và làm trơn ảnh, còn bộ lọc giải thông có hiệu quả làm nổi biên. Về biên sẽ được trình bày kỹ trong các phần sau. Tuy nhiên, dễ nhận thấy, biên là điểm có độ biến thiên nhanh về giá trị mức xám theo quan điểm về tần số tín hiệu. Như vậy, các điểm biên ứng với các thành phần tần số cao. Từ đó, có thể dùng bộ lọc thông cao để cải thiện nhiễu nghĩa là có thể lọc các thành phần tần số thấp và giữ lại các thành phần tần số cao. Vì thế, lọc thông cao thường được dùng làm trơn biên trước khi tiến hành các thao tác với biên ảnh. Dưới đây là một số mặt nạ dùng trong lọc thông cao:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

(1)

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

(2)

$$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

(3)

Các nhân chập thông cao có đặc tính chung là tổng các hệ số của bộ lọc bằng 1. Nguyên nhân chính là ngăn cản sự tăng quá giới hạn của các giá trị mức xám (các giá trị điểm ảnh vẫn giữ được giá trị của nó một cách gần đúng không thay đổi quá nhiều với giá trị thực).

### 1.3 Tổng quan về biên

#### 1.3.1 Biên và các kiểu biên cơ bản

##### 1.3.1.1 Một số khái niệm về biên

Cho tới nay chưa có định nghĩa chính xác về biên và mỗi định nghĩa được sử dụng trong một số trường hợp nhất định. Biên có thể được tạo ra bởi bóng tối, kết cấu hình học... Biên cũng có thể được định nghĩa là không liên tục ở cường độ hình ảnh do sự thay đổi trong cấu trúc hình ảnh. Biên trong một hình ảnh thường xảy ra với độ phân giải hoặc quy mô khác nhau và đại diện cho quá trình chuyển đổi của mức xám khác nhau, hay mức độ gradient. Tuy nhiên, nhìn chung biên có thể được định nghĩa như sau:

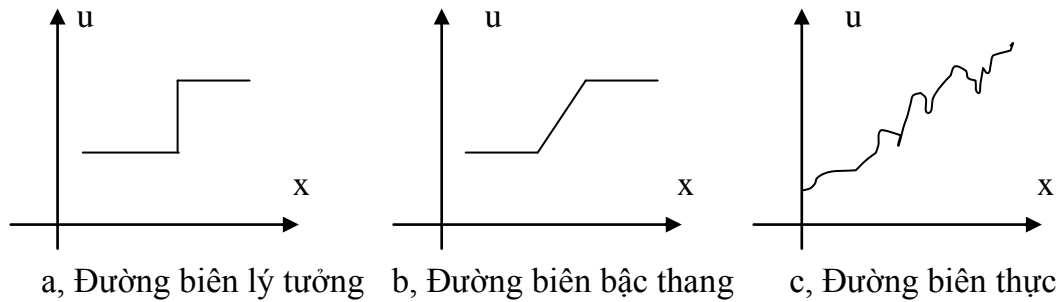
*Điểm biên:* một điểm ảnh được coi là điểm biên nếu có sự thay đổi đột ngột về mức xám. Ví dụ: đối với ảnh đen trắng, một điểm được gọi là điểm biên nếu nó là điểm đen có ít nhất một điểm trắng bên cạnh.

*Đường biên (đường bao của ảnh - boundary) của đối tượng:* được tạo thành bởi một tập các điểm biên liên tiếp.

Mỗi một biên là một thuộc tính gắn liền với một điểm riêng biệt và được tính toán từ những điểm lân cận nó. Đó là một biến Vector bao gồm 2 thành phần:

- Độ lớn của Gradient.
- Hướng của biên với góc  $\phi$ , lệch so với hướng của Gradient  $\psi$  một góc  $-90^\circ$ .

*Mô hình biểu diễn đường biên:* theo toán học, điểm ảnh có sự biến đổi mức xám  $u(x)$  một cách đột ngột theo hình dưới:



Hình 1.6: Đường bao của ảnh

### 1.3.1.2 Các kiểu biên cơ bản

#### a, Biên lý tưởng

Việc phát hiện biên một cách lý tưởng là việc xác định được tất cả các đường bao trong đối tượng. Biên là sự thay đổi đột ngột về mức xám nên sự thay đổi này càng lớn thì càng dễ dàng nhận ra biên.

Một biên được coi là biên lý tưởng khi có sự thay đổi cấp xám lớn giữa các vùng trong ảnh. Biên này thường chỉ xuất hiện khi có sự thay đổi cấp xám qua một điểm ảnh.

#### b, Biên bậc thang (biên dốc)

Biên dốc xuất hiện khi sự thay đổi cấp xám trải rộng qua nhiều điểm ảnh. Vị trí của biên được xem như vị trí chính giữa của đường dốc nối giữa cấp xám thấp và cấp xám cao. Tuy nhiên đây chỉ là đường dốc trong toán học, từ khi ảnh được kỹ thuật số hóa thì đường dốc không còn là đường thẳng mà thành những đường lờm chờm, không trơn.

#### c, Biên thực

Trên thực tế, ảnh thường có biên không lý tưởng, có thể do các nguyên nhân sau:

- Hình dạng không sắc nét.
- Nhiễu: kết quả của nhiễu trên ảnh gây ra một sự biến thiên ngẫu nhiên giữa các điểm ảnh. Sự xuất hiện ngẫu nhiên của các điểm ảnh có mức xám chênh lệch cao làm cho các đường biên dốc trở lên không trơn chu mà trở thành các đường biên gồ ghề, mập mờ, không hẳn, đây chính là đường biên trên thực tế.

### 1.3.2 Vai trò của biên trong nhận dạng

Đường biên là một loại đặc trưng cục bộ tiêu biểu trong phân tích nhận dạng ảnh. Người ta sử dụng đường biên làm phân cách các vùng xám (màu) cách biệt. Ngược lại, người ta cũng dùng các vùng ảnh để tìm đường phân cách.

Như đã đề cập tới ở phần tổng quan về một hệ thống nhận dạng và xử lý ảnh, quá trình nhận dạng có hai giai đoạn cần thực hiện:

– Giai đoạn học: Các đặc điểm của đối tượng mẫu được lưu trữ (gọi là học mẫu) và tập các phần tử mẫu được chia thành các lớp.

– Giai đoạn nhận dạng: Khi có đối tượng cần nhận dạng, các đặc điểm của đối tượng sẽ được trích chọn và sử dụng hàm quyết định để xác định đối tượng cần nhận dạng thuộc lớp nào.

## CHƯƠNG II: PHƯƠNG PHÁP PHÁT HIỆN BIÊN CHO ẢNH ĐA MỨC XÁM

### 2.1 Giới thiệu

Một biên trong một hình ảnh nói chung có thể được định nghĩa như là đường bao của ảnh hoặc đường viền ngăn cách khu vực ảnh liền kề có đặc điểm tương đối khác biệt theo một số tính năng quan tâm, một trong số đó là độ đo về sự thay đổi đột ngột về cấp xám. Biên có thông tin quan trọng góp phần hướng tới phân tích và giải thích thông tin hình ảnh (nhận dạng). Các giai đoạn của phát hiện biên là làm mịn hay nâng cao chất lượng ảnh (quá trình loại bỏ nhiễu) và phát hiện biên.

Có nhiều phương pháp để phát hiện biên cơ bản, nhưng hầu hết trong số đó có thể chia thành hai loại:

\* **Phương pháp phát hiện biên trực tiếp:** tìm biên dựa vào sự biến thiên về giá trị độ sáng (cấp xám) của ảnh. Chủ yếu dựa vào kỹ thuật lấy đạo hàm. Ở phương pháp này ta có thể nhóm thành 2 loại:

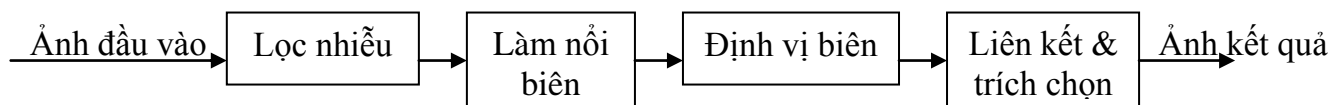
– Phương pháp Gradient: chúng ta tính toán ước lượng độ lớn gradient bằng cách sử dụng bộ lọc làm mịn và sử dụng dự đoán tính toán để xác định vị trí của biên. Nói cách khác phương pháp Gradient là phương pháp dò biên cục bộ bằng cách tìm kiếm cực đại và cực tiểu khi lấy đạo hàm bậc nhất của ảnh.

– Phương pháp Laplace: chúng ta lấy đạo hàm bậc hai của các tín hiệu và biên độ đạo hàm là cực đại khi đạo hàm bậc hai bằng 0. Trong ngắn hạn, phương pháp Laplace tìm kiếm toán tử chéo không tại đạo hàm bậc hai của ảnh để tìm biên.

\* **Phương pháp 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.

Ngoài ra còn các phương pháp nâng cao khác như phương pháp Canny, Wavelet..

### 2.2 Quy trình phát hiện biên



Hình 2.1: Quy trình phát hiện biên

*Bước 1: Lọc nhiễu do ảnh ghi vào thường có nhiễu*

*Bước 2: Làm nổi biên sử dụng các toán tử phát hiện biên*

*Bước 3: Định vị biên (cần loại bỏ biên giả)*

*Bước 4: Liên kết và trích chọn biên*

## 2.3 Phương pháp phát hiện biên cơ bản

### 2.3.1 Phương pháp phát hiện biên Gradient

Phương pháp Gradient là phương pháp dò biên cục bộ bằng cách tìm kiếm cực đại và cực tiểu khi lấy đạo hàm bậc nhất của ảnh trong không gian hai chiều.

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

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

Trong đó  $dx, dy$  là khoảng cách giữa 2 điểm lân cận theo hướng  $x$  và  $y$  tương ứng. Trong những hình ảnh rời rạc, ta có thể xem xét  $dx, dy$  về số lượng điểm ảnh giữa hai điểm.

Thực tế ta chọn  $dy = dx = 1$  (khoảng cách điểm ảnh) là điểm mà tại đó tọa độ điểm ảnh là  $(i, j)$ , do đó:

$$\Delta x = f(i+1, j) - f(i, j)$$

$$\Delta y = f(i, j+1) - f(i, j)$$

Do tính chất phức tạp trong tính toán khi áp dụng phương pháp Gradient trong xử lý ảnh, người ta sử dụng kỹ thuật Gradient dùng cặp mặt nạ  $H_1, H_2$  trực giao (theo 2 hướng vuông góc). Nếu định nghĩa  $g_1, g_2$  là Gradient theo hai hướng  $x, y$  tương ứng thì biên độ  $g(m, n)$  tại điểm  $(m, n)$  được tính:

$$g(m, n) = \sqrt{g_1^2(m, n) + g_2^2(m, n)} = A_0$$

$$\theta_r(m, n) = \arctan(g_2(m, n) / g_1(m, n))$$

Để giảm độ phức tạp tính toán,  $A_0$  được tính gần đúng như sau:

$$A_0 = |g_1(m, n)| + |g_2(m, n)|$$

Việc xấp xỉ đạo hàm bậc nhất theo các hướng  $x$  và  $y$  được thực hiện thông qua 2 mặt nạ nhân chập tương ứng sẽ cho ta các kỹ thuật phát hiện biên khác nhau (Roberts, Sobel, Prewitt,...).



\* Thuật toán:

– Với mỗi điểm ảnh  $I(x,y)$  tính:

$$\text{Grad}(I(x,y)) = \sqrt{(I(x,y) * H_1)^2 + (I(x,y) * H_2)^2}$$

– Phân ngưỡng:

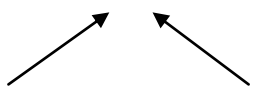
$$I(x,y) = \begin{cases} 1 & \text{if } I(x,y) \geq \theta \\ 0 & \text{if } I(x,y) < \theta \end{cases}$$

### 2.3.1.1 Toán tử Robert – Phương pháp phát hiện biên Robert

Lý do chính để sử dụng toán tử Robert là nó rất tính toán rất nhanh. Chỉ có bốn điểm ảnh đầu vào cần phải được kiểm tra để xác định giá trị của mỗi điểm ảnh đầu ra, và chỉ phép trừ và phép cộng được sử dụng trong tính toán. Trong phép cộng không có tập tham số nào.

Nhược điểm chính của toán tử Robert là sử dụng một mặt nạ nhỏ, và rất nhạy cảm với nhiễu. Nó cũng tạo ra phản hồi rất yếu với biên thực, trừ khi các biên này rất sắc nét.

Toán tử Robert bao gồm một cặp mặt nạ nhân chập  $2 \times 2$  là  $H_x$  và  $H_y$ , những mặt nạ này được thiết kế có thể đáp ứng tối đa để biên hoạt động theo hai hướng  $-45^\circ$  và  $+45^\circ$ :

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


Hướng của gradient

Chiều dài đường biên ảnh có thể rút ra bằng cách dùng bất kỳ phép xử lý không tuyến tính nào sau đây:

$$f(i,j) = \sqrt{f_1(i,j)^2 + f_2(i,j)^2}$$

$$f(i,j) = \max \left\{ |f_1(i,j)|, |f_2(i,j)| \right\}$$

$$f(i,j) = |f_1(i,j)| + |f_2(i,j)|$$

Trong đó:  $f_1(i,j)$  và  $f_2(i,j)$  là đáp ứng rút ra từ mẫu  $G_x$  và  $G_y$

Hướng của đường biên  $\theta(i,j)$  tính theo phương nằm ngang, có thể rút ra bởi:

$$\theta(i,j) = \frac{\pi}{4} \tan^{-1} \frac{f_2(i,j)}{f_1(i,j)}$$

Các phần tử trong mặt nạ gọi là các trọng số. Di chuyển lần lượt các mặt nạ trên ảnh đang xét sao cho phần tử đầu tiên của mặt nạ trùng với phần tử  $(i,j)$  đang xét trên ảnh. Từ đó cho ra kết quả của ảnh mới theo phương pháp này.

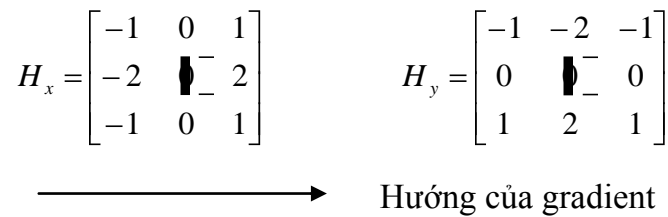
Như vậy, ta có kết quả Gradient tại một điểm ảnh  $(i,j)$ :

$$f(i, j) = |f(i, j+1) - f(i+1, j)| + |f(i, j) + f(i+1, j+1)|$$

### 2.3.1.2 Toán tử Sobel - Phương pháp phát hiện biên Sobel

Ở đây chúng ta sử dụng hai mặt nạ nhân chập  $3 \times 3$  theo hướng x, y với hình ảnh đa mức xám.

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



\* Thuật toán:

+ Bước 1: tính  $I(x,y) \otimes H_x = I_1$

và  $I(x,y) \otimes H_y = I_2$

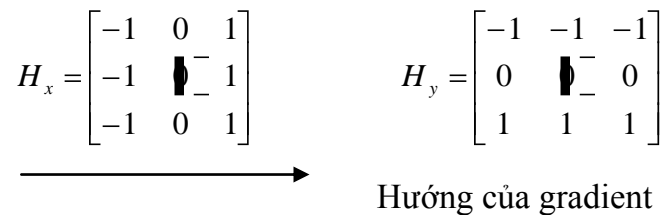
+ Bước 2: Tính  $|I_1| + |I_2| = I_s$

+ Bước 3: Hiệu chỉnh  $I(x,y) = I_s \geq \theta ? 1 : 0$

### 2.3.1.3 Toán tử Prewitt - Phương pháp phát hiện biên Prewitt

Toán tử Prewitt làm việc giống với cách làm của toán tử Sobel, nhưng sử dụng mặt nạ nhân chập khác so với toán tử Sobel. Mặt nạ này cho kết quả giống như Sobel và cùng hướng x, y:

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



\* Thuật toán: Các bước tính toán tương tự như Sobel

### 2.3.2 Phương pháp phát hiện biên Laplace

Để khắc phục hạn chế và nhược điểm của phương pháp Gradient, trong đó sử dụng đạo hàm riêng bậc nhất người ta nghĩ đến việc sử dụng đạo hàm riêng bậc hai hay toán tử Laplace. Phương pháp dò biên theo toán tử Laplace hiệu quả hơn phương pháp toán tử Gradient trong trường hợp mức xám biến đổi chậm, miền chuyển đổi mức xám có độ trải rộng.

Toán tử Laplace được định nghĩa như sau:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Toán tử Laplace dùng một số kiểu mặt nạ khác nhau nhằm tính gần đúng đạo hàm riêng bậc hai. Các dạng mặt nạ theo toán tử Laplace bậc 3x3 hay dùng:

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

\* Ghi chú: Mặt nạ  $H_1$  còn cải biên bằng việc lấy giá trị ở tâm bằng 8 thay vì giá trị 4.

Để thấy rõ việc xấp xỉ đạo hàm riêng bậc 2 trong không gian 2 chiều với mặt nạ  $H_1$  làm ví dụ, ta có thể tính gần đúng như sau:

$$\frac{\partial^2 f}{\partial x^2} = 2f(x, y) - f(x-1, y) - f(x+1, y)$$

$$\frac{\partial^2 f}{\partial y^2} = 2f(x, y) - f(x, y-1) - f(x, y+1)$$

$$\rightarrow \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 4f(x, y) - f(x-1, y) - f(x, y-1) - f(x+1, y) - f(x, y+1)$$

Kỹ thuật theo toán tử Laplace tạo đường biên mảnh (có độ rộng 1 pixel). Nhược điểm của kỹ thuật này rất nhạy với nhiễu, do vậy đường biên thu được thường kém ổn định. Để khắc phục nhược điểm này người ta mở rộng toán tử Laplace – dùng hàm Gauss (Laplace of Gauss) để giảm nhiễu cho ảnh (làm trơn ảnh).

#### \* Zero – crossing và toán tử Laplace:

Biên được nhận dạng bằng việc tìm các giao điểm không (chéo – không) trong đạo hàm bậc hai của toán tử Laplace.

Bộ dò “chéo-không” tìm những nơi mà giá trị chỉ toán tử Laplace đi qua điểm 0 trong toán tử Laplace của một ảnh.

#### \* Phương pháp Laplace of Gauss (LoG):

Phương pháp này bao gồm 2 bước:

– Bước 1: Làm trơn ảnh với bộ lọc Gauss (lọc nhiễu). (Nhân chập ảnh gốc với kích thước  $N \times M$  với toán tử LoG kích thước  $M^2$ )

– Bước 2: Tìm điểm “chéo-không”. (Áp dụng bộ lọc Laplace)

Quá trình thực hiện:

*Bước 1: Làm trơn ảnh (lọc nhiễu) với bộ lọc Gauss*

Để giảm bớt nhiễu cho ảnh, ta tiến hành làm trơn ảnh bằng bộ lọc Gauss.

\* *Hàm Gauss:*

– Trong 1D

$$g(x) = e^{-\frac{x^2}{2\sigma^2}}$$

Đạo hàm bậc nhất của Gauss:

$$g'(x) = -\frac{1}{2\sigma^2} 2xe^{-\frac{x^2}{2\sigma^2}} = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$$

Đạo hàm bậc hai của Gauss:

$$g''(x) = \left(\frac{x^2}{\sigma^3} - \frac{1}{\sigma}\right) e^{-\frac{x^2}{2\sigma^2}}$$

– Trong 2D

$$g(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$\log_{\sigma} = \frac{1}{\pi\sigma^4} \left[ 1 - \frac{x^2+y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}}$$

trong đó:  $\sigma$  là ngưỡng chuẩn, thể hiện chiều rộng của việc phân phối Gauss. Vùng phân bố thể hiện chất lượng làm mịn nhiều được thực hiện, và được kiểm soát bằng cách biến đổi  $\sigma$ .

Việc làm trơn ảnh được thực hiện bằng cách tiến hành nhân chập ảnh với bộ lọc Gauss.

$$I_{Lap} = I \otimes g(x, y) = g(x, y) \otimes I$$

Như vậy để thực hiện bước 1: Ta tiến hành nhân chập ảnh cần xử lý với bộ lọc Gauss giống như cách thức nhân chập của các toán tử ta đã thực hiện ở trên. Kết quả sau khi tiến hành bước này ảnh được làm trơn. Điều này đặc biệt có tác dụng tốt với ảnh nhiễu, bước này làm mờ đi ảnh cần xử lý, nghĩa là các điểm nhiễu cũng bị làm mờ đi trong khi những biên chính thì vẫn tồn tại.

*Bước 2: Tiến hành đạo hàm bậc hai*

Kết quả nhận được ở bước 1 là ảnh sau khi làm trơn đã loại bỏ bớt được nhiễu. Để thực hiện phương pháp Laplace of Gauss, ta tiến hành đạo hàm bậc hai kết quả ở bước 1. Việc đạo hàm kết quả này tương tự như việc ta tiến hành nhân chập ảnh kết quả ở bước 1 với một mặt nạ nhân chập được tính toán từ phương pháp Laplace.

$$\Delta^2 I_{Lap} = \frac{\partial^2 I_{Lap}}{\partial x^2} + \frac{\partial^2 I_{Lap}}{\partial y^2}$$

$$\Delta^2 I_{Lap} = \Delta^2 (g * I) = (\Delta^2 g) * I$$

$$\Delta^2 g = \frac{1}{\sqrt{2\pi}\sigma^3} \left[ 2 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

Song ở bước thứ nhất ta tiến hành nhân ảnh với bộ lọc Gauss, nên ở đây ta chỉ phải đạo hàm ảnh  $I_{Lap}$  (không sử dụng công thức đạo hàm Gauss). Vì vậy, ta tiến hành nhân ảnh sau khi làm trơn với một trong ba mặt nạ nhân chập của Laplace.

\* Thuật toán:

- Với mỗi điểm ảnh  $I(x,y)$  tính  $I(x,y) \otimes H = I_{Lap}$
- Lấy  $|I_{Lap}| = \Delta^2 I_{Lap}$
- Hiệu chỉnh  $I(x,y) = I_{Lap} \geq \theta ? 1 : 0$

## 2.4 Phương pháp phát hiện biên nâng cao

### 2.4.1 Phương pháp phát hiện biên Canny

#### 2.4.1.1 Giới thiệu

Mục đích của việc phát hiện biên nói chung là giảm thiểu đáng kể số lượng dữ liệu trong một hình ảnh, trong khi các đặc tính cấu trúc vẫn được giữ để sử dụng cho việc xử lý hình ảnh hơn nữa. Có rất nhiều phương pháp phát hiện biên đã được đề cập ở trên nhưng phương pháp Canny – được phát triển bởi John F. Canny (JFC) vào năm 1986, là một trong những công cụ xử lý hình ảnh thông dụng nhất. Mặc dù nó khá là cũ, song đã trở thành một trong những phương pháp phát hiện biên tiêu chuẩn và vẫn được sử dụng trong nghiên cứu.

Mục tiêu (ràng buộc) của JFC để phát triển thuật toán đó là tối ưu những vấn đề liên quan đến các tiêu chuẩn sau:

- Phát hiện: Xác suất phát hiện những điểm biên thực sự phải là cực đại hóa, trong khi xác suất lỗi phát hiện những điểm không phải biên cần được giảm thiểu. Điều này tương ứng với tối đa hóa tỷ lệ báo hiệu nhiều.
- Cục bộ hóa: Các biên phát hiện nên càng gần biên thực càng tốt. Có nghĩa là độ chênh lệch cấp xám giữa các điểm trên cùng một biên càng nhỏ càng tốt.
- Số lượng trả lời – hiệu suất: biên không được nhận ra nhiều, trong khi chỉ có một biên tồn tại → giảm số lượng biên được phát hiện không được nhận ra.

#### 2.4.1.2 Thuật toán phương pháp phát hiện biên Canny

##### a) Các bước thực hiện

Thuật toán được tiến hành qua 5 bước riêng biệt sau:

- *Bước 1:* Giảm nhiễu: làm trơn ảnh để loại bỏ nhiễu bằng cách nhân chập ảnh với bộ lọc Gauss.
- *Bước 2:* Tìm gradient: Tính toán góc và chiều dài của gradient. Biên nên được đánh dấu nơi mà gradient của ảnh có chiều dài lớn.
- *Bước 3:* Thực hiện “Non-maximum suppression”: Chỉ cực đại cục bộ những điểm được đánh dấu là biên (có mức xám cao).
- *Bước 4:* Sử dụng ngưỡng đôi: Những biên tiềm năng được xác định bởi ngưỡng cao và ngưỡng thấp.

## b) Miêu tả các bước thực hiện

### \* *Bước 1: Giảm nhiễu*

Đây là điều không thể tránh khỏi vì tất cả các hình ảnh chụp từ máy quay sẽ chứa một số nhiễu. Để ngăn nhầm lẫn nhiễu với các biên, nhiễu phải được giảm bớt. Do đó ảnh trước tiên được làm mịn bằng cách áp dụng một bộ lọc Gauss. Cách thức tiến hành giống như ở Laplace of Gauss.

Nhân của bộ lọc Gauss với độ lệch chuẩn  $\sigma = 1,4$  được thể hiện trong phương trình sau (ở đây ta sử dụng một bộ lọc  $5 \times 5$ ):

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

### \* *Bước 2: Tìm Gradient*

Các thuật toán Canny về cơ bản tìm thấy các biên nơi mà cường độ mức xám của hình ảnh thay đổi nhiều nhất. Những vùng này được tìm thấy bằng cách xác định gradient của ảnh. Gradient tại mỗi điểm ảnh trong ảnh được làm mịn được xác định bằng cách áp dụng những phương pháp dựa theo toán tử Sobel. Bước đầu tiên là đạo hàm các kết quả ở bước 1 theo hướng x và y với mặt nạ  $3 \times 3$ :

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

Các biên độ gradient (còn được gọi là những cường độ biên) sau đó có thể được xác định như là một thước đo khoảng cách Euclide bằng cách áp dụng luật của Pythagoras như trong phương trình sau:

$$S' = \sqrt{S_x^2 + S_y^2}$$

Đôi khi nó được đơn giản hóa bằng cách áp dụng thước đo khoảng cách Manhattan để giảm bớt sự phức tạp tính toán:

$$|S'| = |S_x| + |S_y|$$

trong đó  $S_x$  và  $S_y$  là gradient theo 2 hướng x và y tương ứng và hướng của biên  $\theta$  như sau:

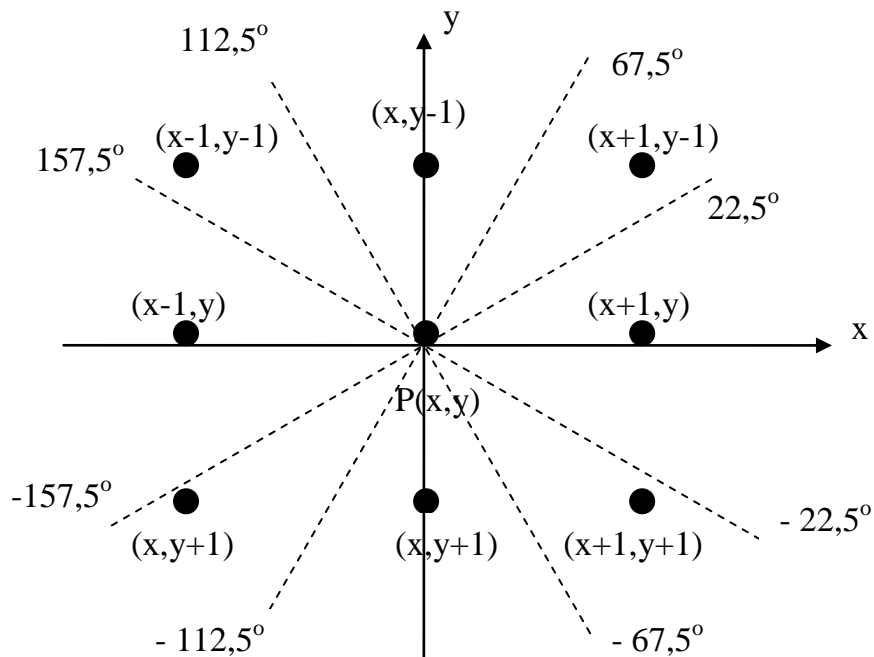
$$\theta = \tan^{-1} \frac{S_y}{S_x}$$

Ảnh  $S'$  tìm được là kết quả của bước thứ 2.

*\* Bước 3: Tiến hành “Non-maximum suppression”*

Tức là loại bỏ một số biên dư thừa: Đối với mỗi điểm ảnh trên ảnh  $S'$  ta tiến hành so sánh giá trị của điểm đó với giá trị của hai điểm lân cận điểm đó. Hai điểm lân cận này là hai điểm nằm trên đường thẳng chứa hướng của đường biên  $\theta$  (công thức tính hướng của đường biên  $\theta$  được tính ở bước 2).

Giả sử với điểm biên đang xét tại vị trí  $P(x, y)$ , ta có 8 điểm biên lân cận điểm biên này:



Hình 2.2: Hình mô tả các điểm biên lân cận của P

Tại điểm biên đó ta tiến hành tính giá trị góc của hướng đường biên  $\theta$ . Nếu hướng của đường biên  $\theta \leq 22,5^\circ$  hoặc  $\theta > 157,5^\circ$  thì đặt giá trị của  $\theta = 0^\circ$  khi đó hai điểm biên lân cận điểm biên này tại vị trí  $(x-1, y)$  và  $(x+1, y)$ .

Tương tự ta có kết quả hai điểm biên lân cận theo các hướng biên khác nhau như bảng dưới đây:

Giá trị $\theta$	Phương hướng	Điểm ảnh
$\theta \leq 22,5^0$ hoặc $\theta > 157,5^0$	$\theta = 0^0$	$(x - 1, y); (x + 1, y)$
$22,5^0 < \theta \leq 67,5^0$	$\theta = 45^0$	$(x - 1, y - 1); (x + 1, y + 1)$
$67,5^0 < \theta \leq 112,5^0$	$\theta = 90^0$	$(x - 1, y - 1); (x + 1, y - 1)$
$112,5^0 < \theta \leq 157,5^0$	$\theta = 135^0$	$(x, y + 1); (x, y - 1)$

Nếu điểm ảnh  $P(x, y)$  có cường độ gradient lớn nhất trong ba điểm ảnh kiểm tra thì được giữ lại điểm biên này. Nếu một trong hai điểm ảnh khác có cường độ gradient cao hơn thì điểm ảnh  $P(x, y)$  này không có trong "trung tâm" của biên và không nên được phân loại như là một điểm biên (tức là loại đi – cho giá trị điểm này = 0).

*\* Bước 4: Ngưỡng đôi*

Các điểm biên còn lại sau khi thực hiện “Non-maximum suppression” vẫn được đánh dấu với cường độ pixel-by-pixel của chúng. Nhiều điểm trong số điểm biên này có lẽ sẽ là biên đúng trong ảnh, nhưng một số có thể được gây ra bởi biến đổi nhiễu hay màu sắc do bề mặt thô ráp. Cách đơn giản nhất để phân biệt những điểm biên này với nhau là sử dụng một ngưỡng, vì vậy chỉ có biên mạnh hơn một giá trị nhất định sẽ được bảo toàn.

Thuật toán phương pháp phát hiện biên Canny sử dụng ngưỡng đôi: ngưỡng thấp  $T_l$  và ngưỡng cao  $T_h$ :

- Xét điểm ảnh  $P(x, y)$
- So sánh  $P(x, y)$  với hai ngưỡng  $T_h$  và  $T_l$ .
- + Nếu  $P(x, y) \geq T_h$ : đánh dấu và giữ lại điểm biên này (đặt giá trị bằng 1).
- + Nếu  $P(x, y) < T_l$ : Loại bỏ điểm biên này (đặt giá trị bằng 0).
- + Nếu  $T_l \leq P(x, y) < T_h$ : ta tiến hành so sánh giá trị  $P(x, y)$  với giá trị của các của 8 điểm lân cận :

Nếu một trong 8 điểm lân cận có giá trị  $> T_h$ : Tiến hành đánh dấu và giữ lại điểm biên này.

Ngược lại: Loại bỏ điểm biên này (đặt giá trị bằng 0).

## 2.4.2 Phương pháp phát hiện biên Wavelet

Phát hiện biên là 1 bước quan trọng trong nhận dạng, phân đoạn ảnh và phân tích cảnh. Các phương pháp tiếp cận để phát hiện biên ko đặt trước nhiễu trong hình ảnh và có thể gây ra các vấn đề trong nhiều ứng dụng. Nhưng nhiễu bị giảm rất hiệu quả bởi phương pháp lọc wavelet mà không làm giảm đáng kể độ phân giải hình ảnh. Không giống như phát hiện biên canny trong đó bước đầu tiên là làm mịn ảnh bằng phương pháp lọc Gaussian để giảm tác động của nhiễu



và bước tiếp theo là xác định biên. Trong Wavelet hai bước này được kết hợp thành một bước và do đó kỹ thuật dựa trên wavelet tính toán hiệu quả hơn. Đó là thực nghiệm đã chứng minh rằng phương pháp phát hiện biên dựa trên Wavelet cho kết quả tốt hơn so với kỹ thuật truyền thống với các ảnh nhiễu. Trong các chương tiếp theo sẽ giải thích rõ hơn về phương pháp này.

### CHƯƠNG III: PHƯƠNG PHÁP PHÁT HIỆN BIÊN CHO ẢNH MÀU

#### 3.1 Cơ sở phát hiện biên ảnh màu

Màu sắc là một tính năng quan trọng mô tả nội dung bên trong một cảnh hình ảnh, và nó được cho là một thuộc tính có độ tin cậy cần được sử dụng để phát hiện biên. Các giả thuyết chính được sử dụng trong việc phát hiện biên màu là có một sự thay đổi trong sắc độ hay cường độ của điểm ảnh ở các đường biên của các đối tượng. Do đó, phát hiện đường biên có thể được thực hiện bằng cách tìm kiếm điểm gián đoạn đột ngột trong tính năng màu.

Những vấn đề thực hiện chính liên quan đến bộ dò biên là khả năng chính xác trích xuất các biên, độ bền của nhiễu, và độ phức tạp tính toán. Một bộ dò biên tối ưu nên giải quyết các vấn đề sau đây:

- *Nhận diện*: không một biên ảnh nào có thể bỏ qua, và biên không được tạo sẵn không được phát hiện.
- *Cục bộ hóa*: biên nên được cục bộ hoá, có nghĩa là khoảng cách giữa các biên thực và biên đã phát hiện nên được giảm thiểu.
- *Độ nhạy*: một bộ dò biên tối ưu chỉ nên có một đáp ứng với một biên duy nhất

Trong chụp ảnh xám đen trắng, biên thường được định nghĩa như là những điểm gián đoạn cường độ sắc nét (mức xám), hay biên vật lý thường trùng với những nơi chiếu sáng mạnh và thay đổi sự phản chiếu. Định nghĩa của một biên trong ảnh màu có nhiều khó khăn hơn. Trong trường hợp ảnh màu, đại diện trong không gian màu ba chiều (3-D), biên có thể được định nghĩa là không liên tục trong lĩnh vực đại diện cho ảnh vector màu. Bằng cách này, biên chia các hình ảnh trong khoảng màu sắc khác nhau, và sự biến thiên về màu sắc đã được kể đến.

Các phương pháp đơn giản để phát hiện biên màu được dựa trên phần mở rộng của các toán tử đơn sắc. Những kỹ thuật này được áp dụng cho ba kênh màu độc lập, thực hiện phát hiện biên xám trên mỗi kênh màu. Các kết quả này sau đó được kết hợp để cung cấp một sơ đồ biên duy nhất. Nói chung, để phát hiện được biên ảnh màu, các kỹ thuật dựa trên biên độ được mở rộng bằng cách lấy biên độ tối đa của gradient, tổng trọng số của chúng, các góc nghĩa là ô của tổng bình phương biên độ của các sơ đồ biên, hoặc bằng cách thực hiện các phép biến đổi riêng lẻ kết quả đầu ra.

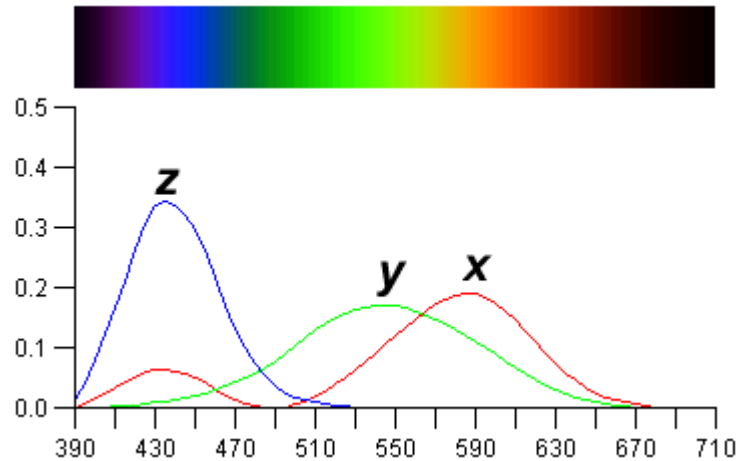
#### 3.2 Mô hình ba màu

Để xác định màu sắc một cách định lượng ta xuất phát từ điểm 3 màu (tức là mỗi màu đều được tạo ra bởi 3 màu cơ bản).

Mô hình 3 màu trở nên rõ rệt ở trong mô hình trộn màu. Sự trộn màu cộng diễn ra khi xuất hiện nhiều kích thích khác nhau về màu, chúng đều được thể hiện qua một thành phần màu cơ bản nào đó để hình thành màu mới.

Do không có tập hợp 3 màu cơ bản nào mà từ đó có thể tạo ra tất cả các màu nên Ủy ban quốc tế về chiếu sáng CIE đã đưa ra mô hình trong đó mỗi màu được đặc trưng bằng đường cong phân bố năng lượng của chúng.

Tiêu chuẩn của CIE ("Commission Internationale d'Eclairage") về các bước sóng như sau:

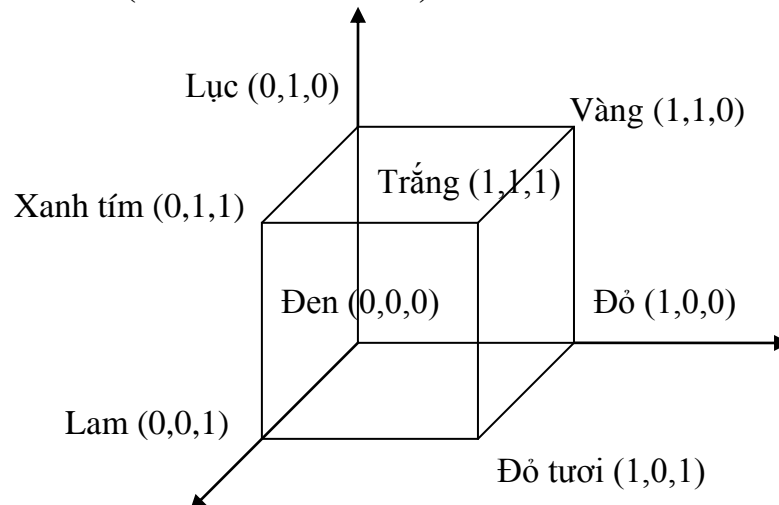


Hình 3.1: Các màu quang phổ theo tiêu chuẩn CIE

trong đó x – đỏ ( $\lambda_x = 700$  nm), y – lục ( $\lambda_y = 546,1$  nm), z – lam ( $\lambda_z = 435,8$  nm) là 3 thành phần của 1 màu nào đó trong các màu cơ bản của CIE. Khi đó các thành phần màu x, y, z được tính như sau trong bản màu CIE 2D:

$$x = \frac{x}{x + y + z} \quad y = \frac{y}{x + y + z} \quad z = \frac{z}{x + y + z} = 1 - x - y$$

**\* Mô hình màu RGB (Red – Green - Blue):**



Hình 3.2: Biểu diễn mô hình màu RGB trên trục tọa độ

Theo Thomas Young (1801) ánh được tổ hợp từ 3 màu cơ bản red, green, blue; từng đôi sẽ cho ra 3 màu thứ cấp yellow (vàng), cyan (lục lam) và magenta (đỏ lila). Màu trắng thu được khi kết hợp cả 3 màu cơ bản. Sự thay đổi cường độ của các màu thành phần sẽ tạo được giá trị mảy bất kỳ trong phổ màu.

Ta có:

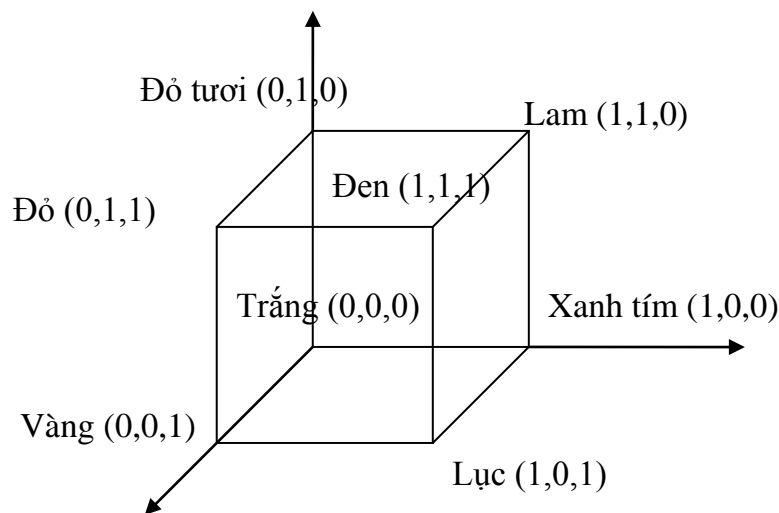
$$C = rR + gG + bB$$

trong đó:  $C$  : màu hay kết quả của ánh sáng kết hợp.

$(r, g, b)$  : các tọa độ màu trong  $[0,1]$ , cường độ của ánh sáng chiếu hay bộ ba giá trị kích thích RGB.

$(R, G, B)$  : các màu chính red, green, blue.

**\* Mô hình màu CMY (Cyan – Megenta – Yellow):**



Hình 3.3: Biểu diễn mô hình màu CMY trên trục tọa độ

Mô hình màu CMY hiển thị ánh sáng và màu sắc phản xạ từ mực in. Nói cách khác, dùng để xuất thông tin màu ra các thiết bị như máy in. Để mô tả ảnh màu ba màu sẽ được tách ra và in chồng lên nhau. Sự chuyển đổi giữa 2 hệ RGB và CMY:

$$\text{Color} = cC + mM + yY \quad \begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

**\* Mô hình màu YIQ:**

Mô hình màu YIQ là mô hình màu được ứng dụng trong truyền hình màu bằng tần rộng tại Mỹ, do đó nó có mối quan hệ chặt chẽ với màn hình đồ họa raster.

YIQ là sự biến đổi tương ứng của tín hiệu RGB nhằm đạt được một hiệu suất truyền phát tốt hơn và giữ nguyên tính tương thích với ti vi đen trắng thế hệ trước. Tín hiệu truyền sử dụng trong hệ thống NTSC (National Television System Committee).

Sự biến đổi RGB thành YIQ được xác định theo công thức:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

### 3.3 Các phương pháp phát hiện biên ảnh màu

Như ta đã biết ảnh màu tổ hợp từ 3 màu cơ bản được thu nhận trên băng giải tần khác nhau và để lưu trữ ảnh màu ta có thể lưu trữ từng màu riêng biệt, mỗi màu lưu trữ như một ảnh đa cấp xám. Vì vậy để phát hiện biên ảnh màu ta có thể dùng phương pháp đơn giản nhất là:

– Bước 1: Chuyển đổi ảnh màu thành ảnh đa mức xám theo công thức:

$$G = 0.299*R + 0.587*G + 0.114*B$$

$$\text{Hay: } G = (R + G + B) / 3$$

– Bước 2: Dùng các phương pháp phát hiện biên cho ảnh đa mức xám để tìm biên.

Ngoài ra ta cũng có thể dùng các phương pháp sau:

#### 3.3.1 Toán tử Vector Gradient

Phương pháp thường dùng để xấp xỉ các đạo hàm là sai số giữa các pixel trong các lân cận nhỏ trong một ảnh. Gọi z là giá trị pixel trong một lân cận kích thước 3 x 3.

$Z_1$	$Z_2$	$Z_3$
$Z_4$	$Z_5$	$Z_6$
$Z_7$	$Z_8$	$Z_9$

Một xấp xỉ của đạo hàm riêng phần theo x (phương dọc) tại tâm điểm của vùng ( $z_5$ ) được cho bởi độ chênh lệch:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

Tương tự đạo hàm theo phương y được xấp xỉ bởi độ chênh lệch

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

Các con số này được tính dễ dàng tại mọi điểm trong ảnh bằng phép nhân chập ảnh lần lượt với từng mặt nạ nhân chập (các mặt nạ Sobel, Prewitt). Sau đó xấp xỉ của đáp ứng gradient ảnh thu được bằng cách cộng hai trị tuyệt đối của hai

ảnh được chấp. Ở chương trước Gradient được tính toán theo cách được mô tả là một trong những phương pháp dùng để tách biên ảnh đa mức xám, nhưng ở chương này ta quan tâm tới việc tính gradient trong không gian màu RGB. Tuy nhiên phương pháp này có thể ứng dụng được trong không gian 2D nhưng không mở rộng được cho không gian có chiều lớn hơn. Cách duy nhất để sử dụng nó đối với các ảnh RGB là tính gradient của mỗi thành phần màu sau đó kết hợp các kết quả.

Do đó Vector gradient (biên độ và hướng) được định nghĩa:

$$c(x, y) = \begin{bmatrix} c_R(x, y) \\ c_G(x, y) \\ c_B(x, y) \end{bmatrix} = \begin{bmatrix} R(x, y) \\ G(x, y) \\ B(x, y) \end{bmatrix}$$

Sau đây là một trong những cách mà khái niệm gradient được mở rộng với các hàm vector. Hàm vô hướng  $f(x, y)$  gradient là một vector mà nó chỉ hướng của tỷ lệ thay đổi cực đại tại tọa độ  $(x, y)$ .

Lấy  $r, g, b$  là các vector đơn vị theo các trục  $R, G, B$  của không gian màu RGB và định nghĩa các vector:

$$u = \frac{\partial R}{\partial x} r + \frac{\partial G}{\partial x} g + \frac{\partial B}{\partial x} b$$

$$v = \frac{\partial R}{\partial y} r + \frac{\partial G}{\partial y} g + \frac{\partial B}{\partial y} b$$

Đặt các số  $g_{xx}, g_{xy}, g_{yy}$  được định nghĩa như sau:

$$g_{xx} = u \cdot u = u^T u = \left| \frac{\partial R}{\partial x} \right|^2 + \left| \frac{\partial G}{\partial x} \right|^2 + \left| \frac{\partial B}{\partial x} \right|^2$$

$$g_{yy} = v \cdot v = v^T v = \left| \frac{\partial R}{\partial y} \right|^2 + \left| \frac{\partial G}{\partial y} \right|^2 + \left| \frac{\partial B}{\partial y} \right|^2$$

$$g_{xy} = u \cdot v = u^T v = \frac{\partial R}{\partial x} \cdot \frac{\partial R}{\partial y} + \frac{\partial G}{\partial x} \cdot \frac{\partial G}{\partial y} + \frac{\partial B}{\partial x} \cdot \frac{\partial B}{\partial y}$$

Chú ý:  $R, G, B$  và các giá trị  $g$  là các hàm của  $x$  và  $y$ . Chú ý này được thể hiện trong Di Zenzo[1986] nói rằng hướng của tỷ lệ thay đổi cực đại của  $c(x, y)$  là một hàm  $(x, y)$ :

$$\theta(x, y) = \frac{1}{2} \tan^{-1} \left[ \frac{2g_{xy}}{g_{xx} - g_{yy}} \right]$$

và giá trị của tỷ lệ thay đổi (biên độ của gradient) theo các hướng được cho bởi các phần tử của:

$$F_{\theta}(x, y) = \left\{ \frac{1}{2} (g_{xx} + g_{yy}) \pm (g_{xx} - g_{yy}) \cos 2\theta + 2g_{xy} \sin 2\theta \right\}^{1/2}$$

Chú ý:  $\theta(x, y)$  và  $F_{\theta}(x, y)$  là các ảnh có cùng kích cỡ với ảnh đầu vào. Các phần tử  $\theta(x, y)$  là các góc tại mỗi điểm gradient được tính và  $F_{\theta}(x, y)$  là ảnh gradient.

Vì  $\tan = \tan(\alpha \pm \pi)$  nếu  $\theta_0$  là một nghiệm của phương trình  $\tan^{-1}$

$$\rightarrow \theta_0 = \pm \pi/2$$

$F_{\theta}(x, y) = F_{\theta+\pi}(x, y) \rightarrow F$  cần được tính chỉ với các giá trị của  $\theta$  trong nửa khoảng  $[0, \pi)$ .

Sự thật là phương trình  $\tan^{-1}$  cho 2 giá trị lệch nhau  $90^\circ$  nghĩa là phương trình này nghiệm đúng với mỗi cặp  $(x, y)$  có hướng vuông góc nhau. Dọc theo một trong những hướng này  $F$  là cực đại, và nó sẽ là cực tiểu nếu theo hướng còn lại.  $\rightarrow$  kết quả nhận được bằng cách chọn giá trị lớn nhất tại mỗi điểm.

\* Thuật toán:

- Input: Ảnh RGB

- OutPut: ảnh đã tách biên.

- Các bước thực hiện:

+ Tính đạo hàm  $x, y$  của 3 thành phần ảnh với mặt nạ được chọn.

+ Tính các tham số vector gradient  $g_{xx}, g_{yy}, g_{xy} \rightarrow \theta_1, \theta_2$

+ Tính  $F_{\theta}(x, y) = \max[F_{\theta_1}(x, y), F_{\theta_2}(x, y)]$  và chuẩn hóa trong khoảng  $[0, 1]$

+ So sánh với ngưỡng, lấy giá trị lớn nhất tại mỗi điểm  $\rightarrow$  được ảnh tách biên

### 3.3.2 Phương pháp Vector Field

Sự khác biệt giữa hai màu giống nhau tại vị trí  $\chi_1$  và  $\chi_2$  của vector Field được gọi là sự tương phản cục bộ, được cho bởi  $\Delta x = x_{\chi_1} - x_{\chi_2}$ . Khi khoảng cách giữa hai vector hướng về 0, sự khác biệt trở thành :

$$dx = \sum_{j=1}^2 \frac{\partial x}{\partial \chi_j} d\chi_j$$

$$dx^2 = \sum_{k=1}^2 \sum_{l=1}^2 \frac{\partial x}{\partial \chi_k} * \frac{\partial x}{\partial \chi_l} d\chi_k d\chi_l = \sum_{k=1}^2 \sum_{l=1}^2 g_{kl} d\chi_k d\chi_l = \begin{bmatrix} d\chi_1 \\ d\chi_2 \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} d\chi_1 \\ d\chi_2 \end{bmatrix}$$

$$g_{kl} = \frac{\partial x}{\partial \chi_k} \cdot \frac{\partial x}{\partial \chi_l} = \sum_{j=1}^3 \frac{\partial x^j}{\partial \chi_k} \cdot \frac{\partial x^j}{\partial \chi_l}$$

Đối với một vector đơn vị  $n = (\cos \theta, \sin \theta)$ ,  $dx^2(n)$  cho thấy tỷ lệ thay đổi của hình ảnh theo hướng  $n$ :

$$dx^2 = E \cos^2 \theta + 2F \cos \theta \sin \theta + G \sin^2 \theta = (E + G + H \cos 2\theta + 2F \sin 2\theta)/2$$

với:  $E = g_{11}$ ;  $F = g_{12} - g_{21}$ ;  $G = g_{22}$ ;  $H = g_{11} - g_{22}$

Các giá trị riêng tương ứng và các vector đặc trưng là:

$$\lambda_{\pm} = \frac{1}{2} \left( E + G \pm \sqrt{H^2 + 4F^2} \right)$$

$$n_{\pm} = \left( \cos \theta_{\pm}, \sin \theta_{\pm} \right) \quad \text{với: } \theta_{-} = \theta_{+} \pm \frac{\pi}{2}.$$

$$\theta_{+} = \frac{1}{2} \arctan \left\{ \frac{2F}{H} \right\}$$

Hai đáp số tương ứng với cực đại và cực tiểu của  $dx^2$ , và  $\theta_{+}$ ,  $\theta_{-}$  định nghĩa theo hai hướng: cùng một trong số  $x$  đó họ đạt tỷ lệ biến đổi cực đại của nó, cùng một số khác đạt đến mức cực tiểu của nó.

Thay vì sử dụng hàm lượng giác, các Vector đặc trưng  $n = [n^1, n^2]$  có thể được tính bằng cách sử dụng vector  $v = [H, 2F]$ , với tính chất:  $v = \lambda_{+} - \lambda_{-}$

$$\text{Sau đó: } n = \left[ \sqrt{(1+b)/2}, \text{sign } F \sqrt{(1-b)/2} \right] \quad \text{với } b = H / \|v\|$$

Cường độ của một biên trong trường hợp giá trị vector được miêu tả  $\lambda^{+}$  bởi sự khác biệt giữa các cực trị:

$$\lambda^{+} - \lambda^{-} = (E + G)/2$$

$$\lambda^{+} + \lambda^{-} = \sqrt{H^2 + 4F^2}$$

### 3.3.3 Bộ dò biên Vector Order-Statistic

Toán tử dựa trên Order-Statistic đóng vai trò quan trọng trong việc xử lý ảnh và đã được sử dụng rộng rãi trong xử lý ảnh đơn sắc và ảnh màu. Các loại bộ dò biên dựa trên Order-Statistic được xây dựng như là một sự kết hợp tuyến tính của các mẫu vector tới của sổ xử lý. Tập các hệ số khác nhau của sự kết hợp tuyến tính tạo ra bộ dò biên khác nhau về hiệu suất và hiệu quả.

Toán tử phát hiện biên đơn giản nhất dựa trên Order-Statistic là bộ dò vector vùng (VRD – vector range detector). Toán tử này xác định độ lệch giữa trung bình vector cấp thấp nhất với vector cấp cao nhất. Cho các vector hình ảnh trong một cửa sổ  $W$  ký hiệu  $x_i$ ,  $i = [1, n]$  khi đó:

$$VRD = \rho(x_1, x_n)$$

Tốc độ và hiệu suất tính toán của công thức trên đạt yêu cầu trong trường hợp ảnh không bị ảnh hưởng bởi nhiễu. Trong môi trường nhiễu, các xung có mặt trong ảnh được phát hiện sai biên. Để làm giảm bớt nhược điểm này, các biện pháp tán sắc được biết như là ước lượng chắc chắn trước nhiễu có thể được áp dụng.



Bộ dò vector tán sắc (VDD – vector dispersion detector) được định nghĩa:

$$VDD = \left\| \sum_j^n \psi_j x_j \right\| \quad \text{với } j = [1, n]$$

trong đó  $\psi_j$  là hệ trọng số.

VRD là một trường hợp đặc biệt của VDD nếu:

$$\psi_1 = -1, \psi_n = 1 \text{ và } \psi_j = [2, n-1]$$

Định nghĩa này có thể được khái quát bằng cách sử dụng một tập hợp các hệ số  $k$  và kết hợp các kết quả độ lớn vector là một cách mong đợi. Ví dụ, các bộ dò vector tán sắc tổng quát (GVDD – generalized vector dispersion detector) sử dụng một toán tử min làm giảm thiểu ảnh hưởng của xung nhiễu:

$$GVDD = \min_j \left\| \sum_{l=1}^n \psi_{jl} x_l \right\| \quad \text{với } j = [1, k], k \text{ là tham số của bộ dò.}$$

Cụ thể, bộ dò biên màu có thể thu được từ GVDD bằng cách chọn tập hợp các hệ số  $\psi_{jl}$ . Một thành viên đặc biệt của họ GVDD là bộ dò vector tán sắc tối thiểu (MVDD – minimin vector dispersion detector), được định nghĩa như sau:

$$MVDD = \min_j \rho_{(n-j+1), \hat{x}_{(\alpha)}} \quad \text{với } j = [1, k] \text{ và } \alpha < n$$

Trong đó:

- $\hat{x}_{(\alpha)}$  là trung bình cộng của các vector được sắp xếp từ  $\alpha$  đầu tiên ( $\alpha$ VMF)
- $k, \alpha$  kiểm soát việc điều hòa giữa tính toán phức tạp và hiệu quả ngăn chặn nhiễu.

Các vector thứ bậc cao có thể sẽ đại diện cho điểm ảnh bị xáo trộn bởi nhiễu, ấn định tham số  $j$  dẫn tới một bộ dò biên mạnh với nhiễu xung:

$$MVDD^* = \rho_{(n-k+1), \hat{x}_{(\alpha)}} \quad k < n$$

đó cũng là miễn nhiễm với ảnh hưởng của nhiễu Gaussian, do toán tử trung bình thực hiện trong  $\hat{x}_{(\alpha)}$ .

Một phác họa thay thế của toán tử GVDD dùng khái niệm lọc lân cận gần nhất tương ứng. Bộ dò vùng vector lân cận gần nhất (NNVRD – nearest-neighbor vector range detector) được định nghĩa như sau:

$$NNVRD = \rho \left( x_{(n)}, \sum_{j=1}^n \psi_j x_{(j)} \right), \quad \psi_j = \frac{1}{n} \cdot \frac{R_{(n)} - R_{(j)}}{R_{(n)} - \hat{R}_{(n)}}$$

Với:  $\psi_j$  : hệ số hiệu chỉnh bù được xác định tương ứng.

$R_{(n)}$  : giá trị trung bình số học của các khoảng cách gộp  $R_j$  với  $j = [1, n]$

Toán tử MVDD cũng có thể được kết hợp với toán tử NNVRD để cải thiện hiệu quả của nó khi có nhiễu xung. Bộ dò phân tán vector lân cận gần nhất

tối thiểu (NNMVDD – nearest-neighbor minimum vector dispersion detector) được định nghĩa như sau:

$$NNMVDD = \min_j \rho \left( x_{(n-j+1)}, \sum_{l=1}^n \psi_l x_{(l)} \right), \quad \text{for } j = [k] \text{ and } k < n$$

## CHƯƠNG IV: PHƯƠNG PHÁP PHÁT HIỆN BIÊN THEO WAVELET

### 4.1 Giới thiệu

Phát hiện biên là 1 bước quan trọng trong nhận dạng, phân đoạn ảnh và phân tích cảnh. Các phương pháp tiếp cận để phát hiện biên ko đặt trước nhiều trong hình ảnh và có thể gây ra các vấn đề trong nhiều ứng dụng. Nhưng nhiều bị giảm rất hiệu quả bởi phương pháp lọc wavelet mà không làm giảm đáng kể độ phân giải hình ảnh. Không giống như phát hiện biên canny trong đó bước đầu tiên là làm mịn ảnh bằng bộ lọc Gauss để giảm tác động của nhiễu và bước tiếp theo là xác định biên. Trong Wavelet hai bước này được kết hợp thành một bước và do đó kỹ thuật dựa trên wavelet tính toán hiệu quả hơn. Đó là thực nghiệm đã chứng minh rằng phương pháp phát hiện biên dựa trên Wavelet cho kết quả tốt hơn so với kỹ thuật truyền thống với các ảnh nhiễu.

Việc phân tích Wavelet là phân tích cục bộ và phù hợp cho việc phân tích tần suất thời gian. Trong kỹ thuật phát hiện biên Wavelet, biến đổi được sử dụng là biến đổi Wavelet rời rạc (DWT – Discrete Wavelet Transform) và bộ lọc là một trong những tìm kiếm cho cực đại địa phương trong 1 miền wavelet. Các biến đổi Wavelet cung cấp bộ phân tích đa tỉ lệ, có thể được áp dụng để phát hiện biên.

Mallat và Zhong sử dụng cực đại địa phương giá trị tuyệt đối của các biến đổi wavelet để thực hiện phát hiện biên. Peytavin nghiên cứu phương pháp phát hiện biên nhiều hướng và nhiều độ phân giải bằng cách sử dụng một wavelet, đó là lấy đạo hàm thứ nhất của hàm làm mịn. Barlaud sử dụng biến đổi wavelet song trực giao (biorthogonal wavelet transform) để phát hiện biên.

Một sở hữu đáng kể của các biến đổi wavelet là khả năng mô tả đặc điểm các quy luật địa phương của các hàm. Đối với một hình ảnh  $f(x, y)$ , các biên của nó tương ứng với điểm kỳ dị của  $f(x, y)$ , và do đó có liên quan đến các cực đại địa phương của mô đun biến đổi wavelet. Do đó, biến đổi wavelet là một phương pháp hiệu quả để phát hiện biên.

Biến đổi wavelet hơi tương tự với biến đổi Fourier ngắn hạn (Short Time Fourier Transform – STFT) ngoại trừ cửa sổ không cố định như trong STFT. Việc biến đổi wavelet linh hoạt hơn và nó có thể có bất cứ hàm lựa chọn nào, có thể được co lại và giãn ra để phân tích các tín hiệu. Các Wavelet có thể được hiểu là các làn sóng nhỏ trừu tượng thể hiện trong một giá trị trung bình số không. Các biến đổi wavelet ánh xạ hàm thời gian vào một hàm hai chiều của  $a$  và  $b$ . Các tham số  $a$  được gọi là thang tỉ lệ và nó tỉ lệ một hàm bằng cách nén hoặc kéo dài nó. Các tham số  $b$  được gọi là biến đổi của hàm Wavelet dọc theo trục thời gian.

Hàm Wavelet  $\psi(t)$  được viết như sau:

$$\psi_a(t-b) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (4.1)$$

trong đó:  $a$  : là hệ số tỉ lệ.

$b$  : là biến đổi dọc theo trục thời gian  $t$ .

$\frac{1}{\sqrt{|a|}}$  : hệ số được chuẩn hóa.

## 4.2 Biến đổi Wavelet liên tục (The Continuous Wavelet Transform – CWT)

Trong biến đổi wavelet liên tục, một hàm  $\psi$  (“psi”) được sử dụng để tạo ra một họ các wavelet  $\psi(at + b)$ , với  $a, b$  là các số thực và tham số  $a$  làm giãn nở các hàm  $\psi$  và  $b$  dịch nó. Lưu ý rằng từ “liên tục” đề cập đến biến đổi chứ ko phải wavelet.

Các biến đổi wavelet liên tục biến một tín hiệu  $s(t)$  vào trong một hàm với hai biến số (tỉ lệ và thời gian) như trong phương trình:

$$CWT(a, b) = \int s(t) \psi(at + b) dt \quad (4.2)$$

Với  $\psi$  là wavelet gốc và  $c$  là hệ số của biến đổi wavelet.

Trong bất kỳ ứng dụng ảnh, mã hóa hoàn thiện tái tạo là quan trọng nhất. Ngay cả sau khi xử lý ảnh các ảnh được tái tạo sẽ chỉ được một xấp xỉ của ảnh gốc. Nếu chuyển đổi không phải là khả nghịch nó có thể được cho là đã biến đổi các mô hình khác nhau xử lý giống hệt nhau.

Định lý 1: cho  $s(t) \in L^2(\mathbb{R})$  và CWT của  $s(t)$ , liên quan đến  $\psi(t)$ , ta có:

$$CWT(a, b) = \frac{1}{\sqrt{|a|}} \int s(t) \psi\left(\frac{t-b}{a}\right) dt \quad (4.3)$$

Nếu  $\psi(t)$  là vậy, mà biến đổi này là khả nghịch thì:

$$s(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{a>0} CWT(a, b) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \frac{1}{a^2} da db \quad (4.4)$$

Trong đó  $c_\psi$  là một hằng số chỉ phụ thuộc vào  $\psi(t)$  và  $a$  dương. Liên tục có giá trị:

$$c_\psi = \int_0^\infty \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty \quad (4.5)$$

Từ đó đặt ra một điều kiện thỏa mãn  $\psi(t)$ . Để  $c_\psi < \infty$ ,  $\psi(t)$  phải thỏa mãn:

$$|\psi(\omega)| < \infty, \text{ for any } \omega \quad (4.6)$$

Và  $\psi(0) = 0$ , có nghĩa là  $\int \psi(t) dt = 0$ .

Định lý này khẳng định việc biến đổi một chiều tín hiệu  $s(t)$  thành miền wavelet hai chiều của  $(a,b)$  là nghịch nếu việc biến đổi một đẳng cự lên một hằng số  $c_\psi$ .

### 4.3 MRA – Multi-resolution Analysis

MRA là sự phân hoạch một tín hiệu  $s(t)$  thành các thành phần của tỉ lệ khác nhau  $2^{-m}$ , trong đó  $m$  là một số nguyên. Liên kết với mỗi tỉ lệ là một không gian con  $V_m$ . Các không gian con là các hàm thời gian, đáp ứng các điều kiện sau đây:

1.) Liên tiếp được lồng nhau. Tức là cho tất cả  $j$ .

$$V_j \subset V_{j+1}$$

$$\dots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset V_{-3} \dots \rightarrow L^2(R)$$

Nếu hàm  $s(t)$  trong  $V_j$ , thì  $s(2t)$  trong  $V_{j-1}$  và ngược lại.

2.) Sự tồn tại của hàm tỉ lệ trực giao.

Có sự tồn tại của một hàm tỉ lệ còn được gọi là wavelet cha,  $\phi(t) \in V_0$  sao cho tập:

$$\left\{ \phi_{mn}(t) = 2^{\frac{-m}{2}} \phi(2^m t - n) \right\} : n = \text{interger} \quad (4.7)$$

là một trực giao tiêu chuẩn trải rộng trên  $V_m$ .

Các hợp của không gian này là hẹp, tức là cho tất cả  $f$  trong  $L^2(R)$

$$\lim_{j \rightarrow +\infty} \|f - P_j^0 f\|_{L^2} = 0 \quad (4.8)$$

trong đó  $P_j^0$  là phép chiếu trực giao lên  $V_j$

3.) Hàm cơ sở được định nghĩa bởi hàm sai phân 2 tỉ lệ.

$\phi_{0n}(t)$  kéo dài  $V_0$  và  $\phi_{-1n}(t)$  kéo dài  $V_{-1}$  và  $V_{-1}$  bao gồm  $V_0$ ,  $\phi_{00}(t) = \phi(t)$  là một tổ hợp tuyến tính của  $\phi_{-1n}(t) = \sqrt{2} \phi(2t - n)$ , nghĩa là:

$$\phi(t) = 2 \sum_{l=0}^{p-1} g(l) \phi(2t - l) \quad (4.9)$$

trong đó  $g(l)$  là hệ số kết hợp.

4.) Giao điểm của các không gian là giảm tới hàm NULL.

$$\lim_{j \rightarrow -\infty} \|P_j^0 f\|_{L^2} = 0 \quad (4.10)$$

5.) Tồn tại một hàm  $\phi \in V_0$  sao cho họ  $\phi(-k), k \in Z$  là cơ sở Riesz của  $V_0$ .

#### 4.4 Xây dựng một Wavelet

Phương trình tỉ lệ 2 được cho bởi:

$$\begin{aligned} \phi(x) &= \sum_k h_k \phi_{1,k}(x) = \sum_k h_k \sqrt{2} \phi(\sqrt{2}x - k) \\ h_k &= \langle \phi, \phi_{1,k} \rangle \end{aligned} \quad (4.11)$$

$$\text{Giả sử } g(x) \in V_1 \text{ và } g(x) = \sum_k g_k \sqrt{2} \phi(\sqrt{2}x - k) \quad (4.12)$$

$$\begin{aligned} \text{với } g_k &= (-1)^k h_{1-k} \\ \rightarrow g(x) &= \sum_k (-1)^k h_{1-k} \phi_{1,k}(x) \end{aligned} \quad (4.13)$$

$$\text{Ta định nghĩa: } g_{jk}(x) = 2^{\frac{j}{2}} g(2^j x - k) \quad (4.14)$$

vì vậy  $\langle \phi_{0k}, \phi_{0l} \rangle = 0$  cho tất cả số nguyên  $k$  và  $l$ .

Định nghĩa  $W_0$  là không gian mở rộng bởi  $g_{0k}$ , trong đó cho thấy  $W_0$  là phần bù trực giao của  $V_0$  trong  $V_1$ .

$$V_1 = V_0 \oplus W_0$$

Tương tự cho  $W_j$  là không gian mở rộng bởi  $g_{jk}$ , trong đó cho thấy  $W_j$  là phần bù trực giao của  $V_j$  trong  $V_{j+1}$ .

$$\rightarrow V_j = V_0 \oplus W_0 \oplus W_1 \oplus \dots \oplus W_{j-1}$$

Vì thế,  $g_{jk}$  là cơ sở trực giao cho  $L^2(R)$ . Chúng cũng là trực giao.

Do đó chứng minh rằng  $\psi(x) = g(x)$  là một Wavelet

#### 4.5 Phân hoạch và tái thiết wavelet

Xét  $c_{j,k}$  như là hệ số wavelet cho  $\phi$

$$c_{j,k} = \langle \phi, \phi_{j,k} \rangle = \int \phi(x) \phi_{j,k}(x) dx \quad (4.15)$$

Chúng ta có thể sử dụng các ký hiệu trên cho bất kỳ hàm  $f$  mà chúng ta quan tâm

$$\rightarrow c_{j,k} = \langle f, \phi_{j,k} \rangle = \int f(x) \phi_{j,k}(x) dx \quad (4.16)$$

$$\text{tương tự } d_{j,k} = \langle f, \psi_{j,k} \rangle = \int f(x) \psi_{j,k}(x) dx \quad (4.17)$$

Giả sử chúng ta có hệ số cấp cao  $c_{j,k}$  cho một hàm  $f$  đưa ra, chúng ta thu được các hệ số cấp độ thấp hơn, chi tiết.

$$c_{j-1,k} = \sum_l h_{l-2k} c_{j,l} \quad (4.18)$$

$$d_{j-1,k} = \sum_l g_{l-2k} c_{j,l} \quad (4.19)$$

Nhắc lại cấu trúc của không gian MRA, xét hệ số theo cách tương tự:

$$\dots c_{j-2} \leftarrow c_{j-1} \leftarrow c_j \dots$$

$$\dots d_{j-2} \quad d_{j-1} \quad \dots$$

Quá trình trên được gọi là downsampling. Các hệ số downsampled sẽ chính là một nửa đại lượng các hệ số trước đó. Ta có thể viết:

$$\sum_k c_{j,k} \phi_{j,k} \leftarrow \sum_k c_{j-1,k} \phi_{j-1,k} \leftarrow \sum_k d_{j-1,k} \psi_{j-1,k} \quad (4.20)$$

nếu chúng ta tiếp tục với chuỗi MRA như trên, chúng ta có thể thấy rằng

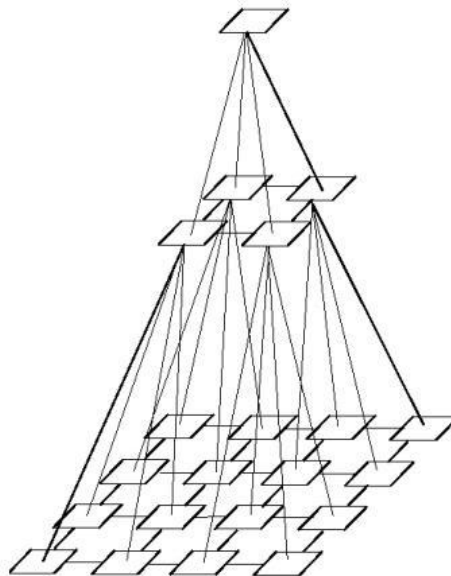
$$\sum_k c_{j,k} \phi_{j,k} \leftarrow \sum_k c_{j0,k} \phi_{j0,k} + \sum_{j=j0}^{j-1} \sum_k d_{j,k} \psi_{j,k} \quad (4.21)$$

Phương trình trên là diễn giải của biến đổi wavelet rời rạc (DWT)

Tái thiết wavelet rất tương tự như sự phân hoạch wavelet. Chúng ta sử dụng mô hình chuỗi MRA để upsample các hệ số. Trong tái thiết chúng ta xét các hệ số mức độ thấp hơn, hệ số chi tiết và kết hợp chúng để có được hệ số mức độ cao hơn. Mỗi lần chúng ta thực hiện bổ sung các hệ số số lượng là gấp đôi số lượng các hệ số của hệ số mức độ thấp trước đó.

#### 4.6 Biến đổi Wavelet rời rạc (DWT – Discrete Wavelet Transform)

Không giống như biến đổi Fourier, DWT không phải là một đối tượng duy nhất. Khái niệm về DWT lần đầu tiên được giới thiệu bởi Stromberg từ Littlewood-Paley phân tách các toán tử một chức năng. Hai chiều DWT giảm từ sơ đồ kim tự tháp Laplace của Burt và Adelson như trong hình sau:



Hình 4.1: Sơ đồ kim tự tháp Laplace phát triển bởi Burt và Adelson

Nếu hàm ngày càng mở rộng là một dãy các con số, như các mẫu của một hàm liên tục  $f(x)$ , các hệ số kết quả được gọi là biến đổi Wavelet rời rạc (DWT) của  $f(x)$ . DWT biến đổi một tín hiệu thời gian rời rạc trên một phép biểu diễn Wavelet rời rạc.

Chúng ta không tính toán DWT về ma trận do vấn đề lưu trữ, thay vì vậy chúng ta sử dụng bộ lọc để tính toán DWT.

Cho một dãy  $a = \{\dots, a_{-1}, a_0, a_1, \dots\}$  và hai bộ lọc xung  $h$  và  $g$ .

Chúng ta định nghĩa phép nhân chập như sau:

$$(Ha)_k = \sum_l h_{l-2k} a_k$$

$$(Ga)_k = \sum_l g_{l-2k} a_k$$

trong đó  $H$  và  $G$  tương ứng với một bước của DWT (lên hoặc xuống).

Trong phân tách, hàm được nhân chập lần lượt với hai bộ lọc  $H$  (thông thấp) và  $G$  (thông cao). Mỗi hàm kết quả bị mất đi bởi việc loại bỏ một trong hai mẫu. Tín hiệu thông cao bị loại bỏ, và ta lặp lại với tín hiệu thông thấp. Tái giả thiết, chúng ta khôi phục lại mẫu bằng cách chèn 0 giữa mỗi mẫu, sau đó chúng ta nhân chập với các bộ lọc liên hợp  $\tilde{H}$  và  $\tilde{G}$ , thêm các hàm kết quả và nhân kết quả với 2. Chúng ta lặp lại cho đến tỉ lệ nhỏ nhất.

Giải thuật 2D dựa trên những biến số riêng biệt để ưu tiên  $x$  và hướng. Hàm tỉ lệ được xác định như sau:

$$\phi(x, y) \Rightarrow \phi(x) \phi(y)$$

Tín hiệu chi tiết được lấy từ 3 Wavelet:

\* Wavelet dọc:

$$\psi^1(x, y) \Rightarrow \phi(x) \psi(y)$$

\* Wavelet ngang:

$$\psi^2(x, y) \Rightarrow \psi(x) \phi(y)$$

\* Wavelet chéo:

$$\psi^3(x, y) \Rightarrow \psi(x) \psi(y)$$



$f^{(2)}$	H.D $j = 2$	Horizontal Details $j = 1$	Horizontal Details $j = 0$
V. D $j = 2$	D.D $j = 2$		
Vertical Details $j = 1$		Diagonal Details $j = 1$	
Vertical Details $j = 0$			Diagonal Details $j = 0$

Hình 4.2: DWT của hai chiều tín hiệu

#### 4.7 Phương pháp phát hiện biên DWT

Cường độ thay đổi đáng kể trong một hình ảnh thường xảy ra ở độ phân giải và tỉ lệ có trong không gian khác nhau. Bộ dò biên thông thường lựa chọn một mặt nạ không gian đặc biệt, phát hiện các biên ở độ phân giải cụ thể. Bộ dò biên với mặt nạ nhỏ nhạy cảm với nhiễu và tạo ra các biên giả. Ngược lại, với một mặt nạ lớn lại tương đối mạnh trước nhiễu, nhưng làm biến dạng các biên và có thể không phát hiện một số chi tiết tốt hơn. Vì vậy rất khó khăn để phát hiện các biên với một mặt nạ biên không gian duy nhất.

Việc bảo toàn biên loại bỏ nhiễu mượn chính nó cho một quy trình dựa trên Wavelet để phát hiện biên. Các biên trong một tín hiệu cho phép tăng tới đỉnh cao vượt qua kết quả đầu ra của bộ lọc hoặc các băng con cụ thể tại các vị trí phù hợp. Đây là một đặc tính của DWT. Nói một cách khác, biên làm gia tăng đỉnh trên nhiều mức của các chi tiết tại các giá trị tọa độ mà di chuyển sang trái bằng một nửa yếu tố tại mọi quá trình chuyển đổi từ một tỉ lệ tốt hơn tới một tỉ lệ thô hơn. Các biên mạnh, cao hơn các đỉnh trong DWT này. Do đó, một biên nào đó có thể được tìm thấy từ các biến đổi wavelet của đỉnh xác định tại các vị trí phù hợp.

Các biến đổi wavelet của  $f(x)$  với tỉ lệ  $s$  và vị trí  $x$ , được tính liên quan tới wavelet  $\psi^a(x)$  được định nghĩa bởi:

$$W_s^a f(x) = f * \psi_s^a(x)$$

Biến đổi wavelet với  $\psi^b(x)$ :

$$W_s^b f(x) = f * \psi_s^b(x)$$

trong đó:  $\psi(x)$  là hàm Wavelet tại vị trí x.

a, b : 2 số thực tùy ý

Sau khi lấy đạo hàm ta được:

$$W_s^a f(x) = s \frac{d}{dx} (f * \theta_s)(x)$$

$$W_s^b f(x) = s^2 \frac{d^2}{dx^2} (f * \theta_s)(x)$$

Các biến đổi wavelet  $W_s^a f(x)$  và  $W_s^b f(x)$  tương ứng với đạo hàm bậc nhất và đạo hàm bậc hai của dấu hiệu được làm mịn trong s. Hàm làm mịn 2-D được định nghĩa như bất cứ hàm  $\theta(x,y)$  nào có tích phân qua x và y bằng 1 và đồng qui từ 0 tới  $\infty$ . Ảnh  $f(x,y)$  được làm mịn ở các tỉ lệ s khác nhau bằng việc nhân chập với  $\theta(x,y)$ .

Vector gradient tính được cho bởi  $\vec{\nabla} (f * \theta_s)(x, y)$ . Biên được định nghĩa là các điểm  $(x_0, y_0)$  mà mô đun của vector gradient là cực đại về hướng.

Chúng ta định nghĩa hai hàm wavelet như sau:

$$\psi_s^1(x, y) = \frac{\partial \theta(x, y)}{\partial x}$$

$$\psi_s^2(x, y) = \frac{\partial \theta(x, y)}{\partial y}$$

$$\text{với: } \psi_s^1(x, y) = \frac{1}{s^2} \psi^1\left(\frac{x}{s}, \frac{y}{s}\right); \psi_s^2(x, y) = \frac{1}{s^2} \psi^2\left(\frac{x}{s}, \frac{y}{s}\right) \text{ và } f(x, y) \in L^2(R^2)$$

→ Các biến đổi Wavelet được định nghĩa:

$$\psi_s^1 f(x, y) = f * \psi_s^1(x, y)$$

$$\psi_s^2 f(x, y) = f * \psi_s^2(x, y)$$

Vậy ta có:

$$\begin{pmatrix} \psi_s^1 f(x, y) \\ \psi_s^2 f(x, y) \end{pmatrix} = s \vec{\nabla} (f * \theta_s)(x, y)$$

Do đó các điểm biên có thể được xác định từ các thành phần  $W_s^1 f(x)$  và  $W_s^2 f(x)$  của biến đổi wavelet.

## **CHƯƠNG V:      CÀI ĐẶT THỬ NGHIỆM VÀ NHẬN XÉT ĐÁNH GIÁ CÁC PHƯƠNG PHÁP**

### **4.1 So sánh các phương pháp phát hiện biên (Gradient, Laplace, Canny)**



a) Ảnh gốc



b) Gradient (ngưỡng 20)



c) Sobel (ngưỡng 128)



d) Prewitt (ngưỡng 128)



e) Laplace (ngưỡng 50)



f) Canny (ngưỡng 20)

Hình 5.1: Biên ảnh theo Gradient, Laplace, Canny.

Để so sánh sự khác nhau giữa các phương pháp ta sử dụng kết quả của phương pháp Sobel, Laplace và Canny. Từ kết quả thu được ở các hình trên ta có thể thấy, khi phát hiện biên, các biên không quan trọng nên được loại bỏ.

- Phương pháp Sobel vẫn phát hiện được biên nhưng các biên mờ, không được rõ nét, do trong ảnh có những vùng có mức xám thấp, sự thay đổi giữa các mức xám nhỏ.
- Phương pháp Laplace cho kết quả rõ nét hơn (do phương pháp này sử dụng phương pháp đạo hàm bậc hai, các điểm biên là các điểm chéo không). Tuy vậy do ảnh có rất nhiều điểm biên nhỏ nên các biên ảnh ở trên qua phương pháp này rất nhiều và rối, chúng ta nên loại bỏ các điểm biên thừa.
- Còn đối với phương pháp Canny, do quá trình “Non maximum Suppression” và do quá trình áp dụng ngưỡng mà các điểm biên phụ bị loại bớt đi, các biên chính được giữ lại nên biên rõ nét hơn.

Đối với ảnh có nhiều có mức xám nhỏ, sự biến thiên các mức xám là thấp ta nên sử dụng phương pháp Laplace, song nếu ảnh đó có quá nhiều biên thì ta nên sử dụng phương pháp Canny để loại bỏ bớt các biên không cần quan tâm đi.

#### 4.2 Cài đặt thử nghiệm chương trình Wavelet Transform

Trong chương trình, em xin trình bày quá trình thực hiện biến đổi Wavelet:

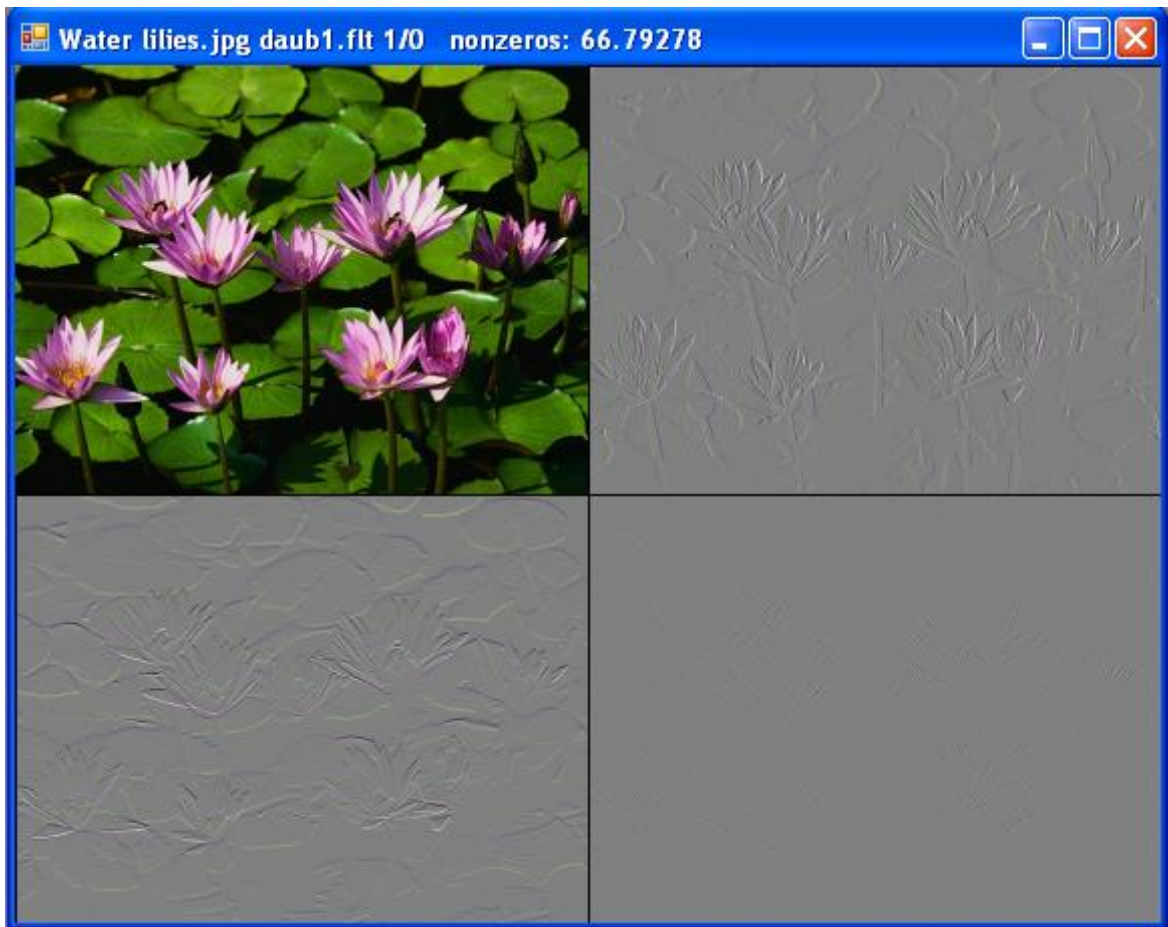
- Ảnh đầu vào: \*.bmp, \*.jpg, \*.jpeg, \*.png, \*.tif, \*.tiff.
- Ảnh đầu ra: Ảnh đã tìm được biên.
- Các bước thực hiện:

**Bước 1:** Mở hình ảnh và chuyển đổi (*Transform* → *DWT2D RGB Transform*) ảnh với tỉ lệ 1, 2 hoặc 3. Có thể thêm vào các ngưỡng để loại bỏ nhiễu. Dưới đây, bộ lọc daub1 được chọn với 1 quy mô chuyển đổi mà không cần loại nhiễu:



Hình 5.2: Biến đổi ảnh với tỉ lệ 1 và bộ lọc daub1.

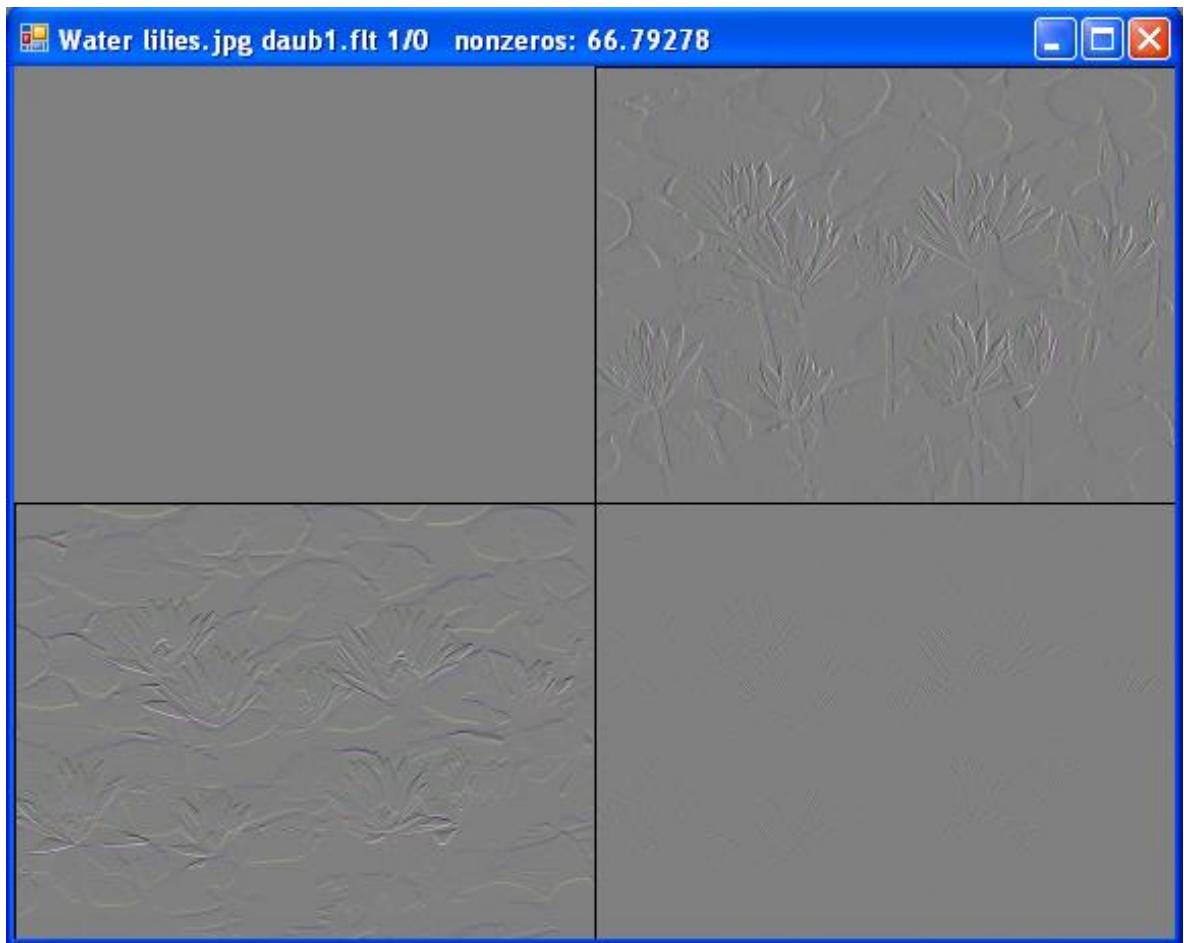
Ảnh biến đổi (co lại) sau khi chọn tỉ lệ 1 và bộ lọc daub1:



Hình 5.3: DWT hai chiều tín hiệu.



**Bước 2:** Loại bỏ thành phần tần số thấp (*Transform* → *Denoise*)



Hình 5.4: Loại bỏ thành phần tần số thấp.

Hàm tương ứng trong lớp BaseDWT2D:

```
void BaseDWT2D::remove_LLband()
{
    if (m_status <= 0)
        return;

    unsigned int width = m_width / (unsigned
int)(pow(2.0f, (float)getJ()));
    unsigned int height = m_height / (unsigned
int)(pow(2.0f, (float)getJ()));
```

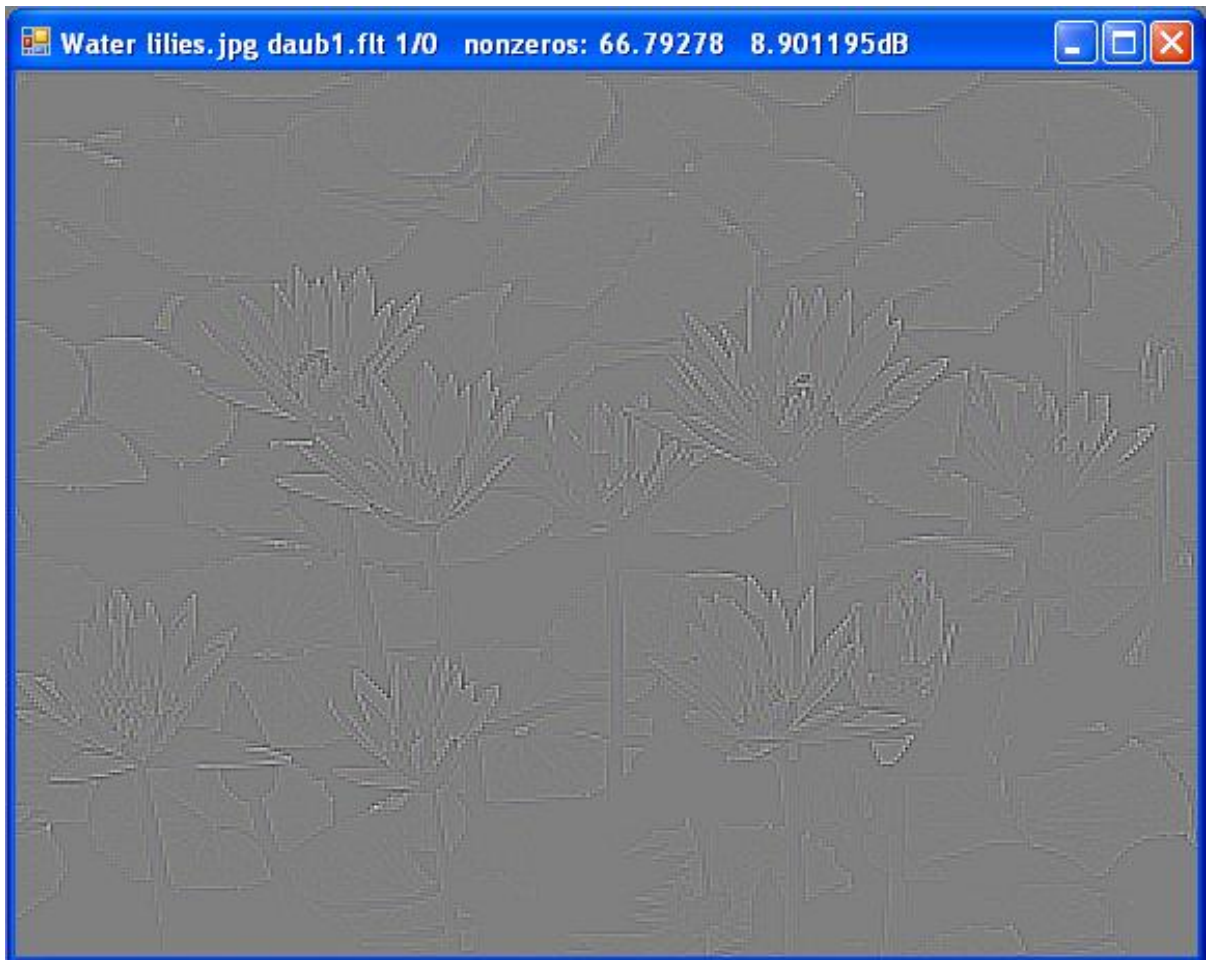
---

```

    for (unsigned int y = 0; y < height; y++)
        for (unsigned int x = 0; x < width; x++)
            spec2d[y][x] = 0;
}

```

**Bước 3:** Tái thiết lại hình ảnh (giãn ảnh với tỉ lệ 1). (*Transform* → *DWT2D RGB synth*)



Hình 5.5: Ảnh biến đổi sau khi xây dựng lại

**Bước 4:** Làm nổi biên (Abs values) bằng cách trừ đi 128 từ ảnh và tính giá trị tuyệt đối:

```

if (mr && mg && mb) {
    for (int i = 0; i < width * height; i++) {
        mr[i] = abs((int)mr[i] - 128);
    }
}

```



```

        mg[i] = abs((int)mg[i] - 128);
        mb[i] = abs((int)mb[i] - 128);
    }
    setRGB(m_clone, mr, mg, mb);
}

```



Hình 5.6: Ảnh sau khi được làm nổi biên.

#### **Bước 5:** Khuếch đại biên. (*Transform->Contrast stretch*)

Ảnh nhận đc đã nổi được biên, tuy nhiên biên còn chưa rõ ràng. Vì vậy ta trải dài độ tương phản – bình thường hóa ảnh vào phạm vi 0...255, nhưng có thể có 1 vài điểm ảnh ở giới hạn trên của phạm vi đó do vậy không thực sự cải thiện được tình hình. Sự lựa chọn tốt hơn là bình thường hóa phi tuyến tính giống như tỉ lệ logarit, nhưng ở đây, ta chỉ cần nhân các dữ liệu điểm ảnh bằng một số giá trị và có được biên rõ hơn. Đối với tỉ lệ biến đổi 1, ta nhân với 7 hoạt động tốt và không vượt khỏi giới hạn 255 cho phần lớn các điểm ảnh. Nhưng với tỉ lệ 2 hoặc 3 ta có thể giảm số lượng nhân.

```

void normalize(unsigned char* data, unsigned int
size, float a, float b) {

```

```

float min = (float)data[0];
float max = (float)data[0];
for (unsigned int i = 0; i < size; i++) {
    if ((float)data[i] < min)
        min = data[i];
    if ((float)data[i] > max)
        max = data[i];
}
for (unsigned int i = 0; i < size; i++) {
    float val = (float)data[i] * 7.0f;
    if (val < 0.0f)    val = 0.0f;
    if (val > 255.0f) val = 255.0f;
    data[i] = (unsigned char)val;
}
}

```



Hình 5.7: Biên ảnh sau khi được khuếch đại.

Với việc phát hiện biên sử dụng phương pháp Wavelet và sự kết hợp giữa chúng đạt được kết quả rất tốt. Khi sử dụng chức năng lọc nó cân bằng với việc nhận ra biên của phương pháp Canny. Thuật toán của sự biến đổi Wavelet giống sự biến đổi biên của Canny. Bằng những hình ảnh thực tế ta thấy phát hiện biên sử dụng Wavelet có thể thu được rất nhiều thông tin về biên.

Bên cạnh việc phát hiện tìm ra biên tốt thì phương pháp này cũng bộ lộ nhược điểm là mức độ tính toán phức tạp. Tốc độ xử lý chậm hơn so với các phương pháp đã trình bày.

## KẾT LUẬN

Trong quá trình tìm hiểu tài liệu và thực hiện luận văn dưới sự định hướng của thầy hướng dẫn luận văn đã đạt được một số kết quả như sau:

- Tìm hiểu được một cách tổng quan các vấn đề về XLA và phát hiện biên ảnh. Hệ thống hoá các phương pháp phát hiện biên. Đưa ra nhận xét, đánh giá các phương pháp phát hiện biên và có lựa chọn phương pháp phù hợp với từng loại ảnh.

- Đặc biệt việc phát hiện biên sử dụng phương pháp Wavelet và sự kết hợp giữa chúng đã phần nào đạt kết quả tốt hơn trong việc phát hiện biên nhưng chúng vẫn còn bộc lộ một số nhược điểm cần phải khắc phục.

- Bằng cách hiểu bản chất của từng phương pháp, em đã cài đặt và chạy thử nghiệm đối với một số phương pháp phát hiện biên.

- Ngoài ra, trong quá trình tìm hiểu em cũng tự tích lũy thêm cho mình các kiến thức về toán học, về kỹ thuật lập trình,... Tuy mới chỉ là bước đầu, nhưng những kết quả này sẽ giúp ích cho tôi trong những tìm hiểu sau này để thu được những kết quả tốt hơn. Dựa trên những kết quả bước đầu đã đạt được trong luận văn, tiếp tục tìm hiểu đề xuất một số cải tiến phương pháp phát hiện biên hiệu quả hơn trong tương lai.

- Đối với phương pháp Wavelet sẽ tiếp tục tìm hiểu để giảm độ phức tạp tính toán của phương pháp Wavelet. Vấn đề này có thể được giải quyết bởi việc chỉ sử dụng những cặp điểm liên quan nhất trong quá trình so sánh. Tiếp tục tìm hiểu các đặc điểm về hình dạng, các đặc điểm màu sắc và kết cấu để tăng hiệu quả cho đánh dấu tìm đường biên.

- Xây dựng một ứng dụng xử lý ảnh hoàn chỉnh dựa theo các phương pháp phát hiện biên đã trình bày trong luận văn

---

## TÀI LIỆU THAM KHẢO:

**Tiếng Việt:**

- [1]. Ts. Đỗ Năng Toàn & Ts. Phạm Việt Bình – “*Giáo trình môn học Xử lý ảnh*” – Đại Học Thái Nguyên – 2007.
- [2]. Lương Mạnh Bá & Nguyễn Thanh Thủy – “*Nhập môn Xử lý ảnh số*” – Nxb Khoa học và Kỹ thuật.
- [3]. Nguyễn Quang Sơn – “*Nghiên cứu một số phương pháp phát hiện biên*” – Luận văn thạc sĩ CNTT – Đại học Thái Nguyên.

Một số địa chỉ khác trên internet....

**Tiếng Anh:**

- [4]. John Canny – “*A computational approach to edge detection*” – Pattern Analysis and Machine Intelligence, IEEE Transactions on., PAMI-8(6):679–698, Nov. 1986
- [5]. O. Rebecca Vincent & Olusegun Folorunso – “*A Descriptive Algorithm for Sobel Image Edge Detection*” – Processings of Informing Science & IT Education Conference (InSITE) 2009
- [6]. <http://www.google.com.vn> + **Keyword:** *Canny Edge Detection 09gr820 March 23, 2009*
- [7]. Rastislav Lukac, Konstantinos N. Plataniotis – “*Color Image Processing Methods and Applications*” (2007) – University of Toronto; Toronto, Ontario, Canada – CRC Press, Taylor & Francis Group
- [8]. Plataniotis, Konstantinos N. – “*Comprehensive Analysis of Edge Detection in Color Image Processing*” (1999) – Electrical and Computer Engineering Publications and Research. Paper 1 – <http://digitalcommons.ryerson.ca/ee/1>
- [9]. Jun Li – “*A Wavelet Approach to Edge Detection*” (2003) – Master of Science (Mathematics & Statistics), Sam Houston State University. Huntsville, Texas.
- [10]. Chaganti, Venkata RaviKiran – “*Edge Detection of Noisy Images Using 2-D Discrete Wavelet Transform*” (2005) – Master of Science(Electrical & Computer Engineering), The Florida State University.
- [11]. <http://www.codeproject.com> + **Keyword:** *Edge Detection in Images with Wavelet Transform*