**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA CÔNG NGHỆ PHẦN MỀM**

**logo**

**KHÓA LUẬN TỐT NGHIỆP**

**Xây dựng ỨNG DỤNG NHẬN DẠNG BIỂN BÁO GIAO THÔNG TRÊN THIẾT BỊ DI ĐỘNG**

Giáo viên hướng dẫn: **ThS. PHAN NGUYỆT MINH**

Sinh viên thực hiện: **NGUYỄN BÁ CHUNG**

**ĐỖ TRƯỜNG GIANG**

Lớp : **CNPM02**

Khoá: **02**

***TP. Hồ Chí Minh, tháng 2 năm 2012***

# MỞ ĐẦU

Ngày nay với các tiến bộ của khoa học kỹ thuật thì mọi công việc hầu như đều có thể tiến hành trên máy tính một cách tự động hóa hoàn toàn hoặc một phần. Một trong những sự thay đổi lớn đó là cách thức chúng ta thu nhận và xử lý dữ liệu. Các công cụ nhập liệu như bàn phím hay máy scan dần bị thay thế bằng các thiết bị tiện lợi hơn như màn hình cảm ứng, camera…

Hơn thế nữa, các máy tính để bàn không còn là công cụ duy nhất có thể hỗ trợ cho con người. Chúng ta bước sang thế kỷ 21 với sự phát triển mạnh mẽ của các thiết bị di động, giải trí cầm tay hay smartphone. Với kích thước ngày càng nhỏ gọn và hiệu suất làm việc thì không ngừng được cải tiến, các công cụ mini này hứa hẹn sẽ là một phần không thể thiếu trong xã hội hiện đại. Và do đó, phát triển các ứng dụng trên các thiết bị này cũng là một xu thế tất yếu.

Công nghệ nhận dạng là một trong các công nghệ đang được áp dụng cho các thiết bị di động hiện nay. Nhận dạng có thể bao gồm nhận dạng âm thanh, hình ảnh. Các đối tượng nhận dạng có nhiều kiểu như tiếng nói, chữ viết, khuôn mặt, mã vạch … và biển báo giao thông cùng là một trong số đó. Chương trình nhận dạng biển báo giao thông thường phức tạp và được cài đặt trên những hệ thống có bộ xử lý lớn, camera chất lượng cao. Mục tiêu của khóa luận là cải tiến công nghệ nhận dạng này và mang nó cài đặt trên các thiết bị di động, giúp chúng ta phát hiện biển báo và nhận dạng nó một cách nhanh nhất.

Khóa luận **“Xây dựng ứng dụng nhận dạng biển báo giao thông trên thiết bị di động”** bao gồm tất cả 4 chương.

**Chương I: Giới thiệu:** Giới thiệu khái quát về khóa luận và mục đích của khóa luận.

**Chương II: Nền tảng và công nghệ:** Giới thiệu đầy đủ về các kiến thức nền tảng cũng như công nghệ và phần mềm được sử dụng trong khóa luận bao gồm kiến thức về xử lý ảnh, lý thuyết mạng noron, môi trường hệ điều hành dành cho di động Android, thư viện xử lý ảnh OpenCV.

**Chương III: Xây dựng ứng dụng:** Trình bày mô hình giải quyết bài toán nhận dạng trên thiết bị di động, các sơ đồ chức năng và thiết kế giao diện của chương trình.

**Chương IV: Đánh giá kết quả và kết luận:** Tổng kết quá trình thực hiện khóa luận và rút ra hướng phát triển sau này.

# LỜI CẢM ƠN

Trong suốt thời gian thực hiện khóa luận tốt nghiệp, chúng em đã nhận được sự giúp đỡ, chỉ bảo tận tình của các thầy cô Trường ĐH CNTT – ĐHQGTPHCM. Chúng em xin gửi lời cảm ơn sâu sắc đến quý thầy cô. Đặc biệt xin chân thành cảm ơn cô **Phan Nguyệt Minh** – người đã trực tiếp hướng dẫn và tạo mọi điều kiện thuận lợi giúp đỡ chúng em hoàn thành khóa luận này.

Chúng em cũng xin cảm ơn chân thành tới gia đình và bạn bè, công ty đã tạo điều kiện, giúp đỡ và động viên chúng em hoàn thành khóa luận đúng thời hạn.

Mặc dù đã cố gắng hết khả năng nhưng khóa luận không thể nào tránh khỏi những thiếu xót. Rất mong nhận được sự góp ý quý báu của quý thầy cô để khóa luận có thể hoàn chỉnh hơn.

Nhóm sinh viên thực hiện

**Nguyễn Bá Chung – Đỗ Trường Giang**

Tháng 2 – 2012

# NHẬN XÉT

**(Của giảng viên hướng dẫn)**

# NHẬN XÉT

**(Của giảng viên phản biện)**

# MỤC LỤC

# DANH MỤC BẢNG BIỂU

# DANH MỤC HÌNH VẼ

# 

# GIỚI THIỆU

Chương này trình bày các vấn đề sau:

* 1. Giới thiệu đề tài
  2. Mục tiêu của đề tài

C:\Users\HoaMua\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\O7N40L5F\MC900059919[1].gif

## GIỚI THIỆU ĐỀ TÀI

Ngày nay, những tiến bộ mới trong khoa học kỹ thuật công nghệ đã giúp ích rất nhiều cho cuộc sống của con người. Mọi thứ hầu như đều được tự động và hiệu suất công việc được nâng cao hơn với sự trợ giúp của máy móc, thiết bị. Một trong những công nghệ tiên tiến đang được áp dụng rộng rãi trong đời sống chính là công nghệ nhận dạng.



Hình 1.1 – Một số thiết bị hay được sử dụng trong nhận dạng

Nhận dạng dữ liệu bao gồm có nhận dạng âm thanh và nhận dạng hình ảnh. Các đối tượng của bài toán nhận dạng thì rất phong phú, ví dụ như nhận dạng khuôn mặt, tiếng nói, nhận dạng chữ viết tay, nhận dạng mã vạch … Biển báo giao thông cũng là một trong số đó. Đây là kiểu đối tượng có tính chất hình học đặc trưng, thường bắt gặp trong đời sống hằng ngày với công dụng là đưa ra những cảnh báo thông tin cho người tham gia giao thông. Tuy nhiên các biển cáo giao thông thì không có quy luật mà chỉ là hệ thống các ký hiệu với ý nghĩa qui ước kèm theo. Việc ghi nhớ hình dạng và ý nghĩa của tất cả các loại biển báo đối với chúng ta sẽ là một khó khăn lớn, do đó chúng ta thường hay có nhu cầu tra cứu tìm hiểu trực quan.

Bài toán nhận dạng nói chung và nhận dạng biển báo giao thông nói riêng hiện vẫn còn là một trong những chủ đề được các nhà khoa học nghiên cứu. Hiện tại đã có một số hệ thống tiên tiến của nước ngoài có khả năng nhận dạng biển báo giao thông nhưng hầu hết các hệ thống này đều đòi hỏi một khả năng xử lý mạnh mẽ, đi kèm với nó là camera có chất lượng cao.

Quay trở lại vấn đề, ngày nay máy tính không còn là công cụ trợ giúp độc tôn dành cho con người. Hầu hết chúng ta ai cũng biết đến sự phát triển mạnh mẽ của các loại thiết bị giải trí cầm tay nhỏ gọn. Đó chính là smartphone. Với ưu điểm là kích thước bé, đi kèm với nó là các chíp xử lý thông minh tốc độ cao, smartphone có khả năng đảm đương rất nhiều tác vụ giống y như đang thao tác trên máy tính. Phát triển phần mềm cho smartphone hiện cũng là xu thế tất yếu.



Hình 1.2 – Smartphone đang là xu hướng phát triển mới trong giai đoạn này

Ứng dụng công nghệ nhận dạng trên smartphone chính là ý tưởng mà nhóm hướng tới khi thực hiện khóa luận này. Bài toán nhóm sẽ giải quyết là làm thế nào xây dựng một hệ thống thông minh cho phép phát hiện và nhận dạng biển báo giao thông trên thiết bị di động.

## MỤC TIÊU CỦA ĐỀ TÀI

Mục tiêu của đề tài là nghiên cứu bài toán nhận dạng nói chung và nhận dạng biển báo giao thông nói riêng, từ đó cải tiến áp dụng cho việc xây dựng hệ thống trên các thiết bị di động. Mặc dù smartphone có những cải tiến vượt trội nhưng tất nhiên nhưng khác biệt về phần cứng như chip xử lý hay camera sẽ không thể so sánh với máy tính được. Do đó chương trình nhóm xây dựng sẽ tìm cách tối ưu hóa làm sao cho có thể tận dụng được những khả năng vốn có của smartphone.

Ngoài ra hệ thống biển báo giao thông của nước ta khá lớn, do đó nhóm sẽ xây dựng chương trình hoàn chỉnh nhưng sẽ thu nhỏ tập dữ liệu lại, coi đây như là một tập dữ liệu demo áp dụng cho khóa luận này.

Chương trình sẽ được xây dựng trên nền tảng Android, một trong những nền tảng di động phát triển mạnh nhất hiện nay. Ngoài ra nhóm sử dụng thư viện OpenCv hỗ trợ cho việc xử lý ảnh.



Hình 1.3 – Những nền tảng sẽ sử dụng trong khóa luận

# NỀN TẢNG VÀ CÔNG NGHỆ

Chương này trình bày các vấn đề sau:

* 1. Lý thuyết xử lý ảnh
  2. Lý thuyết mạng nơron
  3. Nền tảng Android
  4. Thư viện xử lý ảnh OpenCV

C:\Users\HoaMua\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\O7N40L5F\MC900059919[1].gif

## LÝ THUYẾT XỬ LÝ ẢNH

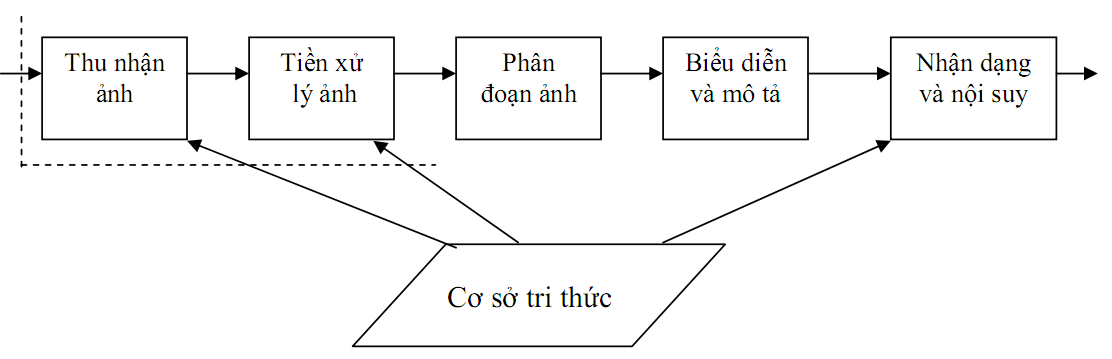
### Tổng quan về xử lý ảnh

#### 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. Những năm trở lại đây với sự phát triển của phần cứng máy tính, xử lý ảnh và đồ hoạ đã phát triển một cách mạnh mẽ và có nhiều ứng dụng trong cuộc sống. Xử lý ảnh và đồ hoạ đóng một vai trò quan trọng trong tương tác người máy.

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

Có thể hiểu một cách khác, xử lý ảnh hay cao cấp hơn nữa là thị giác máy tính (Computer Vision) bao gồm tất cả các lý thuyết và kỹ thuật liên quan, cho phép tạo lập một hệ thống có khả năng tiếp nhận thông tin từ các hình ảnh thu được, lưu trữ và xử lý theo nhu cầu.



Hình 2.1 – Các bước cơ bản trong xử lý ảnh

* ***Thu nhận ảnh***: Quá trình tiếp nhận thông tin từ vật thể thông qua camera màu hoặc trắng đen, ảnh thu nhận được có thể là ảnh tương tự hoặc ảnh đã số hóa.
* ***Tiền xử lý ảnh***: Sau bộ thu nhận, ảnh có thể nhiễu độ tương phản thấp nên cần đưa vào bộ tiền xử lý để nâng cao chất lượng. Chức năng chính của bộ tiền xử lý là lọc nhiễu, nâng độ tương phản để làm ảnh rõ hơn, nét hơn.
* ***Phân đoạn ảnh***: Là tách một ảnh đầu vào thành các vùng thành phần để biểu diễn phân tích, nhận dạng ảnh. Ví dụ: để nhận dạng chữ (hoặc mã vạch) trên phong bì thư cho mục đích phân loại bưu phẩm, cần chia các câu, chữ về địa chỉ hoặc tên người thành các từ, các chữ, các số (hoặc các vạch) riêng biệt để nhận dạng. Đây là phần phức tạp khó khăn nhất trong xử lý ảnh và cũng dễ gây lỗi, làm mất độ chính xác của ảnh. Kết quả nhận dạng ảnh phụ thuộc rất nhiều vào công đoạn này.
* ***Biểu diễn ảnh*:** Đầu ra ảnh sau phân đoạn chứa các điểm ảnh của vùng ảnh (ảnh đã phân đoạn) cộng với mã liên kết với các vùng lận cận. Việc biến đổi các số liệu này thành dạng thích hợp là cần thiết cho xử lý tiếp theo bằng máy tính. Việc chọn các tính chất để thể hiện ảnh gọi là trích chọn đặc trưng (Feature Selection) gắn với việc tách các đặc tính của ảnh dưới dạng các thông tin định lượng hoặc làm cơ sở để phân biệt lớp đối tượng này với đối tượng khác trong phạm vi ảnh nhận được. Ví dụ: trong nhận dạng ký tự trên phong bì thư, chúng ta miêu tả các đặc trưng của từng ký tự giúp phân biệt ký tự này với ký tự khác.
* ***Nhận dạng và nội suy ảnh*:** Nhận dạng ảnh là quá trình xác định ảnh. Quá trình này thường thu được bằng cách so sánh với mẫu chuẩn đã được học (hoặc lưu) từ trước. Nội suy là phán đoán theo ý nghĩa trên cơ sở nhận dạng. Ví dụ: một loạt chữ số và nét gạch ngang trên phong bì thư có thể được nội suy thành mã điện thoại.

#### Một số khái niệm trong xử lý ảnh

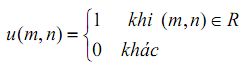
* ***Ảnh và điểm ảnh***: Gốc của ảnh (ảnh tự nhiên) là ảnh liên tục về không gian và độ sáng. Để xử lý bằng máy tính, ảnh cần phải được số hoá. Số hoá ảnh là sự biến đổi gần đúng một ảnh liên tục thành một tập điểm phù hợp với ảnh thật về vị trí (không gian) và độ sáng (mức xám). Khoảng cách giữa các điểm ảnh đó được thiết lập sao cho mắt người không phân biệt được ranh giới giữa chúng. Mỗi một điểm như vậy gọi là điểm ảnh và ảnh được xem như là 1 tập hợp các điểm ảnh
* ***Độ phân giải của ảnh***: Độ phân giải (Resolution) của ảnh là mật độ điểm ảnh được ấn định trên một ảnh số được hiển thị.
* ***Mức xám của ảnh***: Mức xám của điểm ảnh là cường độ sáng của nó được gán bằng giá trị số tại điểm đó. Giá trị mức xám thông thường: 16, 32, 64, 128, 256.
* ***Ảnh đen trắng*:** là ảnh có hai màu đen, trắng (không chứa màu khác) với mức xám ở các điểm ảnh có thể khác nhau.
* ***Ảnh nhị phân*:** là ảnh chỉ có 2 mức đen trắng phân biệt tức dùng 1 bit mô tả 21 mức khác nhau. Nói cách khác: mỗi điểm ảnh của ảnh nhị phân chỉ có thể là 0 hoặc 1.

### Một số phương pháp biểu diễn ảnh

Sau khi thu nhận và số hóa, ảnh sẽ được lưu trữ hay chuyển sang giai đoạn phân tích. Trước khi đề cập đến vấn đề lưu trữ ảnh, cần xem xét ảnh sẽ được biểu diễn ra sao trong bộ nhớ máy tính.

#### Mã loạt dài

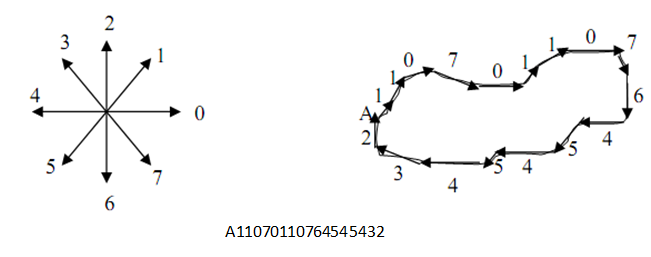
Mã loạt dài (Run-length Code) hay dùng để biểu diễn cho vùng ảnh hay ảnh nhị phân. Một vùng ảnh R có thể biểu diễn đơn giản nhờ một ma trận nhị phân:



Với các biểu diễn trên, một vùng ảnh hay ảnh nhị phân đựoc xem như chuỗi 0 hay 1 đan xen. Các chuỗi này được gọi là mạch (run). Theo phương pháp này, mỗi mạch sẽ được biểu diễn bởi địa chỉ bắt đầu của mạch và chiều dài mạch theo dạng {<hàng,cột>, chiều dài}.

#### Mã xích

Mã xích thường được dùng để biểu diễn biên của ảnh. Thay vì lưu trữ toàn bộ ảnh, người ta lưu trữ dãy các điểm ảnh như A, B…M. Theo phương pháp này, 8 hướng của vectơ nối 2 điểm biên liên tục được mã hóa. Khi đó ảnh được biểu diễn qua điểm ảnh bắt đầu A cùng với chuỗi các từ mã. Điều này được minh họa trong hình dưới đây:



Hình 2.2 – Hướng các điểm biên và mã tương ứng

#### Mã tứ phân

Theo phương pháp mã tứ phân, một vùng ảnh coi như bao kín một hình chứ nhật. Vùng này được chia làm 4 vùng con (Quadrant). Nếu một vùng con gồm toàn điểm đen (1) hay toàn điểm trắng (0) thì không cần chia tiếp. Trong trường hợp ngược lại, vùng con gồm cả điểm đen và trắng gọi là vùng không đồng nhất, ta tiếp tục chia thành 4 vùng con tiếp và kiểm tra tính đồng nhất của các vùng con đó. Quá trình chia dừng lại khi mỗi vùng con chỉ chứa thuần nhất điểm đen hoặc điểm trắng. Quá trình đó tạo thành một cây chia theo bốn phần gọi là cây tứ phân. Như vậy, cây biểu diễn ảnh gồm một chuỗi các ký hiệu b (black), w (white) và g (grey) kèm theo ký hiệu mã hóa 4 vùng con. Biểu diễn theo phương pháp này ưu việt hơn so với các phương pháp trên, nhất là so với mã loạt dài. Tuy nhiên, để tính toán số đo các hình như chu vi, mô men là tương đối khó khăn.

### Phương pháp phát hiện biên ảnh

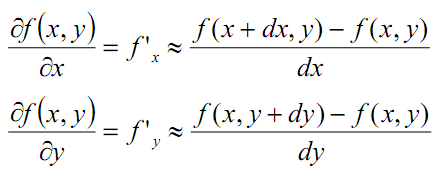
* Điểm Biên: Một điểm ảnh được coi là điểm biên nếu có sự thay đổi nhanh hoặc đột ngột về mức xám (hoặc màu). Ví dụ trong ảnh nhị phân, điểm đen gọi là điểm biên nếu lân cận nó có ít nhất một điểm trắng.
* Đường biên (đường bao: boundary): tập hợp các điểm biên liên tiếp tạo thành một đường biên hay đường bao.
* Ý nghĩa của đường biên : đường biên là một loại đặc trưng cục bộ tiêu biểu trong phân tích, nhận dạng ảnh. Nngười ta sử dụng biên làm phân cách các vùng xám (màu) cách biệt.

#### Phát hiện biên trực tiếp

Phương pháp này làm nổi biên dựa vào sự biến thiên mức xám của ảnh. Kỹ thuật chủ yếu dùng để phát hiện biên ở đây là kỹ thuật lấy đạo hàm. Nếu lấy đạo hàm bậc nhất của ảnh ta có các kỹ thuật Gradient, nếu lấy đạo hàm bậc hai của ảnh ta có kỹ thuật Laplace. Ngoài ra còn có một số cách tiếp cận khác.

##### **Kỹ thuật phát hiện biên Gradient**

Gradient là một vec tơ f(x, y) có các thành phần biểu thị tốc độ thay đổi mức xám của điểm ảnh (theo hai hướng x, y trong bối cảnh xử lý ảnh hai chiều)

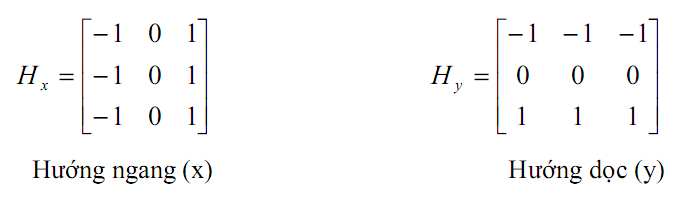


Trong đó, dx, dy là khoảng cách (tính bằng số điểm) theo hướng x và y. Tuy ta nói là lấy đạo hàm nhưng thực chất chỉ là mô phỏng và xấp xỉ đạo hàm bằng các kỹ thuật nhân chập vì ảnh số là tín hiệu rời rạc nên đạo hàm không tồn tại (thực tế chọn dx= dy=1).

Theo định nghĩa về Gradient, nếu áp dụng nó vào xử lý ảnh, việc tính toán sẽ rất phức tạp. Để đơn giản mà không mất tính chất của phương pháp Gradient, người ta sử dụng kỹ thuật Gradient dùng cặp mặt nạ H1, H2 trực giao (theo 2 hướng vuông góc).

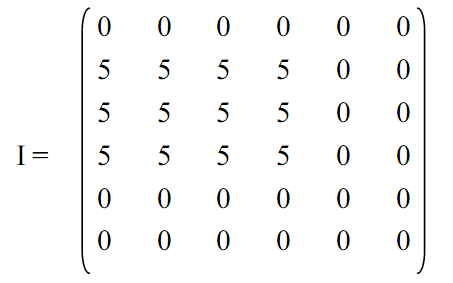
* ***Mặt nạ Prewitt***

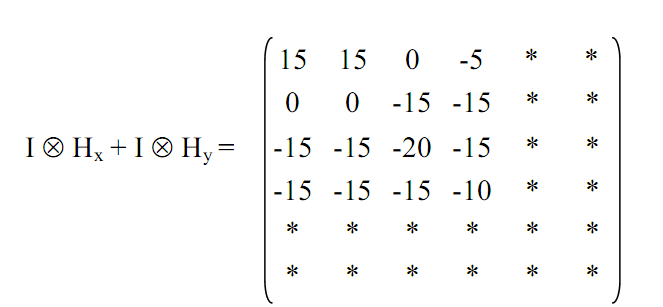
- Kỹ thuật sử dụng 2 mặt nạ nhập chập xấp xỉ đạo hàm theo 2 hướng x và y là:



- Tính I ⊗ Hx + I ⊗ Hy  để ra được kết quả

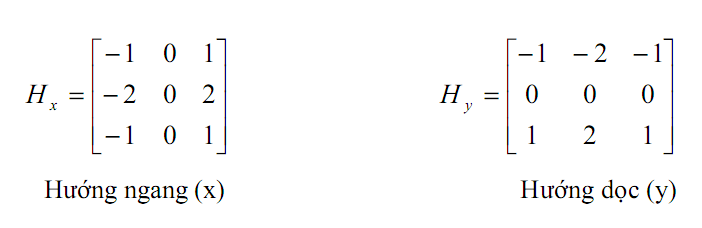
- Ví dụ:





* ***Mặt nạ Sobel***

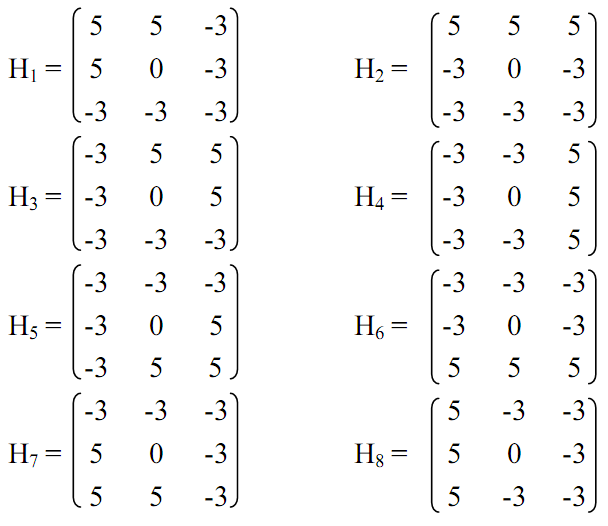
- Kỹ thuật sử dụng 2 mặt nạ nhân chập xấp xỉ đạo hàm theo 2 hướng x và y là:



- Tính I ⊗ Hx + I ⊗ Hy để ra được kết quả.

* ***Kỹ thuật la bàn***

- Kỹ thuật sử dụng 8 mặt nạ nhân chập theo 8 hướng 00, 450, 900, 1350, 1800, 2250, 2700, 3150.



- Kết quả thu được bằng cách tính 

##### **Kỹ thuật phát hiện biên Laplace**

Toán tử Laplace được định nghĩa như sau:

Ta có:







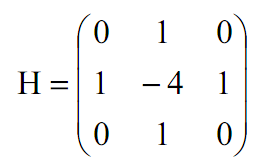
Tương tự



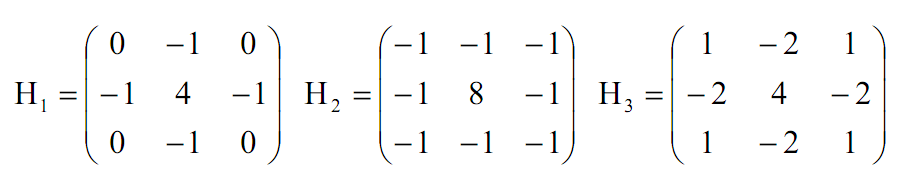


Vậy: ∇2 f= f(x+1,y) + f(x,y+1) - 4f(x,y) + f(x-1,y) + f(x,y-1)

Dẫn tới



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



##### **Kỹ thuật phát hiện biên Canny**

Đây là một thuật toán tương đối tốt, có khả năng đưa ra đường biên mảnh, và phát hiện chính xác điểm biên với điểm nhiễu.

Ta có thuật toán như sau:

* **Bước 1**: Làm trơn ảnh

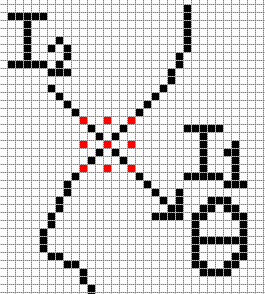
Tính I ⊗ H, với:



Gọi G là kết quả lọc nhiễu: G= I ⊗ H

* **Bước 2**: Tính gradient của ảnh bằng mặt nạ Prewitt, kết quả đặt vào Gx,Gy.

Gx = G ⊗ Hx, Gy = G ⊗ Hy

* **Bước 3**: Tính gradient hướng tại mỗi điểm (i,j) của ảnh. Hướng này sẽ được nguyên hóa để nằm trong 8 hướng [0..7], tương đương với 8 lân cận của một điểm ảnh.
* **Bước 4**: Dùng ràng buộc “loại bỏ những điểm không phải là cực đại” để xóa bỏ những điểm không là biên. Xét (i,j), θ là gradient hướng tại (i,j). I1, I2 là hai điểm lân cận của (i,j) theo hướng θ. Theo định nghĩa điểm biên cục bộ thì (i,j) là biên nếu I(i,j) cực đại địa phương theo hướng gradient 🡪 Nếu I(i,j) > I1 và I(i,j) > I2 thì mới giữ lại I(i,j), ngược lại xóa I(i,j) về điểm ảnh nền.

Hình 2.3 – Minh họa xác định điểm biên

* **Bước 5**: Phân ngưỡng. Với các điểm được giữ lại, thực hiện lấy ngưỡng gradient biên độ lần cuối để xác định các điểm biên thực sự.

#### Phát hiện biên gián tiếp

Nếu bằng cách nào đấy, chúng ta thu đượng các vùng ảnh khác nhau thì đường phân cách giữa các vùng đó chính là biên. Nói cách khác, việc xác định đường bao của ảnh được thực hiện từ ảnh đã được phân vùng. Phương pháp dò biên gián tiếp khó cài đặt nhưng áp dụng tốt khi sự biến thiên độ sáng nhỏ. Để có thể tiến hành xác định biên theo cách gián tiếp này, chúng ta cần giải quyết được bài toán phân vùng ảnh.

### Phân vùng ảnh

Phân vùng ảnh là bước then chốt trong xử lý ảnh. Giai đoạn này nhằm phân tích ảnh thành những thành phần có cùng tính chất nào đó dựa theo biên hay các vùng liên thông. Tiêu chuẩn để xác định các vùng liên thông có thể là cùng mức xám, cùng màu hay cùng độ nhám…

Vùng ảnh là một chi tiết, một thực thể trông toàn cảnh. Nó là một tập hợp các điểm có cùng hoặc gần cùng một tính chất nào đó : mức xám, mức màu, độ nhám… Vùng ảnh là một trong hai thuộc tính của ảnh. Nói đến vùng ảnh là nói đến tính chất bề mặt. Đường bao quanh một vùng ảnh (Boundary) là biên ảnh. Các điểm trong một vùng ảnh có độ biến thiên giá trị mức xám tương đối đồng đều hay tính kết cấu tương đồng.

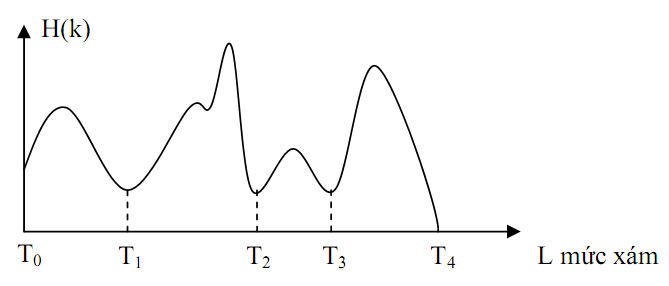
Dựa vào đặc tính vật lý của ảnh, người ta có nhiều kỹ thuật phân vùng : phân vùng dựa theo miền liên thông gọi là phân vùng dựa theo miền đồng nhất hay miền kề, phân vùng dựa vào biên gọi là phân vùng biên. Ngoài ra còn có các kỹ thuật phân vùng khác dựa vào biên độ, phân vùng dựa theo kết cấu.

#### Phân vùng theo ngưỡng biên độ

Đặc tính đơn giản nhất và có thể hữu ích nhất của ảnh đó là biên độ của các tính chất vật lý của ảnh như: độ tương phản, độ truyền sáng, màu sắc hoặc quang phổ.

Như vậy, có thể dùng ngưỡng biên độ để phân vùng khi biên độ đủ lớn đặc trưng cho ảnh. Thí dụ, biên độ trong bộ cảm biến ảnh hồng ngoại có thể phản ánh vùng có nhiệt độ thấp hay vùng có nhiệt độ cao. Kỹ thuật phân ngưỡng theo biên độ rất có lợi đối với ảnh nhị phân như văn bản in, đồ họa, ảnh màu hay ảnh X-quang.

Việc chọn ngưỡng rất quan trọng. Nó bao gồm các bước :

* Xem xét lược đồ xám của ảnh để xác định các đỉnh và các khe. Nếu ảnh có dạng rắn lượn (nhiều đỉnh và khe), các khe có thể dùng để chọn ngưỡng.
* Chọn ngưỡng t sao cho một phần xác định trước η của toàn bộ số mẫu là thấp hơn t.
* Điều chỉnh ngưỡng dựa trên lược đồ xám của các điểm lân cận.
* Chọn ngưỡng theo lược đồ xám của những điểm thỏa mãn tiêu chuẩn chọn. Thí dụ, với ảnh có độ tương phản thấp, lược đồ của những điểm có biên độ Laplace g(m,n) lớn hơn giá trị t định trước (sao cho từ 5% đến 10% số điểm ảnh với Gradient lớn nhất sẽ coi như biên) sẽ cho phép xác định các đặc tính ảnh lưỡng cực tốt hơn ảnh gốc.

Ta xét ví dụ sau về việc phân vùng dựa trên ngưỡng biên độ:

Hình 2.4 – Lược đồ rắn lượn và cách chọn ngưỡng

Giả sử ảnh có lược đồ xám và cách chọn các ngưỡng như hình trên với: T0=Lmin, …, T4=Lmax. Ta có 5 ngưỡng và phân ảnh thành 4 vùng, ký hiệu Ck là vùng thứ k của ảnh, k=1,2,3,4. Cách phân vùng theo nguyên tắc :

*P(m,n) ∈ Ck nếu Tk-1 ≤ P(m,n) < Tk , k=1,2,3,4.*

Khi phân vùng xong, nếu ảnh rõ nét thì việc phân vùng coi như kết thúc. Nếu không, cần điều chỉnh ngưỡng.

#### Phân vùng theo miền đồng nhất

Kỹ thuật phân vùng ảnh thành các miền đồng nhất dựa vào các tính chất quan trọng nào đó của miền ảnh. Việc lựa chọn các tính chất của miền sẽ xác định tiêu chuẩn phân vùng. Tính đồng nhất của một miền ảnh là điểm chủ yếu xác định tính hiệu quả của việc phân vùng. Các tiêu chuẩn hay được dùng là sự thuần nhất về mức xám, màu sắc đối với ảnh màu, kết cấu sợi và chuyển động.

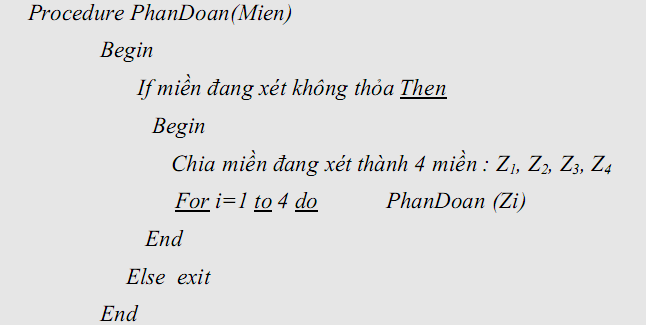
Các phương pháp phân vùng ảnh theo miền đồng nhất thường áp dụng là :

* Phương pháp tách cây tứ phân
* Phương pháp cục bộ
* Phương pháp tổng hợp

##### **Phương pháp tách cây tứ phân**

Về nguyên tắc, phương pháp này kiểm tra tính đúng đắn của tiêu chuẩn đề ra một cách tổng thể trên miền lớn của ảnh. Nếu tiêu chuẩn được thỏa mãn, việc phân đoạn coi như kết thúc. Trong trường hợp ngược lại, chia miền đang xét thành 4 miền nhỏ hơn. Với mỗi miền nhỏ, áp dụng một cách đệ quy phương pháp trên cho đến khi tất cả các miền đều thỏa mãn điều kiện.

Phương pháp này có thể mô tả bằng thuật toán sau :



Tiêu chuẩn xét miền đồng nhất ở đây có thể dựa vào mức xám. Ngoài ra, có thể dựa vào độ lệch chuẩn hay độ chênh giữa giá trị mức xám lớn nhất và giá trị mức xám nhỏ nhất. Giả sử Max và Min là giá trị mức xám lớn nhất và nhỏ nhất trong miền đang xét.

Nếu |Max – Min| < T (ngưỡng) ta coi miền đang xét là đồng nhất. Trường hợp ngược lại, miền đang xét không là miền đồng nhất và sẽ được chia làm 4 phần.

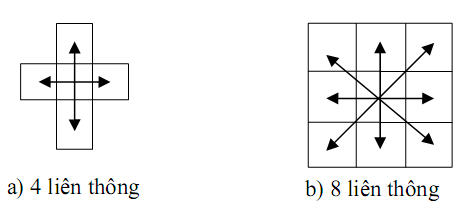
##### **Phương pháp cục bộ**

Ý tưởng của phương pháp là xét ảnh từ các miền nhỏ nhất rồi nối chúng lại nếu thỏa mãn tiêu chuẩn để được một miền đồng nhất lớn hơn. Tiếp tục với các miền thu được cho đến khi không thể nối thêm được nữa. Số miền còn lại cho ta kết quả phân đoạn. Như vậy, miền nhỏ nhất của bước xuất phát là điểm ảnh.

Phương pháp này hoàn toàn ngược với phương pháp tách. Song điều quan trọng ở đây là nguyên lý nối 2 vùng. Việc nối 2 vùng được thực hiện theo nguyên tắc sau :

* Hai vùng phải đáp ứng tiêu chuẩn, thí dụ như cùng màu hay cùng mức xám.
* Hai vùng phải kế cận nhau.

Trong xử lý ảnh, người ta dùng khái niệm liên thông để xác định tính chất kế cận. Có hai khái niệm về liên thông là 4 liên thông và 8 liên thông. Với 4 liên thông một điểm ảnh I(x,y) sẽ có 4 kế cận theo 2 hướng x và y ; trong khi đó với 8 liên thông, điểm I(x,y) sẽ có 4 liên thông theo 2 hướng x, y và 4 liên thông khác theo hướng chéo 45o



Hình 2.5 – Minh họa khái niệm liên thông

Dựa theo nguyên lý của phương pháp nối, ta có 2 thuật toán :

* Thuật toán tô màu (Blob Coloring) : sử dụng khái niệm 4 liên thông, dùng một cửa sổ di chuyển trên ảnh để so sánh với tiêu chuẩn nối.
* Thuật toán đệ quy cục bộ: sử dụng phương pháp tìm kiếm trong một cây để làm tăng kích thước vùng.

##### **Phương pháp tổng hợp**

Hai phương pháp nối (hợp) và tách đều có nhược điểm. Phương pháp tách sẽ tạo nên một cấu trúc phân cấp và thiết lập mối quan hệ giữa các vùng. Tuy nhiên, nó thực hiện việc chia quá chi tiết. Phương pháp hợp cho phép làm giảm số miền liên thông xuống tối thiểu, nhưng cấu trúc hàng ngang dàn trải, không cho ta thấy rõ mối liên hệ giữa các miền.

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

#### Phân vùng theo kết cấu bề mặt

Kết cấu thường được nhận biết trên bề mặt của các đối tượng như gỗ, cát, vải vóc…Kết cấu là thuật ngữ phản ánh sự lặp lại của các phần tử sợi (texel) cơ bản. Sự lặp lại này có thể ngẫu nhiên hay có tính chu kì hoặc gần chu kì. Một texel chứa rất nhiều điểm ảnh. Trong phân tích ảnh, kết cấu được chia làm hai loại chính là: loại thống kê và loại cấu trúc.

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

Nhìn chung, việc phân loại và phân vùng dựa vào kết cấu là một vấn đề phức tạp. Trong thực tế, chúng ta thường chỉ giải quyết vấn đề này bằng cách cho biết trước các loại kết cấu (dựa vào quy luật hay các phân bố của nó).

### Nhận dạng ảnh

Nhận dạng ảnh là giai đoạn cuối của các hệ thống xử lý ảnh. Nhận dạng là quá trình phân loại các đối tượng được biểu diễn theo một mô hình nào đó và gán chúng một tên (gán cho đối tượng một tên gọi, tức là một dạng) dựa theo những quy luật và mẫu chuẩn.

Trong lý thuyết về nhận dạng nói chung và nhận dạng ảnh nói riêng có ba cách tiếp cận khác nhau:

* Nhận dạng dựa vào phân hoạch không gian.
* Nhận dạng dựa vào cấu trúc.
* Nhận dạng dựa vào kỹ thuật mạng nơron.

**Học có thầy**: kỹ thuật phân loại nhờ kiến thức biết trước gọi là học có thầy. Đặc điểm cơ bản của kỹ thuật này là người ta có một thư viện các mẫu chuẩn. Mẫu cần nhận dạng sẽ được đem so sánh với mẫu chuẩn để xem nó thuộc loại nào. Vấn đề chủ yếu là thiết kế một hệ thống để có thể đối sánh đối tượng trong ảnh với mẫu chuẩn và quyết định gán cho chúng vào một lớp. Việc đối sánh nhờ vào các thủ tục ra quyết định dựa trên một công cụ gọi là hàm phân lớp hay hàm ra quyết định.

**Học không có thầy**: kỹ thuật này phải tự định ra các lớp khác nhau và xác định các tham số đặc trưng cho từng lớp. Học không có thầy đương nhiên là gặp khó khăn hơn. Một mặt, do số lớp không được biết trước, mặt khác những đặc trưng của lớp cũng không được biết trước. Kỹ thuật này nhằm tiến hành mọi cách gộp nhóm có thể và chọn lựa cách tốt nhất. Bắt đầu từ tập dữ liệu, nhiều thủ tục xử lý khác nhau nhằm phân lớp và nâng cấp dần để đạt được một phương án phân loại.

## LÝ THUYẾT MẠNG NƠRON