

**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC BÁCH KHOA**

**KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH**



**THỰC TẬP TỐT NGHIỆP**

**NHẬN DẠNG LÁ CÂY THUỐC BẰNG THIẾT BỊ DI ĐỘNG**

**GVHD: Ths. Văn Đức Sơn Hà**

**Ths. Đặng Trần Trí**

**SVTH : Nguyễn Văn Hai (51000798)**

**Lý Ngọc Hiếu (51004182)**

**Tp Hồ Chí Minh 06/2014**

## LỜI CẢM ƠN

*Đầu tiên chúng tôi xin chân thành cảm ơn khoa Khoa Học Và Kỹ Thuật Máy Tính, trường Đại Học Bách Khoa, Đại Học Quốc Gia Tp Hồ Chí Minh đã tạo điều kiện cho chúng tôi trong suốt quá trình nghiên cứu và thực hiện đề tài. Chúng tôi xin chân thành các thầy cô trong khoa Khoa Học Và Kỹ Thuật Máy Tính đã tận tình giảng dạy và cung cấp những kiến thức quý báu trong những năm học vừa qua.*

*Chúng tôi xin gửi lời cảm ơn chân thành và sâu sắc nhất đến thầy Vân Đức Sơn Hà và thầy Đặng Trần Trí, giảng viên hướng dẫn trực tiếp đề tài. Các thầy đã tận tình chỉ dẫn, theo sát, sửa chữa và góp ý để chúng tôi thực hiện đề tài này.*

*Xin chân thành cảm ơn đến gia đình, bạn bè đã sát cánh bên cạnh chúng tôi, giúp đỡ về vật chất cũng như tinh thần trong suốt quá trình thực hiện đề tài này.*

*Mặc dù đã rất cố gắng song không thể tránh những sai sót, rất mong được sự góp ý chân thành của thầy cô và các bạn.*

*Cuối cùng xin chân thành cảm ơn thầy cô và các bạn đã dành thời gian đọc tài liệu này.*

*Tp Hồ Chí Minh tháng 6 năm 2014*

## TÓM TẮT NỘI DUNG

Trong giai đoạn thực tập đề tài chủ yếu bao gồm các nội dung : giới thiệu đề tài, ý nghĩa khoa học và thực tiễn. Tìm hiểu xử lý ảnh và một vài phương pháp xử lý ảnh. Tìm hiểu về các đặc trưng, tính chất của lá. Tìm hiểu về các kỹ thuật liên quan đến nhận dạng : Phân lớp ảnh chụp lá cây bằng phương pháp máy vector hỗ trợ SVM (support vector machine), kỹ thuật PCA (Principal Components Analysis). Demo sơ lược ứng dụng nhận dạng lá cây thuốc trên thiết bị di động. Các hướng phát triển trong tương lai.

# MỤC LỤC

LỜI CẢM ƠN .....	2
TÓM TẮT NỘI DUNG .....	3
MỤC LỤC HÌNH .....	6
DANH SÁCH THUẬT NGỮ .....	7
CHƯƠNG 1 GIỚI THIỆU .....	8
1.1 Giới thiệu đề tài, ý nghĩa khoa học thực tiễn .....	8
1.1.1 Ý nghĩa khoa học thực tiễn .....	8
1.1.2 Giới thiệu đề tài.....	8
1.2 Đối tượng và mục đích nghiên cứu .....	9
1.2.1 Đối tượng .....	9
1.2.2 Mục đích .....	9
CHƯƠNG 2 CƠ SỞ LÝ THUYẾT .....	9
2.1 Xử lý ảnh và một số khái niệm .....	9
2.1.1 Xử lý ảnh.....	9
2.1.2 Các vấn đề cơ bản trong xử lý ảnh .....	10
2.1.3 Thu nhận và biểu diễn ảnh.....	13
2.1.3.1 Thu nhận, các thiết bị thu nhận ảnh .....	13
2.1.3.2 Biểu diễn ảnh .....	13
2.2 Các phương pháp trích chọn đặc trưng .....	15
2.2.1 Khái quát đặc trưng ảnh.....	15
2.2.1.1 Đặc trưng văn bản đi kèm ảnh .....	15
2.2.1.2 Đặc trưng nội dung ảnh.....	16
2.2.2 Các phương pháp trích chọn đặc trưng.....	17
2.2.2.1 Đặc trưng màu sắc .....	18
2.2.2.2 Đặc trưng kết cấu .....	19
2.2.2.3 Đặc trưng hình dạng.....	20
2.2.2.4 Đặc trưng cục bộ bất biến .....	20
CHƯƠNG 3 CÁC NGHIÊN CỨU LIÊN QUAN ĐẾN NHẬN DẠNG .....	25
3.1 Phân lớp ảnh chụp lá cây bằng phương pháp máy vector hỗ trợ SVM (support vector	

machine) .....	25
3.1.1 Giới thiệu SVM .....	25
3.1.2 Phương pháp phân lớp dữ liệu máy vector hỗ trợ (SVM) .....	27
3.1.3 Mô hình phân lớp ảnh chụp lá cây.....	28
3.1.3.1 Giai đoạn tiền xử lý.....	29
3.1.3.2 Giai đoạn rút trích đặc trưng.....	29
3.1.3.3 Chọn thuật toán huấn luyện và phân lớp dữ liệu .....	30
3.1.4 Thực nghiệm.....	30
3.1.5 Kết luận.....	32
3.2 Kỹ thuật PCA (Principal Components Analysis).....	32
3.2.1 Giới thiệu.....	32
3.2.2 Nội dung PCA .....	33
3.2.3 Quá trình rút trích .....	34
CHƯƠNG 4 HIỆN THỰC DEMO.....	38
4.1 Rút trích các đặc trưng của lá.....	38
4.1.1 Rút trích đường viền của lá .....	38
4.1.2 Rút trích các đặc trưng hình học.....	39
4.1.3 Chuẩn hóa dữ liệu.....	42
4.1.4 Tính độ tương đồng .....	42
4.2 Demo .....	43
CHƯƠNG 5 TỔNG KẾT .....	52
5.1 Đánh giá .....	52
5.2 Hướng phát triển luận văn.....	52
TÀI LIỆU THAM KHẢO .....	52

## MỤC LỤC HÌNH

Hình 2.1 Quá trình xử lý ảnh .....	10
Hình 2.2 Các bước cơ bản trong một hệ thống xử lý ảnh .....	10
Hình 2.3 Quá trình hiển thị và chỉnh sửa, lưu trữ ảnh thông qua DIB.....	14
Hình 2.4 Sự biến đổi giữa các mô hình biểu diễn ảnh .....	15
Hình 2.5 Ví dụ hiển thị một ảnh .....	16
Hình 2.6 Ví dụ về một số loại kết cấu.....	17
Hình 2.7 Biểu đồ mô phỏng việc tính toán các DoG ảnh từ các ảnh kê mờ.....	22
Hình 2.8 Mỗi điểm ảnh được so sánh với 26 láng giềng của nó.....	23
Hình 2.9 Quá trình lựa chọn các điểm hấp dẫn. (a) Ảnh gốc. (b) Các điểm hấp dẫn được phát hiện. (c) Ảnh sau khi loại bỏ các điểm hấp dẫn có độ tương phản thấp. (d) Ảnh sau khi loại bỏ các điểm dọc theo cạnh .....	24
Hình 2.10 Biểu diễn các vector đặc trưng.....	25
Hình 3.1 Ví dụ siêu phẳng với lẽ cực đại trong không gian $R^2$ .....	27
Hình 3.2 Mô hình phân lớp ảnh chụp có phải là lá cây .....	29
Hình 3.3 Giai đoạn tiền xử lý(tách ảnh khỏi nền).....	31
Hình 3.4 Giao diện chương trình .....	31
Hình 3.5 Tập dữ liệu chọn để kiểm tra ứng dụng .....	32
Hình 3.6 Quá trình rút trích Eigenface.....	35
Hình 3.7 Chuẩn hóa ma trận ảnh .....	36
Hình 3.8 Trung bình ảnh của các tập dữ liệu học .....	37
Hình 4.1 Rút trích đường viền lá (a) ảnh đầu vào, (b) ảnh xám, (c) ảnh nhị phân, (d) ảnh sau khi rút trích đường viền.....	39
Hình 4.2 a) MBR b) Convex Hull c) Incircle and Excircle .....	39
Hình 4.3 Cơ sở dữ liệu của demo .....	43
Hình 4.4 Màn hình chính của demo.....	44
Hình 4.5 Mở tập dữ liệu ảnh .....	45
Hình 4.6 Chọn vùng lá .....	46
Hình 4.7 Rút trích đường viền .....	47
Hình 4.8 Diện tích Hull.....	48
Hình 4.9 Chu vi Hull.....	49
Hình 4.10 Các đặc trưng của lá.....	50
Hình 4.11 Tính năng Search .....	51

## DANH SÁCH THUẬT NGỮ

Thuật ngữ	Tên đầy đủ	Ghi chú
DIB	Device Independent Bitmap	
RGB	Red Green Blue	
JSD	Jensen-Shannon divergence	
SIFT	Scale Invariant Feature Transform	Đặc trưng cục bộ bất biến
SVM	Support vector machine	Máy vector hỗ trợ
PCA	Principal Components Analysis	
MBR	Minimum Bounding Rectangle	

# CHƯƠNG 1 GIỚI THIỆU

## 1.1 Giới thiệu đề tài, ý nghĩa khoa học thực tiễn

### 1.1.1 Ý nghĩa khoa học thực tiễn

Theo tổ chức Y tế thế giới (WHO) đánh giá cho đến nay 80% dân số trên thế giới dựa vào nền y học cổ truyền để đáp ứng cho nhu cầu chăm sóc sức khỏe ban đầu, trong đó chủ yếu là thuốc từ cây cỏ. Sự quan tâm về các hệ thống y học cổ truyền và đặc biệt là các loại thuốc dược thảo, thực tế là đã ngày càng gia tăng tại các nước phát triển và đang phát triển trong hơn hai thập kỷ qua.

Ở Việt Nam, thực vật rất phong phú và đa dạng. Tổng số loài thực vật đã ghi nhận cho Việt Nam là 10.500 loài, ước đoán hệ thực vật Việt Nam có khoảng 12.000 loài. Trong số này, nguồn tài nguyên cây làm thuốc chiếm khoảng 30%. Kết quả điều tra nguồn tài nguyên dược liệu ở Việt Nam giai đoạn 2001-2005 của Viện Dược liệu (2006) cho biết ở Việt Nam có 3.948 loài thực vật bậc cao, bậc thấp và nấm lớn được dùng làm thuốc. Trong đó nhóm thực vật bậc cao có mạch có 3.870 loài. Những cây thuốc có giá trị sử dụng cao, có khả năng khai thác trong tự nhiên là những cây thuốc nằm trong danh mục 185 cây thuốc và vị thuốc thiết yếu của Bộ Y tế cũng như những cây thuốc đang được thị trường dược liệu quan tâm gồm có 206 loài cây thuốc có khả năng khai thác.

Hiện nay người ta có xu hướng quay trở về với cây thuốc và thuốc có nguồn gốc thiên nhiên tạo ra hơn là hóa chất làm thuốc. Xu hướng này đã tác động đến việc sản xuất, thu hái, chế biến, lưu thông, tiêu thụ và sử dụng dược liệu thảo mộc. Trong khi các tài liệu tra cứu về cây thuốc chủ yếu được viết trên sách, do đó hạn chế đối tượng sử dụng nhất là không phải là nhà chuyên môn muốn tìm hiểu sử dụng cây thuốc. Nhiều cây thuốc mà dân gian có thể bị nhầm lẫn trong sự xác định loài dựa theo tên phổ thông hay những loài có hình dạng giống nhau, rất dễ nhầm lẫn nếu thiếu sự mô tả tỷ mỉ đặc điểm hình thái và giải phẫu.

Với các xu hướng trên thì việc nhận dạng các loại cây thuốc ngày càng đóng vai trò rất quan trọng trong các lĩnh vực xã hội hiện nay (nói chung) và y học (nói riêng). Việc tìm ra các loại cây thuốc thông qua các công nghệ hiện đại sẽ giúp ít rất nhiều trong việc nghiên cứu và phát triển về y học, bảo tồn và phát triển các loại thuốc quý, giúp phân biệt tránh sự nhầm lẫn giữa các loại cây thuốc với các loại cây dại khác.

### 1.1.2 Giới thiệu đề tài

Với sự phát triển vượt bậc của khoa học kỹ thuật như ngày nay thì hàng loạt các công nghệ ngày càng được hoàn thiện và phát triển. Trong đó có công nghệ xử lý ảnh và các công nghệ có liên quan đến nhận dạng ngày càng phát triển. Trong các thiết bị được tập trung phát triển về khả năng nhận dạng thì các thiết bị cảm tay được ưu tiên phát triển



hơn cả. Vì tính di động và linh hoạt cao của các thiết bị cầm tay sẽ giúp cho các công nghệ nhận dạng phát huy được tối đa khả năng.

Đề tài sẽ tìm hiểu các phương pháp nhận dạng các lá cây thông dụng. Tập trung nghiên cứu phương pháp nhận dạng lá cây thuốc qua hình dạng(shape) của lá. Các phương pháp sẽ được phát triển thành ứng dụng và được tích hợp trong các thiết bị cầm tay(đặc trưng là các smartphome) sẽ giúp ít rất nhiều trong việc nhận dạng các loại cây thuốc, hiển thị thông tin các loại cây và thêm các loại cây thuốc mới.

## **1.2 Đối tượng và mục đích nghiên cứu**

### **1.2.1 Đối tượng**

Các loại cây thực vật(bao gồm những loài cây dại và những loại cây được sử dụng làm thuốc). Đặc trưng là các lá của cây. Đề tài sẽ tập trung vào phân tích và xử lý hình ảnh lá của các loại cây thực vật để xác định đó có thuộc một loại cây thuốc hay không? Hiển thị các thông tin cần thiết của một loại cây thuốc nào đó.

### **1.2.2 Mục đích**

Với ý nghĩa khoa học thực tiễn, đề tài tập trung nghiên cứu tìm ra phương pháp nhận dạng các loại thực vật(tập trung vào phần lá của các loại cây đó) để giúp cho việc nhận dạng và phân loại trở nên dễ dàng hơn, tiện ích hơn.

Việc các phương pháp nhận dạng được phát triển thành các ứng dụng được tích hợp vào các thiết bị di động(đặc trưng là các smartphome) sẽ tạo điều kiện thuận lợi và giúp ít trong việc nhận dạng và phân loại các loại cây thuốc mọi lúc mọi nơi, tìm kiếm thông tin cần thiết của một loại cây thuốc.

## **CHƯƠNG 2 CƠ SỞ LÝ THUYẾT**

### **2.1 Xử lý ảnh và một số khái niệm**

#### **2.1.1 Xử lý ảnh**

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

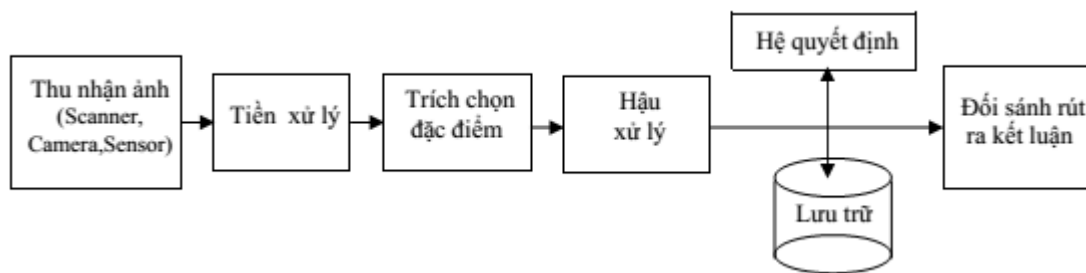
Quá trình xử lý ảnh được xem như là quá trình thao tác ảnh đầu vào nhằm cho ra kết quả mong muốn. Kết quả đầu ra của một quá trình xử lý ảnh có thể là một ảnh “tốt hơn” hoặc một kết luận.



**Hình 2.1 Quá trình xử lý ảnh**

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

Sơ đồ tổng quát của một hệ thống xử lý ảnh :



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

### 2.1.2 Các vấn đề cơ bản trong xử lý ảnh

✚ Khử nhiễu : Có 2 loại nhiễu cơ bản trong quá trình thu nhận

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

✚ Chỉnh mức xám : nhằm khắc phục tính không đồng đều của hệ thống gây ra.  
Thông thường có 2 hướng tiếp cận:

- ✓ Giảm sô mức xám: Thực hiện bằng cách nhóm các mức xám gần nhau thành một bó. Trường hợp chỉ có 2 mức xám thì chính là chuyển về ảnh đen trắng. Ứng dụng: In ảnh màu ra máy in đen trắng.

- ✓ Tăng sốmức xám: Thực hiện nội suy ra các mức xám trung gian bằng kỹ thuật nội suy. Kỹ thuật này nhằm tăng cường độ mịn cho ảnh.

✚ Trích chọn đặc điểm : Các đặc điểm của đối tượng được trích chọn tùy theo mục đích nhận dạng trong quá trình xử lý ảnh. Có thể nêu ra một số đặc điểm của ảnh sau đây:

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

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

✚ Nhận dạng :

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

Hoặc **phân loại có mẫu**(supervised classification), chẳng hạn phân tích phân biệt (discriminant analysis), trong đó mẫu đầu vào được định danh như một thành phần của một lớp đã xác định.

Hoặc **phân loại không có mẫu**(unsupervised classification hay clustering)

trong đó các mẫu được gán vào các lớp khác nhau dựa trên một tiêu chuẩn đồng dạng nào đó. Các lớp này cho đến thời điểm phân loại vẫn chưa biết hay chưa được định danh.

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


- i. Thu nhận dữ liệu và tiền xử lý.
- ii. Biểu diễn dữ liệu.
- iii. Nhận dạng, ra quyết định.

Bốn cách tiếp cận khác nhau trong lý thuyết nhận dạng là :

- i. Đối sánh mẫu dựa trên các đặc trưng trích chọn.
- ii. Phân loại thống kê.
- iii. Đối sánh cấu trúc.
- iv. Phân loại dựa trên mạng Nơ-ron nhân tạo.

Trong các ứng dụng rõ ràng là không thể chỉ dùng có một cách tiếp cận đơn lẻ để phân loại “tối ưu” do vậy cần sử dụng cùng một lúc nhiều phương pháp và cách tiếp cận khác nhau. Do vậy, các phương thức phân loại tổ hợp hay được sử dụng khi nhận dạng và nay đã có những kết quả có triển vọng dựa trên thiết kế các hệ hổng lai (hybrid system) bao gồm nhiều mô hình kết hợp.

Việc giải quyết bài toán nhận dạng trong những ứng dụng mới, nảy sinh trong cuộc sống không chỉ tạo ra những thách thức về thuật giải, mà còn đặt ra những yêu cầu về tốc độ tính toán. Đặc điểm chung của tất cả những ứng dụng đó là những đặc điểm đặc trưng cần thiết thường là nhiều, không thể do chuyên gia đề xuất, mà phải được trích chọn dựa trên các thủ tục phân tích dữ liệu.

 **Nén ảnh** : Nhằm giảm thiểu không gian lưu trữ. Thường được tiến hành theo cả hai cách khuynh hướng là nén có bảo toàn và không bảo toàn thông tin. Nén không bảo toàn thì thường có khả năng nén cao hơn nhưng khả năng phục hồi thì kém hơn. Trên cơ sở hai khuynh hướng, có 4 cách tiếp cận cơ bản trong nén ảnh:

- ✓ **Nén ảnh thống kê**: Kỹ thuật nén này dựa vào việc thống kê tần suất xuất hiện của giá trị các điểm ảnh, trên cơ sở đó mà có chiến lược mã hóa thích hợp. Một ví dụ điển hình cho kỹ thuật mã hóa này là \*.TIF.
- ✓ **Nén ảnh không gian**: Kỹ thuật này dựa vào vị trí không gian của các điểm ảnh để tiến hành mã hóa. Kỹ thuật lợi dụng sự giống nhau của các điểm

ảnh trong các vùng gần nhau. Ví dụ cho kỹ thuật này là mã nén \*.PCX.

- ✓ Nén ảnh sử dụng phép biến đổi: Đây là kỹ thuật tiếp cận theo hướng nén không bảo toàn và do vậy, kỹ thuật thường nén hiệu quả hơn. \*.JPG chính là tiếp cận theo kỹ thuật nén này.
- ✓ Nén ảnh Fractal: Sử dụng tính chất Fractal của các đối tượng ảnh, thể hiện sự lặp lại của các chi tiết. Kỹ thuật nén sẽ tính toán để chỉ cần lưu trữ phần gốc ảnh và quy luật sinh ra ảnh theo nguyên lý Fractal.

### **2.1.3 Thu nhận và biểu diễn ảnh**

#### **2.1.3.1 Thu nhận, các thiết bị thu nhận ảnh**

Các thiết bị thu nhận ảnh bao gồm camera, scanner các thiết bị thu nhận này có thể cho ảnh đen trắng.

Các thiết bị thu nhận ảnh có 2 loại chính, ứng với 2 loại ảnh thông dụng Raster, Vector.

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

Nhìn chung các hệ thống thu nhận ảnh thực hiện quá trình :

- Cảm biến: biến đổi năng lượng quang học thành năng lượng điện.
- Tổng hợp năng lượng điện thành ảnh.

#### **2.1.3.2 Biểu diễn ảnh**

Ảnh trên máy tính là kết quả thu nhận theo các phương pháp số hoá được nhúng trong các thiết bị kỹ thuật khác nhau. Quá trình lưu trữ ảnh nhằm 2 mục đích:

- Tiết kiệm bộ nhớ
- Giảm thời gian xử lý

Việc lưu trữ thông tin trong bộ nhớ có ảnh hưởng rất lớn đến việc hiển thị, in ấn và xử lý ảnh được xem như là 1 tập hợp các điểm với cùng kích thước nếu sử dụng càng nhiều điểm ảnh thì bức ảnh càng đẹp, càng mịn và càng thể hiện rõ hơn chi tiết của ảnh người ta gọi đặc điểm này là độ phân giải.

Việc lựa chọn độ phân giải thích hợp tùy thuộc vào nhu cầu sử dụng và đặc trưng của mỗi ảnh cụ thể, trên cơ sở đó các ảnh thường được biểu diễn theo 2

mô hình cơ bản.

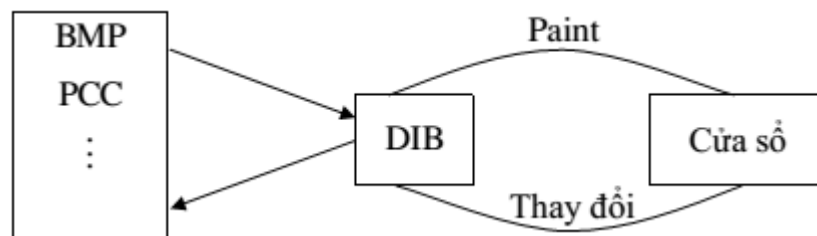
## 1. Mô hình Raster

Đây là cách biểu diễn ảnh thông dụng nhất hiện nay, ảnh được biểu diễn dưới dạng ma trận các điểm (điểm ảnh). Thường thu nhận qua các thiết bị như camera, scanner. Tùy theo yêu cầu thực tế mà mỗi điểm ảnh được biểu diễn qua 1 hay nhiều bit.

Mô hình Raster thuận lợi cho hiển thị và in ấn. Ngày nay công nghệ phần cứng cung cấp những thiết bị thu nhận ảnh Raster phù hợp với tốc độ nhanh và chất lượng cao cho cả đầu vào và đầu ra. Một thuận lợi cho việc hiển thị trong môi trường Windows là Microsoft đưa ra khuôn dạng ảnh DIB (Device Independent Bitmap) làm trung gian. Hình 2.3 thể hiện hình quy trình chung để hiển thị ảnh Raster thông qua DIB.

Một trong những hướng nghiên cứu cơ bản trên mô hình biểu diễn này là kỹ thuật nén ảnh các kỹ thuật nén ảnh lại chia ra theo 2 khuynh hướng là nén bảo toàn và không bảo toàn thông tin nén bảo toàn có khả năng phục hồi hoàn toàn dữ liệu ban đầu còn nếu không bảo toàn chỉ có khả năng phục hồi độ sai số cho phép nào đó. Theo cách tiếp cận này người ta đã đề ra nhiều quy cách khác nhau như BMP, TIF, GIF, PCX...

Hiện nay trên thế giới có trên 50 khuôn dạng ảnh thông dụng bao gồm cả trong đó các kỹ thuật nén có khả năng phục hồi dữ liệu 100% và nén có khả năng phục hồi với độ sai số nhận được.



**Hình 2.2** Quá trình hiển thị và chỉnh sửa, lưu trữ ảnh thông qua DIB

## 2. Mô hình Vector

Biểu diễn ảnh ngoài mục đích tiết kiệm không gian lưu trữ dễ dàng cho hiển thị và in ấn còn đảm bảo dễ dàng trong lựa chọn sao chép di chuyển tìm kiếm... Theo những yêu cầu này kỹ thuật biểu diễn vector tỏ

ra ưu việt hơn.

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

Công nghệ phần cứng cung cấp những thiết bị xử lý với tốc độ nhanh và chất lượng cho cả đầu vào và ra nhưng lại chỉ hỗ trợ cho ảnh Raster.

Do vậy, những nghiên cứu về biểu diễn vector đều tập trung từ chuyển đổi từ ảnh Raster.



Hình 2.3 Sự biến đổi giữa các mô hình biểu diễn ảnh

## 2.2 Các phương pháp trích chọn đặc trưng

### 2.2.1 Khái quát đặc trưng ảnh

#### 2.2.1.1 Đặc trưng văn bản đi kèm ảnh

Mỗi ảnh trên web thường có các văn bản đi kèm như là tên ảnh (title), các thẻ (tags), bình luận (comment),... để mô tả các thông tin về ảnh, đây là các siêu dữ liệu (metadata) về ảnh. Các dữ liệu này thường do người dùng tạo ảnh gắn cho mỗi ảnh, vì vậy chúng đều mang một ý nghĩa nhất định. Độ quan trọng của các loại siêu dữ liệu khác nhau cũng khác nhau. Ví dụ, các thẻ thường quan trọng hơn tên ảnh, tên ảnh quan trọng hơn bình luận. Dưới đây là một ví dụ về văn bản đi kèm một ảnh:

- Title: "Red\_Rose Flower"
- Tags: "redRoseflower, hongkongflowershow, 2009, bokeh, causewaybay, hongkong, jonnoj, jonbinalay, nikond80, interestingness50"
- Description: "HEAVEN SCENT"...FOR THE LOVE OF THE RED RED ROSE...
- Content:



**Hình 2.4 Ví dụ hiển thị một ảnh**

Vì văn bản đi kèm ảnh mang ngữ nghĩa về nội ảnh cho nên hai bức ảnh có nội dung giống nhau thường có tên giống nhau và các thẻ tương tự nhau. Vì vậy, các công cụ tìm kiếm ảnh theo văn bản đi kèm thường tập trung khai thác nội dung của các văn bản này để tìm kiếm và xếp hạng ảnh.

#### **2.2.1.2 Đặc trưng nội dung ảnh**

Một số đặc trưng nội dung ảnh:

- **Đặc trưng màu sắc :** Màu sắc là một đặc trưng nổi bật và được sử dụng phổ biến nhất trong tìm kiếm ảnh theo nội dung. Mỗi một điểm ảnh (thông tin màu sắc) có thể được biểu diễn như một điểm trong không gian màu sắc ba chiều. Các không gian màu sắc thường dùng là: RGB, Munsell, CIE, HSV. Tìm kiếm ảnh theo màu sắc tiến hành tính toán biểu đồ màu cho mỗi ảnh để xác định tỉ trọng các điểm ảnh của ảnh mà chứa các giá trị đặc biệt (màu sắc). Các nghiên cứu gần đây đang cố gắng phân vùng ảnh theo các màu sắc khác nhau và tìm mối quan hệ giữa các vùng này.
- **Đặc trưng kết cấu :** Trích xuất nội dung ảnh theo kết cấu nhằm tìm ra mô hình trực quan của ảnh và cách thức chúng được xác định trong không gian. Kết cấu được biểu diễn bởi các texel mà sau đó được đặt vào một số các tập phụ thuộc vào số kết cấu được phát hiện trong ảnh. Các tập này không chỉ xác định các kết cấu mà còn chỉ rõ vị trí các kết cấu trong ảnh. Việc xác định các kết cấu đặc biệt trong ảnh đạt được chủ yếu bằng cách mô hình các kết cấu như những biến thể cấp độ xám 2 chiều.





**Hình 2.5 Ví dụ về một số loại kết cấu**

- Đặc trưng hình dạng : Hình dạng của một ảnh hay một vùng là một đặc trưng quan trọng trong việc xác định và phân biệt ảnh trong nhận dạng mẫu. Mục tiêu chính của biểu diễn hình dạng trong nhận dạng mẫu là đo thuộc tính hình học của một đối tượng được dùng trong phân lớp, so sánh và nhận dạng đối tượng.

### **2.2.2 Các phương pháp trích chọn đặc trưng**

Việc lựa chọn các đặc trưng và độ đo thích hợp sẽ giúp tăng cả tốc độ và mức độ chính xác của các hệ thống. J.V.Jawahe và cộng sự đã nêu ra các yêu cầu cơ bản đối với thành phần lựa chọn đặc trưng cho ảnh:

- ✓ Thành phần lựa chọn đặc trưng phải lựa chọn được một tập các đặc trưng cung cấp đầu vào tốt nhất cho hệ thống tìm kiếm ảnh. Nếu số lượng các đặc trưng quá nhiều sẽ làm “che khuất” các “tín hiệu”(giảm các “tín hiệu” đối với tỉ lệ nhiễu), mặt khác, nếu số lượng các đặc trưng quá ít sẽ khó phân biệt được ảnh trong tìm kiếm.
- ✓ Nó phải giảm bớt được độ phức tạp trong lúc tính toán tổng thể bằng giảm đa chiều của bài toán phân lớp.
- ✓ Khi người dùng muốn sử dụng các đặc trưng đó cho mọi truy vấn, thì việc sử dụng các đặc trưng này phải hiệu quả. Vì số lượng các đặc trưng có thể là hàng ngàn, do đó thời gian xử lý của module phải tuyến tính với số lượng đặc trưng.
- ✓ Vì thời gian xử lý của thành phần lựa chọn đặc trưng tuyến tính với số

lượng đặc trưng, do đó việc lựa chọn các đặc trưng cũng nên tuyến tính dựa trên phân lớp.

- ✓ Thành phần lựa chọn đặc trưng có thể xử lý được với kích thước tập mẫu nhỏ (khoảng 5 mẫu).

#### 2.2.2.1 Đặc trưng màu sắc

Trích chọn đặc trưng màu sắc là phương pháp được sử dụng phổ biến hiện nay. Đây là phương pháp đơn giản, tốc độ tìm kiếm tương đối nhanh tuy nhiên kết quả tìm kiếm có độ chính xác không cao. Một số lược đồ màu được sử dụng như: lược đồ màu RGB, lược đồ màu HSI, lược đồ HSI cải tiến. Trong đó, lược đồ màu RGB được sử dụng phổ biến nhất.

- Lược đồ màu RGB :

Đối với ảnh 256 màu, lược đồ màu của ảnh tương đương với lược đồ màu của ảnh xám. Đối với ảnh 24 bit màu, lược đồ miêu tả khả năng kết nối về cường độ của ba kênh màu R, G, B. Lược đồ màu này được định nghĩa như sau:

$$h_{R,G,B}[r,g,b] = N * \Pr\{R=r, G=g, B=b\} \quad (1)$$

Trong đó N là số lượng điểm ảnh có trong ảnh

Lược đồ màu này được tính bằng cách rời rạc hóa từng màu trong ảnh, sau đó đếm số điểm ảnh của mỗi màu. Khi mà số lượng màu là có hạn, để thuận tiện hơn, người ta thường chuyển đổi ba kênh màu thành một biến giá trị duy nhất. Một cách khác để tính lược đồ màu của ảnh RGB là ta phân ra làm 3 lược đồ riêng biệt  $h_R[]$ ,  $h_G[]$  và  $h_B[]$ . Khi đó, mỗi lược đồ được tính bằng cách đếm kênh màu tương ứng trong mỗi điểm ảnh.

#### ❖ Độ đo tương đồng về màu sắc :

Một số độ đo tương đồng được sử dụng như: Độ đo khoảng cách Öclit, độ đo Jensen-Shannon divergence (JSD).

Gọi  $h(I)$  và  $h(M)$  tương ứng là 2 lược đồ màu của hai ảnh I và ảnh M. Khi đó các loại độ đo màu được định nghĩa là một số nguyên (hoặc số thực) theo các loại độ đo tương ứng như sau:

- Khoảng cách Öclit:

Đây là khoảng cách Öclit thông thường giữa các K bin:

$$Intersection(h(I), h(M)) = \sum_{j=1}^K \sqrt{(h(I) - h(M))^2} \quad (2)$$

Hoặc :

$$Intersection(h(I), h(M)) = \sum_{j=1}^K |h(I) - h(M)| \quad (3)$$

- Độ đo Jensen-Shannon divergence (JSD) :

Độ đo Jensen-Shannon divergence sử dụng lược độ màu RGB để tính toán độ tương đồng về màu sắc giữa 2 ảnh :

$$d_{JSD}(H, H') = \sum_{m=1}^M H_m \log \frac{2H_m}{H_m + H'_m} + H'_m \log \frac{2H'_m}{H'_m + H_m} \quad (4)$$

Trong đó : H và H' là 2 biểu đồ màu được so sánh,  $H_m$  là bin thứ m của biểu đồ H.

#### 2.2.2.2 Đặc trưng kết cấu

Hiện tại, vẫn chưa có một định nghĩa chính thức cụ thể về kết cấu. Kết cấu là một đối tượng dùng để phân hoạch ảnh ra thành những vùng quan tâm để phân lớp những vùng đó.

Kết cấu cung cấp thông tin về sự sắp xếp về mặt không gian của màu sắc và cường độ một ảnh. Kết cấu được đặc trưng bởi sự phân bố không gian của những mức cường độ trong một khu vực láng giềng với nhau. Kết cấu gồm các kết cấu gốc hay nhiều kết cấu gộp lại đôi khi gọi là texel.

Một số phương pháp dùng để trích xuất các đặc trưng kết cấu như :

- Kim tự tháp “có thể lái được” (the steerable pyramid).
- Biến đổi đường viền (the cotourlet transform).
- Biến đổi sóng Gabor (The Gabor Wavelet transform).
- Biểu diễn ma trận đồng hiện (co-occurrence matrix)
- Hệ thống bộ lọc định hướng phức tạp (The complex directional filter bank).

#### ❖ Đo độ tương đồng cho kết cấu ảnh :

Để đo độ tương đồng theo kết cấu giữa các ảnh, người ta thường sử dụng độ đo Öclit. Kết cấu được trích xuất từ các bức ảnh sẽ được biểu diễn thành các vector nhiều chiều và khoảng cách Öclit được dùng để đo độ tương đồng giữa các đặc trưng của ảnh truy vấn với đặc trưng của ảnh trong cơ sở dữ liệu.

### 2.2.2.3 Đặc trưng hình dạng

Màu sắc và kết cấu là những thuộc tính có khái niệm toàn cục trong một ảnh. Trong khi đó, hình dạng không phải là một thuộc tính của ảnh. Nói tới hình dạng không phải là nhắc đến hình dạng của một ảnh. Thay vì vậy, hình dạng có khuynh hướng chỉ đến một khu vực đặc biệt trong ảnh, hay hình dạng chỉ là biên của một đối tượng nào đó trong ảnh.

Trong tìm kiếm ảnh theo nội dung, hình dạng là một cấp cao hơn so với màu sắc và kết cấu. Nó đòi hỏi sự phân biệt giữa các vùng để tiến hành xử lý về độ đo của hình dạng. Các hệ thống tìm kiếm ảnh theo nội dung thường khai thác hai nhóm biểu diễn hình dạng sau :

- Biểu diễn hình dạng theo đường biên (cotour-based descriptor): Biểu diễn các đường biên bao bên ngoài.
- Biểu diễn theo vùng (region-based descriptor): Biểu diễn một vùng toàn vẹn.

#### ❖ Đo độ tương đồng cho hình dạng :

Độ đo về hình dạng rất nhiều trong phạm vi lý thuyết của bộ môn xử lý ảnh. Chúng trải rộng từ những độ đo toàn cục dạng thô với sự trợ giúp của việc nhận dạng đối tượng, cho tới những độ đo chi tiết tự động tìm kiếm những hình dạng đặc biệt. Lược đồ hình dạng là một ví dụ của độ đo đơn giản. Kỹ thuật dùng đường biên hiệu quả hơn phương pháp trước, chúng tìm kiếm những hình dạng đối tượng gần giống với đường biên nhất. Phương pháp vẽ phác họa là phương pháp có nhiều đặc trưng rõ ràng hơn, không chỉ tìm kiếm những đường biên đối tượng đơn, mà còn đối với tập những đối tượng đã được phân đoạn trong một ảnh mà người dùng vẽ hay cung cấp.

### 2.2.2.4 Đặc trưng cục bộ bất biến

Người ta thường chia đặc trưng cục bộ thành 2 loại là những điểm trích xuất được từ điểm “nhô ra” (salient points) của ảnh và đặc trưng SIFT được trích chọn từ các điểm hấp dẫn Harris (interest points). Ở đây trình bày chi tiết về việc trích chọn các đặc trưng cục bộ bất biến (Scale Invariant Feature Transform) của ảnh.

Các đặc trưng bất biến với việc thay đổi tỉ lệ ảnh, quay ảnh, đôi khi là thay đổi điểm nhìn và thêm nhiễu ảnh hay thay đổi cường độ chiếu sáng của ảnh. Phương pháp được lựa chọn có tên là Scale-Invariant Feature Transform (SIFT) và đặc trưng rút trích được gọi là đặc trưng SIFT (SIFT Feature). Các đặc trưng SIFT này được rút trích ra từ các điểm hấp dẫn cục bộ.

**Điểm hấp dẫn (Interest Point (Keypoint)):** Là vị trí (điểm ảnh) “hấp

dẫn” trên ảnh. “Hấp dẫn” ở đây có nghĩa là điểm đó có thể có các đặc trưng bất biến với việc quay ảnh, co giãn ảnh hay thay đổi cường độ chiếu sáng của ảnh.

Phương pháp rút trích các đặc trưng bất biến SIFT được tiếp cận theo phương pháp thác lọc, theo đó phương pháp được thực hiện lần lượt theo các bước sau:

- **Phát hiện các điểm cực trị Scale-Space** (Scale-Space extrema detection): Bước đầu tiên này tiến hành tìm kiếm các điểm hấp dẫn trên tất cả các tỉ lệ và vị trí của ảnh. Nó sử dụng hàm different-of-Gaussian để xác định tất cả các điểm hấp dẫn tiềm năng mà bất biến với quy mô và hướng của ảnh.
- **Định vị các điểm hấp dẫn** (keypoint localization): Một hàm kiểm tra sẽ được đưa ra để quyết định xem các điểm hấp dẫn tiềm năng có được lựa chọn hay không?
- **Xác định hướng cho các điểm hấp dẫn** (Orientation assignment): Xác định hướng cho các điểm hấp dẫn được chọn.
- **Mô tả các điểm hấp dẫn** (Keypoint descriptor): Các điểm hấp dẫn sau khi được xác định hướng sẽ được mô tả dưới dạng các vector đặc trưng nhiều chiều.

## 1. Phát hiện điểm cực trị Scale-Space

Các điểm hấp dẫn với đặc trưng SIFT tương thích với các cực trị địa phương của bộ lọc difference-of-Gaussian (DoG) ở các tỉ lệ khác nhau. Định nghĩa không gian tỉ lệ của một hình ảnh là hàm  $L(x,y,k\sigma)$  được mô tả như sau :

$$L(x,y,k\sigma) = G(x,y,k\sigma) * I(x,y) \quad (5)$$

Với  $G(x,y,k\sigma)$  : biến tỉ lệ Gaussian (variable scale Gaussian)

$I(x,y)$  : ảnh đầu vào

\* là phép nhân chập giữa  $x$  và  $y$ .

Và 
$$G(x,y,k\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (6)$$

Để phát hiện các điểm hấp dẫn, ta đi tìm các cực trị của hàm DoG được định nghĩa:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y)$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (7)$$

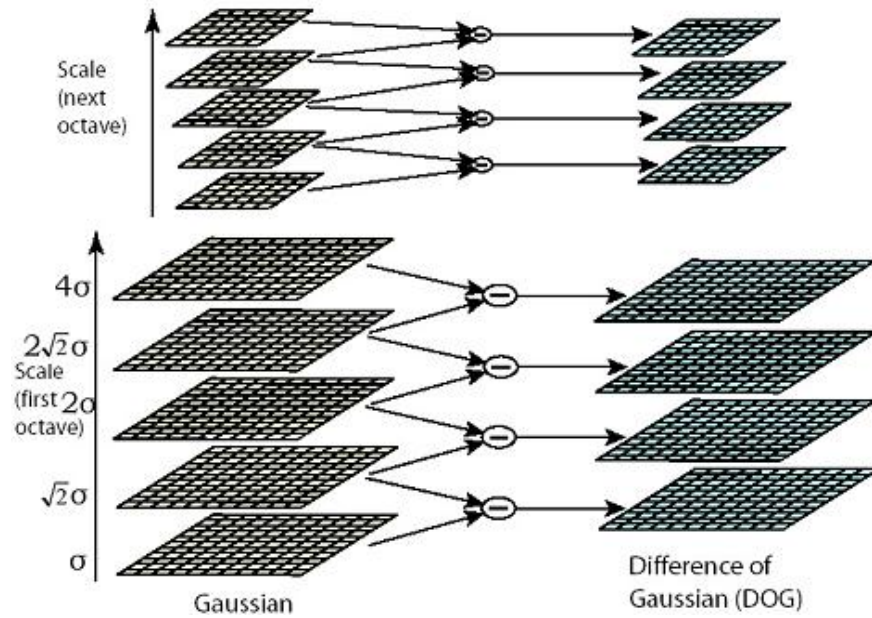
Giá trị hàm DoG được tính xấp xỉ dựa vào giá trị scale-normalized Laplacian of Gaussian ( $\sigma^2 \nabla^2 G$ ) thông qua các phương trình (5)(6)(7).

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G \quad (8)$$

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \quad (9)$$

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k - 1)\sigma^2 \nabla^2 G \quad (10)$$

Như vậy bước đầu tiên của giải thuật SIFT phát hiện các điểm hấp dẫn với bộ lọc Gaussian ở các tỉ lệ khác nhau và các ảnh DoG từ sự khác nhau của các ảnh mờ.

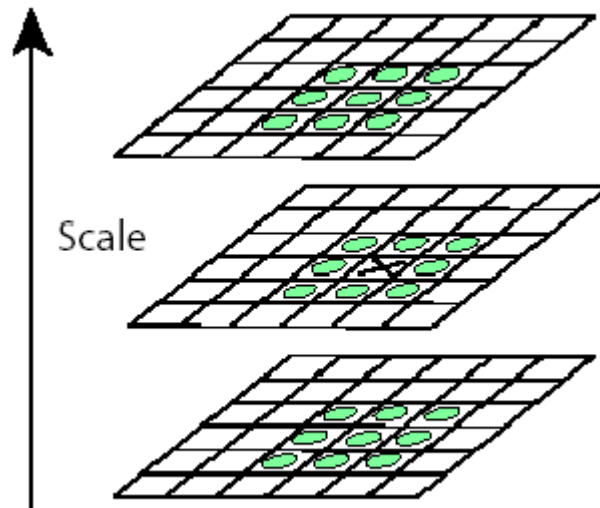


**Hình 2.6 Biểu đồ mô phỏng việc tính toán các DoG ảnh từ các ảnh mờ**

Các ảnh cuộc được nhóm thành các octave (mỗi octave tương ứng với giá trị gấp đôi của  $\sigma$ ). Giá trị của  $k$  được chọn sao cho số lượng ảnh mờ (blurred images) cho mỗi octave là cố định. Điều này đảm bảo cho số lượng các ảnh DoG cho mỗi octave không thay đổi.

Các điểm hấp dẫn được xác định là các cực đại hoặc cực tiểu của các ảnh DoG qua các tỉ lệ. Mỗi điểm ảnh trong DoG được so sánh với 8 điểm ảnh láng giềng của nó ở cùng tỉ lệ đó và 9 láng giềng kề ở các tỉ lệ ngay trước và sau nó. Nếu điểm ảnh đó đạt giá trị cực tiểu hoặc cực đại thì sẽ được chọn làm các điểm

hấp dẫn ứng viên.



**Hình 2.7** Mỗi điểm ảnh được so sánh với 26 láng giềng của nó

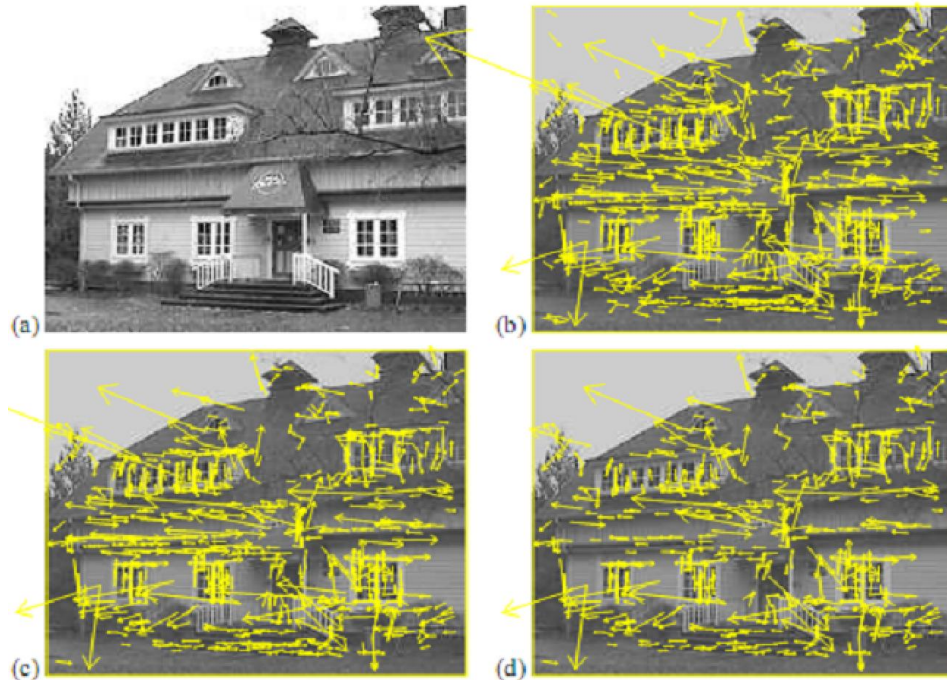
## **2. Định vị điểm hấp dẫn**

Mỗi điểm hấp dẫn ứng viên sau khi được chọn sẽ được đánh giá xem có được giữ lại hay không:

- Loại bỏ các điểm hấp dẫn có độ tương phản thấp
- Một số điểm hấp dẫn dọc theo các cạnh không giữ được tính ổn định khi ảnh bị nhiễu cũng bị loại bỏ.

Các điểm hấp dẫn còn lại sẽ được xác định hướng.





**Hình 2.8** Quá trình lựa chọn các điểm hấp dẫn. (a) Ảnh gốc. (b) Các điểm hấp dẫn được phát hiện. (c) Ảnh sau khi loại bỏ các điểm hấp dẫn có độ tương phản thấp. (d) Ảnh sau khi loại bỏ các điểm dọc theo cạnh

### 3. Xác định hướng cho các điểm hấp dẫn

Để xác định hướng cho các điểm hấp dẫn, người ta tính toán biểu đồ hướng Gradient trong vùng láng giềng của điểm hấp dẫn. Độ lớn và hướng của các điểm hấp dẫn được xác định theo công thức:

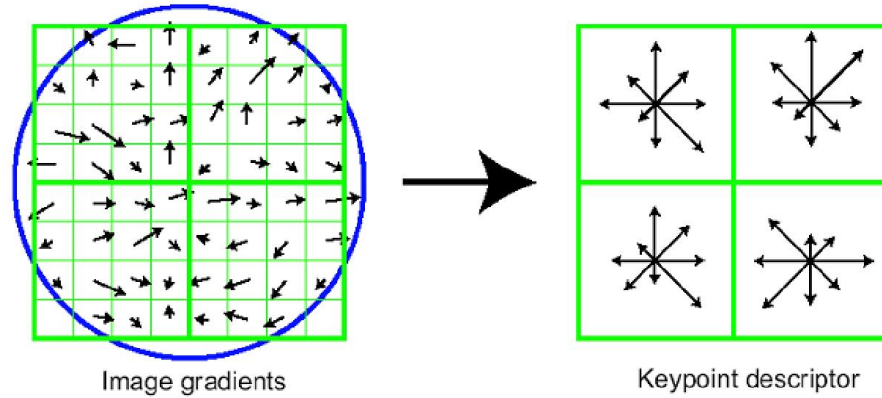
$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad (11)$$

$$\theta(x, y) = \tan^{-1}((L(x + 1, y) - L(x - 1, y)) / (L(x, y + 1) - L(x, y - 1))) \quad (12)$$

### 4. Biểu diễn vector cho điểm hấp dẫn

Điểm hấp dẫn sau khi được xác định hướng sẽ được biểu diễn dưới dạng các vector  $4 \times 4 \times 8 = 128$  chiều.





**Hình 2.9 Biểu diễn các vector đặc trưng**

❖ **Độ đo tương đồng cho đặc trưng cục bộ bất biến**

Một số độ đo tương đồng cho ảnh sử dụng đặc trưng SIFT như :

- Độ đo Cosin :

$$d(x, y) = \frac{x \cdot y}{\|x\| \cdot \|y\|} \quad (13)$$

- Khoảng cách góc :

$$d(x, y) = \cos^{-1}(x \cdot y) \quad (14)$$

- Độ đo Euclide :

$$d(x, y) = \sqrt{\sum_{i=1}^n |x_i - y_i|^2} \quad (15)$$

- Độ đo Jensen-Shannon divergence :

$$d_{JSD}(H, H') = \sum_{m=1}^M H_m \log \frac{2H_m}{H_m + H'_m} + H'_m \log \frac{2H'_m}{H'_m + H_m} \quad (16)$$

## **CHƯƠNG 3 CÁC NGHIÊN CỨU LIÊN QUAN ĐẾN NHẬN DẠNG**

### **3.1 Phân lớp ảnh chụp lá cây bằng phương pháp máy vector hỗ trợ SVM (support vector machine)**

#### **3.1.1 Giới thiệu SVM**

Thị giác máy tính là sử dụng các thiết bị điện tử (camera, sensor,...) thu nhận ảnh từ thực tế, sau đó sử dụng máy tính tự động phân tích hình ảnh qua đó nhận biết được đối tượng. Kỹ thuật này được nghiên cứu và ứng dụng rộng rãi vào nhiều lĩnh vực trong nông

ngiệp. Trong chọn giống cây nông nghiệp, thực tiễn chứng minh, áp dụng kỹ thuật thị giác máy tính thông qua lựa chọn thông số đặc trưng ngoại hình để tiến hành phân loại và kiểm tra chất lượng hạt giống đạt kết quả cao.

Từ giữa thập kỉ 80 của thế kỷ XX, thị giác máy tính được ứng dụng trong tự động hóa thu hoạch nông sản phẩm. Năm 1996, Zhang and Takahashi thông qua mô hình phân biệt thực hiện kiểm tra đo lường định vị quả táo, tiến tới có thể tự động hái táo. Nhóm đã tính toán để phân biệt quả táo từ hình ảnh ban đầu của cây táo và tìm ra vị trí của quả táo. Một số ứng dụng khác cũng sử dụng kỹ thuật thị giác máy tính như nhận dạng loài hoa (Nilsback and Zisserman, 2008), xác định loài cây (Berrinet al., 2012), nhận dạng loài cây thông qua ảnh chụp lá cây (Neeraj et al., 2012).

Việc nhận dạng loài cây dựa vào ảnh chụp của lá hoặc phát hiện bệnh cây dựa vào các biểu hiện của lá cây thông qua ảnh chụp của chúng, nếu không có những công cụ nhận dạng tự động, để nhận dạng một loài cây hoặc phát hiện một loại bệnh cây chúng ta sẽ phải tự tìm hiểu, tra cứu từ nhiều nguồn tài liệu, dữ liệu.

Hệ thống nhận dạng lá cây hoặc phát hiện bệnh cây sẽ được thực hiện qua một số bước :

- Bước đầu tiên: phân lớp ảnh, xác định xem ảnh có chứa một lá hợp lệ hay không? Để quyết định xem nó có khả dụng trong quá trình xử lý tiếp theo hay không?
- Bước thứ hai: thực hiện phân vùng ảnh để thu được ảnh nhị phân và tách lá ra khỏi nền.
- Bước thứ ba: rút trích những đặc trưng của lá từ ảnh chụp.
- Bước cuối cùng: Bước cuối cùng so sánh những đặc trưng vừa được trích chọn với những ảnh từ ảnh lá trong cơ sở dữ liệu đã được gán nhãn và trả lại các loài cây (hoặc loại bệnh cây) với sự tương đồng gần nhất.

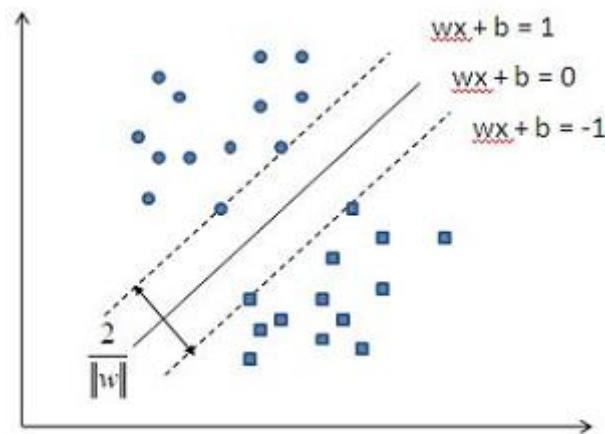
Trong lĩnh vực thị giác máy tính có thể sử dụng nhiều phương pháp phân lớp dữ liệu như Naïve Bayes [Alberto Tellaecche et al., (2008)], cây quyết định (decision tree) [John et al. (1996)], mạng nơron [Liao et al. (1994)], máy vector hỗ trợ SVM [Neeraj Kumar et al. (2012)]... Trong những năm gần đây, phương pháp phân loại máy vector hỗ trợ (SVM) được quan tâm và sử dụng nhiều trong những lĩnh vực nhận dạng và phân loại.

SVM là một họ các phương pháp phân loại dựa trên cơ sở các hàm nhân (kernel) để tối thiểu hóa rủi ro ước lượng. Phương pháp SVM ra đời từ lý thuyết học thống kê do Vapnik và Chervonenkis xây dựng (Vapnik and Chervonenkis, 1974), Vapnik (1999) và có nhiều tiềm năng phát triển về mặt lý thuyết cũng như ứng dụng trong thực tiễn. Các thử nghiệm thực tế cho thấy, phương pháp SVM có khả năng phân loại khá tốt đối với

nhiều ứng dụng khác nhau như bài toán phân loại văn bản, nhận dạng chữ viết tay, phát hiện mặt người trong các ảnh, ước lượng hồi quy,...

### 3.1.2 Phương pháp phân lớp dữ liệu máy vector hỗ trợ (SVM)

SVM là một giải thuật máy học dựa trên lý thuyết học thống kê do Vapnik and Chervonenkis (1974), Vapnik (1999) xây dựng. Bài toán cơ bản của SVM là bài toán phân loại hai lớp: Cho trước  $r$  điểm trong không gian  $n$  chiều (mỗi điểm thuộc vào một lớp kí hiệu là  $+1$  hoặc  $-1$ ), mục đích của giải thuật SVM là tìm một siêu phẳng (hyperplane) phân hoạch tối ưu cho phép chia các điểm này thành hai phần sao cho các điểm cùng một lớp nằm về một phía với siêu phẳng này. Hình 3.1 minh họa phân lớp với SVM trong mặt phẳng.



Hình 10 Ví dụ siêu phẳng với lề cực đại trong không gian  $R^2$

Xét tập  $r$  mẫu huấn luyện  $\{(x_1, y_1), (x_2, y_2) \dots (x_r, y_r)\}$ . Trong đó  $x_i$  là một vector đầu vào được biểu diễn trong không gian  $x \subseteq R^n$ ,  $y_i$  là một nhãn lớp;  $y_i \in \{1, -1\}$ . Siêu phẳng tối ưu phân tập dữ liệu này thành hai lớp là siêu phẳng có thể tách rời dữ liệu thành hai lớp riêng biệt với lề (margin) lớn nhất. Tức là, cần tìm siêu phẳng  $H_0 : y = w \cdot x + b = 0$ , và hai siêu phẳng  $H_+$ ,  $H_-$  hỗ trợ song song với  $H_0$  và có cùng khoảng cách đến  $H_0$ . Với điều kiện không có phần tử nào của tập mẫu nằm giữa  $H_+$  và  $H_-$ , khi đó :

$$H_+ : w \cdot x + b \geq +1 \text{ với } y = +1$$

$$H_- : w \cdot x + b \leq -1 \text{ với } y = -1$$

Kết hợp hai điều kiện trên, có  $y(w \cdot x + b) \geq 1$ .

Khoảng cách của siêu phẳng  $H_+$  và  $H_-$  đến  $H_0$  là  $\frac{1}{\|w\|}$ . Cần tìm siêu phẳng  $H_0$  với lề lớn nhất, là giải bài toán tối ưu tìm  $w$  và  $b$  sao cho:  $\frac{2}{\|w\|}$  đạt cực đại với ràng

buộc  $y_i(w \cdot x_i + b) \geq 1$ . Tương đương với bài toán cực tiểu hóa :  $\frac{w \cdot w}{2}$  với điều kiện:

$y_i(w \cdot x_i + b) \geq 1, \forall_i i = 1 \dots r$ . Lời giải cho bài toán tối ưu này là cực tiểu hóa hàm Lagrange:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^r \alpha_i [y_i(\|w \cdot x_i\| + b) - 1] \quad (17)$$

Trong đó  $\alpha$  là các hệ số Lagrange,  $\alpha \geq 0$ .

Lời giải tìm siêu phẳng tối ưu trên có thể mở rộng trong trường hợp dữ liệu không thể tách rời tuyến tính (Letsche and Berry, 2001) bằng cách ánh xạ dữ liệu vào một không gian có số chiều lớn hơn bằng cách sử dụng một hàm nhân K (kernel). Một số hàm nhân thường dùng như:

- Hàm tuyến tính có dạng  $K(x, y) = x \cdot y$
- Hàm đa thức có dạng  $K(x, y) = (x \cdot y + 1)^d$
- Hàm RBF (Radial Basis Function) có dạng  $K(x, y) = e^{-\frac{|x-y|^2}{2\sigma^2}}$

Với khả năng vượt trội của SVM về tính hiệu quả, độ chính xác, khả năng xử lý các bộ dữ liệu một cách linh hoạt, việc sử dụng máy vectơ hỗ trợ SVM đã và đang là sự lựa chọn tối ưu nhất trong việc giải quyết các bài toán phân loại/dự báo trong một số các ngành khoa học.

### 3.1.3 Mô hình phân lớp ảnh chụp lá cây

Trong các hệ thống nhận dạng loài cây, khâu đầu tiên là xác định xem ảnh đưa vào có chứa lá cây hay không trước khi chuyển tiếp sang giai đoạn nhận dạng lá của cây gì. Khâu này quan trọng vì nó ảnh hưởng đến độ chính xác nhận dạng của toàn hệ thống. Sử dụng phương pháp SVM như đã trình bày ở trên để thực hiện bài toán này. Quá trình phân lớp được thực hiện theo hai bước:

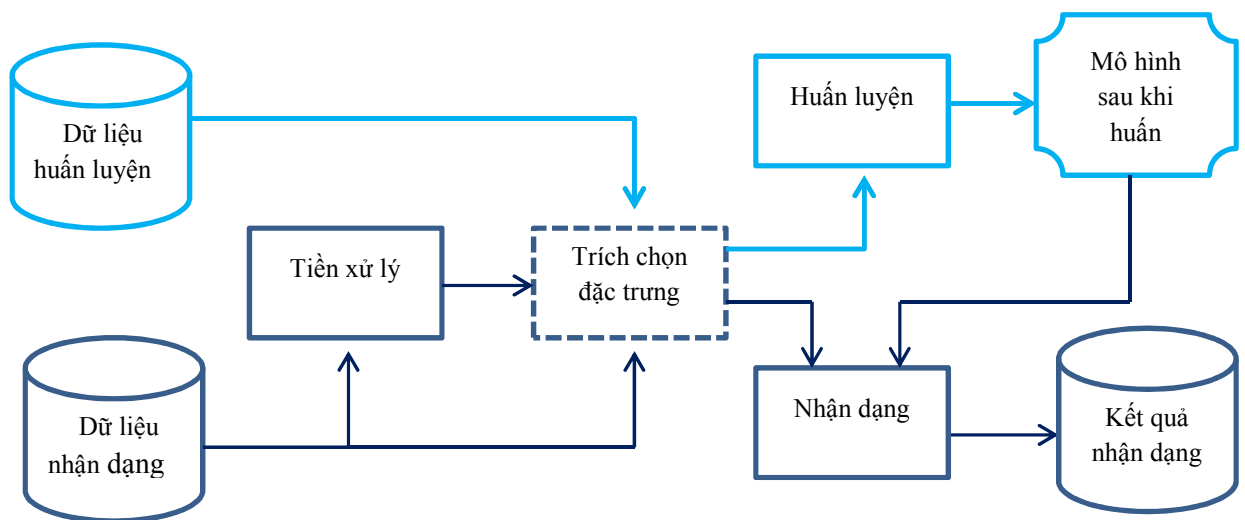
- Bước 1 : Huấn luyện mô hình phân lớp dữ liệu

Chuẩn bị dữ liệu cho quá trình huấn luyện (lựa chọn tập hợp các ảnh chụp lá cây thích hợp; một ảnh đạt yêu cầu là ảnh có chứa lá xác định và nền ảnh màu trắng hoặc đen hoàn toàn hoặc một màu nền đồng nhất làm nổi bật lá). Tập dữ liệu được trích chọn đặc trưng và đưa vào máy huấn luyện phân lớp SVM. Sau khi kết thúc quá trình huấn luyện, hệ thống sẽ lưu lại giá trị các tham số này (các tham số quyết định phân lớp - mô hình sau khi huấn luyện) để phục vụ cho quá trình nhận dạng sau này. Quá trình huấn luyện dữ liệu nhanh hay chậm phụ thuộc vào số lượng mẫu dữ liệu tham gia huấn luyện, thuật toán chọn để huấn luyện dữ

liệu(chọn thuật toán phân lớp SVN).

- **Bước 2 : Phân lớp dữ liệu thử, nghiệm mô hình**

Sau khi đã xây dựng xong mô hình huấn luyện dữ liệu (bước 1): tiến hành phân lớp với một mẫu dữ liệu mới x. Dữ liệu x (có thể được thực hiện tiền xử lý hoặc không tùy thuộc vào người dùng ứng dụng quyết định) được trích chọn đặc trưng, sau đó sẽ được đưa vào tính toán thông qua các tham số của hàm quyết định (các tham số của mô hình huấn luyện) để xác định lớp của mẫu dữ liệu x (Hình 3.2). Bước phân lớp dữ liệu này gồm nhiều giai đoạn:



**Hình 11 Mô hình phân lớp ảnh chụp có phải là lá cây**

#### **3.1.3.1 Giai đoạn tiền xử lý**

Vì ảnh đầu vào (dữ liệu nhận dạng / phân lớp) có thể là ảnh chứa lá hợp lệ, có thể là ảnh không phải là lá cây, ảnh lá cây có chứa nhiễu (ánh sáng, màu nền không thích hợp, ảnh chứa những chi tiết không thuộc lá...) nên dữ liệu này có thể được tiến hành tiền xử lý trước khi nhận dạng (quyết định tiền xử lý phụ thuộc vào người sử dụng ứng dụng).

#### **3.1.3.2 Giai đoạn rút trích đặc trưng**

Để thực hiện quá trình phân lớp, bước trích chọn đặc trưng ảnh có vai trò rất quan trọng. Đặc trưng ảnh ở đây chính là đặc trưng nội dung ảnh, là phân tích nội dung thực sự của các bức ảnh. Nội dung ảnh được thể hiện bằng màu sắc, hình dạng, kết cấu (texture), các đặc trưng cục bộ (local features)... hay bất cứ thông tin nào có từ chính nội dung ảnh. Ở đây sử dụng phương pháp trích chọn đặc trưng toàn cục Gist [Oliva and Torralba (2001)] và Gist descriptor [Douze et al.

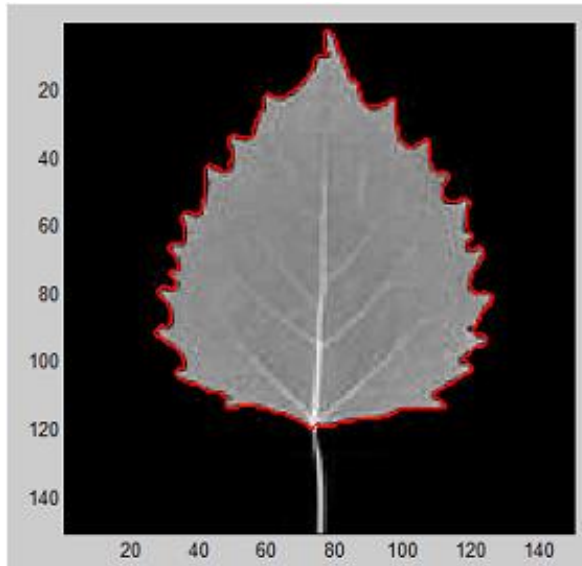
(2009)] được sử dụng để trích chọn đặc trưng Gist cho ảnh. Đặc trưng Gist sử dụng bộ lọc Gabor với kích thước là 32 (gồm 4 thang chia và 8 hướng). Bộ lọc Gabor được tính toán trước và lưu lại. Ảnh được chia ra thành 16 cửa sổ, trong mỗi cửa sổ bộ lọc được áp dụng để tính vector đặc trưng. Do vậy, sẽ có tổng số là  $32 \times 16 = 512$  đặc trưng trong vector đặc trưng để biểu diễn một ảnh.

#### ***3.1.3.3 Chọn thuật toán huấn luyện và phân lớp dữ liệu***

Sau khi hoàn thành giai đoạn trích chọn đặc trưng ảnh, tiếp theo là chọn thuật toán huấn luyện và phân lớp dữ liệu ảnh. Trong phần cài đặt thực nghiệm, lựa chọn thuật toán huấn luyện và nhận dạng dữ liệu SVM [Cortes and Vapnik (1995)]. Để phát triển ứng dụng nhận dạng ảnh chụp lá cây sử dụng thư viện mã nguồn mở LibSVM (Chang and Lin (2011)).

#### **3.1.4 Thực nghiệm**

Sử dụng Matlab để cài đặt ứng dụng và thực nghiệm trên tập dữ liệu thu thập từ trang web của hệ thống Leafsnap [<http://leafsnap.com/species/>]. Dữ liệu ảnh được thu thập bao gồm ảnh các loại ảnh lá cây, ảnh không phải là lá cây, ảnh lá cây có chứa nhiễu (ánh sáng, màu nền không thích hợp, ảnh chứa những chi tiết không thuộc lá...). Các ảnh lá cây phù hợp là 814 ảnh, không phù hợp là 343 ảnh, tổng số lấy 744 mẫu ảnh lá phù hợp, 301 mẫu ảnh lá không phù hợp cho tập huấn luyện. Đối với tập mẫu để kiểm tra, chọn ngẫu nhiên 150 ảnh (Hình 3.5), trong đó có 100 ảnh lá phù hợp, 50 ảnh lá không phù hợp từ 1157 ảnh ban đầu. Chương trình ứng dụng cho phép đưa một ảnh bất kì vào, cho phép người dùng tùy chọn bước tiền xử lý ảnh đầu vào hay thực hiện kiểm tra trực tiếp ảnh và cho ra kết quả phân lớp của hệ thống, ảnh đó có phải là lá hợp lệ hay không, có đạt yêu cầu cho các bước nhận dạng tiếp theo hay không.



**Hình 12 Giai đoạn tiền xử lý(tách ảnh khỏi nền)**



**Hình 13 Giao diện chương trình**





**Hình 14 Tập dữ liệu chọn để kiểm tra ứng dụng**

### 3.1.5 Kết luận

Ứng dụng phân lớp ảnh chụp lá cây dựa trên phương pháp máy vector hỗ trợ SVM bước đầu đã phân biệt được một ảnh có chứa lá hay không, xác định được cả những ảnh chứa nhiều không thích hợp cho các bước trong quá trình nhận dạng tiếp theo với độ chính xác tương đối cao (với trường hợp chọn hàm nhân phù hợp). Một ứng dụng hoàn thiện nhận dạng loài cây dựa vào ảnh chụp của lá là bước tiếp theo để phát triển ứng dụng phân lớp trên.

## 3.2 Kỹ thuật PCA (Principal Components Analysis)

### 3.2.1 Giới thiệu

PCA (Principle Components Analysis) là một thuật toán để tạo ra một ảnh mới từ ảnh ban đầu. Ảnh mới này có kích thước nhỏ hơn rất nhiều so với ảnh đầu vào nhưng vẫn mang những đặc trưng cơ bản nhất cần nhận dạng.

Mục tiêu của PCA là ảnh mới được tạo ra có kích thước nhỏ nhất nhưng vẫn chứa nhiều nhất các đặc trưng ảnh đầu vào. Về bản chất, PCA tìm ra một không gian mới theo hướng biến thiên mạnh nhất của một tập hợp các vector trong một không gian cho trước. Trong không gian mới, người ta hy vọng rằng việc phân loại sẽ mang lại kết quả tốt hơn so với không gian ban đầu.

❖ Ưu điểm của PCA :



- Tìm được các đặc trưng tiêu biểu của ảnh mà không cần phải xác định các thành phần và mối liên hệ giữa các thành phần trên ảnh đó.
- Thuật toán thực hiện tốt với các ảnh đầu vào có độ phân giải cao, do PCA sẽ thu gọn ảnh đầu vào thành một ảnh mới có kích thước rất nhỏ.
- Khối lượng tính toán không nhiều.
- PCA có tính mở và có thể kết hợp được với các phương pháp khác (như mạng Nơ-rôn, Support Vector Machine– SVM,...) để mang lại hiệu quả nhận dạng cao hơn.

❖ Nhược điểm PCA :

- PCA phân loại theo chiều phân bố lớn nhất của tập vector. Tuy nhiên, không phải bao giờ chiều phân bố lớn nhất lại mang lại hiệu quả cao nhất cho nhận dạng. Đây là nhược điểm cơ bản của PCA.
- PCA rất nhạy với nhiễu, vì vậy nhiệm vụ tiền xử lý (lọc nhiễu, chuẩn hóa,...) đòi hỏi phải thực hiện rất kỹ.

### 3.2.2 Nội dung PCA

Phương pháp PCA coi một ảnh xám  $R \times C$  pixels là một vector trong không gian  $N = R \times C$  chiều. Không gian mới được tạo ra bởi PCA được cấu thành từ  $K$  vector đơn vị có chiều là  $N$ ,  $N \ll R \times C$ .

Mỗi vector được gọi là Eigenface.

$$\text{Phép biến đổi : } A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \rightarrow w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} (k \ll n) \quad (18)$$

Theo công thức :  $W = T.A$  với  $T$  là ma trận chuyển đổi và có kích thước là  $K \times N$ . Gọi  $M$  là số ảnh đầu vào, mỗi ảnh đã được chuyển thành một vector  $N$  chiều. Ta có tập hợp đầu vào  $X = \{x_1, x_2, \dots, x_m\} (x_i \in \mathbb{R}^N)$ .

Trung bình của các vector đầu vào (tâm của tập hợp đầu vào) ;

$$\bar{x} = \frac{1}{M} \sum_{i=1}^M x_i \quad (19)$$

Sai lệch so với tâm :  $\Phi_i = x_i - \bar{x} \quad (20)$

Gọi  $A = \{\Phi_1, \Phi_2, \dots, \Phi_M\}$  ta có ma trận hiệp phương sai của  $A$  :

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = A.A^T \quad (20)$$

Gọi các giá trị riêng của  $C$  là  $\lambda_1, \lambda_2, \dots, \lambda_n$  sắp xếp theo thứ tự giảm dần, tương ứng với  $N$  vector riêng  $u_1, u_2, \dots, u_n$ . Các vector riêng này trực giao từng đôi một. Mỗi

vector riêng  $u_i$  được gọi là một eigenface. Tập hợp vector ban đầu được biểu diễn trong không gian tạo bởi N eigenface theo mô tả sau:

$$x - \bar{x} = w_1 u_1 + w_2 u_2 + \dots + w_n u_n = \sum_{i=1}^n w_i u_i \quad (21)$$

Chọn lấy K vector riêng  $u$  tương ứng với K giá trị riêng  $\lambda$  lớn nhất. Ta có:

$$\hat{x} - \bar{x} = w_1 u_1 + w_2 u_2 + \dots + w_k u_k = \sum_{i=1}^k w_i u_i \text{ với } K \ll N \quad (22)$$

Vector các hệ số khai triển  $[w_1, w_2, \dots, w_k]$  chính là biểu diễn mới của ảnh được tạo ra trong không gian PCA. Ảnh mới  $\hat{x}$  vẫn giữ được các đặc điểm chính của ảnh đầu vào. Vector  $[w_1, w_2, \dots, w_k]$  được tính theo công thức sau :

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_k^T \end{bmatrix} (x - \bar{x}) = U^T (x - \bar{x}) \quad (23)$$

Vấn đề cần giải quyết ở đây là ma trận covariance  $C = A.A^T$  có kích thước  $N^2$ . Với  $N=200 \times 200=40000$ , khối lượng tính toán sẽ rất lớn. Do đó, để tính được các eigenface mà không cần phải tính cả ma trận C, người ta đưa ra phương pháp tính nhanh dựa vào tính vector riêng và giá trị riêng của ma trận  $L = A^T.A$  có kích thước  $M \times M$  với M là số ảnh đầu vào. Ta có thể chứng minh như sau: gọi  $v_i, \mu_i$  lần lượt là vector riêng và giá trị riêng của ma trận L:

$$A^T.A.v_i = v_i.\mu_i.$$

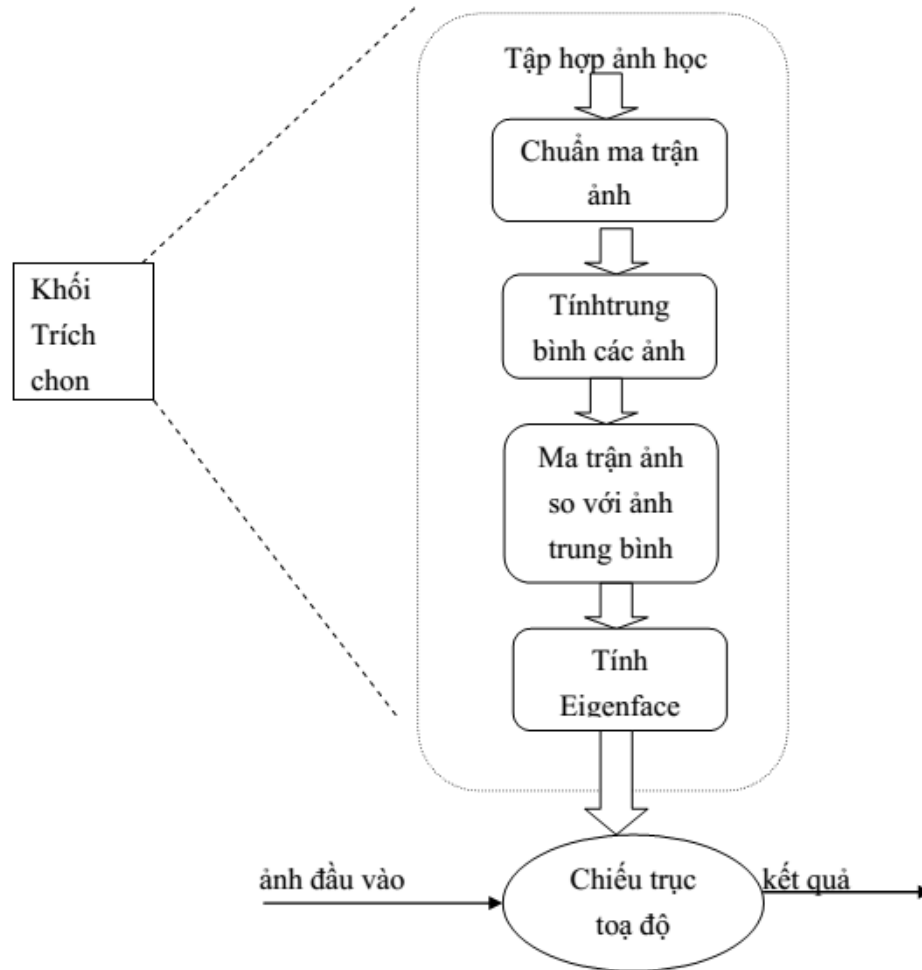
$$\text{Nhân cả hai vế với } A \text{ ta có : } A.A^T.A.v_i = v_i.\mu_i.A$$

Ta thấy  $A.v_i$  chính là vector riêng của  $C = A.A^T$  ứng với giá trị riêng  $\mu_i$ .

### 3.2.3 Quá trình rút trích

- **Giai đoạn 1 :** Tạo ra các eigenfaces thông qua quá trình học các ảnh mẫu.
- **Giai đoạn 2 :** Chiếu vector ảnh cần nhận dạng vào không gian eigenface. Các vector đặc trưng của ảnh đầu vào là phép chiếu từ ảnh ban đầu vào không gian tạo bởi các eigenface.

Quá trình rút trích đặc trưng được thể hiện qua lưu đồ sau :



Hình 15 Quá trình rút trích Eigenface

#### ❖ Chuẩn hóa ma trận ảnh :

Chuẩn hóa vector là khâu đầu tiên cần thực hiện khi sử dụng thuật toán PCA. Sự phân bố của một vec-tơ trong không gian được thể hiện qua trung bình (mean) và chuẩn (standart) của vector đó.

Cho vector  $X = [x_1, x_2, \dots, x_n]$  trung bình và chuẩn của vector X được tính như sau :

$$Mean(X) = \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i \quad (24)$$

$$Std(X) = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{X})^2}{N-1}} \quad (25)$$

Từ một ảnh  $\Gamma$  ban đầu, ta có trung bình  $\bar{\Gamma}$  và chuẩn  $\text{Std}(\Gamma)$  của  $\Gamma$ .

Phép chuẩn hóa ma trận ảnh được thực hiện theo công thức sau:

$$\Gamma'(x, y) = \frac{\Gamma(x, y) - \bar{\Gamma}}{\text{std}(\Gamma)} \quad (26)$$

Trong đó,  $\Gamma$  là ma trận ảnh đã được chuẩn hóa,  $(x, y)$  là tọa độ điểm ảnh. Kết quả quá trình chuẩn hóa được thể hiện như sau:



*Ảnh đầu vào*



*Ảnh sau chuẩn hóa*

**Hình 16 Chuẩn hóa ma trận ảnh**

#### ❖ Tính trung bình tập ảnh đầu vào :

Khâu này cho ta xác định được gốc tọa độ của không gian vector biểu diễn các ảnh. Ảnh trung bình được tính bằng trung bình cộng của tất cả các ảnh. Ảnh trung bình này thể hiện những điểm giống nhau nhất giữa tất cả các ảnh. Gọi số ảnh học là  $N$ , mỗi ảnh có kích thước  $R \times C$ .

$$S = \{\Gamma_1, \Gamma_2, \dots, \Gamma_M\} \quad (27)$$

Trung bình  $N$  ảnh trên được tính theo công thức:

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma^i \quad (28)$$



**Hình 17 Trung bình ảnh của các tập dữ liệu học**

Sau khi tính được ảnh trung bình của bộ ảnh học, các ảnh đầu vào sẽ được so sánh với ảnh trung bình này. Ảnh mới sẽ chứa những điểm khác nhau giữa ảnh đầu vào so với các ảnh còn lại. Ta có tập vec-tơ đầu vào mới sau khi so sánh với trung bình :

$$A = [\Phi_1, \Phi_2, \dots, \Phi_M]$$

$$\Phi_i = I_i - \Psi, \text{ với } i = 1, M$$

#### ❖ Tính eigenface

Ma trận Covariance (covariance matrix) thể hiện sự tương quan của từng vector đối với các vector còn lại trong không gian.

Cho 2 vector N chiều  $X = [x_1, x_2, \dots, x_n]$  và  $Y = [y_1, y_2, \dots, y_n]$  ta có tương quan giữa 2 vector trên được tính theo công thức.

$$cov(X, Y) = \frac{\sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y})}{N-1} \quad (29)$$

Nếu 2 vector được chuẩn hóa trước, ứng với :

$$cov(X, Y) = \frac{\sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y})}{N-1} = \frac{1}{N-1} (X^T \cdot Y) \quad (30)$$

Cho một tập M vector N chiều  $[\Phi_1, \Phi_2, \dots, \Phi_M]$  xếp thành một ma trận A kích thước N x M, ta có covariance của ma trận A là một ma trận C có kích thước NxN, phần tử  $C(i, j)$  được tính như sau :

$$C(i, j) = cov(\Phi_i, \Phi_j) \quad (31)$$

Vector riêng (eigenvector) và giá trị riêng (eigenvalue): Cho ma trận vuông

$C \in \mathbb{R}^{N \times N}$  vector riêng  $u$  và giá trị riêng  $\lambda$  khi và chỉ khi :  $C \cdot u = \lambda \cdot u$ .

Theo lý thuyết, eigenface chính là các vector riêng  $u$  ứng với các giá trị riêng  $v$  có

giá trị lớn nhất của ma trận covariance:

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T \quad (32)$$

Để tính nhanh các eigenface này, người ta chuyển sang tính toán với ma trận  $L=A^T.A$ . Lựa chọn K vector riêng v ứng với K giá trị riêng lớn nhất của L. Ta có công thức chuyển đổi:  $u = A.v$ . Các giá trị riêng của L cũng chính là các giá trị riêng của C.

Gọi các giá trị riêng của L là:  $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$  sắp xếp theo thứ tự giảm dần K giá trị riêng lớn nhất được lựa chọn theo công thức :

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^m \lambda_i} \geq \varepsilon \quad (33)$$

Trong đó,  $\varepsilon$  là ngưỡng chọn. đề án lựa chọn  $\varepsilon = 0,97$ .

## CHƯƠNG 4 HIỆN THỰC DEMO

### 4.1 Rút trích các đặc trưng của lá

#### 4.1.1 Rút trích đường viền của lá

Phần lớn lá có màu xanh, và màu sắc của lá sẽ thay đổi theo mùa. Vì vậy nó sẽ ảnh hưởng tới kết quả nhận biết lá. Việc chuyển màu sắc của lá đầu vào là bước đầu tiên trong nhận biết đường viền của lá. Nó sẽ làm giảm việc nhận biết những lá không hợp lệ.

Do đó chúng ta cần chuyển màu sắc của lá đầu vào thành hình ảnh xám (gray scale image) theo công thức :

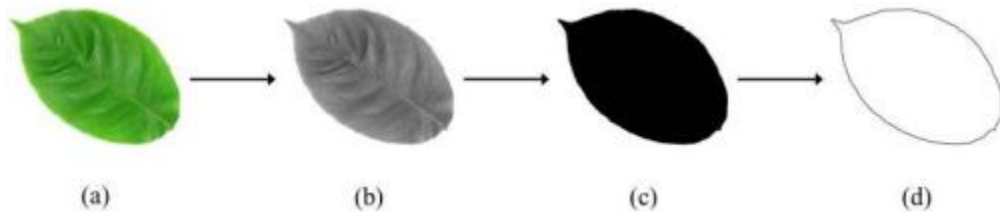
$$Gray = 0.299*R + 0.587*G + 0.114*B \quad (34)$$

Sau đó chúng ta sẽ chuyển lá màu xám (gray scale image) thành lá nhị phân theo công thức:

$$B(x,y) = \begin{cases} 0 & \text{if } f(x,y) \leq T \\ 255 & \text{if } f(x,y) > T \end{cases} \quad (35)$$

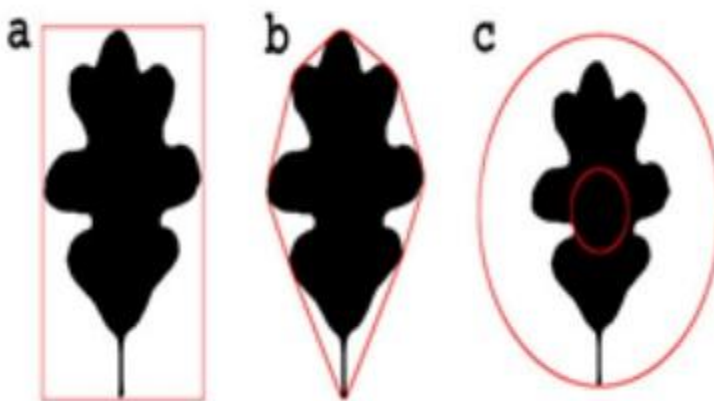
Với  $B(x,y)$  và  $f(x,y)$  là giá trị cường độ của ảnh nhị phân và ảnh xám tại vị trí  $(x,y)$ , và  $T$  là một ngưỡng giá trị nào đó.

Hình vẽ sau khi rút trích đường viền của lá



**Hình 4.1 Rút trích đường viền lá (a) ảnh đầu vào, (b) ảnh xám, (c) ảnh nhị phân, (d) ảnh sau khi rút trích đường viền**

#### 4.1.2 Rút trích các đặc trưng hình học



**Hình 4.2 a) MBR b) Convex Hull c) Incircle and Excircle**

##### ✚ Aspect Ratio :

Là tỷ lệ giữa chiều cao với chiều rộng của hình chữ nhật nhỏ nhất bao quanh chiếc lá (MBR).

$$AR = \frac{\text{Chiều cao}}{\text{Chiều rộng}} \quad (36)$$

##### ✚ Rectangularity

Biểu diễn cho tỷ lệ giữa diện tích của cái lá ( $A_{ROI}$ ) và diện tích của hình chữ nhật nhỏ nhất bao quanh (MBR)

$$R = \frac{A_{ROI}}{A_{MBR}} \quad (37)$$

##### ✚ Area ratio of Convex Hull

Biểu diễn cho tỷ lệ giữa diện tích của cái lá ( $A_{ROI}$ ) và diện tích của convex hull ( $A_{CV}$ ).

$$CA = \frac{A_{ROI}}{A_{CV}} \quad (38)$$

#### Perimeter ratio of convex hull

Biểu diễn cho chu vi của cái lá ( $P_{ROI}$ ) và chu vi của convex hull ( $P_{CV}$ ).

$$CP = \frac{P_{ROI}}{P_{CV}} \quad (39)$$

#### Sphericity

Biểu diễn cho tỷ lệ giữa bán kính đường tròn trong ( $r_i$ ) và bán kính đường tròn ngoài bao quanh chiếc lá ( $r_c$ )

$$S = \frac{r_i}{r_c} \quad (40)$$

#### Circularity

Được định nghĩa bởi tất cả các điểm trên đường viền của lá.

$$C = \frac{\mu_R}{\sigma_R} \quad (41)$$

Với  $\mu_R$  là khoảng cách trung bình giữa trong tâm tới tất cả các đường viền, và  $\sigma_R$  là bình phương độ lệch với trung bình khoảng cách.

$$C(x_{centroid}, y_{centroid}) = C\left(\frac{1}{N} \sum_1^N x_n, \frac{1}{N} \sum_1^N y_n\right)$$

$$\mu_R = \frac{1}{N} \sum_1^N ||(x_i, y_j) - (x_{centroid}, y_{centroid})|| \quad (42)$$

$$\sigma_R = \frac{1}{N} \sum_1^N (|| (x_i, y_j) - (x_{centroid}, y_{centroid}) || - \mu_R)^2$$

#### Eccentricity

Được định nghĩa là chiều dài trục chính của đường tròn trong  $E_A$  với chiều dài trục phụ đường tròn ngoài  $E_B$ .

$$E = \frac{E_A}{E_B} \quad (43)$$

#### Form Factor

Được sử dụng phổ biến để nhận mô tả hình dạng. Được xác định bởi công



thức:

$$F = \frac{4\pi A_{ROI}}{P_{ROI}^2} \quad (44)$$

### 🚧 Những giá trị bất biến

Chúng ta sử dụng 7 giá trị bất biến của HU. 7 giá trị này sẽ không thay đổi khi chúng ta dịch chuyển, tỷ lệ, và quay chiếc lá.

Những giá trị được tính từ đường viền theo công thức của Chen's.

$$M_{pq} = \int_C x^p y^q ds$$

Với  $p, q = 0, 1, 2, 3, \dots$

$\int_C$  là tích phân đường trên đường viền  $C$

$$ds = \sqrt{(dx)^2 + (dy)^2}$$

Hoặc có thể tính  $M_{pq}$  theo công thức

$$M_{pq} = \sum_x \sum_y x^p y^q$$

Trọng tâm tạm thời trên đường viền có thể được tính theo công thức

$$\begin{aligned} \mu_{pq} &= \int_C (x - \bar{x})^p (y - \bar{y})^q ds, \\ \bar{x} &= \frac{M_{10}}{M_{00}}, \quad \bar{y} = \frac{M_{01}}{M_{00}}. \end{aligned} \quad (45)$$

Hoặc

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q, \quad (46)$$

Và chuẩn hóa trọng tâm tạm thời bằng công thức

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}. \quad (47)$$

Với  $\gamma = \frac{p+q}{2} + 1$ ;  $p+q = 2, 3, 4, \dots$

Tập 7 giá trị bất biến được tính theo công thức

$$\begin{aligned} Hu1 &= \eta_{20} + \eta_{02}, \\ Hu2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\ Hu3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2, \\ Hu4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2, \\ Hu5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2], \\ Hu6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}), \\ Hu7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] \\ &\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2]. \end{aligned} \quad (48)$$

#### 4.1.3 Chuẩn hóa dữ liệu

Sau khi tính được các đặc trưng hình học theo nhiều kỹ thuật khác nhau. Chúng ta cần chuẩn hóa dữ liệu trước khi sử dụng để tính độ tương đồng giữa các vecto. Để tránh các vecto đặc thù ảnh hưởng tới kết quả chung. Ta sẽ chuẩn hóa từng vecto theo công thức sau

$$v_{ijnormalized} = \frac{v_{ij} - \frac{\sum_1^n v_{ij}}{n}}{\max_{1 \leq i \leq n} v_{ij}} \quad j = 1, \dots, m$$

Với  $+ v_{ij}$  là vecto đặc trưng thứ j của lá thứ i

$+ v_{ijnormalized}$  là vecto  $v_{ij}$  được chuẩn hóa

$+ n$  là số lá

$+ m$  số vecto đặc trưng

#### 4.1.4 Tính độ tương đồng

Sau khi chuẩn hóa các vecto đặc trưng. Chúng ta sẽ sử dụng để tính độ tương đồng giữa các lá. Độ tương đồng tính theo khoảng cách Euclidean

$$distance_e = \sqrt{\sum_1^n (A_i - T_i)^2}$$

Những lá có khoảng cách càng nhỏ thì độ tương đồng càng cao.

## 4.2 Demo

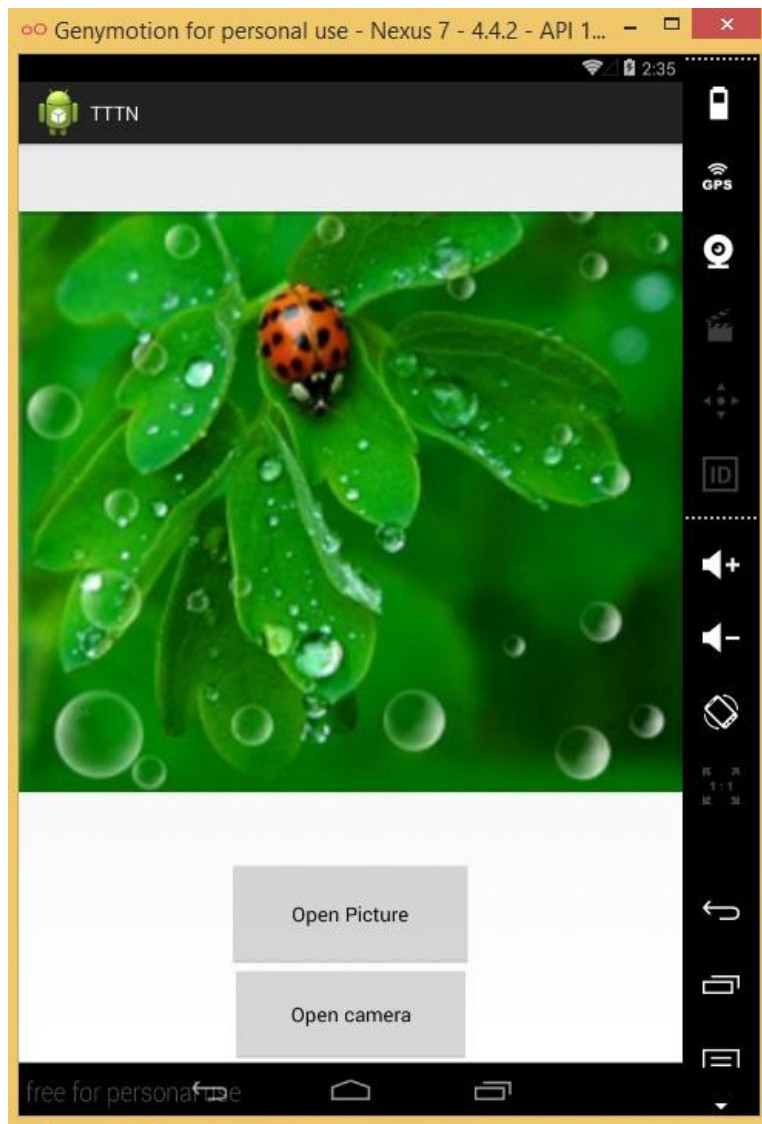
- Đã hoàn thiện các đặc trưng cơ bản.
- Chưa hiện thực 7 giá trị bất biến nên tạm thời chỉ xử lý được những chiếc lá theo hướng thẳng đứng.
- Cơ sở dữ liệu



leaf.information	
id	int(11)
name	varchar(20)
description	varchar(200)
time	timestamp
uses	varchar(200)
image	mediumtext
aspectRatio	double
rectangularity	double
areaRatioCV	double
preRatioCV	double
narrowFactor	double
wdd	double
formFactor	double
roundness	double
spherity	double
circularity	double
eccentricity	double

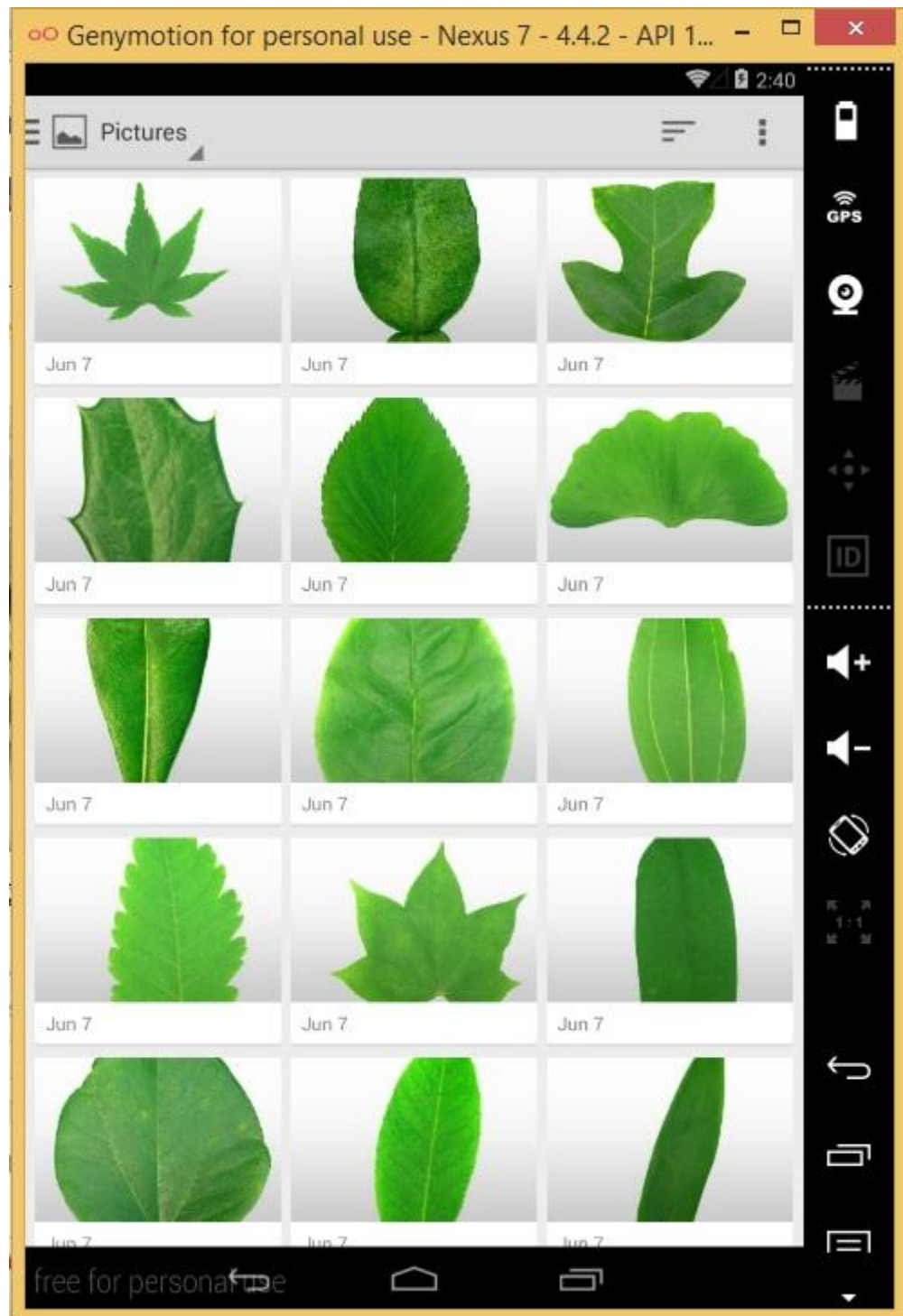
**Hình 4.3 Cơ sở dữ liệu của demo**

- Màn hình chính của ứng dụng có 2 chế độ là mở ảnh và sử dụng camera:



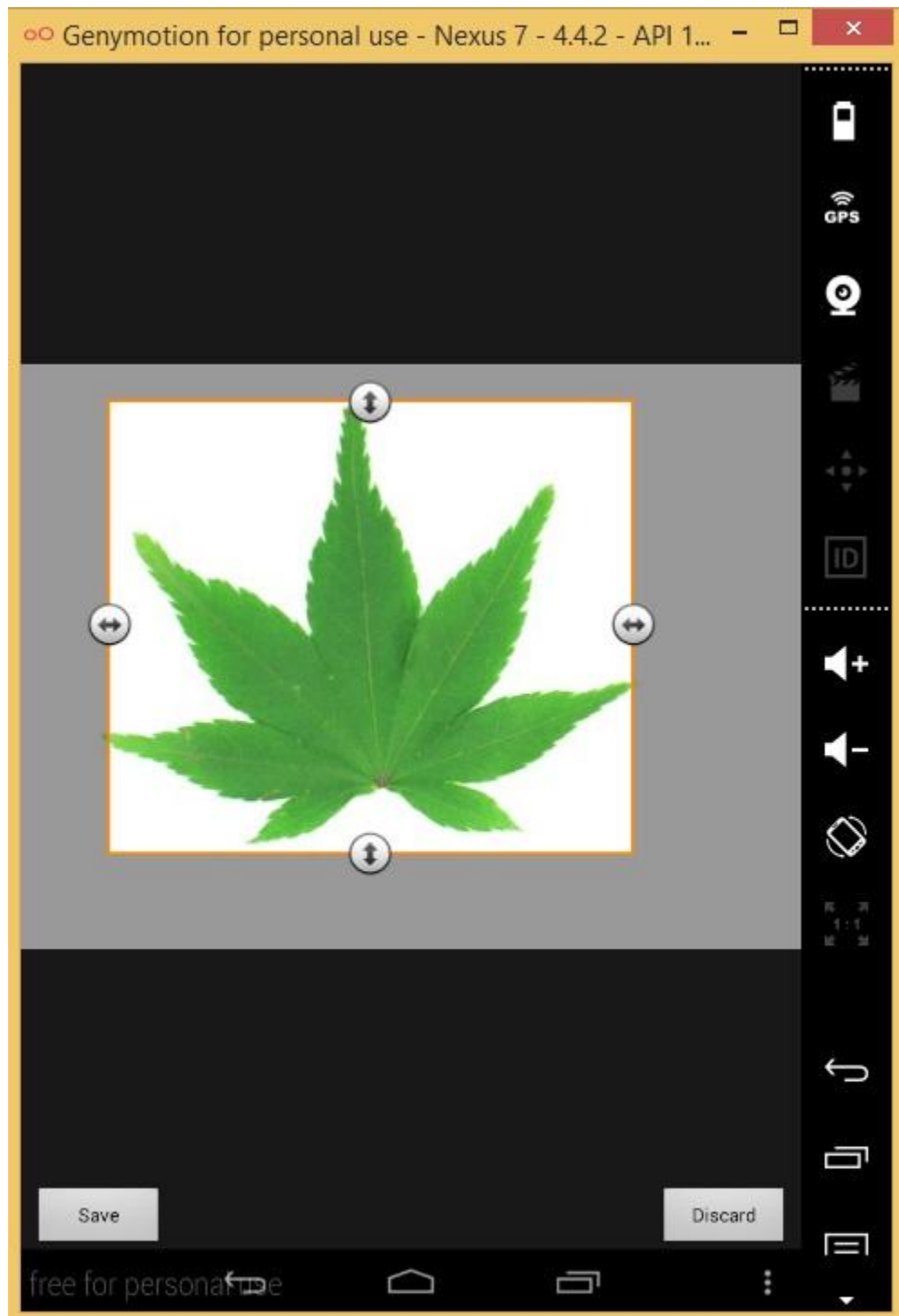
**Hình 4.4 Màn hình chính của demo**

- Mở ảnh



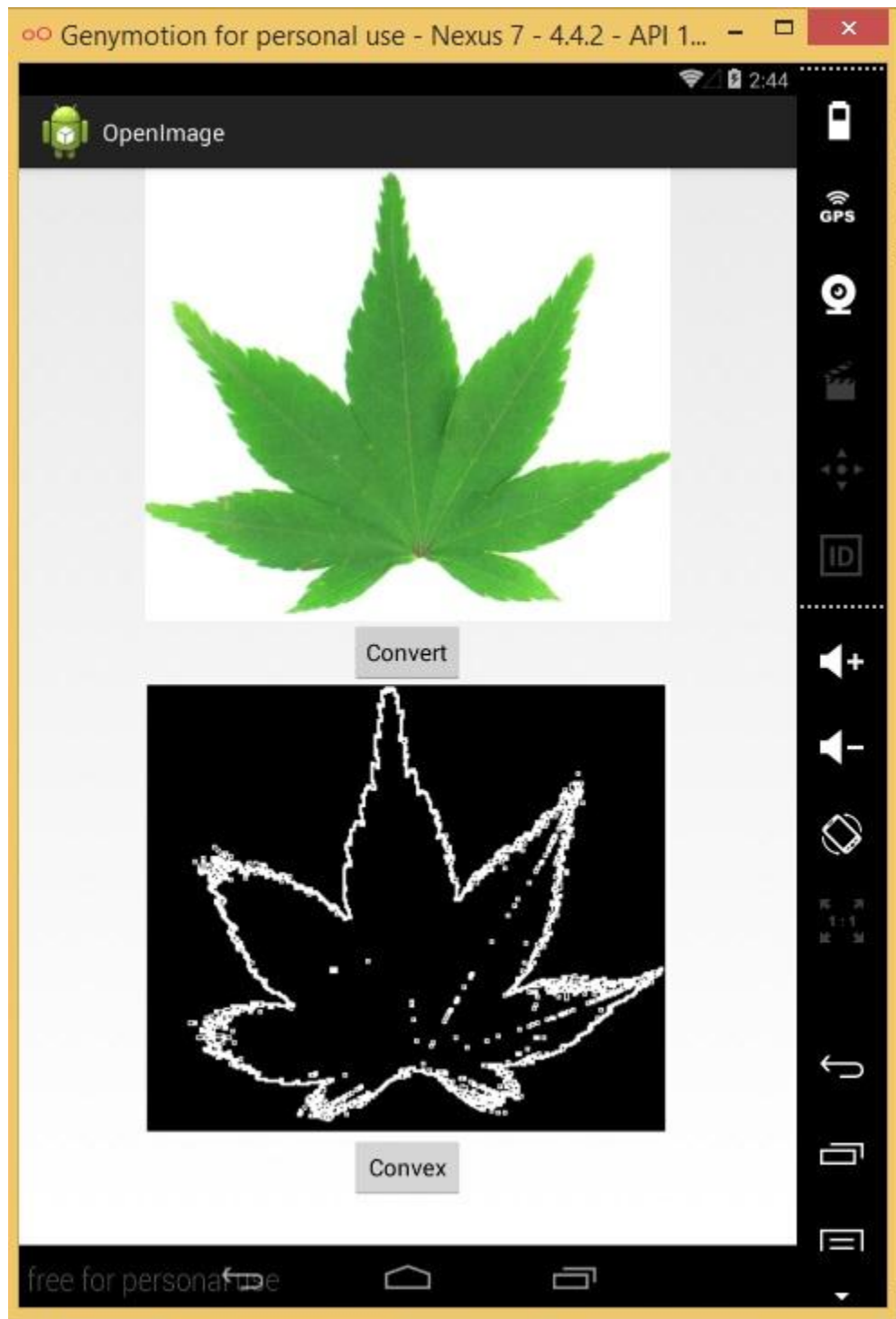
**Hình 4.5 Mở tập dữ liệu ảnh**

- Chọn vùng ảnh cái lá



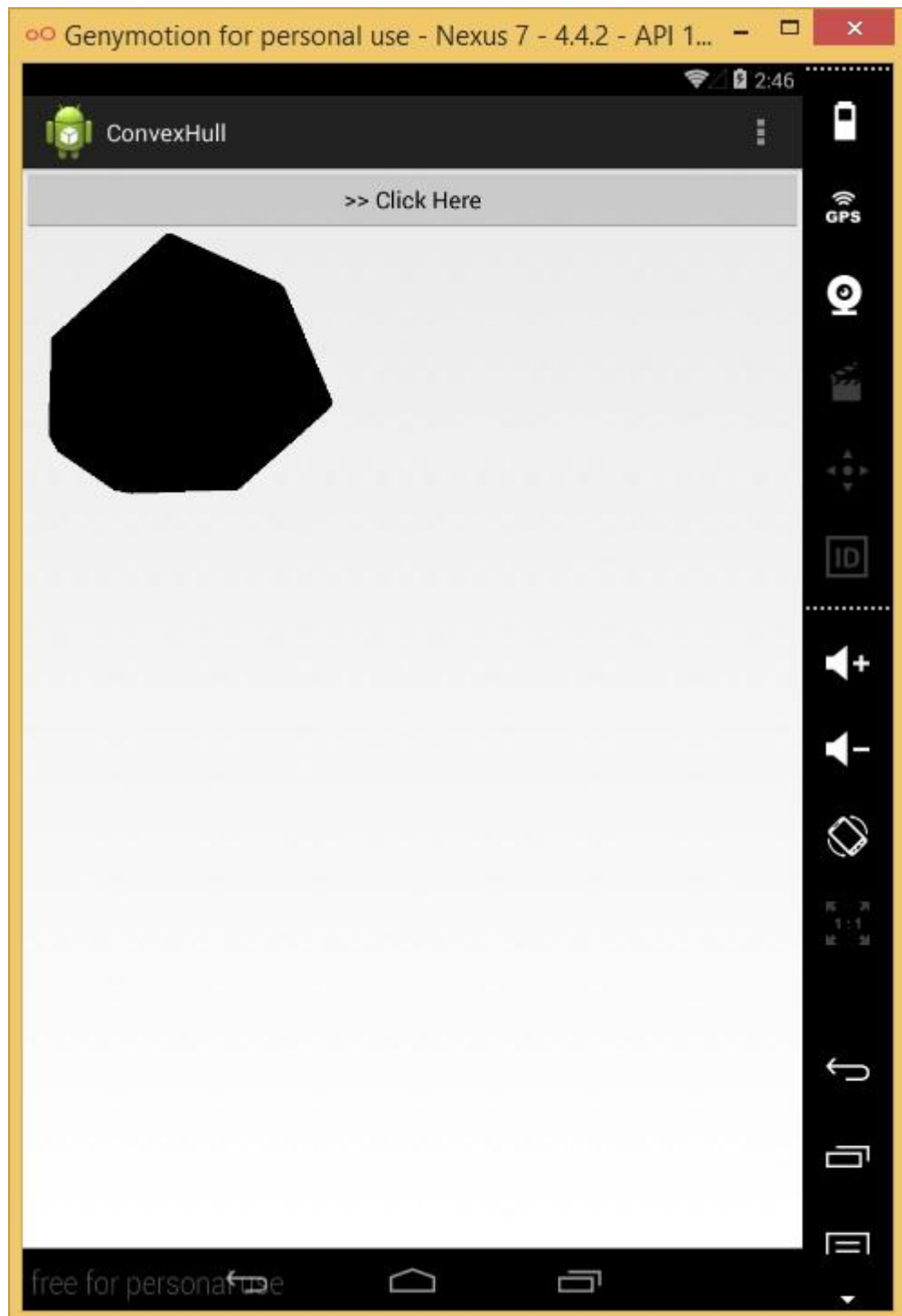
**Hình 4.6 Chọn vùng lá**

- Rút trích đường viền



**Hình 4.7 Rút trích đường viền**

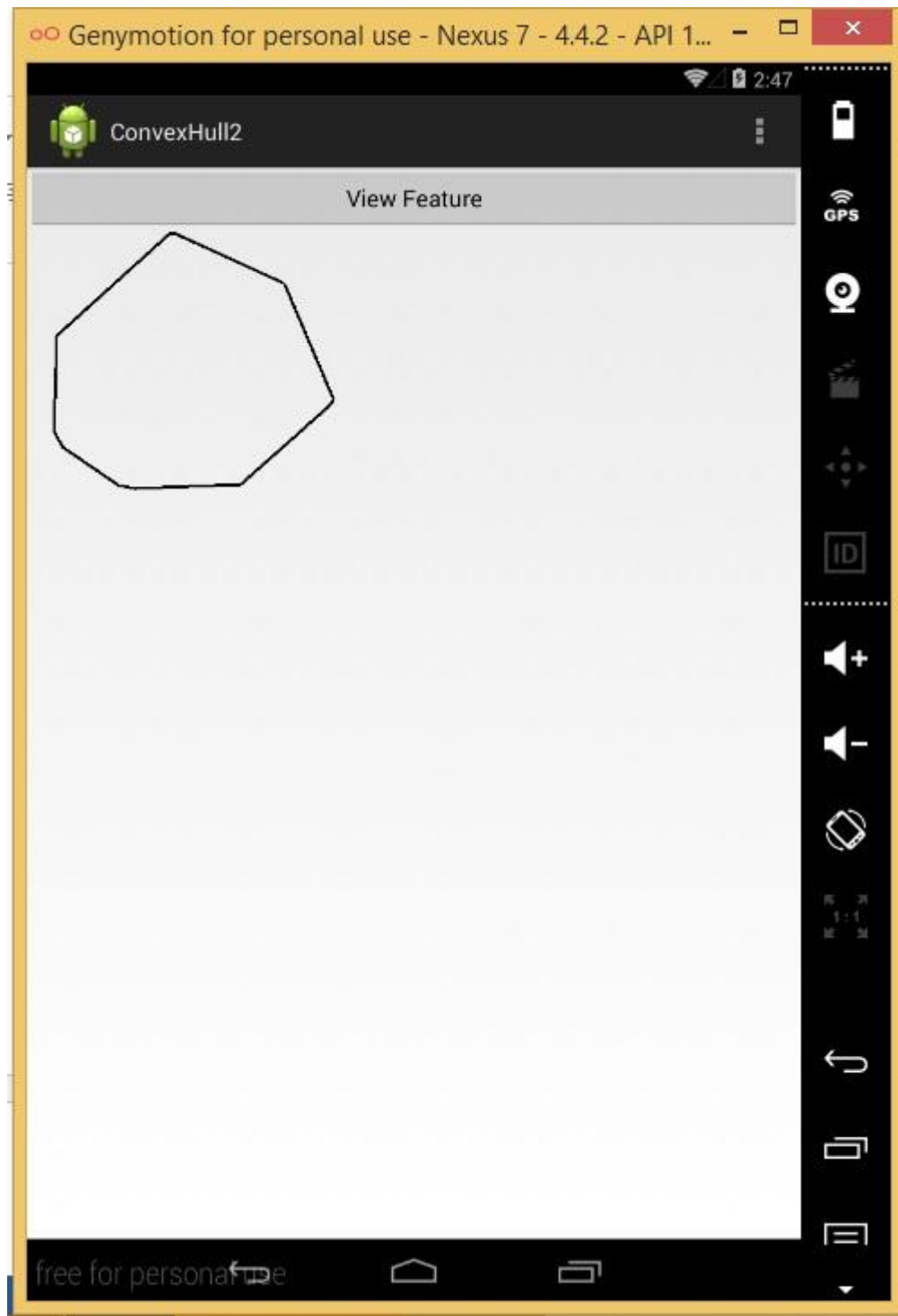
- Tính diện tích convex hull



**Hình 4.8 Diện tích Hull**

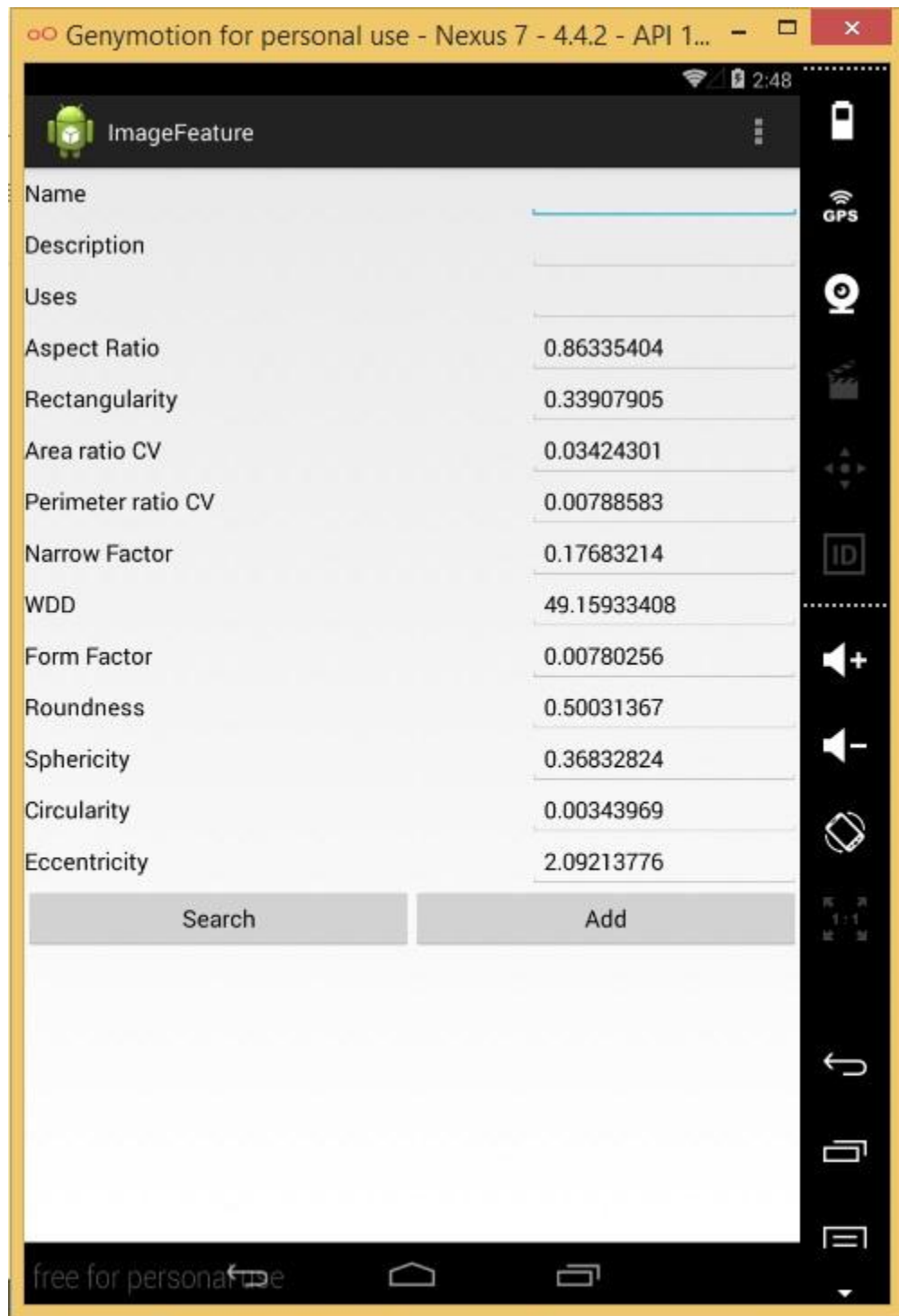
- Tính chu vi convex hull





**Hình 4.9 Chu vi Hull**

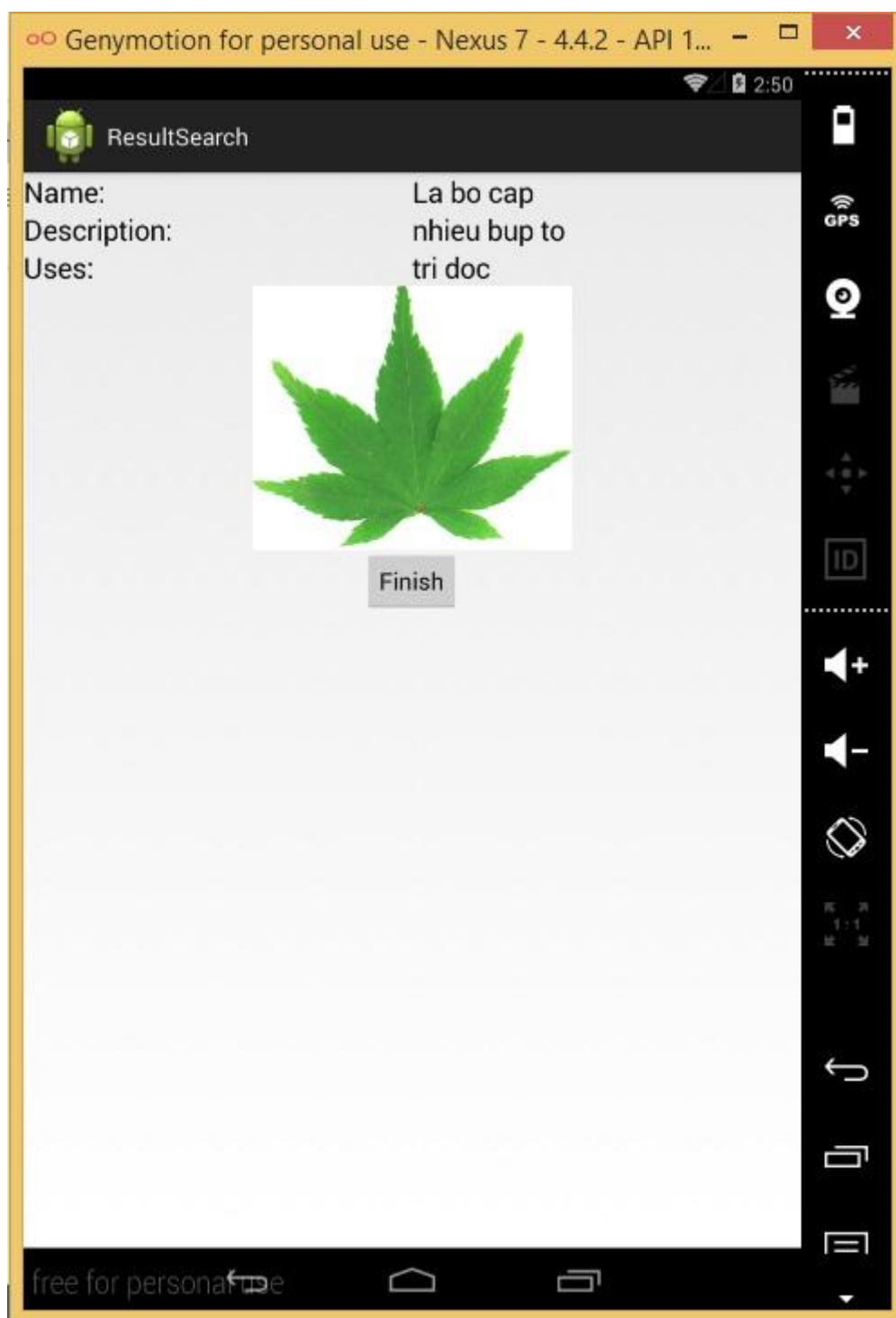
- Các đặc trưng của chiếc lá



**Hình 4.10 Các đặc trưng của lá**

- Tạm thời chức năng add được dùng để thêm dữ liệu vào database. Sau này phát triển phần mềm nên cao sẽ bố trí hợp lý hơn.

- Ta chọn tính năng Search. Và kết quả như sau



**Hình 4.11 Tính năng Search**

## CHƯƠNG 5 TỔNG KẾT

### 5.1 Đánh giá

Do hạn chế về thời gian nên nhóm dù nghiên cứu nhiều phương pháp nhưng chỉ có thể hiện thực được phương pháp cơ bản về nhận dạng lá cây.

Cơ sở dữ liệu còn quá ít, ứng dụng còn sơ xài.

Tính đến thời điểm hiện tại vs cơ sở dữ liệu gồm 15 lá có hình dạng khác nhau nên chưa xảy ra sai sót. Và kết quả đúng 15/15 lá.

Chỉ xử lý những lá theo hướng thẳng đứng. Chưa xử lý được ảnh chụp đầu vào bị nghiêng.

### 5.2 Hướng phát triển luận văn

Hướng xây dựng ứng dụng nhận dạng lá cây đã có những nghiên cứu và phát triển bởi một số các nhà phát triển nhưng đa phần được hiện thực trên công cụ Matlab. Về các ứng dụng trên Android chưa được chú ý phát triển nhiều. Nhóm sẽ tiếp tục nghiên cứu để tiếp tục phát triển ứng dụng trên Android.

Nghiên cứu thêm các lá cây mới để làm giàu thêm cơ sở dữ liệu, cải thiện giao diện chương trình và thêm một số chức năng mới, cải thiện thời gian nhận dạng ảnh, xử lý các chức năng chưa hiện thực được.

## TÀI LIỆU THAM KHẢO

- [1]. Tee Cheng Siew(2008).Feature selection for content-based image retrieval using statistical discriminant analysis. *PhD thesis Faculty of Computer Science and Information System Universiti Teknologi Malaysia. 2008.*
- [2]. Website : [http://www.freeseamlesstextures.com/texture\\_gallery/index.htm](http://www.freeseamlesstextures.com/texture_gallery/index.htm)
- [3]. C. V. Jawahar, P. J. Narayanan, and S. Rakshit(2000).A flexible scheme for representation, matching, and retrieval of images. *ICVGIP 2000, pages 271–277. Allied Publishers Ltd., 2000.*
- [4]. Michele Saad (2008). Low-Level Color and Texture Feature Extraction for Content-Based Image Retrieval . *EE 381K: Multi-Dimensional. Digital Signal Processing.*
- [5]. Yossi Rubner, an Puzicha, Carlo Tomasi and Joachim M. Buhmann Empirical: Evaluation of Dissimilarity Measures for Color and Texture. *Computer Vision and Image Understanding, vol 84, issue 1. Elsvier Science Ins.*

- [6]. Kamarul Hawari Ghazali(2007). Feature Extraction technique using SIFT keypoints descriptors. *The International Conference on Electrical and Engineering and Informatics Institut technology Bandung, Indonesia, june 17-19, 2007*
- [7]. Mohamed Aly(2006).Face Recognition using SIFT Features. *AlyCNS186 Term Project Winter.*
- [8]. [http://en.wikipedia.org/wiki/Support\\_vector\\_machine](http://en.wikipedia.org/wiki/Support_vector_machine)
- [9]. Bài báo khoa học: Leaf Image Classification Using Support Vector Machine
- [10]. [http://en.wikipedia.org/wiki/Principal\\_component\\_analysis](http://en.wikipedia.org/wiki/Principal_component_analysis)
- [11]. A. Ehsanirad, “Plant classification based on leaf recognition,” *International Journal of Computer Science and Information Security*, vol. 8,no. 4, pp. 78–81, 2010.
- [12]. T. Beghin, J. S. Cope, P. Remagnino, and S. Barman, “Shape and texture based plant leaf classification,” in *ACIVS* (2), 2010, pp. 345–353.
- [13]. Ji-Xiang Du, Xiao-Feng Wang, D.S. Huang, Automatic plant leaves recognition system based on image processing techniques, Technical Report, Institute of Intelligent Machines, Chinese Academy of Sciences, October, 2004.
- [14]. S. Abbasi, F. Mokhtarian, J. Kittler, ‘Reliable classification of chrysanthemum leaves through curvature scale space’, in: *International Conference on Scale-Space Theory in Computer Vision*, Netherlands, 1997, pp. 284–295.