

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN
BỘ MÔN CÔNG NGHỆ TRI THỨC**

NGUYỄN QUỐC UY - 9912745

TÌM KIẾM ẢNH DỰA VÀO NỘI DUNG

LUẬN VĂN CỬ NHÂN TIN HỌC

GIÁO VIÊN HƯỚNG DẪN

Th.s LÝ QUỐC NGỌC

TP.HCM, 07/2003

LỜI CẢM ƠN

Trước tiên, em xin chân thành gửi lời cảm ơn đến Thầy Lý Quốc Ngọc, người đã nhiệt tình, tận tâm, hướng dẫn về kiến thức để em có thể làm được luận văn này. Đúng lúc tưởng chừng đã bị bế tắc trong công việc, Thầy đã động viên giúp em mạnh dạng hoàn thành luận văn này. Một lần nữa em xin tỏ lòng biết ơn Thầy.

Sau hết, xin gửi lời cảm ơn đến gia đình bạn bè, người thân đã ủng hộ động viên tinh thần để luận văn được hoàn thành.

Xin chân thành cảm ơn tất cả !

TPHCM, 07/2003

Người thực hiện

Nguyễn Quốc Uy

Khoa CNTT - ĐH KH. TP. HCM

MỤC LỤC

LỜI MỞ ĐẦU

PHẦN 1: TỔNG QUAN

1. Sự hình thành bài toán	3
2. Cách tiếp cận:	3
2.1. Đặc trưng màu sắc:	4
2.2. Đặc trưng vân:	4
2.3. Đặc trưng hình dáng:	4
2.4. Độ đo:	4
2.5. Mô hình giao diện:	5

PHẦN 2: CÁC PHƯƠNG PHÁP TÌM KIẾM ẢNH DỰA VÀO NỘI DUNG

Chương 1: Tìm kiếm ảnh dựa vào màu sắc.....	9
1. Màu sắc:	10
1.1. Một số đặc tính vật lý đặc biệt của màu sắc.....	10
1.2. Hệ thống màu chuẩn RGB	10
1.3. Hệ thống màu CMY	12
1.4. Hệ thống màu L^*a^*b	12
1.5. Hệ thống màu HSI.....	12
2. Tìm kiếm ảnh dựa vào màu sắc	14
2.1. Lướt đồ màu	14
2.2. Các loại độ đo màu sắc.....	19
Chương 2: Tìm kiếm ảnh dựa vào vân.....	22

1. Vân.....	23
1.1. Vân là gì?	23
1.2. Một số loại vân tiêu biểu.....	24
2. Tìm kiếm ảnh dựa vào vân.....	25
2.1. Mật độ của đường biên và hướng của biên.....	25
2.2. Phân hoạch vùng nhị phân cục bộ.....	27
2.3. Ma trận đồng hiện và đối tượng đồng hiện	28
2.4. Độ đo năng lượng của vân dựa vào luật đo	31
2.5. Tương quan tự động và quang phổ năng lượng.....	33
2.6. Phân đoạn vân (Texture segmentation)	34
Chương 3: Tìm kiếm ảnh dựa vào hình dạng.....	35
1. Hình dạng.....	36
1.1. Khái niệm về hình dạng.....	36
1.2. Đặc điểm hình dạng đối với việc tìm kiếm ảnh	36
2. Tìm kiếm ảnh dựa vào hình dạng.....	37
2.1. Lướt đồ hình dạng.....	37
2.2. Độ so khớp đường biên của hình dạng.....	38
2.3. So khớp với ảnh phát họa.....	40
PHẦN 3 CÀI ĐẶT VÀ KẾT QUẢ THỬ NGHIỆM	
Chương 1: Cài đặt.....	44
1. Chương trình	45
2. Phần Màu sắc.....	46

3. Phần Văn.....	50
4. Phần Hình dạng.....	53
Chương 2: Kết quả thử nghiệm	54
1. Phần Màu sắc.....	55
2. Phần Văn.....	73
3. Phần Hình dạng.....	77
PHẦN 4 KẾT LUẬN	
Đánh giá kết quả đạt được	80
Hướng phát triển.....	80
Tài liệu tham khảo	81

Khoa CNTT - ĐH KHTN TP.HCM

TÓM TẮT LUẬN VĂN

Tên đề tài : Tìm kiếm ảnh dựa vào nội dung

Tìm kiếm ảnh dựa vào nội dung là một đề tài nghiên cứu về khả năng để xây dựng một chương trình cho phép tìm kiếm ảnh trên một cơ sở dữ liệu ảnh cho trước. Việc tìm kiếm này dựa trên nội dung của những bức ảnh trong cơ sở dữ liệu. Nội dung của một bức ảnh có thể là màu sắc của bức ảnh, vân của bức ảnh, hay những kiểu hình dạng.

Luận văn này tập trung vào việc nghiên cứu những lý thuyết để xây dựng nên một ứng dụng như vậy, đồng thời ứng dụng lý thuyết đó để xây dựng nên một số demo, cho thấy khả năng có thể phát triển được thành ứng dụng hoàn thiện được.

Đề tài gồm có ba phần: tìm kiếm ảnh dựa vào màu sắc, vân và hình dạng.

Tìm kiếm ảnh dựa vào màu sắc: cách tiếp cận chính là dựa vào lượt đồ màu (colour histogram)

Tìm kiếm ảnh dựa vào vân: cách tiếp cận chính là dựa vào ma trận đồng hiện (co-occurrence matrix)

Tìm kiếm ảnh dựa vào hình dạng: cách tiếp cận chính là dựa vào sự so khớp ảnh phát họa (sketch matching)

LỜI MỞ ĐẦU

Sự mở rộng của multimedia, cùng với khối lượng hình ảnh và phim lớn, sự phát triển của những xa lộ thông tin hiện tại đã thu hút ngày càng nhiều những chuyên gia đi vào nghiên cứu những công cụ cung cấp cho việc lấy thông tin từ dữ liệu ảnh từ nội dung của chúng. Lấy thông tin từ dữ liệu ảnh có liên quan đến rất nhiều các lĩnh vực khác, từ những phòng trưng bày tranh nghệ thuật cho tới những nơi lưu trữ tranh nghệ thuật lớn như viện bảo tàng, kho lưu trữ ảnh chụp, kho lưu trữ ảnh tội phạm, cơ sở dữ liệu ảnh về địa lý, y học, ... điều đó làm cho lĩnh vực nghiên cứu này phát triển nhanh nhất trong công nghệ thông tin.

Lấy thông tin từ dữ liệu ảnh đặt ra nhiều thách thức nghiên cứu mới cho các khoa học gia và các kỹ sư. Phân tích ảnh, xử lý ảnh, nhận dạng mẫu, giao tiếp giữa người và máy là những lĩnh vực nghiên cứu quan trọng góp phần vào phạm vi nghiên cứu mới này.

Khía cạnh tiêu biểu của lấy thông tin từ dữ liệu ảnh dựa trên những công bố có sẵn như là những đối tượng nhận thức như màu sắc, vân (texture), hình dáng, cấu trúc, quan hệ không gian, hay thuộc về ngữ nghĩa căn bản như: đối tượng, vai trò hay sự kiện hay liên quan đến thông tin về ngữ nghĩa quan hệ như cảm giác, cảm xúc, nghĩa của ảnh. Thật ra phân tích ảnh, nhận dạng mẫu, hay xử lý ảnh đóng một vai trò căn bản trong hệ thống lấy thông tin từ ảnh. Chúng cho phép sự trích rút tự động hầu hết những thông tin về nhận thức, thông qua phân tích sự phân bố điểm ảnh và sự phân tích độ đo.

Tìm kiếm theo cách thông thường dựa trên văn bản giờ đây được bổ sung bởi truy vấn dựa vào nội dung, nhắm vào khía cạnh nhận thức thông tin ảnh. Thực hiện truy vấn ở mức nhận thức đòi hỏi những phương thức mới, cho phép chỉ định đến những thuộc tính liên quan đến thị giác cần tìm. Giao diện đồ họa phải hỗ trợ đặc tả những ví dụ đó như là những mẫu có sẵn. Khi đó người dùng trong một vòng lặp, mô hình giao diện sao cho người dùng có thể truy cập vào sự giống nhau giữa những đối tượng.

PHẦN 1

TỔNG QUAN

Khoa CNTT - ĐH KHTN TP.HCM

1. Sự hình thành bài toán:

Bên cạnh kho dữ liệu văn bản, kho dữ liệu ảnh ngày càng trở nên khổng lồ vượt quá sự kiểm soát của con người. Khi có nhu cầu tìm kiếm một vài tấm ảnh nào đó trong một cơ sở dữ liệu hàng trăm ngàn ảnh, điều này khó có thể thực hiện được khi ta tìm kiếm bằng tay theo cách thông thường, nghĩa là xem lần lượt từng tấm ảnh một cho đến khi tìm thấy ảnh có nội dung cần tìm. Song song với sự phát triển của những phương tiện kỹ thuật số, trong tương lai, số lượng ảnh sẽ còn tăng nhanh hơn nữa, nhiều hơn nữa. Do đó, nhu cầu thật sự đòi hỏi phải có một công cụ hỗ trợ cho việc tìm kiếm này càng sớm càng tốt. Vì vậy đề tài “tìm kiếm dữ liệu ảnh dựa vào nội dung” ra đời để góp phần đáp ứng nhu cầu này.

“Tìm kiếm dữ liệu ảnh dựa vào nội dung” là gì? Đây là một chủ đề nghiên cứu mới trong công nghệ thông tin. Mục đích chính của nó là lấy những ảnh từ cơ sở dữ liệu phù hợp với tiêu chí truy vấn.

Thế hệ đầu tiên của hệ thống tìm kiếm ảnh dựa vào nội dung cho phép truy cập trực tiếp đến ảnh thông qua thuộc tính chuỗi. Những tìm kiếm đặc thù cho những hệ thống dạng này là “tìm tất cả những tranh vẽ của trường Florentine trong thế kỷ thứ XV” hay “tìm kiếm tất cả những tấm ảnh về đất đai của Cezanne”. Metadata của hệ thống trong thế hệ đầu tiên dựa trên chuỗi, sơ đồ trình bày, mô hình quan hệ, cấu trúc khung. Xem hình 1.1.

Thế hệ mới của hệ thống tìm kiếm ảnh hỗ trợ đầy đủ việc lấy thông tin dựa vào nội dung thuộc về thị giác. Chúng cho phép phân tích đối tượng, tự động trích rút đặc trưng. Xem hình 1.2.

Những phần tử thuộc về thị giác như là màu sắc, vân, hình dạng đối tượng, quan hệ không gian có liên quan trực tiếp đến khía cạnh nhận thức của nội dung ảnh. Ta trực tiếp đánh vào những phần tử này trong việc tìm kiếm.

Do đó, nội dung của đề tài sẽ giải quyết như sau: hệ thống tìm kiếm ảnh dựa vào nội dung, ba đặc trưng là màu sắc, vân và hình dạng. Đề tài chỉ tập chung vào demo một số phương pháp làm của từng phần.

2. Cách tiếp cận:

Đề tài tiếp cận theo mô hình thống kê.

Có ba loại tìm kiếm ảnh đề tài tập chung giải quyết là :

- Tìm kiếm ảnh dựa vào màu sắc.
- Tìm kiếm ảnh dựa vào vân.
- Tìm kiếm ảnh dựa vào hình dạng.

2.1. Đặc trưng màu sắc:

Màu sắc là vấn đề cần tập chung giải quyết nhiều nhất, vì một ảnh màu thì thông tin quan trọng nhất trong ảnh chính là màu sắc. Hơn nữa thông tin về màu sắc là thông tin người dùng quan tâm nhất; qua đặc trưng màu sắc, có thể lọc được rất nhiều lớp ảnh, thông qua vị trí, không gian, định lượng của màu trong ảnh.

2.2. Đặc trưng vân:

Có những lớp ảnh mà màu sắc không thể giải quyết được, đòi hỏi phải dùng đặc trưng vân. Ví dụ như những ảnh liên quan đến cấu trúc của điểm ảnh như: cỏ, mây, đá, sợi.

Vân sẽ giải quyết tốt cho việc tìm kiếm đối với lớp ảnh này.

2.3. Đặc trưng hình dáng:

Đối với những lớp ảnh cần tìm mà liên quan đến hình dạng của đối tượng thì đặc trưng vân và màu không thể giải quyết được. Ví dụ như tìm một vật có hình dạng ellipse hay hình tròn trong ảnh.

Tìm kiếm theo hình dáng thật sự là một cái đích của hệ thống tìm kiếm dựa vào nội dung muốn đạt tới.

2.4. Độ đo:

Có ý nghĩa quan trọng trong tìm kiếm ảnh dựa vào nội dung. Độ đo mang ý nghĩa quyết định kết quả tìm kiếm sẽ như thế nào, mức độ chính xác.

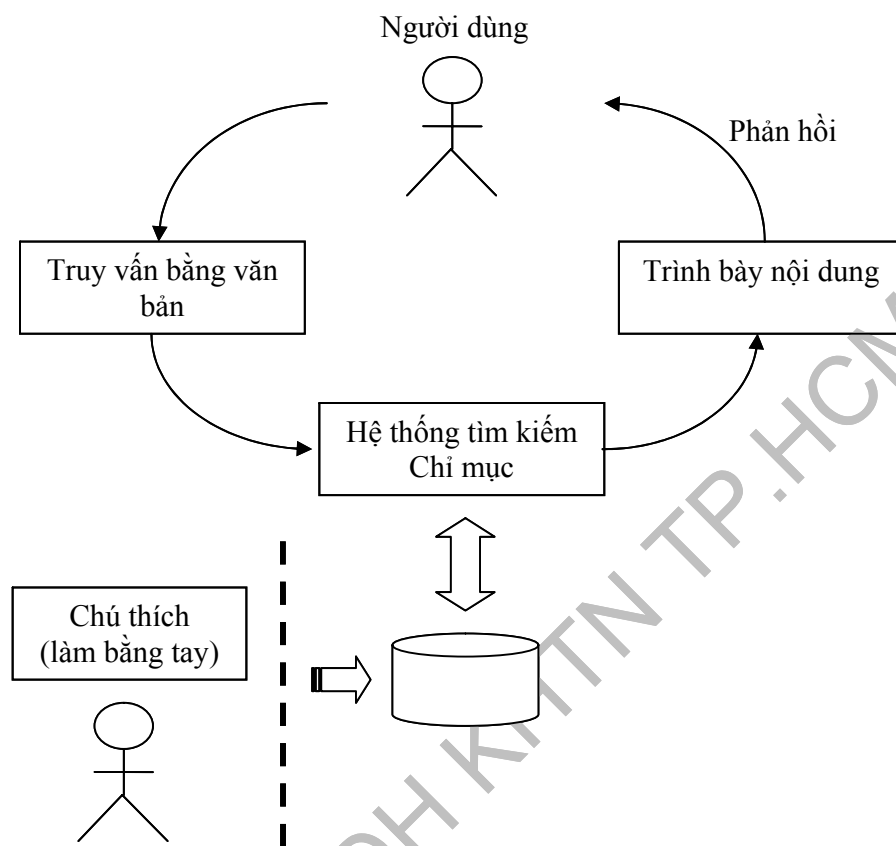
2.5. Mô hình giao diện:

Giao diện cũng là một vấn đề đáng quan tâm của bài toán tìm kiếm ảnh. Mô hình giao diện dựa trên mô hình tìm kiếm cho phép lặp.

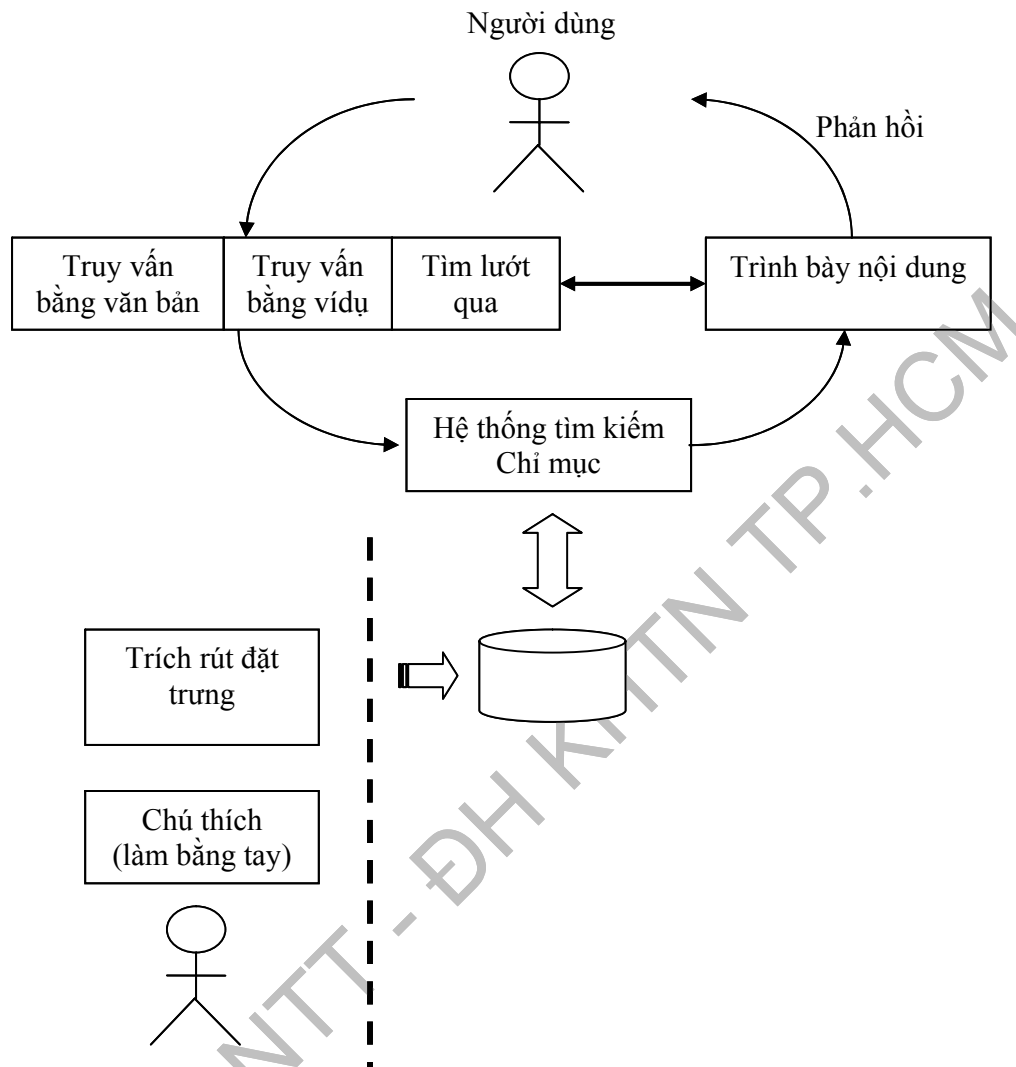
Đây mô hình người dùng có thể đặc tả yêu cầu tìm kiếm với nhiều cấp độ. Quá trình tìm kiếm có thể là một quá trình lặp đi lặp lại nhiều lần: tìm kiếm, lọc, chỉnh sửa ngưỡng cho đến khi đạt đến kết quả như mong muốn. Do đó giao diện chương trình phải hỗ trợ chức năng cho phép chọn lọc kết quả, loại bỏ những kết quả không phù hợp, lưu lại những kết quả đã vừa ý, tìm kiếm trong kết quả vừa tìm được, thay đổi mức độ chính xác của việc tìm kiếm.

Hai vấn đề được đề tài quan tâm về giao diện tìm kiếm là: mô hình tìm kiếm theo không gian toàn cục và cục bộ. Đối với đặc trưng màu: giao diện cho phép giao tiếp cả toàn cục và cục bộ. Đặc trưng văn và hình dạng đề tài chỉ dừng lại ở mức toàn cục.

Khoa CNTT - ĐH KHTN - HCM



Hình 1.1 Mô hình của hệ thống tìm kiếm thế hệ đầu tiên



Hình 1.2 Mô hình của hệ thống tìm kiếm thế hệ mới

PHẦN 2
CÁC PHƯƠNG PHÁP TÌM KIẾM
ẢNH DỰA VÀO NỘI DUNG

Khoa CNTT - ĐH KHTN TP.HCM

Chương 1:

Tìm kiếm ảnh dựa vào màu sắc

1. Màu sắc

1.1. Một số đặc tính vật lý đặc biệt của màu sắc

1.2. Hệ thống màu chuẩn RGB

1.3. Hệ thống màu CMY

1.4. Hệ thống màu L^*a^*b

1.5. Hệ thống màu HSI

2. Tìm kiếm ảnh dựa vào màu sắc

2.1. Lướt đồ màu

2.2. Các loại độ đo màu sắc

Khoa CNTT - ĐHQHTN TP.HCM

1. Màu sắc:

Sự nhận thức về màu sắc là quá trình quan trọng của con người. Sự nhận thức về màu sắc phụ thuộc vào cả tính chất vật lý của ánh sáng và quá trình xử lý của thị giác với sự góp phần rất quan trọng của kinh nghiệm. Con người dùng thông tin màu sắc để phân biệt đối tượng, vật liệu, đồ ăn, vị trí và ngay cả thời gian của ngày.

Với sự phát triển mạnh mẽ của các thiết bị kinh tế, máy móc xử lý màu sắc trở nên thông dụng: Chúng ta có các thiết bị như máy quay phim màu, thiết bị chiếu màu và những phần mềm xử lý ảnh màu. Máy móc có thể dùng màu sắc cho những mục đích như là con người. Đặc biệt, màu sắc thuận tiện bởi vì nó cung cấp phép đo lường đa dạng tại mỗi điểm ảnh đơn, có thể phân lớp, phân loại mà không cần đến những sự xử lý không gian phức tạp để đưa đến quyết định.

1.1. Một số đặc tính vật lý đặc biệt về màu sắc:

Con người chỉ có khả năng nhận thức được ánh sáng có bức xạ điện từ với bước sóng trong khoảng 400 – 700 nanomet. Cơ quan thị giác cảm nhận được ánh sáng là do bề mặt đối tượng phát ra ánh sáng, là kết quả của sự tương tác giữa năng lượng chiếu sáng và những phân tử của bề mặt đối tượng. Một đối tượng màu xanh dương sẽ có bề mặt màu xanh dương khi chiếu ánh sáng trắng vào. Nhưng đối tượng đó sẽ có màu tím khi chiếu ánh sáng đỏ vào.

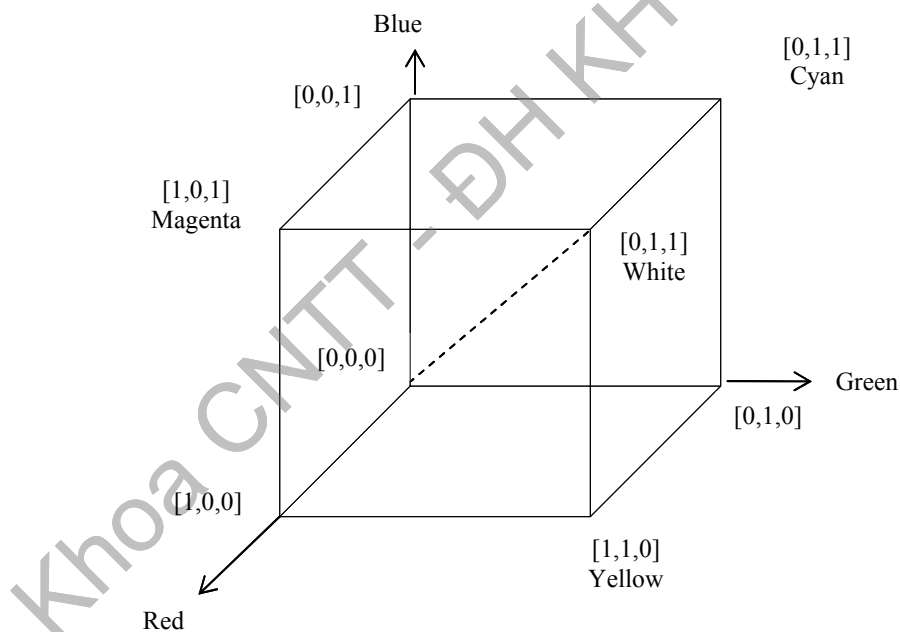
1.2. Hệ thống màu chuẩn RGB:

Mắt người có thể phân biệt hàng ngàn màu sắc khác nhau, những con số chính xác hơn vẫn còn đang được bàn cãi nhiều. Ba màu RGB (Red-Green-Blue) mã hóa hệ thống đồ họa sử dụng ba byte $(2^8)^3$ hay khoảng chừng 16 triệu màu phân biệt. Máy tính có thể phân biệt bất kỳ màu gì sau khi được mã hóa, nhưng việc mã hóa có thể không trình bày được những sự khác biệt

trong thế giới thực. Mỗi điểm ảnh RGB bao gồm một byte cho màu R, một byte cho màu G và một byte cho màu B.

Việc mã hóa một màu tùy ý trong dãy hiển thị được làm bằng cách tổ hợp ba màu chính. Ví dụ: Red(255,0,0), Green(0,255,0), Blue(0,0,255), Black(0,0,0) Hệ thống màu RGB là một hệ thống màu cộng vào bởi vì mỗi màu được tạo nên bằng cách cộng thêm các phần tử vào màu đen(0,0,0)

Khuôn dạng của không gian màu RGB là định dạng phổ biến nhất của ảnh số, lý do chính là tính tương thích với màn hình hiển thị chính là màn hình vi tính. Tuy nhiên không gian màu RGB có hạn chế lớn nhất là không phù hợp với cách con người cảm nhận về màu sắc. Do đó không phù hợp cho việc ứng dụng vào tìm kiếm ảnh.



Hình 1: Khối màu

1.3. Hệ thống màu CMY:

Hệ thống màu CMY theo mô hình in trên giấy trắng và theo khuôn mẫu trừ từ màu trắng thay vì thêm vào từ màu đen như hệ thống màu RGB.

CMY là viết tắt của Cyan-Magenta-Yellow (màu lục lam, màu đỏ tươi, màu vàng), đó là ba màu chính tương ứng với ba màu mực in. Cyan hấp thụ sự chiếu sáng của màu đỏ, Magenta hấp thụ màu xanh lục, Yellow hấp thụ màu xanh dương. Do đó, tạo ra sự phản ánh tương ứng như khi in ảnh được chiếu sáng với ánh sáng trắng. Hệ thống dưới dạng âm tính vì mã hóa theo dạng hấp thụ màu. Có một số mã hóa như sau: trắng (0,0,0) vì không có ánh sáng trắng được hấp thụ, đen (255,255,255) vì tất cả các thành phần của màu trắng đều được hấp thụ.

Hệ thống màu CMY dường như là một sự đảo ngược của hệ thống màu RGB. Đặc tính của nó là sự đơn giản, ứng dụng nhiều trong thực tế. Tuy nhiên khuyết điểm của nó cũng tương tự như không gian màu RGB, tức là cách mã hóa khác với cách mà con người cảm nhận về màu sắc. Không thích hợp cho bài toán tìm kiếm ảnh dựa vào nội dung.

1.4. Hệ thống màu L*a*b:

Mô hình L*a*b được đề cử bởi CIE cho việc lượng hóa sự khác biệt của màu sắc trong vật chiếu sáng của ánh sáng ban ngày. Tuy nhiên nó có một sự chuyển đổi được ghi vào để mà tính toán cho việc thích nghi với những nguồn sáng.

Đây là hệ thống màu có sự tách riêng ánh sáng và màu sắc ra riêng. Do đó, cũng có khả năng lớn cho việc tìm kiếm dựa vào nội dung.

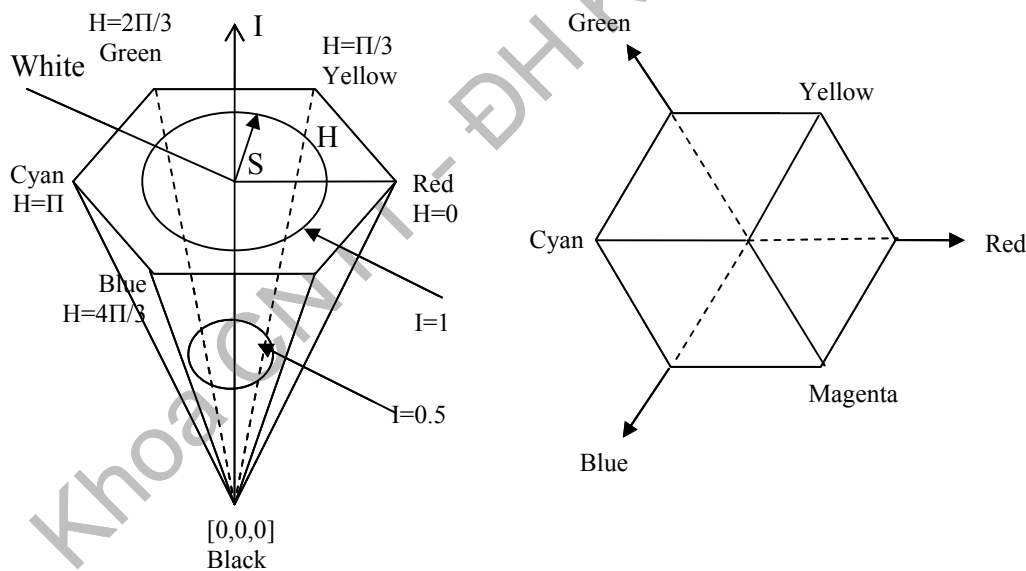
1.5. Hệ thống màu HSI: Hue-Saturation-Intensity

Hệ thống màu HSI mã hóa thông tin màu sắc bằng cách chia giá trị intensity I từ hai giá trị được mã hóa thuộc về độ hội tụ của màu- hue H và saturation S.

Thành phần không gian màu HSI gồm có ba phần: Hue được định nghĩa có giá trị $0-2\pi$, mang thông tin về màu sắc. Saturation có giá trị $0-1$, mang giá trị về độ thuần khiết của thành phần Hue. Intensity (Value) mang thông tin về độ sáng của điểm ảnh. Ta có thể hình dung không gian màu HSI như là vật hình nón. Với trục chính biểu thị cường độ sáng Intensity. Khoảng cách đến trục biểu thị độ tập chung Saturation. Góc xung quanh trục biểu thị cho sắc màu Hue.

Đôi khi, hệ thống màu HSI được coi như là hệ thống màu HSV dùng *Value* thay vì *Intensity*.

Hệ thống màu HSI thì thích hợp hơn với một số thiết kế đồ họa bởi vì nó cung cấp sự điều khiển trực tiếp đến ánh sáng và hue. Hệ thống màu HSI cũng hỗ trợ tốt hơn cho những thuật toán xử lý ảnh vì sự tiêu chuẩn hóa về ánh sáng và tập chung vào hai tham số về độ hội tụ màu, và cường độ màu.



Hình 2: Khối nón màu minh họa hệ thống màu HSI

Hệ thống màu HSI có sự phân chia rõ rệt giữa ánh sáng và màu sắc. Do đó có khả năng rất lớn được áp dụng cho việc tính đặc trưng và so sánh sự giống

nhau về màu sắc của hai ảnh. Do đó nó rất thích hợp cho việc tìm kiếm ảnh dựa vào màu.

Sự giống và khác nhau giữa hai ảnh về mặt màu sắc đối với mắt người chỉ mang ý nghĩa tương đối. Do đó khi áp dụng vào bài toán này trên máy tính thì ta cũng giả lập sự tương đối này.

Phương pháp chính của việc tìm kiếm theo màu sắc là dùng lượt đồ màu để làm đặc trưng cho từng ảnh. Do những đặc điểm riêng của mô hình màu HSI và đặc trưng của việc tìm kiếm nên tính lượt đồ màu cũng được dùng một mô hình rất đặc biệt để phù hợp cho những đặc điểm riêng này.

2. Tìm kiếm ảnh dựa vào màu sắc:

Phương pháp phổ biến để tìm kiếm ảnh trong một tập những ảnh hỗn tạp cho trước là dựa vào lượt đồ màu của chúng. Đây là cách làm khá đơn giản, tốc độ tìm kiếm tương đối nhanh nhưng khuyết điểm là kết quả tìm kiếm lại có độ chính xác không cao. Nhưng đây có thể được xem như là bước lọc đầu tiên cho những tìm kiếm sau. Muốn được kết quả chính xác cao đòi hỏi sự kết hợp đồng thời với vân (texture) và hình dáng (shape).

Cho đến nay, để giải quyết vấn đề về màu sắc, cách tiếp cận chính vẫn là dựa vào lượt đồ màu.

2.1. Lượt đồ màu:

Lượt đồ màu: Như là một bảng tóm tắt thông tin về màu sắc cho một ảnh màu bất kỳ.

Và việc tính lượt đồ màu này được tiến hành một cách rất nhanh chóng trong ảnh mà chỉ qua một lần duyệt qua toàn bộ ảnh. Do đó ứng dụng vào việc tìm kiếm ảnh sẽ có lợi rất lớn về mặt tốc độ.

Một số tính chất cần quan tâm của lượt đồ màu đối với vấn đề truy tìm ảnh:

- Việc tính lượt đồ màu của ảnh diễn ra rất nhanh chóng trong ảnh chỉ qua một lần duyệt qua toàn bộ ảnh.

- Lược đồ màu tương đối bất biến đối với phép tịnh tiến, xoay ảnh, và nhất là sự kéo nhỏ, kéo giãn, thay đổi kích thước của ảnh.
- Lược đồ màu của một ảnh màu có thể là một cách miêu tả rất có ý nghĩa cho việc truy tìm ảnh hay nhận dạng đối tượng trong ảnh

2.1.1. Lược đồ màu thông thường 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 màu 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:

$$h_{R,G,B}[r, g, b] = N * \text{Prob}\{R=r, G=g, B=b\}$$

trong đó N là số lượng điểm có trong ảnh.

Lược đồ màu ở dạng này được tính bằng cách rời rạc hoá từng màu trong ảnh, sau đó là 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 màu duy nhất. Cho một ảnh RGB, một kiểu chuyển đổi thường được sử dụng là:

$$m = r + N_r g + N_r N_g b$$

trong đó N_r, N_g là số lượng bin của màu đỏ và màu xanh lục.

Điều này mang lại một lược đồ đơn duy nhất như sau:

$$h[m] = N * \text{Prob}\{M=m\}$$

Một cách khác để tính lược đồ màu của ảnh RGB là ta phân ra thành 3 lược đồ riêng biệt $h_R[], h_G[], h_B[]$. Khi đó, mỗi lược đồ được tính bằng cách đếm kênh màu tương ứng trong mỗi điểm ảnh.

2.1.2. Lược đồ màu HSI:

Mô hình màu HSI có những ưu điểm lớn cho việc tìm kiếm hình ảnh dựa vào nội dung và nhất là trong công việc tìm kiếm dựa vào màu sắc. Nhưng những ảnh màu thông thường được lưu trữ ở dạng kỹ thuật số trong máy tính

thường theo chuẩn RGB, do đó để có được ảnh màu HSI ta phải qua công đoạn chuyển đổi. Chuyển từ hệ màu RGB sang hệ màu HSI là một thuật giải có rất nhiều trong hầu hết những sách về xử lý ảnh cổ điển.

Khoa CNTT - ĐH KHTN TP.HCM

Gọi R,G,B là giá trị nhập của RGB với giá trị trong khoảng [0,1] hay [0,255]

I : Giá trị xuất của cường độ intensity [0,1]

S : Giá trị xuất của độ bão hòa saturation [0,1]

H : Giá trị xuất của màu sắc hue [0,2 Π]

R,G,B,H,I,S: Tất cả đều là những giá trị số thực.

Procedure RGB_to_HSI(in R,G,B; out H,S,I)

```
{
    I:=Max(R,G,B);
    Min:=Min(R,G,B);
    If (I>=0) then
        S:=(I-Min)/I;
    Else S:=0;
    If (S<=0) then
        {
            H:=-1;
            Return;
        }
    Diff:=I-Min;
    If (R=I) then H:=( $\Pi/3$ )*(G-R) /Diff;
    Else if (G=I) then H:=(2* $\Pi/3$ )+ $\Pi/3$ *(B-R)/diff;
    Else if (B=I) then H:=(4* $\Pi/3$ )+ $\Pi/3$ *(R-G)/diff;
    If (H<=0) H:=H+ $\Pi/2$ ;
}
```

Thuật toán chuyển đổi từ RGB sang HSI

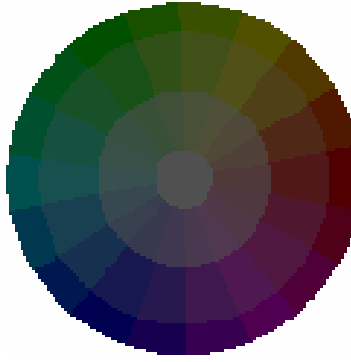
Theo các phương pháp truyền thống, việc tạo lượt đồ màu chỉ đơn giản bằng cách chia không gian màu thành những ngăn riêng biệt, sau đó đếm những pixel có màu sắc phù hợp trong những ngăn này. Cách làm này rất đơn giản, nhưng không phù hợp cho việc so sánh, tìm kiếm ảnh. Vì những thay đổi nhỏ về điều kiện chiếu sáng, sự thay đổi về cường độ có thể gây nên những thay đổi lớn trong lượt đồ màu. Do đó, hai ảnh rất giống nhau về màu sắc có thể có lượt đồ màu hoàn toàn khác nhau.

Ví dụ :không gian màu HSI, thành phần Hue trong nhiều trường hợp mang giá trị nhưng không thể hiện được màu sắc trong hiển thị. Khi giá trị Intensity nằm trong khoảng 0-0.2, Hue mang bất kỳ giá trị nào thì màu mà mắt người nhìn thấy vẫn là màu đen. Vì vậy đối với các cách tính lượt đồ màu thông thường, điểm ảnh A có Hue bằng 2π , Intensity bằng 0.1 và điểm ảnh B có Hue bằng π , Intensity bằng 0.1 sẽ nằm trong những bin khác nhau. Nhưng thật ra, điểm ảnh A và điểm ảnh B đều có giá trị hiển thị là màu đen.

Lượt đồ HSI cải tiến:

Một phương pháp được đề xuất để giải quyết cho trường hợp trên là ta dựa vào Intensity để lọc trước những giá trị mà Hue không thể biểu thị được. Sau đó dùng Saturation để lọc những giá trị có sắc màu xám. Phần còn lại của không gian màu ta sẽ chia đều mỗi thành phần thành những khoảng nhất định có sự tương đồng về màu sắc.

Để giá trị của lượt đồ màu HSI được tính một cách phù hợp nhất đối với việc tìm kiếm, chúng ta cần chia nhỏ không gian 24 bit màu, tương ứng với 2^{24} màu, xuống một con số có thể chấp nhận được. Một con số được đề nghị là 5 giá trị mức xám, 162 cho giá trị sắc màu, tổng cộng là ta chỉ cần lưu trữ 167 bin màu.



Hue và Saturation khi Intensity < 0.33



Hue và Saturation khi Intensity > 0.33

2.2. Các loại độ đo màu:

Bước tiếp theo của quá trình tìm kiếm dữ liệu ảnh dựa vào nội dung là xác định độ trùng khớp của hai lượt đồ màu vừa tính được ở bước trên. Do đó, phát sinh ra một giá trị để biểu thị cho sự trùng khớp này, có nhiều cách để tính giá trị này. Ta gọi những giá trị được tính từ những cách khác nhau này là các loại độ đo màu.

Một cách đơn giản, độ đo màu là được coi một giá trị để biểu thị cho độ so khớp sự trùng khớp của hai lượt đồ màu. Tùy theo từng trường hợp, từng loại độ đo màu giá trị này có thể âm hoặc dương lớn hoặc nhỏ tương ứng với mức độ giống nhau như thế nào của các loại lượt đồ màu.

Mỗi loại độ đo màu có những ưu và khuyết điểm riêng, trong từng trường hợp cụ thể.

Gọi $h(I)$ và $h(M)$ tương ứng là 2 lượt 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:

2.2.1. Độ đo khoảng cách min-max:

Được thực hiện dựa trên ý tưởng lấy phần giao của của hai lượt đồ cần so sánh, ta sẽ được một lượt đồ, tính tổng các giá trị có được từ lượt đồ này cho ta được độ đo min-max.

Đối với độ đo min: ta tính dựa vào giá trị min tại mỗi K bin.

$$\text{Intersection}(h(I), h(M)) = \sum_{j=1}^K \min \{h(I)[j], h(M)[j]\}$$

Đối với độ đo max: ta tính dựa vào giá trị max tại mỗi K bin.

$$\text{Intersection}(h(I), h(M)) = \sum_{j=1}^K \max \{h(I)[j], h(M)[j]\}$$

$$\text{Matching}(h(I), h(M)) = \frac{\text{Intersection}(h(I), h(M))}{\max(\sum_i h(I)[i], \sum_i h(M)[i])}$$

2.2.2. Độ đo khoảng cách euclid:

Đây là cách tính khoảng cách ơclit thông thường giữa các K bin:

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

Hoặc có thể là:

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

2.2.3. Độ đo khoảng cách toàn phương:

$$\text{Intersection}(h(I), h(M)) = \sum_{i=1}^K \sum_{j=1}^K [h(i)-h(j)] a_{ij} [h(i)-h(j)]$$

2.2.3. Độ đo có trọng số:

$$d_{\text{hist}}(I, Q) = (h(I) - h(Q))^T A (h(I) - h(Q))$$

trong đó, $h(I)$ và $h(Q)$ là những lượt đồ tương ứng của ảnh I và Q , và A là ma trận đồng dạng $K \times K$. Trong ma trận này, những màu mà rất giống nhau thì gần với giá trị một, còn những màu rất khác nhau thì sẽ có giá trị gần với không.

Khoa CNTT - ĐH KHTN TP. HCM

Chương 2:

Tìm kiếm ảnh dựa vào vân

1. Vân

1.1. Vân là gì?

1.2. Một số loại vân tiêu biểu.

2. Tìm kiếm ảnh dựa vào vân

2.1. Mật độ của đường biên và hướng của biên

2.2. Phân hoạch vùng nhị phân cục bộ

2.3. Ma trận đồng hiện và đối tượng đồng hiện

2.4. Độ đo năng lượng của vân dựa vào luật đo

2.5. Tương quan tự động và quang phổ năng lượng

1. Vân:

1.1. Vân là gì?

Vân (texture), đến nay vẫn chưa có một định nghĩa chính xác cụ thể về vân, là một đối tượng dùng để phân hoạch ảnh ra thành những vùng được quan tâm và để phân lớp những vùng đó.

Vân cung cấp thông tin về sự sắp xếp về mặt không gian của màu sắc và cường độ của một ảnh.

Vân được đặc trưng bởi sự phân bố không gian của những mức cường độ trong một khu vực láng giềng với nhau.

Vân của ảnh màu và vân đối với ảnh xám là như nhau.

Vân gồm nhiều vân gốc hay vân phần tử gộp lại, đôi khi được gọi là texel.

Xét về vấn đề phân tích vân, có hai đặc trưng chính yếu nhất:

- Cấu trúc vân được định nghĩa như sau: vân là tập hợp những *texel* được sắp xếp theo một số quy luật nhất định hay có cấu trúc không gian lặp đi lặp lại.
- Sự thống kê vân được định nghĩa như sau: vân là một độ đo về số lượng của sự sắp xếp những mức xám hay cường độ sáng trong vùng.

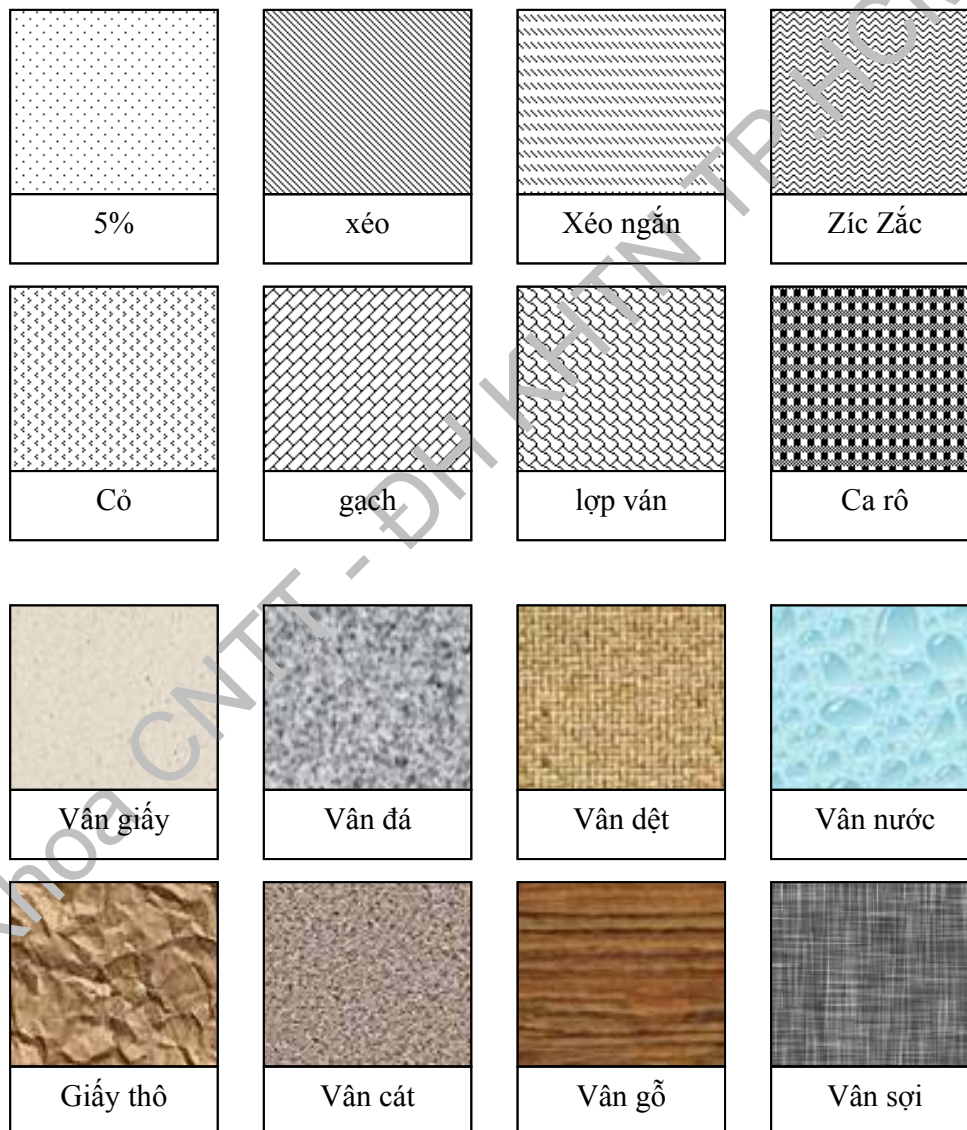
Cấu trúc vân: một vân bất kỳ có thể coi như là một tập của những *texel* thô trong một quan hệ không gian đặc biệt nào đó. Một cấu trúc không gian của một vân bất kỳ sau đó có thể bao gồm một sự mô tả của *texel* và một đặc tả về không gian. Những *texel* đương nhiên phải được phân đoạn và quan hệ không gian phải được tính toán một cách thật hiệu quả. *Texel* là những vùng ảnh có thể trích rút từ một số hàm phân ngưỡng đơn giản. Đặc điểm quan hệ không gian của chúng có thể miêu tả như sau:

Giả sử rằng chúng ta có tập những *texel*, với mỗi phần tử của tập hợp này ta có thể đặc trưng bởi một điểm ý nghĩa nhất, điểm này gọi là trọng tâm. Đặt S là tập của những điểm này. Với mỗi cặp điểm P và Q trong tập S , ta có thể xây dựng đường phân giác trực giao nối chúng lại với nhau. Đường phân giác trực giao này chia mặt phẳng thành hai nửa mặt phẳng, một trong chúng

là tập của những điểm gần với P hơn và cái còn lại là tập những điểm gần với Q hơn. Đặt $H^Q(P)$ là nửa mặt phẳng gần P hơn. Ta có thể lặp lại quá trình này với mỗi điểm Q trong S. Đa giác Voronoi của P là vùng đa giác bao gồm tất cả những điểm gần P hơn những điểm khác của S và được định nghĩa:

$$V(P) = \bigcap_{Q \in S, Q \neq P} H^Q(P)$$

1.2. Một số loại vân tiêu biểu:



2. Tìm kiếm ảnh dựa vào vân:

Trong hầu hết các trường hợp, phân đoạn những ảnh thật ra những texel khó hơn nhiều đối với trường hợp tự nhiên sinh ra những hoa văn thiên nhiên.

Thay vì vậy, việc định lượng về số hay thông tin thống kê bằng số mô tả cho một vân có thể được tính từ chính mức xám, hay mức màu của chúng. Tuy cách tiếp cận này ít trực quan nhưng nó có hiệu suất tính toán cao, hơn nữa cách tiếp cận này cũng phù hợp với đồng thời cho việc phân đoạn vân và phân loại vân.

2.1. Mật độ của đường biên và hướng của biên:

Từ khi phương pháp dò biên được phổ biến rộng và sự đơn giản trong ứng dụng vào quy trình dò đối tượng, nó trở thành là bộ dò biên như là bước tiên quyết trong việc phân tích vân. Số lượng điểm ảnh trong một vùng ảnh xác định trước về mặt kích thước cho ta thấy được một số biểu thị về mật độ điểm trong vùng ảnh đó. Hướng của những đường biên này cũng có thể hữu dụng trong việc mô tả đặc điểm hoa văn của vân.

Xét khu vực gồm có N điểm ảnh. Giả sử rằng bộ dò biên dựa trên gradient áp dụng vào cho vùng ảnh này sinh ra hai kết xuất của của mỗi điểm ảnh p : 1) độ lớn gradient $Mag(p)$ và 2) phương hướng gradient $Dir(p)$. Một trong những đối tượng vẫn rất đơn giản là số đường biên trên một khu vực được định nghĩa như sau:

$$F_{\text{edgeness}} = \frac{|\{p \mid Mag(p) \geq T\}|}{N}$$

với T là ngưỡng định nghĩa trước.

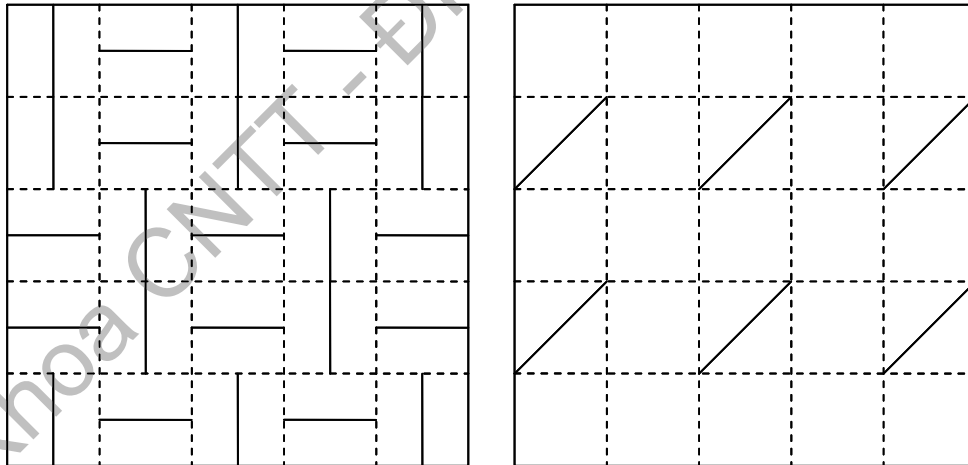
Số đường biên trên một khu vực để đánh giá mật độ, nhưng không đánh giá được phương hướng của vân.

Độ đo này có thể được mở rộng để có thể bao gồm cả mật độ và phương hướng bằng cách dùng lượt đồ cho cả độ lớn gradient và phương hướng gradient. Gọi $H_{\text{mag}}(R)$ biểu thị lượt đồ bình thường của độ lớn gradient của khu vực R , và gọi H_{dir} biểu thị lượt đồ bình thường của phương hướng gradient của khu vực R . Cả hai lượt đồ này có số lượng bin lớn hơn một số cố định, trình bày những nhóm độ lớn gradient và những nhóm của phương hướng gradient. Cả hai được tiêu chuẩn hóa theo kích thước N_R của khu vực ảnh R . Khi đó ta có :

$$F_{\text{mag dir}} = (H_{\text{mag}}(R), H_{\text{dir}}(R))$$

là một mô tả của vận định lượng của khu vực ảnh R .

Xét hai ảnh 5x5 như sau:



Ảnh bên trái có mật độ điểm cao hơn ảnh bên phải. Nó có một cạnh trong mỗi 25 điểm ảnh của nó, vì vậy số cạnh trên một đơn vị khu vực của nó là 1.0. Ảnh bên phải có 6 cạnh trên mỗi 25 điểm ảnh của nó, vì vậy số cạnh trên

một đơn vị khu vực của nó là 0.24. Đối với lượt đồ độ lớn gradient, chúng ta đề dùng hai bin trình bày cho đường biên sáng và đường biên tối. Với lượt đồ phương hướng gradient, chúng ta sẽ dùng ba bin cho đường biên ngang, đường biên dọc và đường biên chéo. Ảnh bên trái có 6 đường biên tối và 19 đường biên sáng, nên lượt đồ độ lớn gradient bình thường của nó là (0.24,0.76), nghĩa là 24 phần trăm của đường biên là đường biên tối, 76 phần trăm của đường biên là đường biên sáng. Nó cũng được xem như có 12 đường biên ngang, 13 đường biên dọc, và không có đường biên chéo, vậy lượt đồ phương hướng gradient bình thường của nó là (0.48,0.52, 0.0), nghĩa là 48 phần trăm đường biên là ngang, 52 phần trăm đường biên là dọc và 0 phần trăm đường biên chéo. Ảnh bên phải không có đường biên tối và có 6 đường biên sáng, nên lượt đồ độ lớn gradient bình thường của nó là (0.0,0.24). Nó không có đường biên ngang, không có đường biên dọc, nhưng có 6 đường biên chéo, vậy lượt đồ phương hướng gradient bình thường của nó là (0.0,0.0, 0.24).

Trong trường hợp hai ảnh này, độ đo đường biên trên một đơn vị khu vực thì thích hợp để phân biệt giữa chúng, nhưng trong trường hợp tổng quát độ đo lượt đồ thường cung cấp một cơ chế mô tả mạnh hơn nhiều. Hai lượt đồ n -bin H_1 và H_2 có thể được so sánh bởi tính khoảng cách L_1 của chúng.

$$L_1(H_1, H_2) = \sum_{i=1}^n |H_1[i] - H_2[i]|$$

2.2. Phân hoạch vùng nhị phân cục bộ:

Một cách khác rất đơn giản, nhưng là độ đo vô hướng dụng là độ phân hoạch vùng nhị phân cục bộ. Đối với mỗi điểm ảnh p trong ảnh, tám điểm lân cận được xem xét rằng cường độ intensity của chúng có lớn hơn của điểm ảnh p hay không. Những kết quả từ tám điểm ảnh lân cận được sử dụng để xây dựng nên một số nhị phân tám chữ số là $b_1b_2b_3b_4b_5b_6b_7b_8$, trong đó $b_i=0$ nếu

cường độ intensity của láng giềng thứ i nhỏ hơn hay bằng với p và $b_i=1$ trong trường hợp ngược lại. Một lượt đồ của những số này được dùng để biểu thị cho vân của ảnh. Hai ảnh hay hai khu vực ảnh được so sánh bởi việc tính toán độ dài $L1$ giữa lượt đồ của chúng được định nghĩa ở trên.

2.3. Ma trận đồng hiện và đối tượng đồng hiện:

Ma trận đồng hiện là một mảng C hai chiều trong đó cả dòng và cột tương ứng với một tập các giá trị có thể có V . Ví dụ như, đối với ảnh mức xám V có thể là tập những mức xám có thể có, đối với ảnh màu V có thể là tập những màu có thể có. Giá trị của $C(i,j)$ cho thấy rằng bao nhiêu lần giá trị i cùng xuất hiện trong với giá trị j trong một số quan hệ không gian được định nghĩa trước. Ví dụ : quan hệ không gian có thể có giá trị i xuất hiện lập tức ngay bên phải của giá trị j . Rõ ràng hơn, chúng ta sẽ xem xét trường hợp riêng trong đó tập V là tập những mức xám và quan hệ không gian được mang lại bởi vectơ d , chỉ định độ dịch chuyển giữa điểm ảnh có giá trị i và điểm ảnh có giá trị j .

Gọi d là một vectơ dịch chuyển (d_r, d_c) ở đó d_r thì thay thế cho hàng và d_c thay thế cho cột. Gọi V là tập của những mức xám. Ma trận đồng hiện mức xám cho ảnh I được định nghĩa bởi:

$$C_d[i,j] = |\{[r,c] \text{ sao cho } I[r,c]=i \text{ và } I[r+d_r,c+d_c]=j\}|$$

		j			
i	1	1	1	0	0
	1	1	1	0	0
	0	0	0	2	2
	0	0	0	2	2

Ảnh 1

		0	1	2
0	4	0	2	
1	2	2	0	
2	0	0	2	

C[0,1]

i	j
---	---

		0	1	2
0	4	0	2	
1	2	2	0	
2	0	0	2	

C[1,0]

i
j

		0	1	2
0	2	0	2	
1	2	1	1	
2	0	0	1	

C[1,1]

i
j

Minh họa nội dung này với một ảnh 4x4 I và ba ma trận đồng hiện cho I là C[0,1], C[1,0] và C[1,1].

Chú thích C[0,1] : vị trí [1,0] có giá trị 2, chỉ định rằng j=0 xuất hiện hai lần trực tiếp bên phải của i=1 trong ảnh. Tuy nhiên vị trí [0,1] có giá trị 0, chỉ định rằng j=1 không bao giờ xuất hiện trực tiếp bên phải của i=0 trong ảnh.

Giá trị lớn nhất trong ma trận đồng hiện là 4 tại vị trí [0,0], chỉ định rằng giá trị 0 xuất hiện 4 lần trực tiếp bên phải của giá trị 0 khác trong ảnh.

Có hai biến tấu quan trọng của ma trận đồng hiện từ ma trận đồng hiện mức xám chuẩn.

Đầu tiên tiêu chuẩn hoá ma trận đồng hiện mức xám N_d được định nghĩa bởi:

$$Nd[i,j] = \frac{Cd[i,j]}{\sum_i \sum_j Cd[i,j]}$$

Tiêu chuẩn hoá những giá trị của ma trận đồng hiện nằm giữa giá trị 0 và 1.

Và do đó có thể xem xét chúng dưới dạng xác suất trong ma trận lớn.

Thứ hai là ma trận đồng hiện mức xám đối xứng S_d định nghĩa bởi :

$$S_d[i,j] = Cd[i,j] + C-d[i,j]$$

thực chất là một cặp của nhóm của những đối xứng kề cận nhau.

Ma trận đồng hiện nắm bắt thuộc tính của vân, nhưng chúng không hữu dụng trực tiếp cho việc phân tích sâu hơn về vân, như là so sánh hai vân. Thay vì vậy, những đối tượng số học của vân được tính toán từ ma trận đồng hiện có thể được dùng để trình bày vân chặt chẽ hơn, súc tích hơn. Dưới đây là những đối tượng chuẩn bắt nguồn từ một ma trận đồng hiện chuẩn.

$$\text{Energy} = \sum_i \sum_j N_d^2[i,j]$$

$$\text{Entropy} = - \sum_i \sum_j N_d[i,j] \log_2 N_d[i,j]$$

$$\text{Contrast} = \sum_i \sum_j (i - j)^2 N_d[i, j]$$

$$\text{Homogeneity} = \sum_i \sum_j \frac{N_d[i, j]}{1 + |i - j|}$$

$$\text{Correlation} = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j) N_d[i, j]}{\sigma_i \sigma_j}$$

Trong đó μ_i, μ_j là giá trị trung bình và σ_i, σ_j là độ lệch chuẩn của hàng và cột i, j .

$$N_d[i] = \sum_j N_d[i, j]$$

$$N_d[j] = \sum_i N_d[i, j]$$

Một vấn đề với độ vãn lệch từ ma trận đồng hiện là bằng cách nào để chọn vectơ d . Một giải pháp được đề nghị bởi Zucker và Terzopoulos là dùng kiểm tra bằng thống kê χ^2 để chọn những giá trị của d mà vãn có cấu trúc nhất; đó là, giá trị cực đại của giá trị:

$$\chi^2(d) = \left(\sum_i \sum_j \frac{N_d^2[i, j]}{N_d[i] N_d[j]} - 1 \right)$$

2.4. Độ đo năng lượng của vãn dựa vào luật đo:

Một cách tiếp cận khác đối với việc sinh ra những đối tượng vãn là dùng những mặt nạ cục bộ để dò tìm những kiểu vãn khác nhau.

Luật đo đưa ra một cách tiếp cận dựa vào năng lượng vãn, đo lường số lượng biến số lớn trong phạm vi một cửa sổ cố định kích thước trước. Một tập của

9 mặt nạ 5x5 được dùng để tính năng lượng vân, được trình bày bởi một vectơ gồm 9 giá trị cho mỗi điểm ảnh của ảnh được mang ra phân tích.

Những mặt nạ được tính từ những vectơ dưới đây :

$$L5 \text{ (Level)} = [1 \quad 4 \quad 6 \quad 4 \quad 1]$$

$$E5 \text{ (Edge)} = [-1 \quad -2 \quad 0 \quad 2 \quad 1]$$

$$S5 \text{ (Spot)} = [-1 \quad 0 \quad 2 \quad 0 \quad -1]$$

$$R5 \text{ (Ripple)} = [1 \quad -4 \quad 6 \quad -4 \quad 1]$$

Tên của những vectơ nói rõ mục đích của chúng. Vectơ L5 mang lại trọng lượng trung tâm trung bình địa phương. Vectơ E5 dò đường biên. Vectơ S5 dò điểm. Vectơ R5 dò sự gợn sóng. Mặt nạ 2 chiều được tính bằng tích những cặp vectơ tương ứng. Ví dụ như mặt nạ E5L5 được tính bằng tích của vectơ E5 và L5 như sau:

$$\begin{bmatrix} -1 \\ -2 \\ 0 \\ 1 \\ 2 \end{bmatrix} \times [1 \ 4 \ 6 \ 4 \ 1] = \begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

Bước đầu tiên trong hàm Luật đo là loại bỏ tất cả những ảnh hưởng của sự chiếu sáng bằng cách di chuyển một cửa sổ nhỏ đi khắp trong ảnh, và trích rút ra trung bình địa phương cho mỗi điểm ảnh, để cho ra kết quả là một ảnh tiền xử lý, trong đó cường độ intensity của mỗi điểm láng giềng xấp xỉ không. Kích thước của cửa sổ phụ thuộc vào lớp ảnh loại nào; Người ta quen dùng cửa sổ kích thước 15x15 cho những khung ảnh tự nhiên. Sau quá trình tiền xử lý, mỗi 16 mặt nạ 5x5 được áp vào ảnh tiền xử lý, sinh ra 16 ảnh lọc. Gọi $F_k[i,j]$ là kết quả lọc với mặt nạ thứ k tại điểm ảnh $[i,j]$. Khi đó bản đồ năng lượng vân E_k cho bộ lọc k được định nghĩa bởi:

$$E_k [r,c] = \sum_{j=c-7}^{c+7} \sum_{i=r-7}^{r+7} |F_k [i, j]|$$

Mỗi bản đồ năng lượng vẫn là một ảnh đầy đủ, trình bày ứng dụng của mặt nạ thứ k tới ảnh nhập vào.

Một khi 16 bản đồ năng lượng được sinh ra, những cặp đối xứng thích hợp được nối lại để sinh ra 9 bản đồ cuối cùng, là sự thay thế mỗi cặp với giá trị trung bình của chúng. Ví dụ như: E5L5 đo được nội dung của đường biên ngang, và L5E5 đo được nội dung đường biên ngang. Trung bình cộng của hai bản đồ đo được toàn bộ nội dung của đường biên của ảnh.

Có tổng cộng 9 bản đồ năng lượng tổng hợp là :

L5E5/E5L5	L5S5/S5L5
L5R5/R5L5	E5E5
E5S5/S5E5	E5R5/R5E5
S5S5	S5R5/R5S5
R5R5	

1.5: Tương quan tự động và quang phổ năng lượng:

Hàm tương quan tự động của một ảnh có thể được dùng để dò tìm những phân tử hoa văn của vân và cũng đồng thời mô tả tính mịn/tính thô của vân. Hàm tương quan tự động $\rho(dr, dc)$ của một ảnh với kích thước $(N+1) \times (N+1)$ với một độ dời $d = (dr, dc)$ là:

$$\begin{aligned} \rho(dr, dc) &= \frac{\sum_{r=0}^N \sum_{c=0}^N I[r, c] I[r + dr, c + dc]}{\sum_{r=0}^N \sum_{c=0}^N I^2[r, c]} \\ &= \frac{I[r, c] o I_d[r, c]}{I[r, c] o I[r, c]} \end{aligned}$$

Nếu vân ở dạng thô, khi đó hàm tương quan tự động sụt giảm chậm; ngược lại nó sẽ sụt giảm rất nhanh. Đối với những vân theo quy tắc, hàm tự động tương quan sẽ yếu. Khi $I[r+dr, c+dc]$ không được xác định rõ là đường bao của ảnh, một phương pháp để tính toán những giá trị ảnh ảo phải được định nghĩa.

Hàm tự động tương tác liên quan với quang phổ năng lượng của phép biến đổi Fourier. Nếu $I[r,c]$ là hàm ảnh và $F(u,v)$ là biến đổi Fourier của nó, $|F(u,v)|^2$ được coi như là quang phổ năng lượng.

2.6. Phân đoạn vân (Texture segmentation):

Bất kỳ độ đo một vân nào, cung cấp một giá trị hay vector giá trị tại mỗi điểm ảnh, mô tả vân trong những điểm láng giềng của điểm ảnh đó, có thể được dùng để phân đoạn ảnh thành những vùng có vân giống nhau. Như những thuật toán phân đoạn ảnh khác, những thuật toán phân đoạn vân có hai loại chính: cách tiếp cận dựa vào vùng và cách tiếp cận dựa vào đường biên. Cách tiếp cận dựa vào vùng có ý tưởng chính là nhóm hay phân chia những điểm ảnh với những đặc tính vân giống nhau. Cách tiếp cận dựa vào đường biên có ý tưởng là đi tìm những đường biên vân giữa những điểm ảnh đến từ những sự phân bố vân khác nhau.

Chương 3:

Tìm kiếm ảnh dựa vào hình dạng

1. Hình dạng

1.1. Khái niệm về hình dạng

1.2. Đặc điểm hình dạng đối với việc tìm kiếm ảnh

2. Tìm kiếm ảnh dựa vào hình dạng

2.1. Lướt đồ hình dạng

2.2. Độ so khớp đường biên của hình dạng

2.3. So khớp với ảnh phát họa

Khoa CNTT - ĐH KHKTN TP HCM

1. Hình dạng:

1.1. Khái niệm về hình dạng:

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

1.2. Đặc điểm hình dạng với việc tìm kiếm ảnh

Hình dạng là một cấp cao hơn màu sắc và vân. Nó đòi hỏi sự phân biệt giữa các vùng để tiến hành xử lý về độ đo của hình dạng. Trong nhiều trường hợp, sự phân biệt này cần thiết phải làm bằng tay. Nhưng sự tự động hóa trong một số trường hợp có thể khả thi. Trong đó, vấn đề chính yếu nhất là quá trình phân đoạn ảnh. Nếu quá trình phân đoạn ảnh được làm một cách chính xác, rõ ràng và nhất là hiệu quả thì sự tìm kiếm thông tin dựa vào hình dạng có thể có hiệu lực rất lớn.

Nhận dạng ảnh hai chiều là một khía cạnh quan trọng của quá trình phân tích ảnh. Tính chất hình dạng toàn cục ám chỉ đến hình dạng ảnh ở mức toàn cục. Hai hình dạng có thể được so sánh với nhau theo tính chất toàn cục bởi những phương pháp nhận dạng theo hoa văn, mẫu vẽ. Sự so khớp hình dạng ảnh cũng có thể dùng những kỹ thuật về cấu trúc, trong đó một ảnh được mô tả bởi những thành phần chính của nó và quan hệ không gian của chúng. Vì sự hiển thị ảnh là một quá trình liên quan đến đồ thị, do đó những phương pháp so khớp về đồ thị có thể được dùng cho việc so sánh hay so khớp. Sự so khớp về đồ thị rất chính xác, vì nó dựa trên những quan hệ không gian hầu như bất biến trong toàn thể các phép biến đổi hai chiều. Tuy nhiên, quá trình so khớp về đồ thị diễn ra rất chậm, thời gian tính toán tăng theo cấp số mũ tương ứng với số lượng các phần tử. Trong việc tìm kiếm dữ liệu ảnh dựa

vào nội dung, ta cần những phương pháp có thể quyết định sự giống và khác nhau một cách nhanh chóng. Thông thường, chúng ta luôn đòi hỏi sự bất biến cả đối với kích thước của ảnh cũng như hướng của ảnh trong không gian. Vì vậy, một đối tượng có thể được xác định trong một số hướng. Tuy nhiên, tính chất này không thường được yêu cầu trong tìm kiếm ảnh. Trong rất nhiều cảnh vật, hướng của đối tượng thường là không đổi. Ví dụ như: cây cối, nhà cửa, ...

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, nó chỉ có thể loại trừ những đối tượng hình dạng không thể so khớp, nhưng điều đó sẽ mang lại khẳng định sai, vì chỉ như là việc làm của lược đồ màu. Kỹ thuật dùng đường biên thì đặc hiệu hơn phương pháp trước, chúng làm việc với sự hiện hữu của đường biên của hình dạng đối tượng và đồng thời cũng tìm kiếm những hình dạng đối tượng gần giống với đường biên nhất. Phương pháp vẽ phác họa có thể 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. Ta sẽ đi vào chi tiết trong chương sau mô tả về hình dạng đối tượng.

2. Tìm kiếm ảnh dựa vào hình dạng:

2.1. Lược đồ hình dạng:

Lược đồ hình dạng được cho rằng là dễ dàng trong tính toán và nhanh trong thi hành. Chúng sử dụng cả sự so sánh về màu sắc và vân. Vấn đề chính là định nghĩa biến cho lược đồ hình dạng được định nghĩa. Xem như hình dạng trong ảnh là một vùng những giá trị một trong ảnh nhị phân, trong khi toàn thể những giá trị khác đều là giá trị không. Một kiểu của so khớp hình dạng ảnh là so khớp hình chiếu

thông qua hình chiếu đứng và hình chiếu nằm của hình dạng. Giả sử rằng hình dạng có n hàng và có m cột. Mỗi hàng và mỗi cột là một bin trong lượt đồ hình dạng. Tổng số được lưu trữ trong một bin là tổng số những giá trị 1 được lưu trữ trong dòng hoặc cột tương ứng đó. Điều này đưa đến một lượt đồ gồm có $m+n$ bin, nhưng điều này cũng chỉ có ý nghĩa khi tất cả những ảnh chúng ta xem xét phải có cùng một kích thước. Để làm cho việc so khớp hình chiếu bất biến đối với kích thước, số lượng bin của dòng và số lượng bin của cột phải ổn định. Bằng cách định nghĩa những bin từ góc trái trên đến góc phải dưới của hình dạng, sự chuyển đổi bất biến đã đạt được. Việc so khớp hình chiếu không bất biến đối với phép xoay ảnh, nhưng nó có thể làm việc tốt với sự xoay nhỏ và sự thiếu chính xác thuộc về hình học ở mức độ nhỏ. Một cách khác để làm nó bất biến đối với phép quay là tính toán theo trục tọa độ êlíp vừa nhất và xoay chúng cho đến khi trục chính là trục nằm ngang. Vì chúng ta không biết nơi đâu là phía trên cùng của hình dạng, xoay hai khả năng có thể xảy ra để thử. Hơn nữa, nếu trục chính và trục phụ có cùng chiều dài, thì 4 khả năng xoay phải được xem xét. So khớp hình chiếu được sử dụng thành công trong tìm kiếm ảnh logo.

Những khả năng khác để xây dựng lượt đồ thông qua góc tiếp tuyến tại mỗi điểm ảnh trên đường bao của hình dạng. Độ đo này thì hoàn toàn tự động về mặt kích thước và bất biến đối với sự dịch chuyển, nhưng nó cũng không bất biến đối với xoay đối tượng, bởi vì góc tiếp tuyến được tính từ hình dạng đối với một hướng xác định. Có một số cách khác nhau để giải quyết vấn đề này. Cách thứ nhất là xoay hình dạng về trục chính như đã mô tả ở trên. Một cách khác đơn giản hơn là xoay lượt đồ hình dạng. Nếu lượt đồ có K bin, thì sẽ có K khả năng xoay. Những vị trí xoay không đúng có thể làm ảnh hưởng tốc độ của việc tính toán, đặc biệt là trong trường hợp lượt đồ và ảnh có kích thước lớn. Hoặc là lượt đồ có thể được tiêu chuẩn hoá bởi cách chọn bin với số đếm lớn nhất là bin đầu tiên. Một vài bin lớn nhất nên được thử vì có thể có sự tồn tại của nhiều.

2.2. Độ so khớp đường biên của hình dạng:

Thuật toán so khớp đường biên đòi hỏi sự trích rút và trình bày đường biên của cả ảnh cần truy vấn và ảnh mang ra so khớp. Đường biên có thể được trình bày bởi

một dãy những điểm ảnh hay có thể được xấp xỉ bởi một đa giác. Đối với một dãy những điểm ảnh, một loại so khớp cổ điển là dùng mô tả Fourier để so sánh hai hình dạng với nhau. Trong toán học hàm liên tục, mô tả Fourier là những hệ số của dãy triển khai Fourier của hàm mà định nghĩa đường biên của hình dạng ảnh. Trong trường hợp đặc biệt, hình dạng được trình bày bởi dãy của m điểm $\langle V_0, V_1, \dots, V_{m-1} \rangle$. Từ những dãy điểm này, một dãy của vector đơn vị :

$$v_k = \frac{V_{k+1} - V_k}{|V_{k+1} - V_k|}$$

và một dãy của những sai phân:

$$l_k = \sum_{i=1}^k |V_i - V_{i-1}|, \quad k > 0$$

$$l_0 = 0$$

có thể được tính.

Mô tả Fourier $\{a_{-M}, \dots, a_0, \dots, a_M\}$ sau đó được xấp xỉ bởi:

$$a_n = \frac{1}{L \left(\frac{n2\pi}{L}\right)^2} \sum_{k=1}^m (v_{k-1} - v_k) e^{-jn(2\pi/L)l_k}$$

Những mô tả này có thể được dùng để định nghĩa độ đo khoảng cách hình dạng.

Giả sử Q là ảnh truy vấn và I là ảnh mang hình dạng được so sánh với Q . Gọi $\{a_n^Q\}$

là dãy của những mô tả Fourier cho ảnh truy vấn, và $\{a_n^I\}$ là mô tả Fourier cho

ảnh. Khi đó độ đo khoảng cách Fourier như sau:

$$d_{Fourier}(I, Q) = \left[\sum_{n=-M}^M |a_n^I - a_n^Q|^2 \right]^{1/2}$$

Như đã mô tả, khoảng cách này chỉ bất biến đối với phép tịnh tiến. Nếu mà những bất biến khác đòi hỏi, có thể dùng sự kết hợp với nhiều hàm số học thể giải quyết vấn đề tỷ lệ, xoay, và điểm bắt đầu để cực tiểu hoá $d_{Fourier}(I, Q)$

Nếu mà đường biên được trình bày dưới dạng một đa giác, chiều dài của các cạnh và góc giữa chúng có thể được tính và dùng để trình bày hình dạng. Một hình dạng có thể được trình bày bởi một dãy những điểm nối liền nhau (X_i, Y_i, α_i) , với cặp đường thẳng giao nhau tại điểm (X_i, Y_i) với góc lớn α_i . Cho một dãy $Q = Q_1, Q_2, \dots, Q_n$ của những điểm nối liền nhau trình bày đường biên của đối tượng truy vấn Q và một dãy tương tự $I = I_1, I_2, \dots, I_m$ trình bày đường biên của đối tượng I , mục tiêu là tìm một ánh xạ từ Q tới I mà ánh xạ từ dãy phân đoạn của ảnh truy vấn tới dãy có chiều dài tương tự phân đoạn của ảnh và sao cho cặp của dãy phân đoạn truy vấn gần kề mà gặp tại một góc đặc biệt α nên ánh xạ tới một cặp những dãy phân đoạn gần kề mà gặp nhau tại một góc α' tương tự.

Một kỹ thuật so khớp đường biên khác là so khớp mềm dẻo (elastic matching) trong đó hình dạng truy vấn được làm biến dạng để trở nên càng giống với hình dạng ảnh mẫu càng tốt. Sự cách biệt giữa hình dạng ảnh truy vấn và hình dạng của ảnh mẫu dựa vào hai thành phần chính: (1) năng lượng đòi hỏi trong quá trình biến dạng từ hình dạng ảnh truy vấn tới mức độ phù hợp nhất với hình dạng trong ảnh mẫu. Và (2) sự đo lường về độ giống nhau giữa hình dạng ảnh truy vấn sau khi bị biến dạng khớp với hình dạng trong ảnh mẫu.

2.3. So khớp ảnh phát họa :

Hệ thống so khớp phát họa cho phép người dùng nhập vào một bản phát họa của những đường biên chính trong một ảnh và sau đó hệ thống sẽ tìm kiếm những ảnh màu hay ảnh xám mà có đường biên khớp nhất. Những ảnh màu được tiền xử lý như sau để đạt được một dạng trung gian gọi là ảnh được trích rút (abstract image).

- Áp dụng phép biến đổi affine để giảm kích thước ảnh về kích thước chỉ định trước. Dùng một mặt nạ trung vị để lọc nhiễu. Kết quả của bước này cho ra một ảnh được tiêu chuẩn hóa

- Dò biên sử dụng thuật toán dò tìm đường biên dựa trên gradient. Dò biên được tiến hành qua hai bước: những biên toàn cục sẽ được tìm thấy trước tiên với một ngưỡng toàn cục dựa trên giá trị trung bình và biến đổi của gradient; sau đó, những đường biên cục bộ sẽ được chọn từ toàn cục theo những ngưỡng cục bộ. Kết quả của bước này cho ra ảnh gọi là ảnh đã được lọc biên (refined edge image).
- Tiến hành làm mảnh và rút ngắn trên ảnh đã được lọc biên. Kết quả cuối cùng được gọi là ảnh được trích rút.

Khi người sử dụng nhập vào một bức ảnh phát họa ở dạng thô như là một ảnh truy vấn, nó cũng lần lượt được đưa qua các giai đoạn chuẩn hóa kích thước, nhị phân hóa, làm mảnh hoá, và rút gọn. Kết quả của quá trình xử lý này cho ra một ảnh gọi là bản phát họa đều nét. Giờ đây, bản phát họa đều nét phải được so khớp với những ảnh được trích rút ở trên. Thuật giải so khớp là thuật giải dựa vào mối tương quan. Hai ảnh sẽ được chia ra làm những hệ thống lưới ô vuông. Đối với mỗi hệ thống lưới ô vuông của ảnh truy vấn, tương quan cục bộ với hệ thống lưới ô vuông tương ứng của ảnh cơ sở dữ liệu sẽ được tính. Để thiết thực hơn, tương quan cục bộ được tính cho một vài những dịch chuyển khác biệt trong vị trí trong hệ thống ô lưới trên những ảnh cơ sở dữ liệu và giá trị tương quan cực đại qua tất cả những dịch chuyển là kết quả của cho hệ thống ô lưới đó. Độ đo sự giống nhau cuối cùng là tổng của mỗi tương quan cục bộ. Độ đo khoảng cách là nghịch đảo của độ đo sự giống nhau này.

Từ những chú giải trên, nó có thể được biểu diễn lại dưới dạng sau:

$$d_{sketch}(I, Q) = \frac{1}{\sum_g \max_n [d_{correlation}(shift_n(A^I(g)), L^Q(g))]}$$

trong đó $A^I(g)$ quy cho hệ thống ô lưới g của ảnh được trích rút được tính từ ảnh cơ sở dữ liệu I , $shift(A^I(g))$ quy cho phiên bản được dịch chuyển của hệ thống lưới g

của cùng ảnh được trích rút, và $L^Q(g)$ quy cho hệ thống ô lưới g của bản họa đều nét có kết quả từ ảnh truy vấn Q .

Khoa CNTT - ĐH KHTN TP.HCM

PHẦN 3
CÀI ĐẶT VÀ
KẾT QUẢ THỬ NGHIỆM

Khoa CNTT - ĐH KHTN TP.HCM

Chương 1:

Cài đặt

- 1. Chương trình**
- 2. Phần Màu sắc**
- 3. Phần Văn**
- 4. Phần Hình dạng**

Khoa CNTT - ĐHQHTN TP.HCM

1. Chương trình:

- Chương trình gồm 3 modul riêng cho mỗi phần: màu sắc, vân và hình dạng.
- Mỗi phần của chương trình được lập trình trên môi trường Visual C++ 7.0, với sự hỗ trợ của thư viện lập trình MFC, một thư viện liên kết tĩnh .dll của intel hỗ trợ cho việc truy xuất ảnh jpeg.
- Mỗi phần là một bản demo cho cơ sở lý thuyết đã được nêu ở phần 2.
- Phần màu sắc, chương trình demo tập chung vào phần màu sắc là nhiều nhất, vì đây là phần khả thi và có thể ứng dụng được liền vào thực tế với tốc độ và kết quả chấp nhận được.
- Phần vân: chương trình làm phần tìm và so sánh vân dựa vào kỹ thuật ma trận đồng hiện và thông kê khác biệt ở lượt đồ xám.
- Phần hình dạng: chương trình làm phần so khớp ảnh phát họa kết hợp với lượt đồ hình dạng.

2. Phần màu sắc:

Quy trình thực hiện như sau:

Bước 1: Nhận vào ảnh mẫu, tính lượt đồ màu HSI cho ảnh mẫu

Bước 2: Đọc ảnh từ cơ sở dữ liệu:

Nếu còn ảnh: Tính lượt đồ màu của ảnh, qua Bước 3

Nếu hết ảnh: Kết thúc

Bước 3: So sánh 2 lượt đồ màu vừa tính được dựa vào độ đo nào đó.

Bước 4: So sánh kết quả độ đo với ngưỡng, thông báo kết quả.

Quay lại bước 2.

Tính lượt đồ màu:

Mô hình màu được ứng dụng vào việc cài đặt là mô hình màu HSI.

Cấu trúc của lượt đồ màu như sau:

```
struct ColorHistogram
{
    unsigned int grey[5];
    unsigned int val[18][3][3];
}
```

(1)

Quá trình tính lượt đồ màu HSI cho việc tìm kiếm:

Bước 1: Nhận vào ảnh RGB, đổi sang ảnh HSI

Bước 2: Đọc giá trị điểm tiếp theo

Nếu không có điểm thì kết thúc.

Nếu tồn tại điểm, qua bước 3.

Bước 3: Đọc thành phần intensity.

Nếu $intensity < 0.3$ thì cho đây là điểm màu đen. Quay lại bước 2

Nếu $intensity \geq 0.3$, qua bước 4

Bước 4: Đọc thành phần Saturation

Nếu $Saturation < 0.08$ thì cho đây là điểm ảnh xám, gán giá trị tương ứng cho lượt đồ màu. Quay lại bước 2.

Nếu $Saturation \geq 0.08$ thì cho đây là điểm ảnh màu, gán giá trị tương ứng cho lượt đồ màu dựa vào thành phần Hue, Saturation, và Intensity.

Quay lại bước 2.

2.1. Tìm kiếm theo tỉ lệ phần trăm về màu sắc:

- Người dùng xác định tỉ lệ phần trăm một số màu cần tìm kiếm, nhưng không quan tâm đến vị trí xuất hiện của những màu này. Chúng có thể xuất hiện khắp nơi trong ảnh hoặc tập chung tại một khu vực nào đó.

Ví dụ: Tìm kiếm những bức ảnh có chứa khoảng 20-30% màu đỏ, 30-50% màu trắng, 10-40% màu đen.

- Cách tiến hành: Không cần phải tính lượt đồ màu mà chỉ cần đếm số lượng những màu được chọn để tìm kiếm.

+ Màu được tìm kiếm được xếp vào một trong 167 bin của (1)

+ Trong quá trình tìm kiếm, nếu điểm ảnh nào nằm trong bin có màu được chọn thì số lượng trong bin đó tăng thêm 1.

+ Nếu tỉ lệ tổng số điểm ảnh trong từng bin được chọn trên tổng số lượng điểm ảnh của ảnh nằm trong khoảng được đặt ra thì ảnh đó thỏa mãn yêu cầu tìm kiếm.

2.2. Tìm kiếm theo lượt đồ màu toàn cục dựa trên ảnh mẫu:

- Người dùng chọn một ảnh, những ảnh tìm được là những ảnh có sắc màu gần giống với ảnh cần tìm nhất theo một ngưỡng tự chọn trước.

- Cách tiến hành: Ý tưởng cách làm là so sánh lượt đồ màu toàn cục của ảnh với những ảnh khác trong cơ sở dữ liệu. Nếu ảnh trong cơ sở dữ liệu có lượt đồ màu phù hợp với độ đo theo một ngưỡng nhất định sẽ được chọn.

+ Tính lượt đồ màu của ảnh gốc như (1).

+ Trong quá trình tìm kiếm, tính lượt đồ màu của từng ảnh. Sau đó so sánh những lượt đồ này với nhau theo các loại độ đo được chọn.

+ Những ảnh thỏa mãn nhu cầu tìm kiếm là những ảnh có sự giao nhau của 2 lượt đồ đạt giá trị thỏa mãn ngưỡng.

Độ đo được dùng : Min-Max, Euclid, độ đo khoảng cách toàn phương.

2.3. Tìm kiếm theo màu chỉ định dạng lưới:

- Người dùng xác định vị trí những khoảng màu sẽ xuất hiện trong tấm ảnh. Những vùng màu được chọn dưới dạng ô lưới. Người dùng có thể xác định độ chia nhỏ của từng ô lưới.

- Cách tiến hành : Dựa vào những ô lưới của người dùng đã ấn định, tiến hành chia ảnh trong cơ sở dữ liệu cần theo những ô lưới. Dùng mỗi màu đại diện cho một ô lưới, tiến hành so sánh ô lưới mà người dùng ấn định với ô lưới trong những ảnh vừa xác lập cho ra được theo một ngưỡng định trước.

+ Xác định mức độ chia nhỏ của mức lưới. Đồng thời xác định những màu được chọn trong các mắt lưới.

+ Tương tự như trên, màu trong những mắt lưới này sẽ được xếp vào 1 trong 167 bin như trong (1).

+ Trong quá trình tìm kiếm, ảnh tìm kiếm sẽ được chia theo đúng tỉ lệ của ô lưới xác lập ở trên.

+ Ứng với những ô lưới có màu được chọn, tính lượt đồ màu của ô lưới tương ứng trên ảnh. Tính màu đại diện trong ô lưới của ảnh. Sau đó tiến hành so sánh màu này với màu được chọn trong ô lưới ban đầu.

+ Tổng những so sánh ô lưới trên so với một ngưỡng sẽ cho được kết quả ảnh có phù hợp hay không.

2.4. Tìm kiếm toàn cục theo dạng so sánh lưới dựa trên ảnh mẫu:

- Người dùng chỉ định ảnh mẫu với yêu cầu là: Tìm những ảnh có sắc màu gần giống với ảnh này. Cách tìm kiếm này khác với cách tìm kiếm dựa trên lượt đồ màu là nó quan tâm đến vị trí của những khoảng màu trong ảnh.

Cách tìm kiếm này không những tìm ra những ảnh có sắc màu gần giống với ảnh mẫu mà còn giống nhau ở vị trí tương ứng những sắc màu đó.

- Cách tiến hành: Ý tưởng phương pháp này giống như phương pháp so sánh lưới cục bộ có tương tác với người dùng.

+ Chia lưới cho ảnh mẫu. Tính những màu đại diện cho từng ô lưới. Tạo thành một mảng những màu đại diện.

+ Trong quá trình tìm kiếm ảnh, mỗi ảnh tìm kiếm ta cũng tiến hành chia thành những ô lưới và tính màu đại diện cho mỗi ô lưới. Ta cũng được một mảng những màu đại diện.

+ Tiến hành so sánh mảng màu đại diện với ngưỡng cho trước.

2.5. Tìm kiếm kết hợp giữa màu chỉ định dạng lưới và toàn cục theo một ảnh mẫu:

- Đây là sự kết hợp giữa phương pháp tìm toàn cục theo dạng lưới và tìm cục bộ theo dạng lưới. Tương tự như tìm toàn cục theo dạng lưới nhưng còn cho phép người dùng thay đổi khoảng màu của từng ô lưới. Những ô lưới không bị thay đổi khoảng màu sẽ dùng khoảng màu của ảnh mẫu đã được chọn.

- Cách tiến hành:

+ Tương tự như trên, tiến hành chia lưới trên ảnh mẫu.

+ Nhận những giá trị của màu đã bị thay đổi trên lưới. Được mảng những màu đại diện.

+ Trong quá trình tìm kiếm, mỗi ảnh tìm kiếm được chia lưới theo tỉ lệ tương ứng như trên. Sau đó, tính mảng màu đại diện.

+ Tiến hành so sánh mảng màu đại diện này với nhau để tìm ra được ảnh phù hợp với yêu cầu tìm kiếm

3. Phần vân:

Quy trình thực hiện như sau:

Bước 1: Nhận vào ảnh mẫu, tính vector vân đặc trưng cho ảnh mẫu

Bước 2: Đọc ảnh từ cơ sở dữ liệu:

Nếu còn ảnh: Tính vector của ảnh, qua Bước 3

Nếu hết ảnh: Kết thúc

Bước 3: So sánh 2 vector vân vừa tính được dựa vào độ đo nào đó.

Bước 4: So sánh kết quả độ đo với ngưỡng, thông báo kết quả.

Quay lại bước 2.

3.1. Tìm kiếm ảnh dựa vào ma trận đồng hiện:

Kỹ thuật dùng ma trận đồng hiện đối với việc tìm kiếm ảnh, qua đó, vân của mỗi ảnh được đặt trưng bởi một vector vân, vector này có 5 thành phần, giá trị của mỗi thành phần tương ứng với những thành phần của các công thức dưới đây.

Công thức được áp dụng:

$$\text{Energy} = \sum_i \sum_j N_d^2[i, j]$$

$$\text{Entropy} = - \sum_i \sum_j N_d[i, j] \log_2 N_d[i, j]$$

$$\text{Contrast} = \sum_i \sum_j (i - j)^2 N_d[i, j]$$

$$\text{Homogeneity} = \sum_i \sum_j \frac{N_d[i, j]}{1 + |i - j|}$$

$$\text{Correlation} = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j) N_d[i, j]}{\sigma_i \sigma_j}$$

Các bước tính ma trận đồng hiện như sau:

Bước 1: Ảnh màu nhận vào được chuyển sang ảnh xám (intensity)

Bước 2: Tính ma trận đồng hiện trên ảnh mức xám này.

Sau bước này có thể có thêm bước 2'.

Bước 2': Chuẩn hoá ma trận đồng hiện.

Bước 3: Tính các giá trị Energy, Entropy, Contrast, Homogeneity, Correlation.

Bước 4: Gán các giá trị vừa tính được cho vector vân đặt trung.

3.2. Tìm kiếm ảnh dựa vào thống kê khác biệt ở lượt đồ xám:

Kỹ thuật này cũng tương tự như đối với ma trận đồng hiện, nhưng nó thường mang lại kết quả tìm kiếm khả quan hơn nhiều, tốc độ cũng nhanh hơn đối với ma trận đồng hiện. Vector vân của trường hợp này cũng có 5 thành phần chính, mỗi thành phần tương ứng với các giá trị của công thức dưới đây:

Gọi $h[I]$ là lượt đồ xám của ảnh.

$$\text{Mean} = \sum_{i=1}^{255} i * h[i]$$

$$\text{Variance} = \sum_{i=1}^{255} (i - \text{mean})^2 * h[i]$$

$$\text{Skewness} = \frac{\sum_{i=1}^{255} (i - \text{mean})^3 * h[i]}{(\sqrt{\text{variance}})^3}$$

$$\text{Kurtosis} = \frac{\sum_{i=1}^{255} (i - \text{mean})^4 * h[i]}{4} - 3$$

$$\text{Entropy} = \sum_{i=1}^{255} -h[i] * \ln(h[i])$$

Các bước tính thống kê khác biệt ở lượt đồ xám như sau:

Bước 1: Ảnh màu được chuyển sang ảnh dạng xám

Bước 2: Tính lượt đồ xám cho ảnh xám vừa tính được

Bước 3: Tính các giá trị mean, variance, skewness, kurtosis, entropy dựa vào các biểu thức ở trên.

Bước 4: Gán các giá trị vừa tính được vào vectơ

4. Phần hình dạng:

Quy trình thực hiện như sau:

Bước 1: Nhận vào ảnh mẫu phát thảo.

Phân đoạn ảnh bằng ngưỡng.

Tính lượt đồ hình dạng cho ảnh vừa phân đoạn

Bước 2: Đọc ảnh từ cơ sở dữ liệu:

Nếu còn ảnh: Phân đoạn ảnh, tính lượt đồ hình dạng, qua Bước 3

Nếu hết ảnh: Kết thúc

Bước 3: So sánh 2 lượt đồ hình dạng vừa tính được dựa vào độ đo nào đó.

Bước 4: So sánh kết quả độ đo với ngưỡng, thông báo kết quả.

Quay lại bước 2.

Khoa CNTT - ĐH KHTN TP. HCM

Chương 2:

Kết quả thử nghiệm

- 1. Phần Màu sắc**
- 2. Phần Vân**
- 3. Phần Hình dạng**

Khoa CNTT - ĐHQG TP.HCM

1.Phần màu sắc:

Một số kết quả thử nghiệm một số chức năng chính của chương trình.

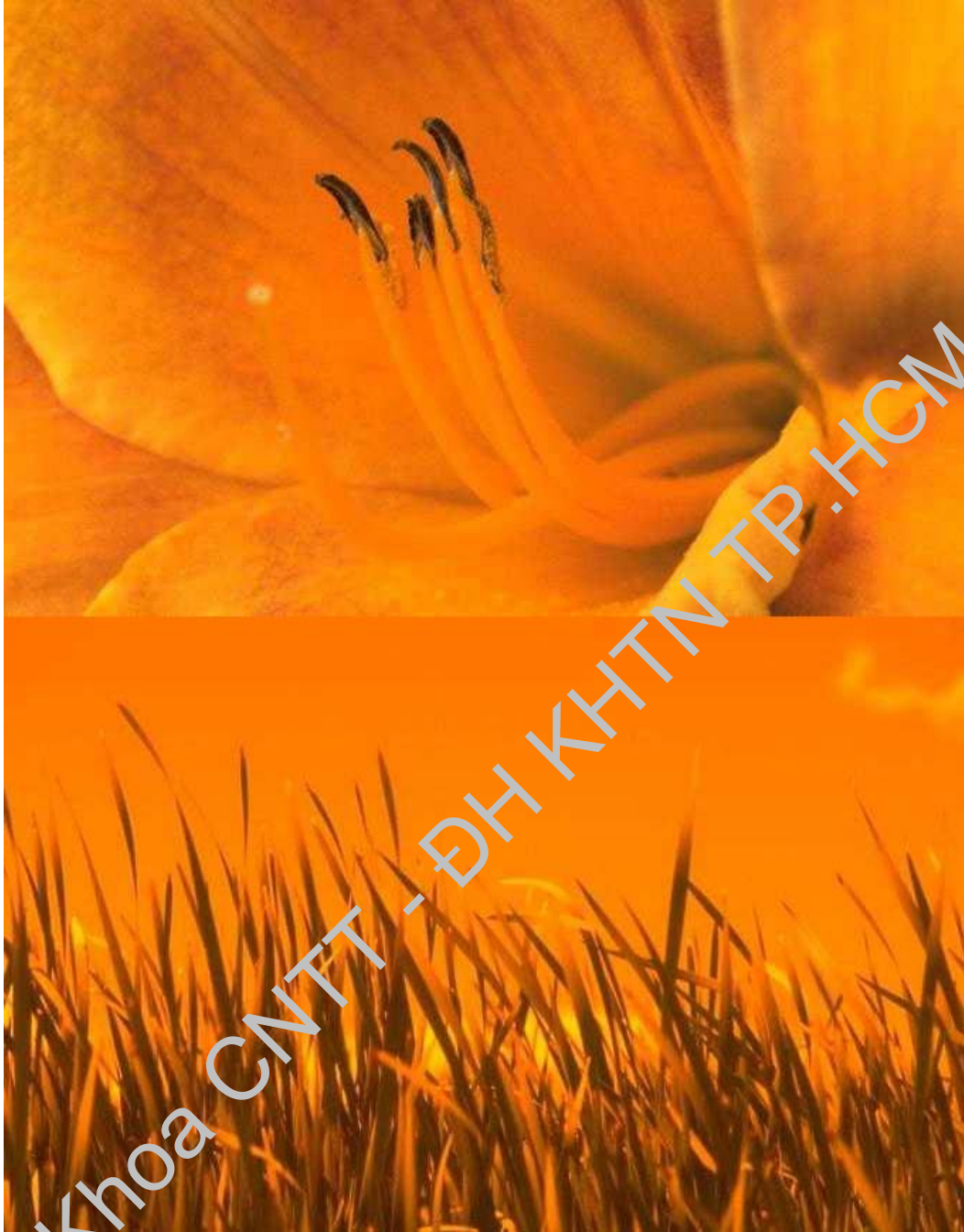
Cơ sở dữ liệu thử nghiệm: hơn 800 ảnh đủ mọi thể loại.

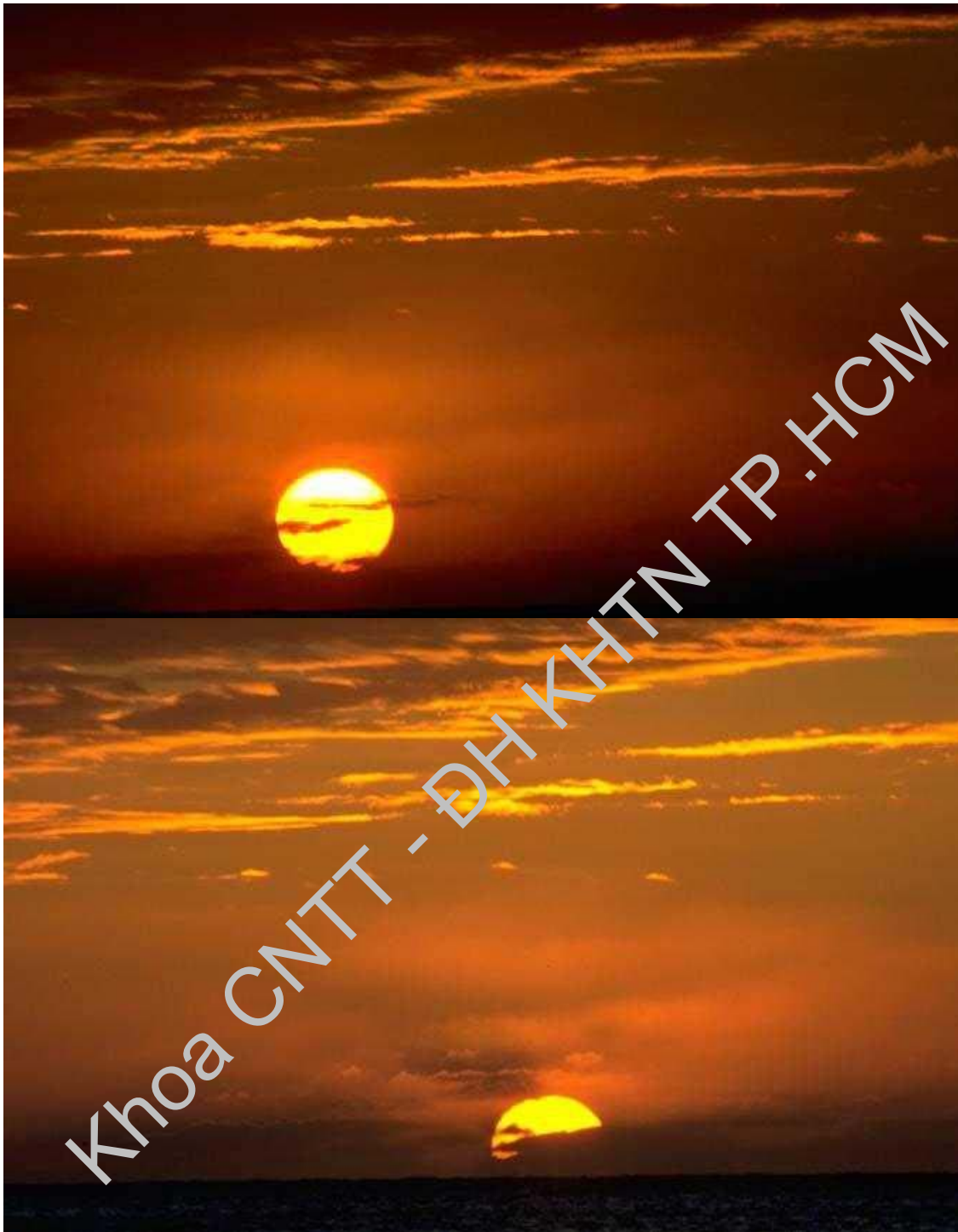
1.1.Tìm kiếm với chức năng tìm kiếm dựa trên lượt đồ toàn cục:

Ảnh đưa vào tìm kiếm:



Ngưỡng ban đầu được chọn là 50, cho ra kết quả tìm kiếm như sau:





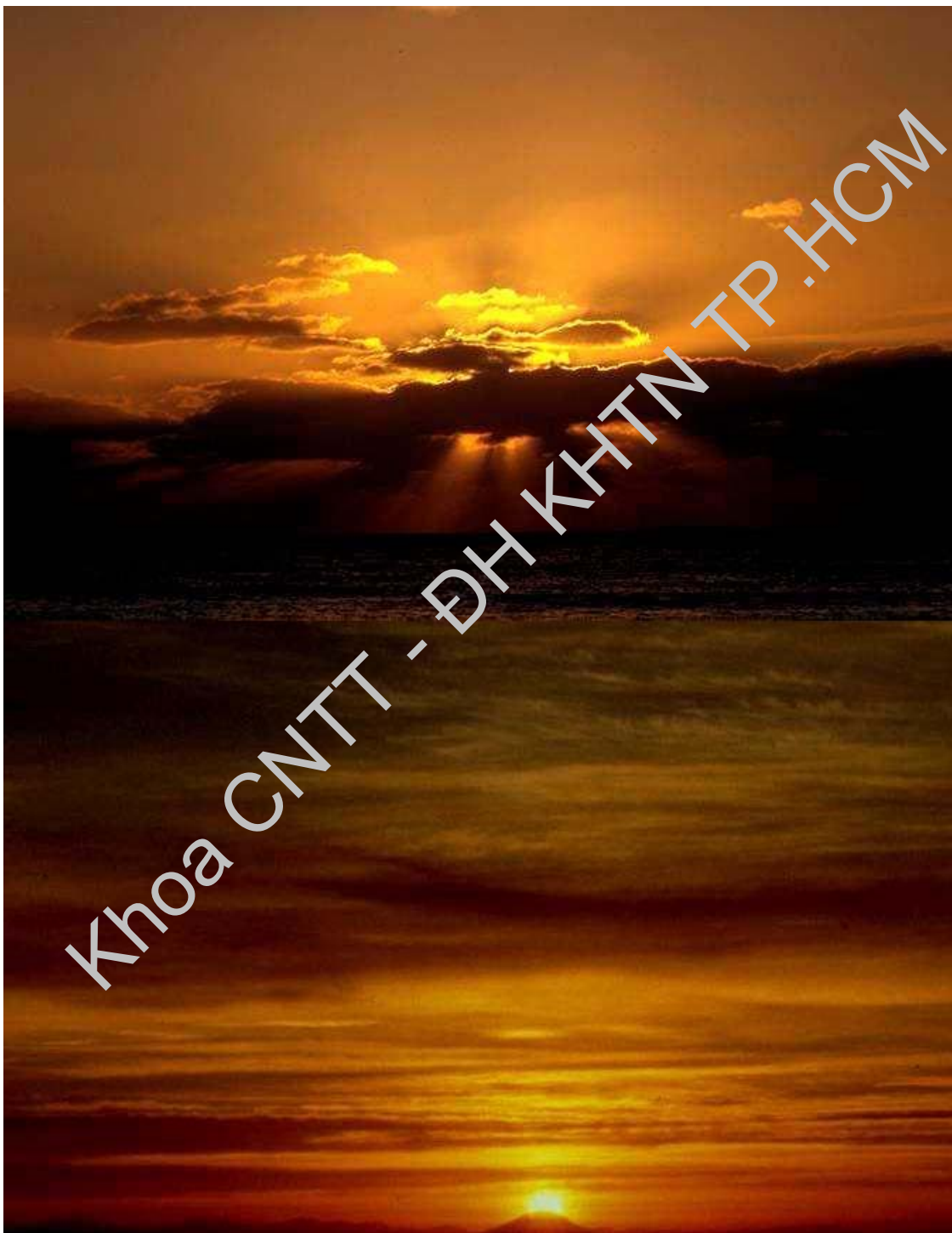








Sau khi đổi ngưỡng lại là 60, và tìm kiếm trong kết quả vừa tìm được, lặp lại quá trình tìm kiếm, ta thu được kết quả như sau:







1.2. Tìm kiếm với chức năng tìm kiếm dựa trên phân trầm màu:

Nhập vào: Chọn màu xanh blue (128,128,255), với tỉ lệ phần trăm xuất hiện trong ảnh là từ 20-40%. Kết quả thu được như sau:









1.3. Tìm kiếm với chức năng tìm kiếm dựa vị trí màu sắc trong không gian:

Chọn ngưỡng là 70, vị trí màu sắc : nửa trên của ảnh là màu trắng

Kết quả thu được như sau:





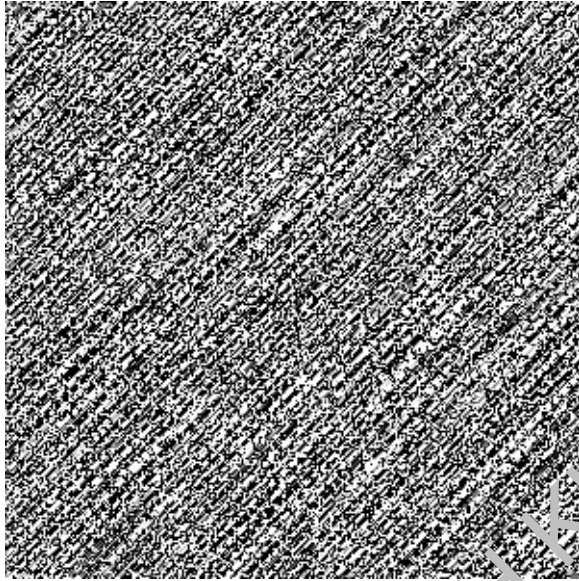




2. Phần vân:

2.1. Ma trận đồng hiện:

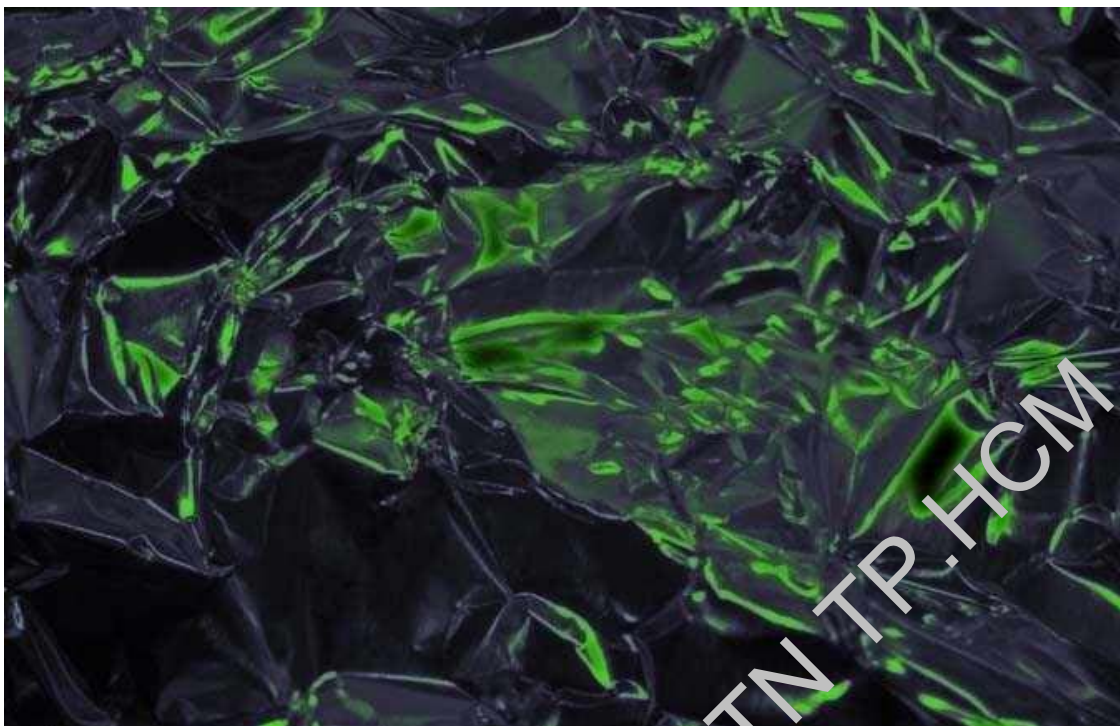
Vân được đưa vào kiểm tra:



Với mức ngưỡng là 0.06

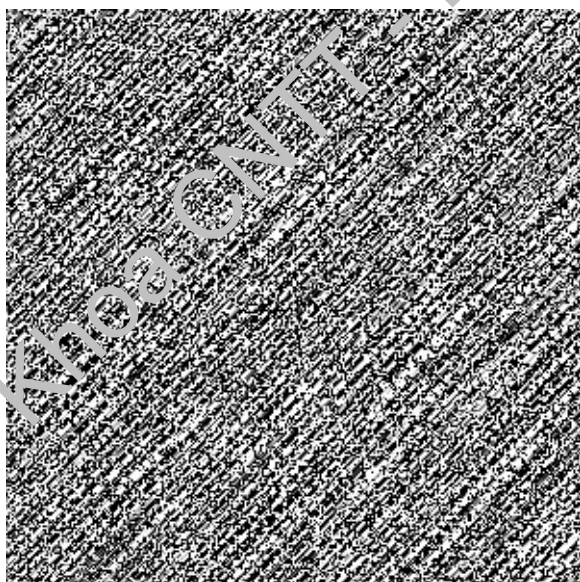
Kết quả thu được là:





2.1. Thống kê khác biệt với lượt đồ xám:

Vân được đưa vào kiểm tra:



Với mức ngưỡng là 0.06

Kết quả thu được là:



3. Phần hình dạng:

Rất tiếc là kết quả thu được không khả quan lắm.

Khoa CNTT - ĐH KHTN TP.HCM

PHẦN 4
KẾT LUẬN

Khoa CNTT - ĐH KHTN TP.HCM

Đánh giá kết quả đạt được

Người dùng có thể dùng chương trình này để tìm kiếm ảnh theo màu sắc, vân, và hình dáng. Việc tìm kiếm ảnh với tiêu chí màu có kết quả tốt. Phù hợp với thị giác người. Việc sử dụng sáng tạo một hệ thống tính toán lượt đồ màu HSI mới đã làm cho việc tìm kiếm ảnh dựa vào màu sắc đạt kết quả tốt.

Có thể tìm kiếm theo màu với rất nhiều tiêu chí đặt ra.

Tìm kiếm ảnh theo vân và hình dạng, chương trình còn sơ sài, chưa đáp ứng được. Vì thực sự đây là hai mảng đề tài rất khó, cần có nhiều sự đầu tư nghiên cứu hơn.

Hướng phát triển

- Kết hợp các chức năng màu sắc, vân, hình dạng lại với nhau để giúp việc tìm kiếm được hiệu quả hơn. Sự kết hợp giữa các chức năng này với nhau sẽ mang lại kết quả khả quan hơn nhiều, vì nó làm tăng khả năng tìm kết quả chính xác, theo như mong muốn.
- Tổ chức thành database ảnh đối với những số lượng ảnh lớn cố định. Điều này sẽ giúp cho tốc độ tìm kiếm đạt giá trị cực tiểu.
- Có thể phát triển thêm để ứng dụng được vào thực tế:
 - Giúp tìm hàng hoá trong siêu thị dựa vào hình ảnh.
 -
 - Tìm ảnh trên mạng.
 -

Tài liệu tham khảo

- [1] Linda G.Shapiro, George C.Stockman
COMPUTER VISION
Upper Saddle River, New Jersey
Prentice-Hall, Inc 2001
- [2] Ratael C.Gonzalez, Richard E. Woods
DIGITAL IMAGE PROCESSING
Addison-Wesley Publishing Company, Inc 1993
- [3] *Marius Tico, Taneli Haverinen, Pauli Kuosmanen*
**A METHOD OF COLOR HISTOGRAM CREATION FOR IMAGE
RETRIEVAL**
Signal Processing Laboratory
Digital Media Institute
Tampere University of Technology
P.O. BOX 553, FIN-33101, Tampere, FINLAND, tico@cs.tut.fi
- [4] Charles E. Jacobs, Adam Finkelstein, David H. Salesin
FAST MULTIREOLUTION IMAGE QUERYING
Department of Computer Science and Engineering
University of Washington
Seattle, Washington 98195

- [5] Alberto Del Bimbo
VISUAL INFORMATION RETRIEVAL
Department of Science of Information
University of Fienze Finenze, Italy
MORGAN KAUFMANN, 1999

Khoa CNTT - ĐH KHTN TP.HCM