

ĐỒ ÁN TỐT NGHIỆP

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG
-----o0o-----**

**NGHIÊN CỨU PHƯƠNG PHÁP TRA CỨU ẢNH DỰA VÀO
NỘI DUNG BIỂU DIỄN MÀU HIỆU QUẢ**

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công Nghệ Thông Tin

**Sinh viên thực hiện: NGUYỄN XUÂN TÙNG
Giáo viên hướng dẫn: PGS - TS. NGÔ QUỐC TẠO
Mã số sinh viên: 1013101008**

HẢI PHÒNG - 2012

Mục Lục

Chương 1: TỔNG QUAN TRÍCH RÚT ĐẶC TRƯNG ẢNH	6
1.1 Tổng quan về xử lý ảnh	6
1.1.1 Những vấn đề cơ bản trong hệ thống xử lý ảnh	6
1.1.2 Khái niệm ảnh đen trắng và ảnh màu	8
1.1.3 Chuyển ảnh màu thành ảnh xám	9
1.2 Tổng quan về tra cứu ảnh và tra cứu ảnh dựa vào nội dung	10
1.3 Khái quát về trích rút đặc trưng ảnh	14
1.3.1 Đặc trưng màu sắc	15
1.3.2 Đặc trưng kết cấu	17
1.3.3 Đặc trưng hình dạng	18
1.3.4 Đặc trưng cục bộ bất biến	18
1.4 Kết luận chương	19
Chương 2: PHƯƠNG PHÁP TÍNH TOÁN BIỂU DIỄN MÀU HIỆU QUẢ	20
2.1 Các không gian màu	20
2.1.1 Không gian màu RGB	20
2.1.2 Không gian màu HSx	21
2.1.3 Các không gian màu YUV và YIQ	22
2.1.4 Các không gian màu CIE XYZ và LUV	22
2.2 Lược đồ màu	23
2.3 Đo khoảng cách giữa các lược đồ màu	23
2.3.1 Khoảng cách dạng Minkowski	23
2.3.2 Khoảng cách dạng toàn phương	24
2.3.3 Khoảng cách trên Histogram màu	25
2.4 Phương pháp tra cứu ảnh sử dụng biểu diễn màu hiệu quả	26
2.4.1 Giới thiệu	26
2.4.2 Mô tả màu trội	26
2.4.3 Tìm kiếm và tra cứu	28
2.5 Kết luận	33

Chương 3: XÂY DỰNG HỆ THỐNG TRA CỨU ẢNH DỰA VÀO NỘI DUNG SỬ DỤNG BIỂU DIỄN MÀU HIỆU QUẢ.....	34
3.1 Giới thiệu bài toán tra cứu ảnh dựa vào nội dung	34
3.2 Mô tả chương trình	34
3.3 Kết luận	42
Chương 4: KẾT LUẬN.....	43

DANH MỤC HÌNH ẢNH

Hình 1.1: Kiến trúc hệ thống tra cứu ảnh dựa vào nội dung

Hình 2.1: Không gian màu RGB được trực quan hóa như một hình khối

Hình 2.2: Không gian màu HSV được trực quan hóa như một hình nón

Hình 2.3: Khoảng cách dạng Minkowski

Hình 2.4: Khoảng cách dạng Quadratic

Hình 2.5: Hai dạng lược đồ mức xám

Hình 2.6: Tìm kiếm và đánh chỉ số cơ bản

Hình 2.7: Mạng tinh thể hai chiều hình lục giác có p là bán kính cực tiểu, p' là bán kính cực đại

Hình 2.8: Độ chính xác và thu hồi cho $K = 50$

Hình 2.9: Cơ chế tìm kiếm trong mặt phẳng hai chiều

KÝ HIỆU TỪ VIẾT TẮT

Từ viết tắt	Tiếng anh	Tiếng Việt
PEL	Picture Element	Điểm ảnh
CGA	Color Graphic Adaptor	Độ phân giải của ảnh trên màn hình
CBIR	Content Based Image Retrieval	Tra cứu ảnh dựa vào nội dung
QBE	Query by Example	Truy vấn bởi ảnh mẫu
QBF	Query by Feature	Truy vấn bởi đặc trưng
RGB	Red Green Blue	Đỏ, Xanh lục, Xanh lơ
HSV	Hue, Saturation, Value	Màu, sắc nét, cường độ
GLA	Generalized Lloyd Algorithm	Thuật toán được điều chỉnh

Chương 1: TỔNG QUAN TRÍCH RÚT ĐẶC TRƯNG ẢNH

1.1 Tổng quan về xử lý ảnh

Xử lý ảnh từ lâu đã là một lĩnh vực mang tính khoa học và công nghệ. Nó là một ngành khoa học mới mẻ so với nhiều ngành khoa học khác nhưng tốc độ phát triển của nó rất nhanh, điều này kích thích các trung tâm nghiên cứu, các ứng dụng và đặc biệt là máy tính chuyên dụng riêng cho nó.

Xử lý ảnh là một môn học liên quan đến nhiều lĩnh vực và cần nhiều kiến thức cơ sở khác nhau. Có thể kể đến đầu tiên là xử lý tín hiệu số, môn học hết sức cơ bản cho xử lý tín hiệu chung. Các khái niệm về tính chập, các biến đổi Fourier, Laplace, các bộ lọc hữu hạn. . . Tiếp đến là các công cụ toán như Đại số tuyến tính, xác suất thống kê. Các kiến thức cần thiết như trí tuệ nhân tạo, mạng nơ ron nhân tạo cũng được đề cập trong quá trình phân tích và nhận dạng ảnh.

Các phương pháp xử lý ảnh bắt đầu từ các ứng dụng chính là nâng cao chất lượng ảnh và phân tích ảnh. Ứng dụng đầu tiên được biết đến là nâng cao chất lượng ảnh báo được truyền qua cáp từ London đến New York từ những năm 1920. Vấn đề nâng cao chất lượng ảnh có liên quan tới phân bố mức sáng và độ phân giải của ảnh. Vấn đề này đã được giải quyết vào những năm 1995. Năm 1964 máy tính đã có khả năng xử lý và nâng cao chất lượng ảnh từ mặt trăng và vệ tinh Ranger 7 của Mỹ bao gồm: làm nổi đường biên, lưu ảnh. Từ năm 1964 đến nay, các phương tiện xử lý nâng cao chất lượng, nhận dạng ảnh không ngừng phát triển và hoàn thiện hơn.

1.1.1 Những vấn đề cơ bản trong hệ thống xử lý ảnh

- **Điểm ảnh (Picture Element)**

Gốc của ảnh (ảnh tự nhiên) là ảnh liên tục về không gian và độ sáng. Để xử lý bằng máy tính (số), ảnh cần được số hóa. Số hóa ảnh là sự biến đổi gần đúng một ảnh liên tục thành một tập điểm phù hợp với ảnh thật về vị trí (không gian) và độ sáng (mức xám). Khoảng cách giữa các điểm ảnh đó được thiết lập sao cho mắt người không phân biệt được ranh giới giữa chúng. Mỗi một điểm như vậy gọi là điểm ảnh (PEL) hay gọi là Pixel. Trong khuôn khổ ảnh hai chiều, mỗi Pixel ứng với cặp tọa độ (x, y) .

Định nghĩa: Điểm ảnh (pixel) là một phần tử của ảnh số tại tọa độ (x, y) với tọa độ xám hoặc màu nhất định. Kích thước và khoảng cách giữa các điểm ảnh đó được chọn thích hợp sao

cho mắt người cảm nhận sự liên tục về không gian và mức xám (hoặc màu) của ảnh số gần như ảnh thật. Mỗi phân tử rong ma trận được gọi là một phân tử ảnh.

- **Độ phân giải của ảnh**

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

Theo định nghĩa, khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ phân bố, đó chính là độ phân giải. Nó được phân bố theo trục x và y trong không gian hai chiều.

Ví dụ: Ở màn hình CGA (Color Graphic Adaptor) là một lưới điểm theo chiều ngang màn hình gồm: 320 điểm chiều dọc * 200 điểm ảnh (320*200). Rõ ràng, ta có thể nhận thấy sự mịn màng hơn hẳn của màn hình 12’’ so với màn hình 17’’ cùng độ phân giải 320*200. Lý do là cùng một mật độ (độ phân giải) nhưng diện tích màn hình rộng hơn thì độ mịn (độ liên tục của các điểm) cũng kém hơn nhiều.

- **Mức xám của ảnh**

Một điểm ảnh (pixel) có hai đặc trưng cơ bản là vị trí (x,y) của điểm ảnh và độ xám của nó

Định nghĩa: Mức xám của điểm ảnh là cường độ sáng của nó được gán bằng giá trị tại điểm đó.

Các thang giá trị mức xám thông thường: 16,32,64,128,256 (Mức 256 là mức phổ dụng). Lý do: từ kỹ thuật máy tính dùng 1 byte (8bit) để biểu diễn mức xám, mức xám dùng 1 byte biểu diễn: $2^8 = 256$ mức, tức là từ 0 – 255.

Ảnh đen trắng: là ảnh có hai màu đen và trắng (không chứa màu sắc khác) với mức xám ở các điểm ảnh có thể khác nhau.

Ảnh nhị phân: ảnh chỉ có 2 mức đen trắng phân biệt, tức là dùng 1 bit mô tả 2^1 mức khác nhau. Tức là mỗi điểm ảnh của ảnh nhị phân chỉ có thể là 0 hoặc 1.

Ảnh màu: trong khuôn khổ lý thuyết ba màu (Red, Blue, Green) để tạo nên thế giới màu, người ta thường dùng 3 byte để mô tả mức màu, khi đó các giá trị màu: $2^{8*3} = 2^{24} \approx 16,7$ triệu màu.

- **Biểu diễn ảnh**

Trong biểu diễn ảnh, người ta thường dùng các phần tử đặc trưng của ảnh là pixel. Nhìn chung có thể xem một hàm hai biến chứa các thông tin như biểu diễn của một ảnh. Các mô hình biểu diễn ảnh cho ta một mô tả logic hay định lượng các tính chất của hàm này. Trong biểu diễn ảnh cần chú ý đến tính trung thực của ảnh hoặc các tiêu chuẩn “thông minh” để đo chất lượng ảnh hoặc tính hiệu quả của các kỹ thuật xử lý.

Việc xử lý ảnh số yêu cầu phải được mẫu hóa và lượng tử hóa. Việc lượng tử hóa ảnh là chuyển đổi tín hiệu tương tự sang tín hiệu số (Analog Digital Convert) của một ảnh đã lấy mẫu sang một số hữu hạn mức xám.

1.1.2 Khái niệm ảnh đen trắng và ảnh màu

Ảnh có thể biểu diễn dưới dạng tín hiệu tương tự hoặc tín hiệu số. Trong biểu diễn số của các ảnh đa mức xám thì một ảnh được biểu diễn dưới dạng một ma trận hai chiều. Mỗi phần tử của ma trận biểu diễn cho mức xám hay cường độ của ảnh tại vị trí đó. Một lưới chia ô vuông được đặt lên ảnh. Độ lớn mỗi ô vuông của lưới xác định kích thước của một điểm ảnh. Mức xám của một điểm được tính bằng cường độ sáng trung bình tại mỗi ô vuông này. Mật lưới càng nhỏ thì chất lượng ảnh càng cao. Trong kỹ thuật truyền hình tiên tiến, hình ảnh cần chất lượng cao với độ phân giải gấp hai lần so với các chuẩn hiện nay.

Trong kỹ thuật tương tự, một bức ảnh thường được biểu diễn dưới dạng các dòng nằm ngang kế tiếp nhau. Mỗi dòng là một tín hiệu tương tự mang theo các thông tin về cường độ sáng dọc theo một đường nằm ngang trong ảnh gốc. Ảnh trên một chiếc tivi được hiện lên qua các dòng quét này. Mặc dù thuật ngữ tương tự được dùng để mô tả cho các ảnh quét liên tiếp nhưng thực tế ảnh chỉ tương tự dọc theo hướng nằm ngang. Nó là rời rạc thì xét theo hướng dọc và chính vì vậy mà tín hiệu ảnh là tín hiệu nửa tương tự nửa số.

- **Ảnh đen trắng**

Ảnh đen trắng chỉ bao gồm 2 màu: màu đen và màu trắng. Người ta phân mức đen trắng đó thành L mức nếu sử dụng số bit $B = 8$ bit để mã hóa mức đen trắng (hay mức xám) thì L được xác định:

$$L = 2^B$$

Nếu $L = 2$, $B = 1$, nghĩa là chỉ có 2 mức: mức 0 và mức 1, còn gọi là ảnh nhị phân. Mức 1 ứng với màu sáng, còn mức 0 ứng với màu tối. Nếu L lớn hơn 2 ta có ảnh đa cấp xám.

Với ảnh nhị phân mỗi điểm ảnh được mã hóa trên 1 bit, còn với ảnh 256 mức, mỗi điểm ảnh được mã hóa trên 8 bit. Như vậy, với ảnh đen trắng: nếu dùng 8 bit (1 byte) để biểu diễn mức xám thì số mức xám có thể biểu diễn sẽ là 256. Mỗi mức xám được biểu diễn dưới dạng là một số nguyên nằm trong khoảng từ 0 – 255, với mức 0 biểu diễn cho mức cường độ đen nhất và mức 255 biểu diễn cho mức cường độ sáng nhất.

Ảnh nhị phân khá đơn giản, các phần tử ảnh có thể coi là các phần tử logic. Ứng dụng chính của nó được dùng để phân biệt đối tượng ảnh với nền hay để phân biệt điểm biên với các điểm khác.

- **Ảnh màu**

Theo lý thuyết của Thomas thì ảnh màu là ảnh tổ hợp từ 3 màu cơ bản là Red, Green, Blue và thu nhận trên các dải băng tần khác nhau. Với ảnh màu thì cách biểu diễn cũng tương tự như ảnh đen trắng, chỉ khác là các số tại mỗi phần tử của ma trận biểu diễn cho ba màu riêng rẽ. Để biểu diễn cho một điểm ảnh màu cần 24 bit. 24 bit này được chia thành ba khoảng 8 bit. Mỗi màu cũng phân thành L cấp màu khác nhau (thường $L = 256$). Mỗi khoảng này biểu diễn cho cường độ sáng của một trong các màu chính. Do đó, để lưu trữ ảnh màu người ta có thể lưu trữ từng màu riêng biệt, mỗi màu lưu trữ một ảnh đa cấp xám. Chính vì vậy không gian nhớ dành cho một ảnh màu lớn gấp 3 lần một ảnh đa cấp xám cùng kích cỡ.

1.1.3 Chuyển ảnh màu thành ảnh xám

Đơn vị tế bào của ảnh số là pixel. Tùy theo mỗi định dạng là ảnh màu hay ảnh xám mà từng pixel có thông số khác nhau. Đối với ảnh màu từng pixel sẽ mang thông tin của ba màu cơ bản

tạo ra bản màu khả kiến là Red, Green và Blue[Thomas 1892]. Trong mỗi pixel của ảnh màu, 3 màu cơ bản R, G, B được bố trí sát nhau và có cường độ sáng khác nhau. Thông thường mỗi màu sẽ được biểu diễn bằng 8 bit ~ 256 mức độ màu khác nhau. Như vậy mỗi pixel ta sẽ có $2^{8 \times 3} = 2^{24}$ màu (~16.78 triệu màu). Đối với ảnh xám thông thường mỗi pixel mang thông tin của 256 mức xám (~8 bit) như vậy ảnh xám hoàn toàn có thể tái hiện đầy đủ cấu trúc của một ảnh màu tương ứng thông qua tám mặt phẳng bit theo độ xám.

Trong hầu hết quá trình xử lý ảnh, chúng ta chỉ quan tâm đến cấu trúc của ảnh và bỏ qua ảnh hưởng của yếu tố màu sắc. Do đó bước chuyển từ ảnh màu thành ảnh xám là một công đoạn phổ biến trong các quá trình xử lý ảnh vì nó làm tăng tốc độ xử lý và làm giảm mức độ phức tạp của các thuật toán trên ảnh.

Chúng ta có công thức chuyển các thông số giá trị màu của một pixel thành mức xám tương ứng như sau:

$$G = \alpha.C_R + \beta.C_G + \delta.C_B \quad (1-1)$$

Trong đó các giá trị C_R , C_G và C_B lần lượt là các mức độ màu R, G, B của pixel màu. Các hệ số α , β và δ là các giá trị thay đổi tùy thuộc hệ màu. Trong phần xây dựng chương trình, sử dụng các hệ số là:

$$\alpha = 0.299, \beta = 0.587 \text{ và } \delta = 0.114.$$

Do đó công thức (1-1) trở thành:

$$G = 0.299.C_R + 0.587.C_G + 0.114.C_B$$

Chúng ta xem ảnh nguồn như một mảng hai chiều, duyệt qua từng pixel phần tử của mảng bằng công thức (1-1).

1.2 Tổng quan về tra cứu ảnh và tra cứu ảnh dựa vào nội dung

Tra cứu ảnh là quá trình tìm kiếm trong một cơ sở dữ liệu ảnh những ảnh thỏa mãn một yêu cầu nào đó. Ví dụ, người sử dụng có thể tìm kiếm tất cả các ảnh về một chủ đề nào đó như biển, trái đất, động vật . . . hoặc người khác lại muốn phân loại cơ sở ảnh của mình thành các bộ

suu tập có chủ đề khác nhau. Hoặc cũng có người muốn tìm tất cả ảnh tương tự với một bức ảnh có sẵn trong cơ sở dữ liệu.

Những kỹ thuật tra cứu ảnh số đã được nghiên cứu từ cuối những năm 1970. Từ đó cho đến nay khả năng ứng dụng cao của các kỹ thuật quản lý cơ sở dữ liệu ảnh đã thu hút sự quan tâm của rất nhiều các nhà nghiên cứu.

Các phương pháp tra cứu ảnh được sử dụng đầu tiên không dựa trên các đặc điểm trực quan của ảnh mà dựa trên các chú thích bằng lời của các bức ảnh, đầu tiên người ta gán cho mỗi ảnh một câu chú thích bằng lời (text) dựa trên một đặc điểm nào đó của ảnh, sau đó sử dụng các kỹ thuật tìm kiếm văn bản thông thường để tìm kiếm ảnh.

Phương pháp tra cứu ảnh dựa trên văn bản như trên sử dụng các kỹ thuật cơ sở dữ liệu truyền thống để quản lý ảnh. Dựa vào các lời chú thích, người ta có thể tổ chức cơ sở dữ liệu ảnh bằng các phân lớp theo chủ đề hay theo ngữ nghĩa và việc duyệt cơ sở dữ liệu ảnh chỉ dựa trên các truy vấn kiểu Bool thông thường. Phương pháp tra cứu ảnh dựa trên chú thích như trên còn được gọi là phương pháp tra cứu ảnh theo từ khóa. Do vậy việc xây dựng các thuật toán có khả năng tự động sinh ra các chú thích cho một cơ sở dữ liệu ảnh có nhiều chủ đề là hết sức khó khăn nên các hệ thống tra cứu ảnh kiểu này vẫn yêu cầu phải chú thích ảnh một cách thủ công, tốn rất nhiều công sức và quan trọng hơn là nó mang tính chủ quan, bị ảnh hưởng bởi hoàn cảnh và không đầy đủ.

Phương pháp tra cứu ảnh theo nội dung ra đời đã khắc phục được những nhược điểm của phương pháp từ khóa. Nội dung chính của phương pháp này dựa trên sự tương tự củ những đặc điểm trực quan của ảnh như màu sắc, hình dạng, kết cấu hay không gian của ảnh để phân loại sắp xếp các ảnh trong một cơ sở dữ liệu ảnh. Tuy nhiên những đặc điểm mà phương pháp này trích chọn để tra cứu vẫn là những đặc điểm ở mức thấp, chưa phản ánh được nội dung mang tính ngữ nghĩa của một đối tượng ảnh. Vì vậy người ta đã đưa ra một số cách tiếp cận mới phát triển phương pháp tra cứu ảnh theo nội dung thành các phương pháp tra cứu ảnh theo đồ thị hay tra cứu ảnh theo bản thể.

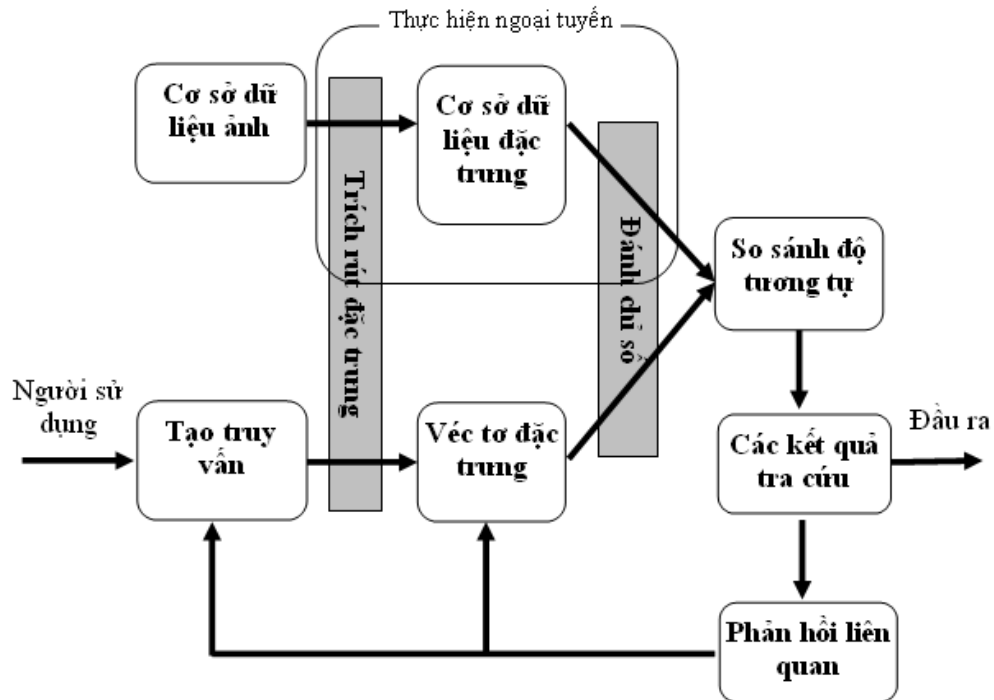
- **Tra cứu ảnh dựa vào nội dung**

Có hai cách tiếp cận chung đối với các giải pháp cho vấn đề tra cứu thông tin dựa trên dạng thông tin trực quan đó là các phương pháp dựa trên đặc trưng và các phương pháp dựa trên văn bản mô tả ảnh. Nhận thức chủ quan và chú thích thiếu chính xác là nguyên nhân tra cứu không chính xác trong các quá trình tra cứu sau của phương pháp dựa trên văn bản mô tả ảnh.

Các vấn đề đối với việc truy cập các ảnh và video dựa vào văn bản đã thúc đẩy nhanh chóng sự quan tâm phát triển các giải pháp dựa vào nội dung. Với giải pháp này, thay vì được chú thích một cách thủ công bởi các từ khóa dựa vào văn bản, các ảnh có thể được trích rút một số đặc trưng trực quan như màu, hình dạng, kết cấu và được đánh chỉ số dựa trên các đặc trưng trực quan này. Cách tiếp cận này dựa chủ yếu vào các kết quả từ thị giác máy. Tuy nhiên, không có đặc trưng riêng lẻ tốt nhất mà cho các kết quả chính xác. Thông thường một sự kết hợp các đặc trưng một cách tùy biến là cần thiết để cung cấp các kết quả tra cứu thích hợp cho ứng dụng tra cứu ảnh dựa vào nội dung.

Một hệ thống tra cứu ảnh dựa vào nội dung (CBIR – Content Based Image Retrieval) tiêu biểu không chỉ xử lý các nguồn thông tin khác nhau ở các khuôn dạng khác nhau (văn bản, hình ảnh, video) mà còn giải quyết các nhu cầu của người sử dụng. Về cơ bản hệ thống phân tích cả các nội dung của nguồn thông tin cũng như các truy vấn sử dụng và sau đó so sánh các nội dung này để tra cứu các mục tin liên quan. Các chức năng chính của hệ thống dựa vào nội dung như sau:

- Phân tích các nội dung của nguồn thông tin, và biểu diễn các nội dung của các nguồn được phân tích theo cách thích hợp cho so sánh các truy vấn sử dụng.
- Phân tích các truy vấn người sử dụng và biểu diễn chúng ở dạng thích hợp cho so sánh với cơ sở dữ liệu nguồn. Bước này tương tự bước trước nhưng chỉ áp dụng với ảnh truy vấn.
- Định nghĩa một chiến lược để so sánh các truy vấn tìm kiếm với thông tin trong cơ sở dữ liệu được lưu trữ. Tra cứu thông tin liên quan một cách hiệu quả. Bước này được thực hiện trực tuyến và được yêu cầu rất nhanh. Các kỹ thuật đánh chỉ số hiện đại có thể được sử dụng để tổ chức lại không gian đặc trưng để tăng tốc quá trình đối sánh.
- Thực hiện các điều chỉnh cần thiết trong hệ thống (thường bằng điều chỉnh các tham số trong máy đối sánh) dựa trên phản hồi từ người sử dụng hoặc các ảnh được tra cứu.



Hình 1.1: Kiến trúc hệ thống tra cứu ảnh dựa vào nội dung

Chúng ta nhận thấy rằng trên một mặt của một hệ thống tra cứu ảnh dựa vào nội dung, có các nguồn thông tin trực quan ở các khuôn dạng khác nhau và trên mặt kia có các truy vấn người sử dụng. Hai mặt này được liên kết thông qua một chuỗi các tác vụ như được minh họa trong Hình 1.1. Hai tác vụ phân tích truy vấn người sử dụng và đánh chỉ số nhiều chiều được tóm lược ở đây trong khi hai tác vụ quan trọng nhất: “Phân tích các nội dung của thông tin nguồn” (trích rút đặc trưng) và “Định nghĩa một chiến lược để đối sánh các truy vấn tìm kiếm với thông tin trong cơ sở dữ liệu được lưu trữ” (các độ đo tương tự).

+ Truy vấn người sử dụng:

Có nhiều cách gửi một truy vấn trực quan. Một phương pháp truy vấn tốt là một phương pháp tự nhiên đối với người sử dụng cũng như thu được đủ thông tin từ người sử dụng để trích rút các kết quả có ý nghĩa. Các phương pháp truy vấn dưới đây được sử dụng phổ biến trong nghiên cứu tra cứu ảnh dựa vào nội dung.

Truy vấn ảnh mẫu (QBE): Trong loại truy vấn này, người sử dụng hệ thống chỉ rõ một ảnh truy vấn đích, dựa trên ảnh đó hệ thống sẽ tìm kiếm trong cơ sở dữ liệu ảnh các ảnh tương tự nhất.

Truy vấn bởi đặc trưng (QBF): Trong hệ thống QBF tiêu biểu, những người sử dụng chỉ rõ các truy vấn bằng việc chỉ rõ các đặc trưng họ quan tâm cho tìm kiếm. Ví dụ: Một người sử dụng có thể truy vấn một cơ sở dữ liệu ảnh bằng việc đưa ra một lệnh tra cứu tất cả các ảnh có góc phần tư phải chứa 25% pixel màu xanh. Truy vấn này được chỉ rõ bằng việc sử dụng các công cụ giao diện đồ họa chuyên dụng. Truy vấn này là bình thường với người sử dụng hệ thống tra cứu ảnh chuyên biệt nhưng sẽ là không thể với người dùng chung.

Các truy vấn dựa vào thuộc tính: Các truy vấn dựa vào thuộc tính sử dụng các chú thích văn bản, trích rút trước bởi nỗ lực của con người, như một khóa tra cứu chính. Phương pháp này nhanh và dễ thực hiện, nhưng có một độ chủ quan và nhập nhằng cao xuất hiện như đã đề cập.

+ Đánh chỉ số nhiều chiều:

Để thực hiện tra cứu ảnh dựa vào nội dung đối với các cơ sở dữ liệu ảnh lớn, các kỹ thuật đánh chỉ số nhiều chiều cần được sử dụng. Có ba cộng đồng nghiên cứu chính đóng góp vào lĩnh vực này: hình học tính toán, quản trị cơ sở dữ liệu và nhận dạng mẫu.

1.3 Khái quát về trích rút đặc trưng ảnh

Trong tìm kiếm ảnh theo nội dung, việc lựa chọn các đặc trưng thích hợp với từng loại truy vấn và miền ứng dụng cùng với các độ đo tương đồng tương ứng là thành phần quan trọng và then chốt nhất. 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. 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).

1.3.1 Đặc trưng màu sắc

Tìm kiếm ảnh theo lược đồ màu là phương pháp phổ biến và được sử dụng nhiều nhất trong các hệ thống tìm kiếm ảnh theo nội dung. 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. Đây có thể xem là bước lọc đầu tiên cho những bước tìm kiếm sau. Một số lược đồ màu được sử dụng như: lược đồ màu RGB, lược đồ màu HIS, lược đồ HIS 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:

$$h_{R, G, B} [r, g, b] = N * Prob\{R = r, G = g, B = b\} \quad (1-2)$$

Trong đó N là số lượng điểm có trong ảnh.

Lượng 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 lượng đồ màu của ảnh RGB là ta phân ra làm 3 lượng đồ riêng biệt $h_R[]$, $h_G[]$, $h_B[]$. Khi đó mỗi lượng 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 lượng 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:

$$Intersection(h(I), h(M)) = \sum_{j=1}^K \sqrt{|h(I)_j - h(M)_j|^2} \quad (1-3)$$

Hoặc:

$$Intersection(h(I), h(M)) = \sum_{j=1}^K |h(I)_j - h(M)_j| \quad (1-4)$$

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

$$d_{JSD}(H, H') = \sum H_m \log \frac{2H_m}{H_m + H'_m} + H'_m \log \frac{2H'_m}{H_m + H'_m} \quad (1-5)$$

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

1.3.2 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 contourlet 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)

Độ đo tương đồng cho kết cấu ảnh:

Để đo độ tương đồng theo kết cấu giữa các ảnh, người ta thường sử dụng độ đo Oclit. Kết cấu được trích xuất từ các bức ảnh sẽ được biểu diễn thành các vector nhiều chiều và khoảng cách Oclit 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.

1.3.3 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 (contour-based descriptor): Biểu diễn các đường biên bao bên ngoài.

+ Biểu diễn theo vùng (region-based descriptor): Biểu diễn một vùng toàn vẹn.

Độ đo 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.

1.3.4 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 Haris (interest points). Phần này trình bày phương pháp trích rút các đặc trưng cục bộ bất biến SIFT của ảnh. Các đặc trưng này 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. Scale-Invariant Feature

Transform (SIFT) và đặc trưng trích rút được gọi là đặc trưng SIFT. Các đặc trưng này được trích rút ra từ các điểm hấp dẫn cục bộ (Local Interest Point).

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 trích rút 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.

1.4 Kết luận chương

Trong chương này, em đã giới thiệu tổng quan về xử lý ảnh, khái quát trích rút các đặc trưng như đặc trưng màu sắc, đặc trưng kết cấu, ... Giới thiệu về tổng quan tra cứu ảnh, đặc biệt là tra cứu ảnh dựa vào nội dung.

Chương 2: PHƯƠNG PHÁP TÍNH TOÁN BIỂU DIỄN MÀU HIỆU QUẢ

2.1 Các không gian màu

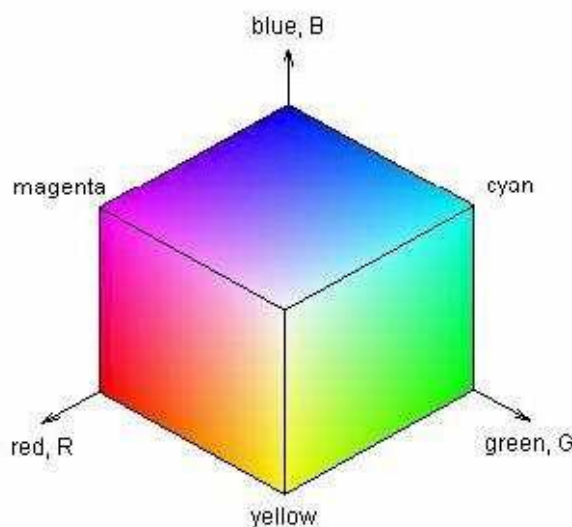
Một không gian màu chỉ rõ các màu như các bộ số, theo các đặc tả nào đó. Các không gian màu có thể dùng để sinh ra các biểu diễn màu, đặc biệt trong các biểu diễn số, như in số hoặc hiển thị điện tử số. Mục đích của không gian màu là để phục vụ đặc tả các màu theo một số cách chuẩn, được chấp nhận chung.

Có thể mô tả các không gian màu sử dụng khái niệm: đồng nhất nhận thức. Đồng nhất nhận thức có nghĩa là hai màu mà có khoảng cách bằng nhau trong không gian màu thì cũng được nhận thức như nhau bởi người quan sát.

Trong phần còn lại, một số không gian màu cùng các lược đồ lượng hóa sẽ được mô tả.

2.1.1 Không gian màu RGB

Không gian màu RGB là không gian màu được sử dụng nhiều nhất cho đồ họa máy tính. Lưu ý rằng R, G và B là viết tắt của các từ đỏ (Red), xanh lục (Green), và xanh lơ (Blue). Đây là không gian màu cộng: đỏ, xanh lục và xanh lơ được kết hợp lại để tạo ra các màu khác. Không gian này không đồng nhất về nhận thức. Không gian màu RGB có thể trực quan hóa thành một khối như được minh họa trong hình 2.1



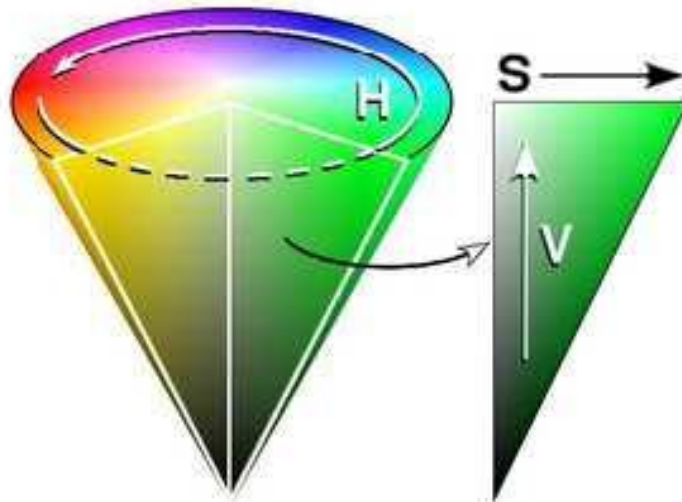
Hình 2.1: Không gian màu RGB được trực quan hóa như một hình khối.

Mỗi trục màu (R, G, B) có độ quan trọng như nhau. Do đó, mỗi trục nên được lượng hóa với cùng một độ chính xác. Khi không gian màu RGB được lượng hóa, số các bin sẽ luôn luôn là một hình khối. Thông thường, 8 (2^3), 64 (4^3), 216 (6^3), 512 (8^3) bin được sử dụng trong lượng hóa không gian màu RGB. Chuyển đổi từ một ảnh RGB sang một ảnh cấp xám nhận được tổng của R, G và B rồi chia kết quả cho ba.

2.1.2 Không gian màu HSx

Các không gian màu HSI, HSV, HSB và HSL (quy ước gọi là HSx) là gần với nhận thức của người hơn không gian màu RGB, nhưng vẫn không là đồng nhất nhận thức.

Các trục từ các không gian màu HSx biểu diễn các đặc trưng màu (hue), độ bão hòa (saturation), và độ sáng (lightness) (cũng được gọi là value, brightness và intensity). Sự khác biệt giữa các không gian màu trong HSx là cách chúng biến đổi từ không gian màu RGB. Chúng thường được biểu diễn bởi các hình khác nhau (ví dụ hình nón, hình trụ). Trong hình 2.2 không gian màu HSV được trực quan hóa như một hình nón.



Hình 2.2: Không gian màu HSV được trực quan hóa như một hình nón.

Hue là thành phần màu của các không gian màu HSx. Hue là một góc giữa một đường tham chiếu và điểm màu trong không gian RGB, phạm vi giá trị này giữa 0^0 và 360^0 , thí dụ Blue là

240⁰. Theo CIE, Hue là “thuộc tính của cảm giác trực quan theo đó một vùng xuất hiện là tương ứng với một vùng màu được nhận thức, đỏ (red), vàng (yellow), xanh lục (green), và xanh lơ (blue), hoặc kết hợp của hai trong chúng”. Nói cách khác, hue là loại màu, như đỏ hoặc xanh lục. Cũng theo CIE, độ bão hòa là “màu đầy của một vùng được xem xét tương ứng với độ sáng của nó”. Trong hình nón, độ bão hòa là khoảng cách từ tâm của mặt cắt ngang tròn hình nón, “chiều cao” nơi mặt cắt ngang này nhận được được xác định bởi giá trị – Value, nó là khoảng cách từ điểm cuối của hình nón. Giá trị là độ sáng của một màu, điều này được định nghĩa bởi CIE như “thuộc tính của cảm giác trực quan theo đó một vùng xuất hiện để phát ra nhiều hoặc ít ánh sáng”. Khi độ bão hòa đạt đến 0, Hue không xác định. Trục giá trị biểu diễn ảnh cấp xám.

Không gian màu HSV có thể được lượng hóa dễ dàng thành 162 bin, ở đây hue nhận 18 bin, độ bão hòa và giá trị nhận 3 bin. Trong vòng tròn hue các màu chính như đỏ, xanh lục, xanh lơ được tách bởi 120⁰. Các màu phụ là yellow, magenta và cyan cũng được tách bởi 120⁰ và 60⁰ cách xa hai màu chính gần nhất.

2.1.3 Các không gian màu YUV và YIQ

Các không gian màu YUV và YIQ được phát triển cho truyền hình vô tuyến. Không gian màu YIQ là giống như YUV, ở đây mặt phẳng I-Q là một mặt phẳng quay 33⁰ của mặt phẳng U-V. Tín hiệu Y biểu diễn độ chói của một pixel và là kênh duy nhất được sử dụng trong tivi đen trắng. U và V cho YUV, I và Q cho YIQ là các thành phần màu.

Kênh Y được định nghĩa bởi các giá trị năng lượng có trọng số của R(0.299), G(0.587) và B(0.144). Các không gian màu YUV và YIQ không là đồng nhất nhận thức. Khi các không gian màu YUV và YIQ được lượng hóa, mỗi trục được lượng hóa với cùng độ chính xác.

2.1.4 Các không gian màu CIE XYZ và LUV

Không gian màu đầu tiên được phát triển bởi CIE là không gian màu XYZ. Thành phần Y là thành phần độ chói được xác định bởi các tổng có trọng số của R(0:212671), G(0:715160), và B(0:072169). X và Z là các thành phần màu. Không gian màu XYZ là không đồng nhất nhận thức. Trong lượng hóa không gian màu XYZ, mỗi trục được lượng hóa với cùng độ chính xác.

Không gian màu CIE LUV là một biến đổi xạ ảnh của không gian màu XYZ và là đồng nhất nhận thức. Kênh L của không gian màu LUV là độ chói của màu. Các kênh U và V là các thành phần màu. Vậy, khi U và V được đặt bằng 0, kênh L biểu diễn một ảnh cấp xám.

Trong lượng hóa không gian LUV, mỗi trục được lượng hóa với cùng độ chính xác. Với cả không gian màu XYZ và LUV, các lược đồ lượng hóa thường được sử dụng $8(2^3)$, $27(3^3)$, $64(4^3)$, $125(5^3)$ bin.

2.2 Lược đồ màu

Lược đồ màu được xác định bởi tập các bin, trong đó mỗi bin biểu thị xác suất của các pixel trong ảnh. Một lược đồ màu H của một ảnh đã cho được xác định bởi véc tơ:

$$H = \{H[0], H[1], H[2], \dots, H[i], \dots, H[N],\} \quad (2-1)$$

Ở đây i biểu diễn một màu trong lược đồ màu và tương ứng với một khối con trong không gian màu RGB, H[i] là số các pixel có màu i trong ảnh, và N là số các bin trong lược đồ màu.

Mỗi pixel trong ảnh sẽ thuộc về một bin của lược đồ màu của ảnh, vì thế với lược đồ màu của một ảnh, giá trị của mỗi bin là số các pixel cùng màu. Để so sánh các ảnh có các kích cỡ khác nhau, các lược đồ màu được chuẩn hóa. Lược đồ màu chuẩn hóa H' được xác định bằng:

$$H' = \{H'[0], H'[1], H'[2], \dots, H'[i], \dots, H'[N],\} \quad (2-2)$$

Ở đây $H'[i] = \frac{H[i]}{P}$, P là tổng số các pixel trong ảnh.

2.3 Đo khoảng cách giữa các lược đồ màu

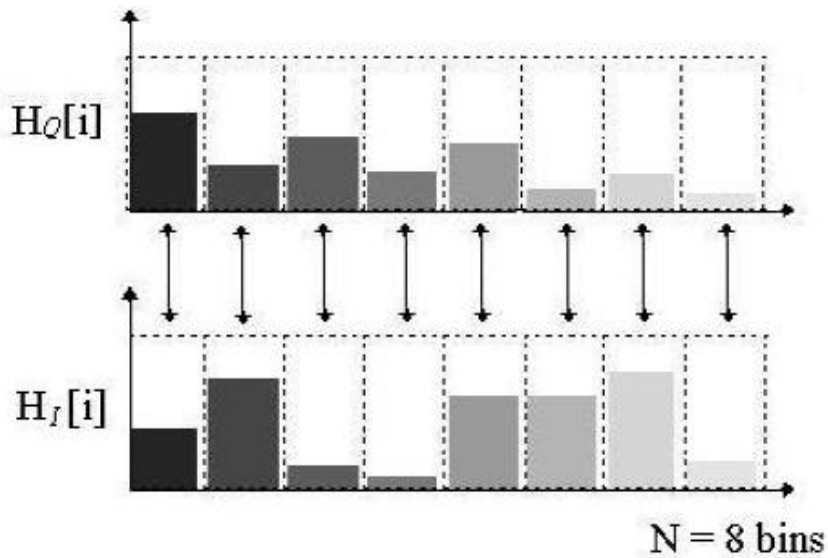
Các độ đo khác nhau đã được đề xuất để tính toán khoảng cách giữa các lược đồ màu. Các độ đo khoảng cách có thể được phân loại thành ba lớp đó là dạng Minkowsk, dạng Quadratic và Histogram.

2.3.1 Khoảng cách dạng Minkowski

Độ đo này chỉ so sánh các bin giống nhau giữa các lược đồ màu (xem hình 2.3) và được xác định:

$$d_{Q,I} = \sum_{i=1}^N |H_Q[i] - H_I[i]| \quad (2-3)$$

Ở đây Q và I là hai ảnh, N là số các bin trong lược đồ màu, $H_Q[i]$ là giá trị của bin i trong lược đồ màu H_Q , và $H_I[i]$ là giá trị của bin i trong lược đồ màu H_I .



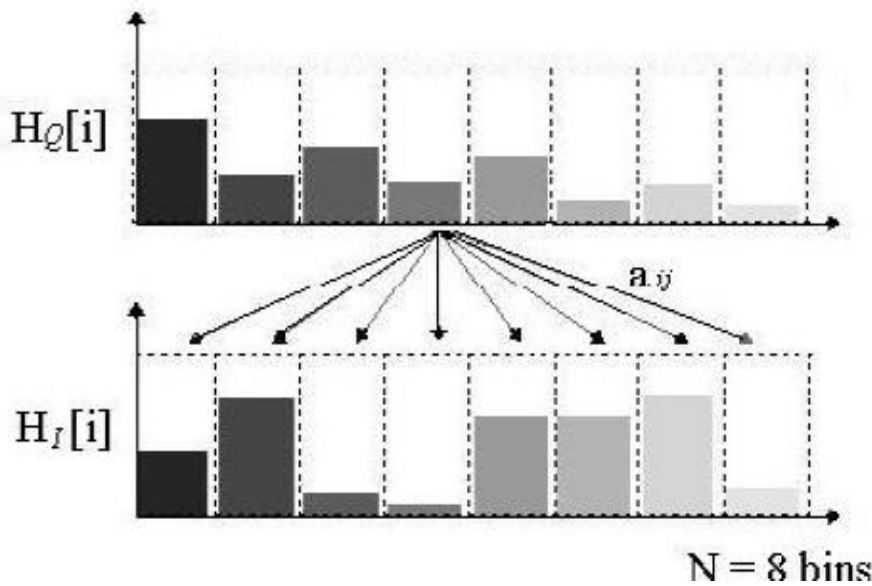
Hình 2.3: Khoảng cách dạng Minkowski

2.3.2 Khoảng cách dạng toàn phương

Độ đo này không chỉ so sánh các bin giống nhau mà so sánh nhiều bin giữa các lược đồ màu (xem hình 2.4) và được xác định:

$$d(Q,I) = (H_Q - H_I)'A(H_Q - H_I) \quad (2-4)$$

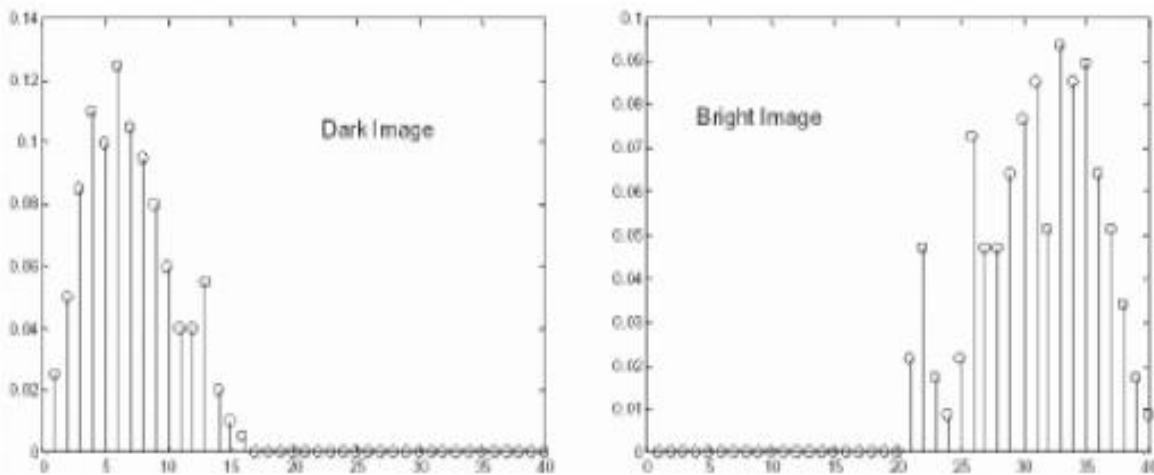
Ở đây, Q và I là ảnh, H_Q là lược đồ màu của ảnh Q , H_I là lược đồ màu của ảnh I , $A = [a_{i,j}]$ là ma trận $N \times N$, N là số các bin trong lược đồ màu, và $a_{i,j}$ biểu thị sự tương tự giữa màu i và màu j .



Hình 2.4: Khoảng cách dạng Quadratic

2.3.3 Khoảng cách trên Histogram màu

Histogram của một ảnh là một hàm cung cấp tần suất xuất hiện của mỗi mức xám. Ta có thể biểu diễn histogram trong hệ trục Oxy với trục hoành Ox biểu diễn số mức xám từ 0 đến L (số mức xám), trục tung Oy biểu diễn số điểm ảnh cho một mức xám tương ứng.



Hình 2.5: Hai dạng lược đồ mức xám

Nghĩa là, histogram của ảnh có mức xám $[0, L]$ là một hàm rời rạc $h(r_k) = n_k$ với r_k : mức xám thứ k và n_k là số pixel có mức k .

Histogram thường sẽ được chuẩn hóa thành $p(r_k) = n_k/n$ với n là tổng số pixel trong ảnh. Và lúc này ta có : $\left(\sum_{k=0}^L p(r_k) = 1 \right)$ (2-5)

Độ đo giữa 2 histogram gọi là độ đo phần giao cặp histogram (histogram intersection measure) của ảnh q và ảnh t là:

$$V_{q,t} = 1 - \frac{\sum_{m=0}^L \min(p^q(r_m), p^t(r_m))}{\min\left(\sum_{m=0}^L p^q(r_m), \sum_{m=0}^L p^t(r_m)\right)} \quad (2-6)$$

2.4 Phương pháp tra cứu ảnh sử dụng biểu diễn màu hiệu quả

2.4.1 Giới thiệu

Phương pháp tra cứu ảnh sử dụng biểu diễn màu hiệu quả là tra cứu độ tương tự trong cơ sở dữ liệu ảnh lớn sử dụng màu sắc. Màu sắc trong một vùng sẽ được chia thành nhiều cụm nhỏ các màu. Mô tả đặc trưng bao gồm biểu diễn màu và tỉ lệ của chúng trong vùng. Phương pháp đo độ tương tự được sử dụng trong đồ án sử dụng độ đo khoảng cách lược đồ màu chính phương, nghĩa là màu sắc biểu diễn có thể được đánh chỉ số trong tọa độ không gian màu ba chiều (3D), điều này có thể giúp chương trình tránh được các vấn đề có thể xảy ra như đối với việc sử dụng lược đồ màu truyền thống. Để tra cứu độ tương tự, mỗi màu biểu diễn trong ảnh truy vấn được kết hợp để đưa ra kết quả cuối cùng. Hiệu quả đánh chỉ số sẽ được đưa ra nhằm tăng tốc độ tra cứu. Thực nghiệm cho thấy rằng phương pháp miêu tả kết hợp này sẽ cho hiệu quả tra cứu cao hơn so với phương pháp lược đồ màu truyền thống.

2.4.2 Mô tả màu trội

Trích rút đặc trưng màu cục bộ bắt đầu với phân đoạn ảnh màu. Với phân đoạn ảnh, chúng ta sử dụng thuật toán luồng biên (edgeflow) [8]. Phân cụm màu được thực hiện trên mỗi vùng được phân đoạn để thu được các màu đại diện của nó.

Sau khi phân cụm, chỉ còn lại một số nhỏ các màu và phần trăm của các màu này được tính toán. Mỗi màu đại diện và phần trăm tương ứng của nó tạo thành một cặp thuộc tính (mô tả các đặc tính màu trong một vùng ảnh). Mô tả màu trội F được xác định là:

$$F = \{ \{c_i, p_i\}, i = 1, \dots, N \} \quad (2-7)$$

Ở đây N là tổng số các cụm màu trong vùng ảnh, c_i là vector màu ba chiều, p_i là phần trăm của nó và $\sum_i p_i = 1$. Lưu ý rằng N có thể thay đổi từ vùng này sang vùng kia.

Phân cụm màu

Chúng ta sử dụng phương pháp phân cụm màu được đề xuất trong [6]. Trong phân cụm này, các giá trị màu của pixel là véc tơ được lượng hóa sử dụng thuật toán được điều chỉnh GLA (Generalized Lloyd Algorithm). Các màu được biểu diễn trong không gian màu đồng nhất nhận thức CIE LUV. Biến dạng D trong mỗi cụm được cho bởi:

$$D_i = \sum_n v(n) \|x(n) - c_i\|^2, \quad x(n) \in C_i \quad (2-8)$$

Ở đây c_i là trọng tâm của cụm C_i , $x(n)$ là véc tơ màu tại pixel n và $v(n)$ là trọng số nhận thức cho pixel n . Các trọng số nhận thức được tính toán từ thông kê pixel cục bộ vào bản miêu tả với thực tế rằng nhận thức trực quan của con người nhạy cảm với các thay đổi trong các vùng trơn hơn là các vùng kết cấu như được miêu tả trong [6]. Luật cập nhật cho độ đo biến dạng này là:

$$c_i = \frac{\sum v(n)x(n)}{\sum v(n)}, \quad x(n) \in C_i \quad (2-9)$$

Phân cụm tích lũy [9] được thực hiện trên các trọng tâm cụm để nhập tiếp các cụm gần sao cho khoảng cách các chực tiêu giữa hai trọng tâm cụm vượt quá một ngưỡng được xác định trước. Ảnh được lượng hóa cuối cùng thu được bằng việc gán mỗi điểm ảnh cho các trọng tâm cụm gần nhất của nó. Nhược điểm đối với cách tiếp cận này là các màu đại diện không được chính xác như trích rút màu trực tiếp từ các pixel và độ phức tạp thuật toán cao.

Tương tự màu

Các màu $\{c_i\}$ và phần trăm của các pixel tương ứng tạo ra mô tả màu được cho bởi (2-7). $F_1 = \{ \{c_i, p_i\}, i = 1, \dots, N_1 \}$ và $F_2 = \{ \{b_j, q_j\}, j = 1, \dots, N_2 \}$ là hai mô tả đặc trưng màu. Khoảng cách giữa F_1 và F_2 được cho bởi:

$$D^2(F_1, F_2) = \sum_{i=1}^{N_1} p_i^2 + \sum_{j=1}^{N_2} q_j^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{i,j} p_i q_j \quad (2-10)$$

Ở đây $a_{i,j}$ là hệ số tương tự giữa các màu c_i và b_j

$$a_{i,j} = \begin{cases} 1 - d_{i,j}/d_{\max}, & d_{i,j} \leq T_d \\ 0, & d_{i,j} > T_d \end{cases} \quad (2-11)$$

Ở đây $d_{i,j}$ là khoảng cách Euclide giữa các màu c_i và b_j

$$d_{i,j} = \|c_i - b_j\| \quad (2-12)$$

và T_d được định nghĩa trước, là khoảng cách cực đại của hai màu được xem là tương tự. $d_{\max} = \alpha T_d$ và α được đặt đến 1.2 trong các thực nghiệm.

Độ đo khoảng cách trên được chỉ ra là tương đương độ đo khoảng cách lược đồ màu bình phương [10]

$$D^2_h(H_1, H_2) = \sum_{i=1}^{N_1} \sum_{k=1}^{N_1} a_{i,k} p_i p_k + \sum_{j=1}^{N_2} \sum_{l=1}^{N_2} a_{j,l} q_j q_l - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{i,j} p_i q_j \quad (2-13)$$

Ở đây H_1 và H_2 là các vector lược đồ truyền thống có độ dài N_h và hệ số của ma trận A là $a_{i,j}$. Trong thực tế, nếu số các bin màu trong véc tơ lược đồ N_h là đủ lớn sao cho tất cả các màu đại diện là các bin màu của phương pháp lược đồ, một véc tơ lược đồ màu có thể được xây dựng sử dụng các giá trị phần trăm p_i . Bỏ qua tất cả các mục không (zero) và viết lại khoảng cách bình phương

$$D^2_h(H_1, H_2) = \sum_{i=1}^{N_1} \sum_{k=1}^{N_1} a_{i,k} p_i p_k + \sum_{j=1}^{N_2} \sum_{l=1}^{N_2} a_{j,l} q_j q_l - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{i,j} p_i q_j \quad (2-14)$$

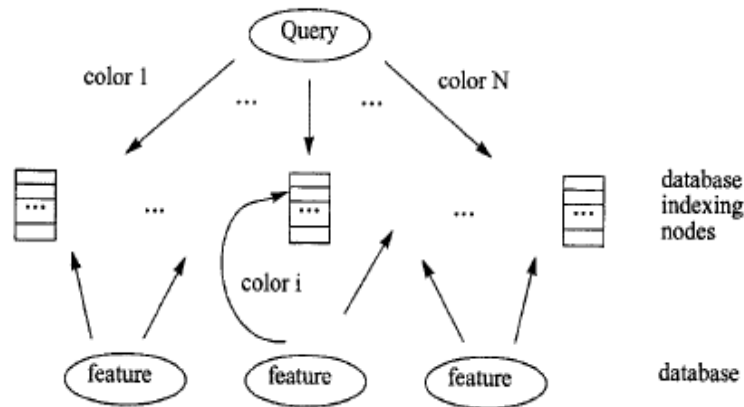
Trong quá trình phân cụm, khoảng cách cực tiểu giữa hai trọng tâm cụm có thể được đặt là T_d như được bàn luận ở trên. Lưu ý rằng:

$$a_{i,k} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \quad \text{và} \quad a_{j,l} = \begin{cases} 1, & j = l \\ 0, & j \neq l \end{cases} \quad (2-15)$$

Ta có thể dễ dàng nhận ra: $D_h(H_1, H_2) = D(F_1, F_2)$.

2.4.3 Tìm kiếm và tra cứu

Mỗi đối tượng hoặc vùng trong cơ sở dữ liệu được biểu diễn sử dụng mô tả màu trội như được định nghĩa trong (2-9). Tiêu biểu, từ ba đến bốn màu cung cấp mô tả đặc tính tốt của các màu vùng. Một ảnh truy vấn đã cho, tra cứu tương tự bao gồm tìm kiếm cơ sở dữ liệu với các phân bố màu tương tự như truy vấn đầu vào. Do số các màu đại diện là nhỏ, cho nên có thể tìm kiếm cơ sở dữ liệu với mỗi màu đại diện tách biệt và kết hợp các kết quả. Hình 2.7 minh họa lược đồ tìm kiếm và đánh chỉ số cơ bản.



Hình 2.6: Tìm kiếm và đánh chỉ số cơ bản.

Tìm kiếm các màu đơn lẻ có thể được thực hiện rất hiệu quả trong không gian màu ba chiều. Ở đây chúng em chỉ xét các truy vấn có phạm vi cố định (T_d [7] giới hạn phạm vi tìm kiếm) Trong khi có thể sử dụng các cấu trúc cây được biết rộng rãi như R-tree [2] và SS tree [5], một biểu diễn một lớp với một phân hoạch không gian màu đồng nhất được thiết kế chuyên biệt cho một phạm vi đã cho là hiệu quả hơn [11]. Một cách tiếp cận để phân hoạch không gian màu thành một số hữu hạn các khối có cỡ như nhau. Tuy nhiên, các cấu trúc lưới hình chữ nhật như thế không phải là hiệu quả nhất cho các truy vấn phạm vi hình cầu.

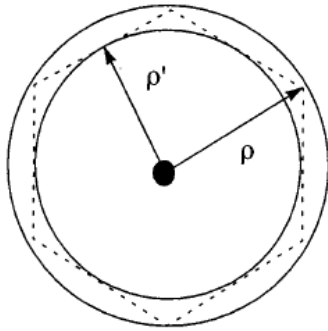
Cấu trúc lưới (Lattice) D_3^*

Lưới D_3^* có một số thuộc tính mong muốn nào đó trong gói hình cầu [7] (liên quan đến các truy vấn có phạm vi hình cầu):

- Độ chính xác: lưới D_3^* đưa ra tổng các sai số lượng hóa bình phương trung bình cực tiểu giữa tất cả các cấu trúc lưới trong không gian ba chiều.
- Hiệu quả: Lưới D_3^* có phủ không gian ba chiều tối ưu nhất.

Cấu trúc của lưới D_3^* là rất đơn giản. Lưới cơ bản bao gồm các điểm (x, y, z) ở đây x, y và z tất cả là các số nguyên chẵn hoặc lẻ. Chẳng hạn, $(0, 0, 0)$, $(1, 1, 3)$ và $(2, 10, 20)$ thuộc về lưới D_3^* . Các điểm này có thể được tỉ lệ và trượt để có các vị trí và các khoảng điểm lưới mong muốn.

Có hai tham số quan trọng trong thiết kế lưới: p : bán kính cực tiểu của hình cầu mà có thể phủ lên tế bào Voronoi và p' : bán kính cực đại của hình cầu mà tế bào Voronoi có thể phủ. Hình 2.8 minh họa các tham số này cho trường hợp hai chiều.



Hình 2.7: Mạng tinh thể hai chiều hình lục giác có p là bán kính cực tiểu, p' là bán kính cực đại

Với lưới D_3^* cơ bản, $p'_{basic} = \|(0,0,0) - (1,1,1)\| / 2 = 0.866$. Tuy nhiên, tính toán của p hơi phức tạp một chút. Trong [7], tỉ lệ giữa hai biến được cung cấp tức là $p'/p=0.7747$. Giá trị của p được đặt trong quá trình thiết kế lưới. Có thể tính toán nhân tố tỉ lệ s mà tỉ lệ D_3^* cơ bản sử dụng quan hệ sau:

$$s = 0.7747 p/p'_{basic} \quad (2-16)$$

Chẳng hạn, các điểm lưới được cho $(1,1,1)$ và $(2,0,0)$, các điểm lưới tỉ lệ tương ứng là (s,s,s) và $(2s,0,0)$.

Cho một điểm c tùy ý trong không gian ba chiều, điểm lưới gần nhất tương ứng b_c có thể được tính dễ dàng như sau.

$$b_e = \text{round}\left(\frac{c}{L}\right).L \quad (2-17)$$

$$\text{Và } b_0 = \text{round}\left(\frac{c-L/2}{L}\right).L + L/2 \quad (2-18)$$

Ở đây $L = 2s$ và L là cỡ bước dọc theo trục tọa độ. Chẳng hạn, $L = 2$ cho lưới cơ bản. Góc được giả thiết là một điểm lưới. Lưu ý rằng b_e là điểm lưới chẵn gần nhất và b_0 là điểm lưới lẻ gần nhất khi cho trước điểm c . Cho $b = \{b_e, b_0\}$. thì:

$$b_c = \arg \min_b \|b - c\| \quad (2-19)$$

Đánh chỉ số

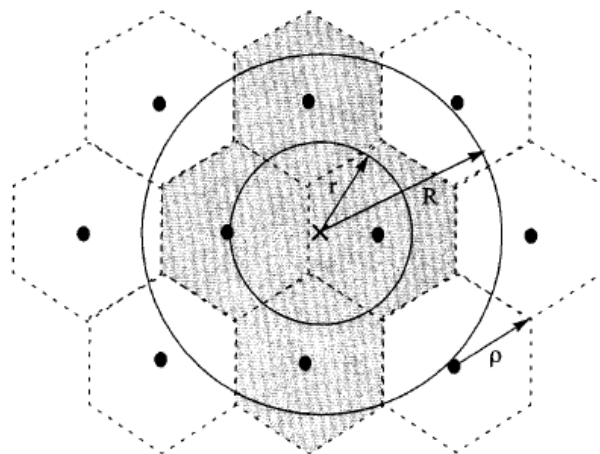
Trong xây dựng cơ sở dữ liệu, mỗi màu vùng được gán đối với điểm lưới gần nhất của nó. Hình 2.9 chỉ ra một ví dụ về cấu trúc dữ liệu của một nút chỉ số. ID vùng ảnh là một nhãn số nguyên duy nhất (định danh mỗi vùng trong cơ sở dữ liệu). Các mục trong mỗi nút chỉ số được

lưu trữ bởi các số ID này. Sau khi nút chỉ số của mục mới được tìm thấy, ID vùng của nó được so sánh đối với danh sách các ID vùng được lưu trữ trong nút và được chèn vào vị trí đúng trong danh sách được lưu trữ. Cơ chế đánh chỉ số được đề xuất cho phép cơ sở dữ liệu là động (nghĩa là các phép chèn và xóa của các mục cơ sở dữ liệu là đơn giản và không cần xây dựng lại toàn bộ cấu trúc đánh chỉ số của cơ sở dữ liệu).

entry	image region ID	3D color vector	percentage of the color in the region
1	ID ₁	c ₁	p ₁
2	ID ₂	c ₂	p ₂
3	ID ₃	c ₃	p ₃
...

Hình 2.8: Độ chính xác và thu hồi cho $K = 50$

Tuy nhiên, trong quá trình tìm kiếm, ngoài điểm lưới gần nhất đối với màu truy vấn, các điểm lưới gần khác nhau cũng được xem xét bởi vì màu truy vấn có thể nằm gần các điểm biên của các tế bào lưới. Hình 2.10 minh họa trường hợp hai chiều này, nơi bán kính tìm kiếm mong muốn là phạm vi tìm kiếm truy vấn và bán kính tìm kiếm thực tế là khoảng cách tìm kiếm cực tiểu cho các điểm lưới sao cho hình cầu mong muốn có bán kính được phủ. Ký hiệu bán kính cực tiểu của một hình cầu có thể phủ một tế bào Voronoi, như được chỉ ra trong hình 2.10. Lưu ý rằng: $R = r + p$.



Hình 2.9: Cơ chế tìm kiếm trong mặt phẳng hai chiều

Do $R > r$ và một nút chỉ số chứa tất cả các mục trong tế bào Voronoi của nó, một phần của không gian tìm kiếm không chứa bất cứ đối sánh liên quan nào. Chẳng hạn, với trường hợp hai chiều đã chỉ ra trong hình 2.10 không gian tìm kiếm thực tế gồm tất cả các diện tích màu đen. Với một phạm vi tìm kiếm r được cho, giá trị của p trong thiết kế lưới là quan trọng đối với hiệu năng tra cứu. Một giá trị nhỏ của p có nghĩa rằng không gian tìm kiếm thực tế chỉ rộng hơn không gian tìm kiếm mong muốn một chút và do đó hầu hết các nút chỉ số được truy cập là liên quan. Tuy nhiên, có một thỏa hiệp bởi vì số các nút chỉ số tăng khi p giảm và bản thân nút chỉ số trở nên kém hiệu quả. Số các nút chỉ số đã truy cập trên màu truy vẫn là $O(R^3/p^3)$ và không phụ thuộc trực tiếp vào cơ sở dữ liệu.

Cấu trúc lưới cũng có thể có nhiều hơn một lớp biểu diễn. Nếu một tế bào Voronoi dày đặc với quá nhiều mục, không gian của nó có thể được chia tiếp thành tập các tế bào con. Các kết quả này trong một cấu trúc lưới phân cấp. Các cấu trúc lưới phân cấp được sử dụng trong mã ảnh dựa vào VQ [12]. Có nhiều phạm vi tìm kiếm cố định, một cho mỗi mức. Một thiết kế cấu trúc phân cấp cẩn thận có thể cải tiến hiệu năng tra cứu. Tuy nhiên, trong các thực nghiệm chúng tôi chỉ sử dụng biểu diễn một lớp.

Thủ tục tìm kiếm

Thủ tục tìm kiếm hoàn chỉnh gồm các bước sau:

Bước 1: Với mỗi màu truy vấn, tìm các vùng đối sánh mà chứa các màu tương tự bởi sử dụng cấu trúc đánh chỉ số lưới. Để loại bỏ nhanh một số đối sánh sai, một ngưỡng T_p được đặt cho sự khác biệt giữa phần trăm truy vấn p_i và phần trăm được tra cứu q_j . Một vùng đối sánh được loại bỏ nếu điều kiện sau không được thỏa mãn:

$$|p_i - q_j| < T_p \quad (2-20)$$

Bước 2: Gộp các kết quả đối sánh từ tất cả các màu truy vấn và loại tất cả các đối sánh sai. Các vùng mà thỏa mãn hai điều kiện sau được xem xét như các ứng viên tra cứu cuối cùng:

$$\sum_i p_i \geq T_t \quad \text{và} \quad \sum_j q_j \geq T_t \quad (2-21)$$

Và ở đây i và j đánh chỉ số các màu được đối sánh. Các mục được đối sánh từng phần trong cơ sở dữ liệu được loại nhanh. T_t được đặt là 0.6 trong các thực nghiệm.

Bước 3: Tính các khoảng cách giữa các tra cứu và truy vấn và phân hạng chúng theo thứ tự. Với chỉ số và độ đo khoảng cách là phù hợp, bán kính tìm kiếm r nên bằng với T_d , khoảng cách cực đại cho hai màu được xét tương tự.

Bước 4: Nếu một truy vấn vùng được sử dụng, tất cả các ứng viên đối sánh với các khoảng cách nhỏ hơn phạm vi cho trước được trả về. Nếu một truy vấn N lân cận gần nhất được sử dụng, N ứng viên ở trên được trả về.

Bước 1 và bước 2 loại nhanh một lượng lớn các đối sánh sai. Bước 3 bao gồm các tính toán phức tạp hơn, nhưng chỉ cho các tra cứu cuối cùng trong cơ sở dữ liệu. Toàn bộ, độ phức tạp tính toán của thủ tục tìm kiếm là thấp. Lưu ý rằng số các truy cập đĩa ngẫu nhiên là $O(mn)$, ở đây n là số các màu trong truy vấn (tiêu biểu khoảng 3-5), m là số các nút đã truy cập trên màu truy vấn. Lưu ý rằng lược đồ màu truyền thống là khó để đánh chỉ số khi độ phức tạp của đánh chỉ số phát triển nhanh với số chiều. Ngược lại, mô tả màu trội được đánh chỉ số trong không gian màu ba chiều.

2.5 Kết luận

Trong chương này, ngoài đã trình bày được các không gian màu, lược đồ màu, đo khoảng cách giữa các lược đồ màu, đặc biệt em đã trình bày thuật toán tra cứu ảnh sử dụng biểu diễn màu hiệu quả để tìm kiếm và phân vùng ảnh cho chương trình thực nghiệm.

Chương 3: XÂY DỰNG HỆ THỐNG TRA CỨU ẢNH DỰA VÀO NỘI DUNG SỬ DỤNG BIỂU DIỄN MÀU HIỆU QUẢ

3.1 Giới thiệu bài toán tra cứu ảnh dựa vào nội dung

Trong thực tế, ảnh tra cứu thường rất đa dạng và phong phú. Số lượng ảnh trên internet là rất lớn. Khi người sử dụng muốn tìm một bộ ảnh tương tự với một ảnh đã biết (ảnh truy vấn). Vấn đề là làm thế nào xác định một tập các ảnh tương tự với một ảnh mà người sử dụng đã biết là không dễ dàng.

Ngày nay với sự phát triển mạnh mẽ của công nghệ thông tin người ta đã có thể thu nhập được bộ ảnh theo mục đích nào đó. Tuy nhiên khi cơ sở dữ liệu trên internet có kích cỡ cực lớn thì việc thu nhập này trở nên cực kỳ khó khăn. Để giải quyết vấn đề này chúng ta có thể sử dụng công nghệ tra cứu ảnh dựa vào nội dung để tìm ra những ảnh tương tự với ảnh truy vấn nhất. Các kết quả tìm được sẽ giúp người sử dụng nhanh chóng thực hiện được mục tiêu của mình.

Thông thường ảnh tra cứu theo nội dung có thể nhiều vùng và màu sắc khá phong phú. Chúng ta cũng nhận thấy rằng để so sánh và phân loại ảnh này thì điểm quan trọng nhất là màu sắc của vùng trong ảnh.

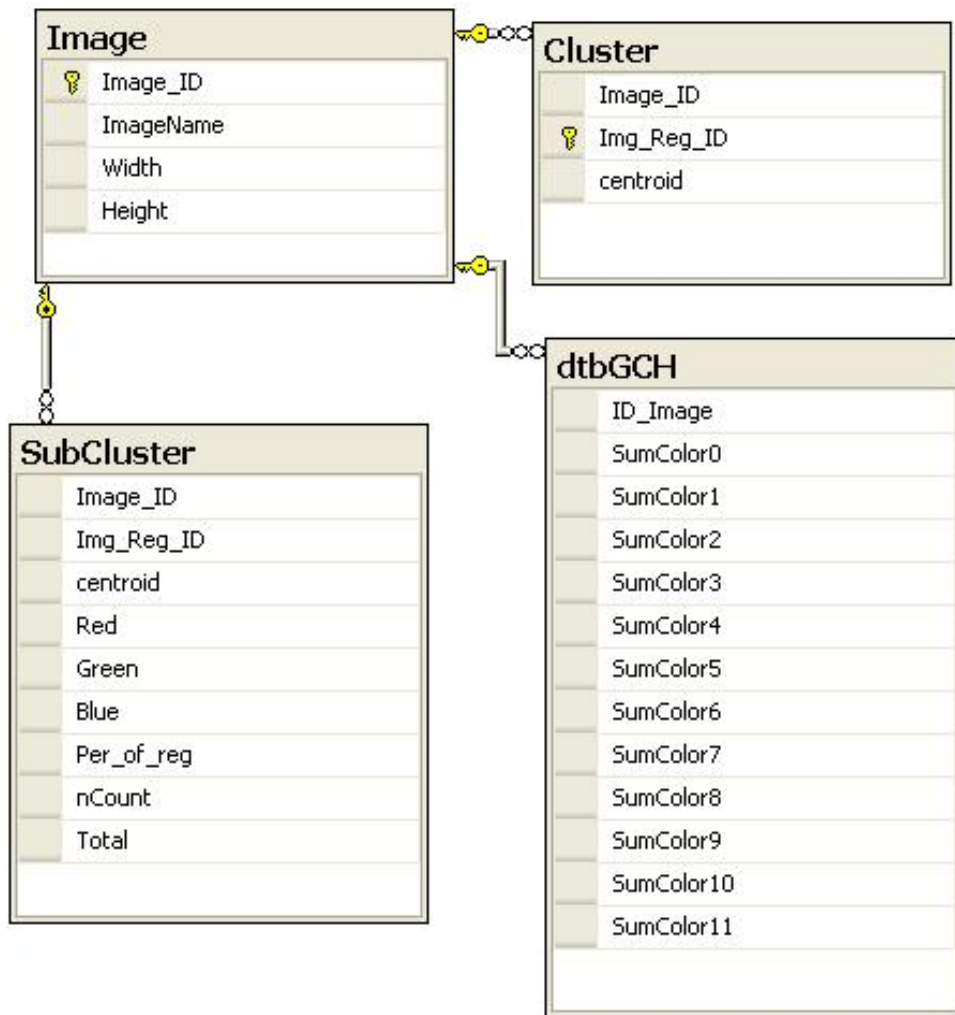
Như vậy, nhiệm vụ của bài toán là xây dựng hệ thống tra cứu ảnh có các chức năng sau:

- Khi người sử dụng sưu tầm được một hoặc một số ảnh, hệ thống cho phép đưa một hoặc một tập ảnh này vào lưu trữ một cách thuận lợi.
- Khi người sử dụng cung cấp một ảnh mẫu cần tra cứu. Hệ thống có nhiệm vụ tìm kiếm trong một cơ sở dữ liệu ảnh đã có và đưa ra danh sách các ảnh tương tự như ảnh mẫu theo thứ tự ảnh nào tương tự với ảnh mẫu hơn thì được phân hạng trước.

3.2 Mô tả chương trình

Hệ thống được thiết kế và thực hiện trên hệ điều hành Windows, sử dụng ngôn ngữ lập trình C# trên môi trường Visual Studio 2010, hệ quản trị cơ sở dữ liệu SQL Server 2008.

Biểu đồ liên kết các bảng trong cơ sở dữ liệu



Thông tin các trường dữ liệu trong bảng cơ sở dữ liệu

+ Bảng Image

Tên trường	Kiểu dữ liệu	Mô tả
Image_ID	Int(Identity)	Lưu Id của ảnh
ImageName	Nvarchar(500)	Lưu tên ảnh
Width	Int	Độ rộng của ảnh
Height	Int	Độ cao của ảnh

+ Bảng Cluster

ĐỒ ÁN TỐT NGHIỆP

Tên trường	Kiểu dữ liệu	Mô tả
Image_ID	Int	Lưu Id của ảnh
Img_reg_ID	Int(Identity)	Lưu id của centroid ảnh
Centroid	Int	Giá trị của các centroid

+ Bảng Subcluster

Tên trường	Kiểu dữ liệu	Mô tả
Image_ID	Int	Lưu Id của ảnh
Img_reg_ID	Int	Lưu id của centroid ảnh
Centroid	Int	Giá trị của các centroid
Red	Numeric(24,12)	Giá trị Red của màu
Green	Numeric(24,12)	Giá trị Green của màu
Blue	Numeric(24,12)	Giá trị Blue của màu
Per_of_reg	Numeric(24,12)	% màu trong vùng
Ncount	Numeric(24,12)	Số pixel trong vùng có màu này
Total	Numeric(24,12)	Tổng số pixel của vùng

+ Bảng btbGCH

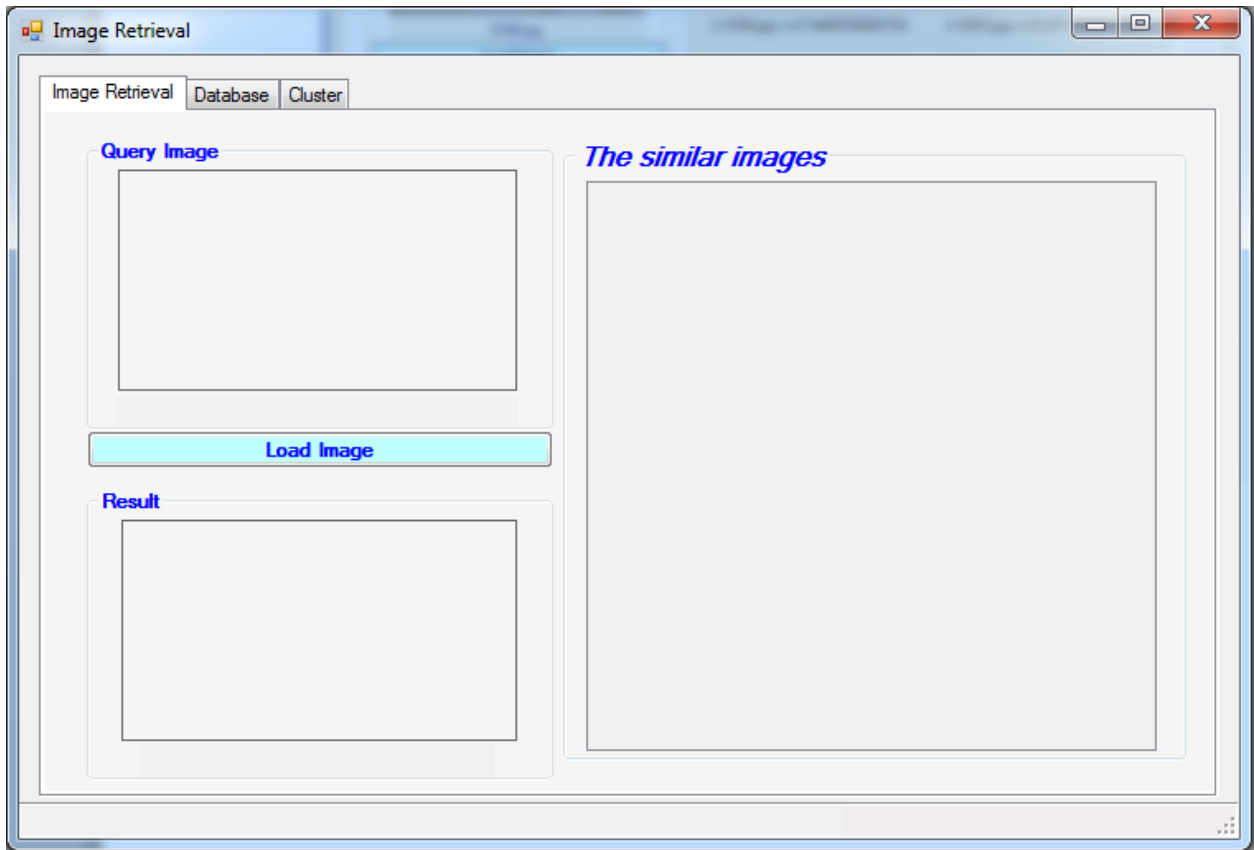
Tên trường	Kiểu dữ liệu	Mô tả
ID_Image	Int	Lưu Id của ảnh
SumColor0	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 0
SumColor1	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 1
SumColor2	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 2

ĐỒ ÁN TỐT NGHIỆP

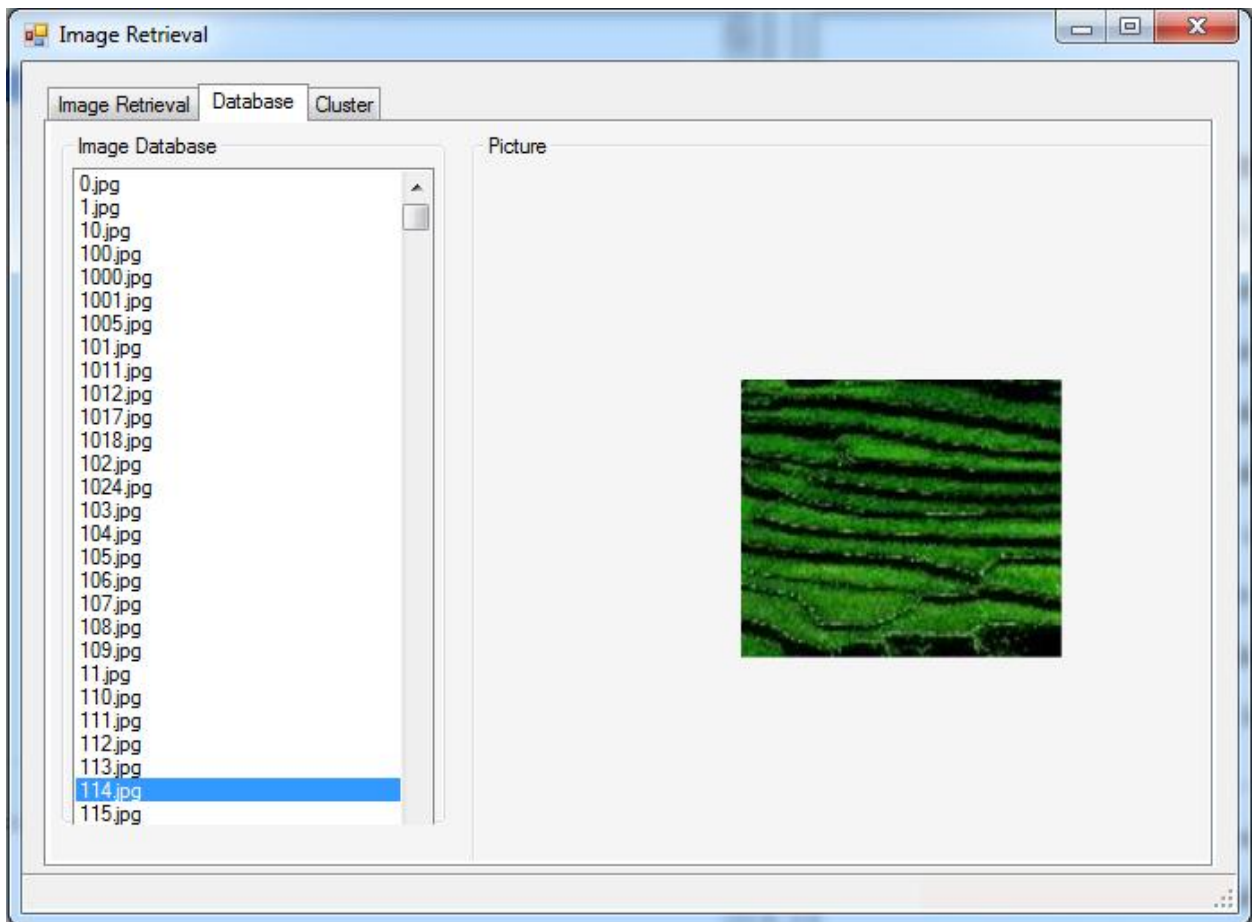
SumColor3	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 3
SumColor4	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 4
SumColor5	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 5
SumColor6	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 6
SumColor7	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 7
SumColor8	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 8
SumColor9	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 9
SumColor10	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 10
SumColor11	Int	Tổng số pixel ảnh có giá trị màu tương ứng với màu 11

Giao diện chương trình

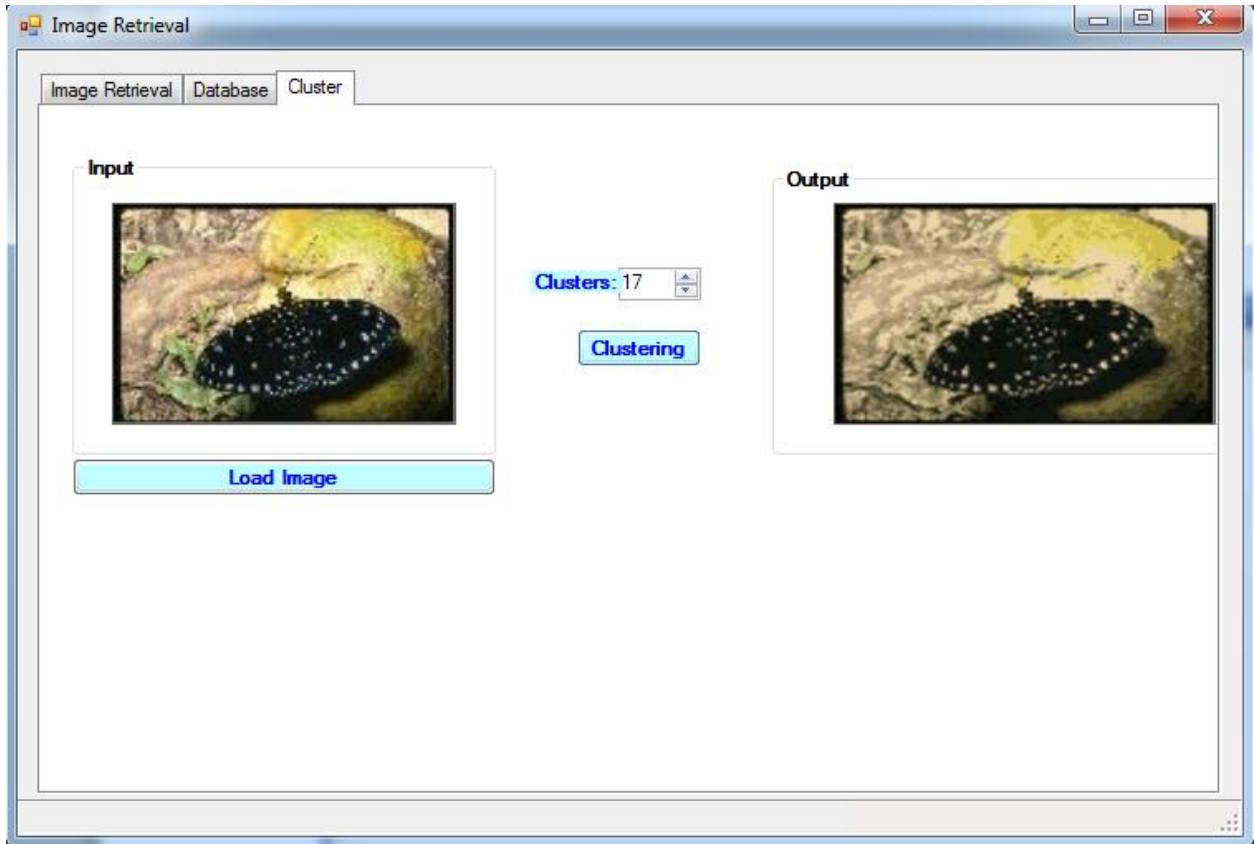
+ Giao diện tra cứu ảnh:



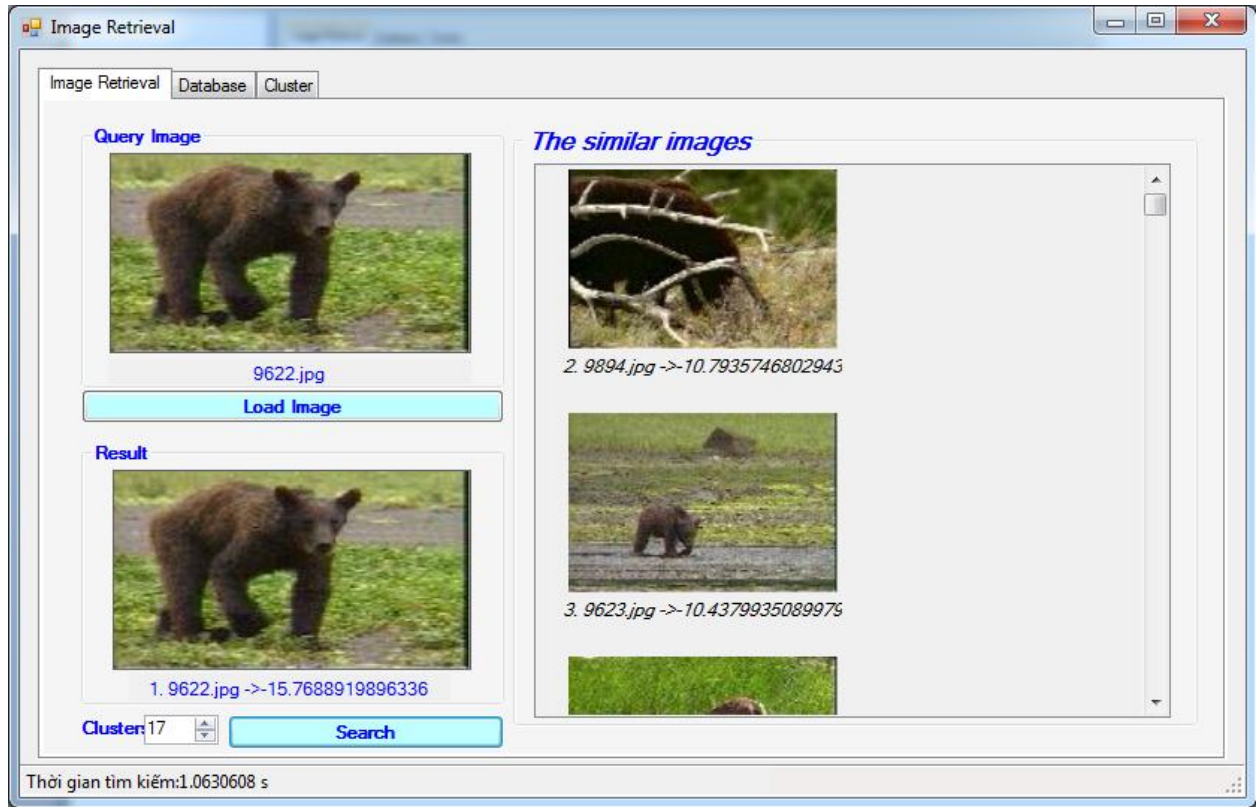
+ Giao diện cơ sở dữ liệu ảnh:



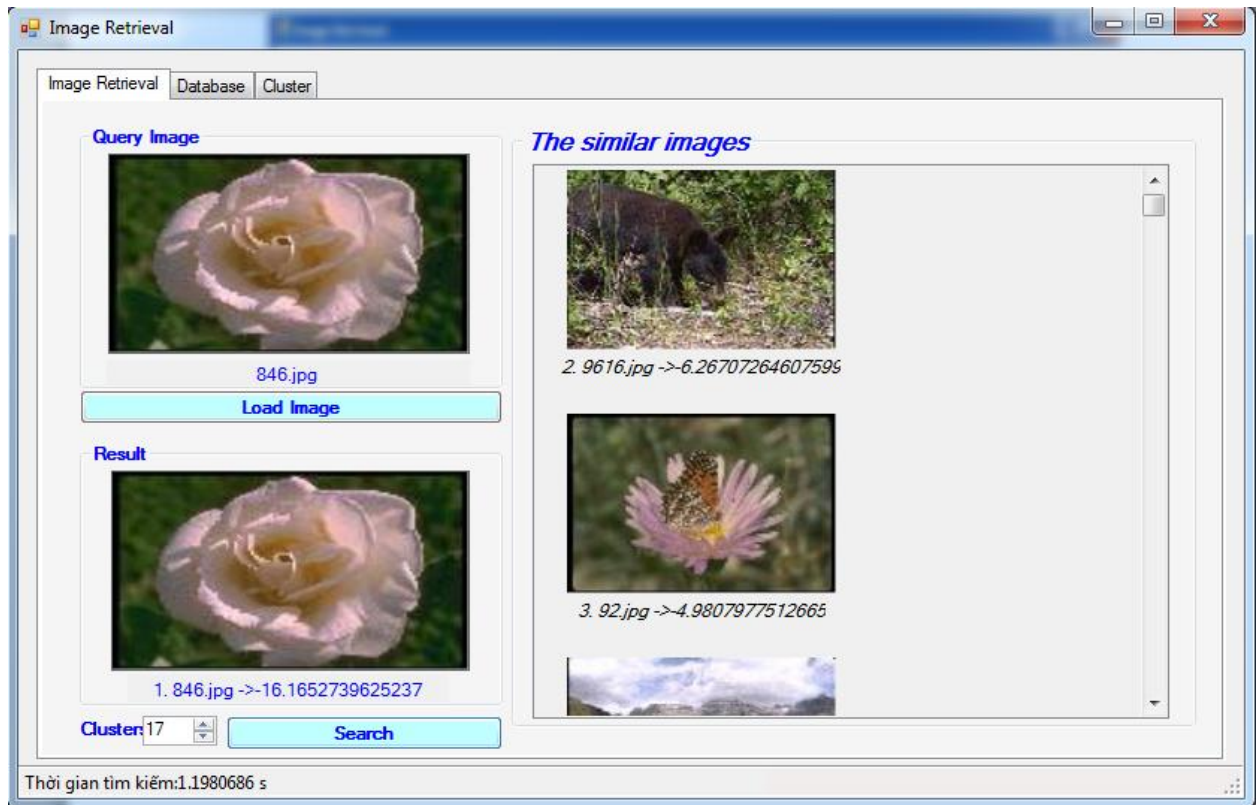
+ Đưa vào một ảnh -> kết quả trả về là ảnh đã phân vùng:



+ Tra cứu với ảnh động vật:



+ Tra cứu với ảnh hoa:



3.3 Kết luận

Trong chương này em đã đưa ra chương trình thực nghiệm, với tập ảnh dữ liệu đưa vào gồm ảnh động vật, hoa. Khi tra cứu ảnh theo từng chủ đề, ảnh sẽ được sắp xếp theo độ thứ tự chính xác gần nhất.

Chương 4: KẾT LUẬN

Thưa các thầy cô trong hội đồng, đến đây em đã hoàn thành đồ án tốt nghiệp với đề tài “Nghiên cứu phương pháp tra cứu ảnh dựa vào nội dung biểu diễn màu hiệu quả”.

Các vấn đề đã được giải quyết trong đồ án:

- Tìm hiểu về xử lý ảnh
- Giới thiệu tổng quan về tra cứu ảnh theo nội dung
- Nghiên cứu và tiến hành cài đặt các phương pháp tra cứu ảnh dựa vào nội dung sử dụng biểu diễn màu hiệu quả.

Các vấn đề còn tồn tại:

Do thời gian tìm hiểu về đề tài chưa được nhiều và những hạn chế về khả năng lập trình đồ họa nên một số mục tiêu đặt ra từ khi bắt tay nghiên cứu em chưa thực hiện được trong chương trình chạy thử này, bao gồm:

- Độ chính xác chưa cao
- Chưa sưu tầm được tập ảnh đủ lớn và mang tính đại diện cho từng chủ đề cao

Hướng phát triển:

Trong thời gian tới em sẽ hoàn thiện và tối ưu thuật toán nhằm nâng cao độ chính xác, đồng thời tăng tốc độ truy vấn. Hoàn thiện thêm về cơ sở dữ liệu ảnh, nhằm làm cho cơ sở dữ liệu ảnh phong phú hơn.

TÀI LIỆU THAM KHẢO

- [1] W. Y. Ma and B. S. Manjunath, “Edge flow: A framework of boundary detection and image segmentation,” in *Proc. IEEE Conf. Computer Vision Pattern Recognition*, 1997, pp. 744-749.
- [2] Y. Deng, C. Kenney, M. S. Moore, and B. S. Manjunath, “Peer group filtering and perceptual color quantization,” in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 4, 1999, pp. 21- 24.
- [3] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1970.
- [4] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer – Verlag, 1993.
- [5] N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger, “The R* - tree, an efficient and robust access method for points and rectangles”, in *Proc. ACM SIGMOD Int. Conf. Management Data*, 1990, pp. 322-331.
- [6] D. A. White and R. Jain, “Similarity indexing with the SS-tree” in *Proc. Int. Conf. Data Engineering*, 1996, pp. 516-523.
- [7] H. Samet, *The Design and Analysis of Spatial Data Structures*. Reading, MA: Addison-Wesley, 1990.
- [8] D. Mukherjee and S. K. Mitra, “Vector set-partitioning with successive refinement Voronoi lattice VQ for embedded wavelet image coding,” in *Proc. IEEE Int. Conf. Image Processing*, 1998.