MỤC LỤC

MỤC LỤC.....
LỜI CẢM ƠN.....
LỜI MỞ ĐẦU.....
CHƯƠNG 1: TỔNG QUAN VỀ TRA CỨU ẢNH DỰA TRÊN NỘI DUNG.....**1**
1.1 Giới thiệu 1
1.2 Tra cứu thông tin thị giác.....1
1.2.1 Những thành phần của một hệ thống tra cứu ảnh 2
1.2.2 Công nghệ tự động trích chọn metadata 3
1.2.3 Giao diện để lấy yêu cầu truy vấn của người sử dụng.....3
1.2.4 Phương pháp để so sánh độ tương tự giữa các ảnh.....4
1.2.5 Công nghệ tạo chỉ số và lưu trữ dữ liệu hiệu quả 4
1.3 Đặc điểm của tra cứu ảnh.....5
1.4 Những ứng dụng cơ bản của tra cứu ảnh 7
1.5 Tra cứu ảnh dựa trên nội dung.....7
1.5.1 Những phương pháp quản lý dữ liệu ảnh truyền thống 8
1.5.2 Các chức năng của hệ thống tra cứu ảnh dựa trên nội dung 9
1.5.3 Trích chọn những đặc điểm 11
1.5.4 Những khoảng cách tương tự.....13
1.6 Các phương pháp tra cứu ảnh dựa trên nội dung.....16
1.6.1 Tra cứu ảnh dựa trên màu sắc 16
1.6.2 Tra cứu ảnh dựa trên kết cấu.....16
1.6.3 Tra cứu ảnh dựa trên hình dạng 17
1.6.4 Tra cứu ảnh bởi các đặc điểm khác.....18
CHƯƠNG 2: TRA CỨU ẢNH DỰA TRÊN KẾT CẤU.....**19**
2.1 Giới thiệu 19
2.2 Kết cấu theo nhận thức của con người.....19
2.3 Phương pháp cho phân tích kết cấu 21
2.3.1 Tiêu chuẩn kết cấu thống kê 21

2.3.2 Mô hình kết cấu ước lượng (Stochastic) 21
2.3.3 Tiêu chuẩn kết cấu cấu trúc 21
2.3.4 Những đặc điểm kết cấu 22
2.4 Những phương pháp phân tích kết cấu 23
2.4.1 Phương pháp Gause Markov Random Field (GMRF) 23
2.4.2 Phương pháp Gray-Level Co-occurrence Matrices 23
2.4.3 Phương pháp Gray-Level Difference (GLD) 25
2.4.4 Phương pháp phân bố kết cấu (Texture spectrum) 25
2.5 Mô hình hình dạng chung dùng trong kết cấu (GS-Gross Shape) 27
2.5.1 Phương pháp Autocorrelation 27
2.5.2 Phương pháp Tamura 28
2.6 Những phương pháp Primitive 29
2.6.1 Phương pháp Primitive đầu tiên (Early primitive) 30
2.6.2 Phương pháp Gabor 30
CHƯƠNG 3: PHƯƠNG PHÁP PHÂN TÍCH KẾT CẤU MÀU 32
3.1 Phương pháp Color auto-corrlegram 32
3.1.1 Giới thiệu: 32
3.1.2 Thước đo khoảng cách điểm ảnh 33
3.1.3 Những đặc điểm thước đo khoảng cách 33
3.2 Phương pháp ma trận đồng mức xám Co-occurrence Matrix 34
3.2.1 Mô tả những đặc điểm 34
3.2.2 Thực hiện cải tiến việc tính toán ma trận Co-occerrence 36
CHƯƠNG 4: CÀI ĐẶT CHƯƠNG TRÌNH THỬ NGHIỆM 38
4.1 Môi trường thực nghiệm 38
4.2 Kết quả thử nghiệm 38
4.2.1 Giao diện chương trình 38
4.2.2 Chọn ảnh cần tìm kiếm 39
4.2.3 Kết quả tìm kiếm ảnh hoàn thiện 39
KẾT LUẬN 40
TÀI LIỆU THAM KHẢO 41

LỜI CẢM ƠN

Trước tiên em xin gửi lời cảm ơn chân thành tới thầy giáo hướng dẫn Ngô Trường Giang, người đã định hướng nghiên cứu và tận tình chỉ bảo, giúp đỡ em trong quá trình thực tập và làm đồ án, giúp em hoàn thành báo cáo thực tập đúng kế hoạch.

Em xin chân thành cảm ơn các thầy cô trong Khoa, trong Trường ĐHDL Hải Phòng đã tận tình giảng dạy, truyền đạt những kiến thức và kinh nghiệm vô cùng quý báu trong những năm học vừa qua.

Cho em gửi lời cảm ơn chân thành đến trường ĐH Công Nghiệp TP Hồ Chí Minh đào tạo từ xa Trường Trung Cấp Nghề Việt Đức đã giảng dạy truyền đạt kiến thức giúp đỡ em trong 3 năm học Cao Đẳng.

Sau cùng là lòng biết ơn sâu sắc đến bố mẹ, anh, chị, bạn bè đã luôn động viên, giúp đỡ, ủng hộ trong suốt những tháng năm ngồi trên ghế giảng đường.

Hà Nội, ngày 25 tháng 10 năm 2010

Sinh viên thực hiện

Đồng Nam Hà

LỜI MỞ ĐẦU

Sự mở rộng của đa phương tiện (multimedia), cùng với khối lượng hình ảnh, phim lớn, sự phát triển của những xa lộ thông tin 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 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 nhà khoa học 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 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 phụ 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 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 vào nội dung, nhằm vào khía cạnh nhận thức thông tin. 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. 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.

CHƯƠNG 1: TỔNG QUAN VỀ TRA CỨU ẢNH DỰA TRÊN NỘI DUNG

1.1 Giới thiệu

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 đó 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 tra cứu ảnh dựa trên nội dung cơ sở dữ liệu là rất cần thiết.

Tra cứu ảnh theo nội dung chính thức xuất hiện năm 1992, đánh dấu bằng Hội thảo về các hệ thống quản lý thông tin trực quan của Quỹ Khoa học Quốc gia của Hoa Kỳ.

Tra cứu ảnh theo nội dung là một quá trình tìm kiếm trong một cơ sở dữ liệu ảnh những ảnh nào thỏa mãn một yếu cầu nào đó. Những tìm kiếm đặc thù tiêu biểu cho hệ thống dạng này là: QBIC, VIR Image, Engine, VisualSEEK, NeTrA, MARS, Viper...

Tra cứu ảnh được sử dụng trong nhiều lĩnh vực khác nhau: y tế, khoa học, hình sự, bảo tồn, ngân hàng... Tra cứu ảnh nhận được sự quan tâm của nhiều nhà nghiên cứu trong việc tìm kiếm. Wikipedia: Hệ thống tra cứu ảnh của một hệ thống máy tính sử dụng để duyệt, tìm kiếm và tra cứu ảnh từ một cơ sở dữ liệu ảnh số lớn.

1.2 Tra cứu thông tin thị giác

Thuật ngữ "Tra cứu thông tin" được đưa ra vào năm 1952 và đã dành được sự quan tâm đặc biệt của hội các nhà nghiên cứu từ năm 1961 [Jones and Willet, 1977]. Chúng ta có thể dễ dàng mô tả một hệ thống tra cứu thông tin như là một hệ thống lưu trữ và tra cứu thông tin. Như là một hệ thống, vì vậy nó gồm một tập hợp các thành phần tương tác lẫn nhau, mỗi thành phần được thiết kế cho một chức năng riêng, có mục đích riêng và tất các các thành phần này có quan hệ với nhau để đạt được mục đích là tìm kiếm thông tin trong một phạm vi nào đó.

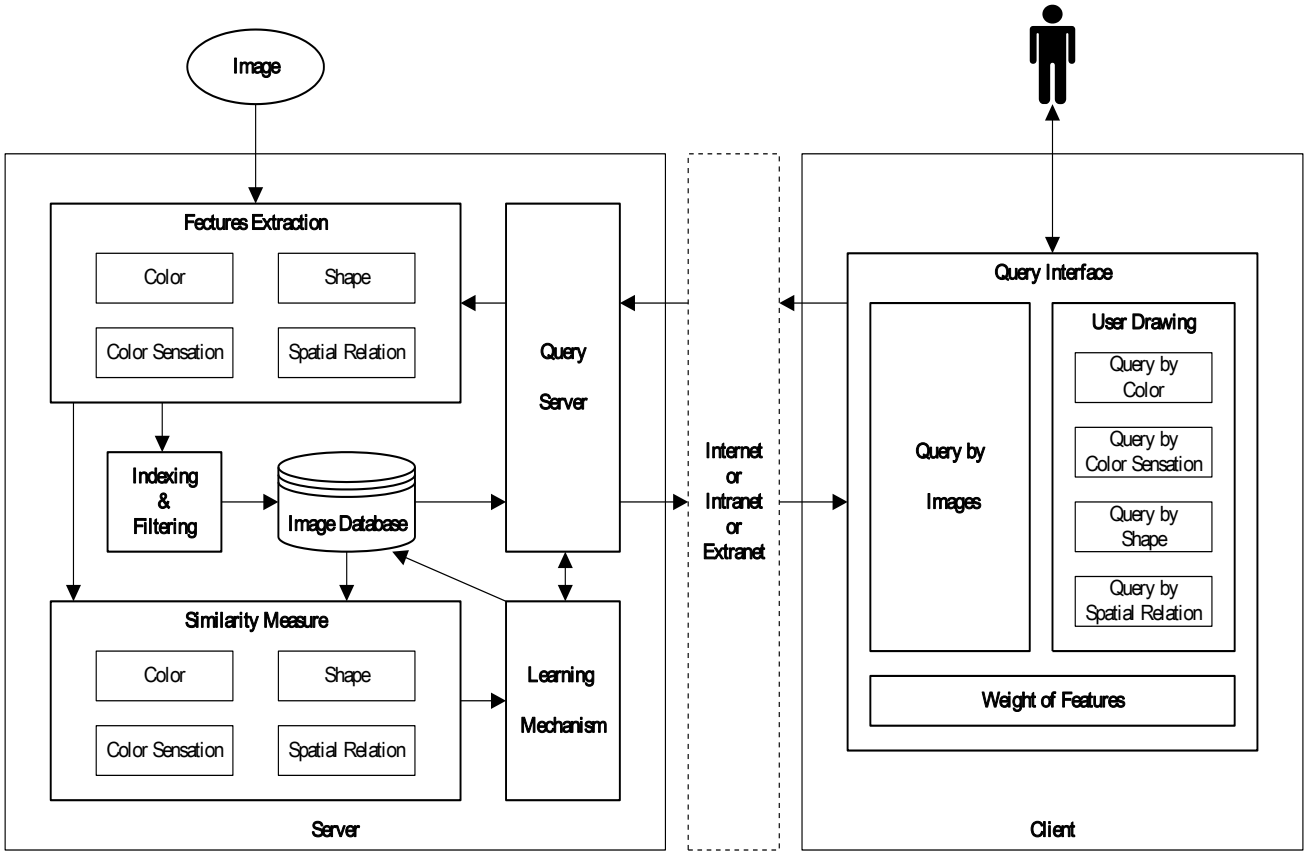
Trước đây, tra cứu thông tin có nghĩa là tra cứu thông tin theo kết cấu, nhưng định nghĩa trên vẫn được giữ khi ứng dụng vào việc tra cứu thông tin thị giác (VIR-Visual Infomation Retrieval). Mặc dù vậy vẫn có sự phân biệt giữa kiểu của thông tin

và nét tự nhiên của tra cứu của văn bản và các đối tương trực quan. Thông tin kết cấu là tuyến tính trong khi ảnh là hai chiều và video là ba chiều. Một cách chính xác hơn là văn bản được cung cấp với một điểm bắt đầu và kết thúc vốn có và với một chuỗi phân tích cú pháp tự nhiên. Chiến lược phân tích cú pháp tự nhiên như vậy không thích hợp với ảnh và video.

Có hai phương pháp chung để giải bài toán tra cứu thông tin thị giác (trực quan) dựa trên những thông tin trực quan đó là: Phương pháp dựa trên những thuộc tính và phương pháp dựa trên những đặc điểm. Phương pháp dựa trên thuộc tính dựa vào tra cứu thông tin kết cấu truyền thống và những phương pháp quản lý cơ sở dữ liệu dựa trên lý trí cũng như là sự can thiệp của con người để trích chọn metadata về đối tượng trực quan và sự chú thích kết cấu. Thật không may là việc phân tích kết cấu đều mất nhiều thời gian và tốn nhiều công sức. Hơn nữa lời chú thích phụ thuộc rất nhiều vào cảm nhận chủ quan của con người, mà sự cảm nhận chủ quan và sự giải thích mơ hồ chính là nguyên nhân của sự ghép đôi không cân xứng trong quá trình xử lý. Vấn đề truy cập ảnh và video dựa trên text đã thúc đẩy quan tâm đến sự phát triển những giải pháp dựa trên đặc điểm. Đó là thay sự giải thích thủ công bằng những từ khoá dựa trên văn bản, ảnh có thể được trích chọn ra bằng cách sử dụng một số đặc điểm thị giác như là màu sắc, kết cấu, hình dạng và được đánh chỉ số dựa trên những đặc điểm thị giác này. Phương pháp này chủ yếu dựa trên kết của của đồ hoạ máy tính. Trong bài luận văn này, sẽ tập trung vào một số đặc điểm cụ thể đặc biệt là những đặc điểm dựa trên màu sắc và kết cấu ứng dụng cho tra cứu ảnh nói chung hoặc cho tra cứu ảnh dựa trên nội dung. Mặc dù vậy không có đặc điểm riêng lẻ nào tốt nhất có thể cho ra những kết quả chính xác trong bất kỳ một thiết lập chung nào. Một kết hợp thông thường của các đặc điểm là cần thiết để cung cấp những kết quả tra cứu thích đáng đối với ứng dụng tra cứu ảnh dựa trên nội dung.

1.2.1 Những thành phần của một hệ thống tra cứu ảnh

Một hệ thống tra cứu ảnh đòi hỏi các thành phần như hình 1.1:



Hình 1.1 Kiến trúc tổng quan của hệ thống tra cứu ảnh

1.2.2 Công nghệ tự động trích chọn metadata

Mỗi đặc điểm nguyên thủy của ảnh có định dạng đặc trưng của nó như biểu đồ màu được sử dụng rộng rãi để biểu thị đặc điểm màu sắc. Một ví dụ khác đặc điểm hình dạng có thể biểu thị bằng một tập các đoạn biên liền nhau. Với metadata thích hợp hệ thống tra cứu ảnh dựa trên nội dung có thể tra cứu ảnh bởi màu sắc, hình dạng, kết cấu và bởi sự kết hợp các đặc tính trên.

1.2.3 Giao diện để lấy yêu cầu truy vấn của người sử dụng

Trong bất kỳ một hệ thống tra cứu nào thì quá trình tra cứu đều bắt đầu từ một yêu cầu tra cứu. Vì vậy, nó là vấn đề cốt yếu để lấy yêu cầu truy vấn của người sử dụng một cách chính xác và dễ dàng. Tra cứu dựa trên text đã được sử dụng rộng rãi trong các hệ thống tra cứu, ví dụ tìm một quyển sách mà mình mong muốn với từ khóa nào đó trong thư viện. Với hệ thống tra cứu ảnh dựa trên nội dung thì quá trình tra cứu thường được thực hiện thông qua một hình ảnh mẫu được cung cấp bởi người sử dụng gọi là truy vấn bởi mẫu. Mặc dù vậy người sử dụng không thể luôn luôn đưa ra một ảnh mẫu cho hệ thống tra cứu. Hệ thống tra cứu ảnh dựa trên nội dung hiện nay giải quyết vấn đề này bằng cách đưa ra một giao diện để chỉ định hoặc chọn một số đặc điểm cơ bản cho việc cung cấp ảnh mẫu. Chẳng hạn như khi sử dụng hệ thống QBIC

của IBM người sử dụng có thể chỉ định truy vấn đặc điểm màu sắc bằng cách chọn ra số lượng thành phần RED, BLUE, GREEN liên quan hoặc là có thể lựa chọn màu sắc ảnh mong muốn từ bảng màu, đồng thời người sử dụng có thể chọn kết cấu mong muốn cho đặc điểm kết cấu và vẽ ra một phác họa cho truy vấn đặc điểm hình dạng.

1.2.4 Phương pháp để so sánh độ tương tự giữa các ảnh

Hệ thống tra cứu ảnh dựa trên nội dung yêu cầu những phương pháp dựa trên những đặc điểm nguyên thủy để so sánh độ tương tự giữa ảnh mẫu và tất cả những hình ảnh trong tập ảnh. Mặc dù vậy sự tương tự hoặc sự khác nhau gữa các ảnh không chỉ xác định theo một cách. Số lượng của ảnh tương tự sẽ thay đổi khi yêu cầu truy vấn thay đổi. Chẳng hạn trong trường hợp hai bức tranh, một là biển xanh mặt trời mọc và trường hợp khác là núi xanh với mặt trời mọc. Khi mặt trời được xem xét thì độ tương tự giữa hai ảnh này là cao nhưng nếu đối tượng quan tâm là biển xanh thì độ tương tự giữa hai ảnh này là thấp. Như vậy rất khó khăn để tìm ra phương pháp đo độ tương tự giữa hai hình ảnh một cách chính xác đối với tất cả các kiểu yêu cầu của truy vấn. Hay nói cách khác mỗi một phương pháp tra cứu sẽ có giới hạn của chính nó. Ví dụ rất khó cho công nghệ tra cứu dựa trên màu sắc để tìm ra điểm khác nhau giữa một ảnh là bầu trời màu xanh với một ảnh là mặt biển xanh. Vì vậy khi đánh giá một công nghệ tra cứu ảnh dựa trên nội dung cần phải biết rằng hiệu quả của công nghệ đó phụ thuộc vào kiểu yêu cầu tra cứu mà người dùng sử dụng.

1.2.5 Công nghệ tạo chỉ số và lưu trữ dữ liệu hiệu quả

Đối với một tập dữ liệu ảnh lớn thì không gian lưu trữ cho metadata là rất cần thiết. Một hệ thống tra cứu ảnh dựa trên nội dung phải có những công nghệ hiệu quả để quản lý metadata đồng thời phải có chuẩn để mô tả nó. Chuẩn MP7 đang là chuẩn quan trọng nhất để mô tả metadata cho cả dữ liệu ảnh và dữ liệu video. Khi một truy vấn được xử lý trên một cơ sở dữ liệu lớn, việc so sánh độ tương tự giữa ảnh truy vấn và tất cả các hình ảnh từng cặp là không thể thực hiện được bởi người dùng chỉ cần những ảnh có độ tương tự cao so với ảnh mẫu. Những chỉ số cấu trúc có thể giúp tránh được việc tìm kiếm tuần tự và cải thiện truy vấn một cách hiệu quả nên được sử dụng trong hệ thống tra cứu ảnh dựa trên nội dung. Hơn nữa với những cơ sở dữ liệu ảnh thường xuyên thay đổi thì chỉ số cấu trúc động là rất cần thiết. Khi nội dung của ảnh được thể hiện bởi các vector low dimension và khoảng cách giữa các ảnh được định nghĩa (chẳng hạn như khoảng không gian được tính toán bằng khoảng cách Euclidean) cây R và các thành phần của nó có thể được sử dụng để đánh chỉ số cho ảnh. Khi khoảng cách không được định nghĩa như không gian vector hoặc khi không gian vector là Hight dimension hoặc khi mà những gì chúng ta có chỉ là một hàm khoảng cách tức là

khoảng không metric thì những phương pháp để đánh chỉ số ảnh dựa trên hàm khoảng cách trong không gian metric là thích hợp.

1.3 Đặc điểm của tra cứu ảnh

Kiểu truy vấn nào thích hợp để người sử dụng đưa vào cơ sở dữ liệu ảnh? Để trả lời câu hỏi này một cách sâu sắc đòi hỏi phải có sự hiểu biết chi tiết về nhu cầu của người sử dụng: Tại sao những người dùng lại tìm kiếm ảnh, họ sử dụng chúng để làm gì, và họ đánh giá lợi ích của hình ảnh mà họ tìm được như thế nào. Cảm giác chung gợi ra rằng ảnh tĩnh được yêu cầu bởi một loạt các lý do gồm:

- Minh họa của những bài báo, truyền đạt thông tin hoặc cảm xúc khó mô tả bằng từ
- Hiển thị dữ liệu chi tiết cho phân tích
- Ghi lại dữ liệu thiết kế cho việc sử dụng sau này.

Truy cập tới một ảnh yêu cầu từ một kho dữ liệu ảnh có thể liên quan đến việc tìm kiếm ảnh mô tả kiểu đặc biệt của đối tượng hoặc đơn giản bao gồm kết cấu hoặc mầu đặc biệt. Vì vậy ảnh có rất nhiều thuộc tính có thể sử dụng cho việc tra cứu bao gồm:

- Sự kết hợp đặc biệt của đặc tính màu sắc, kết cấu, hình dạng (ví dụ những ngôi sao mà xanh)
- Sự xắp xếp của các kiểu riêng biệt của đối tượng (ví dụ những chiếc ghế xung quanh cái bàn)
- Sự mô tả kiểu sự kiện (Trận bóng đá)
- Tên cá nhân , vị trí, sự kiện(ví dụ Nữ hoàng đón nhận vương miện)
- Những cảm xúc chủ quan kết hợp với hình ảnh(ví dụ niềm hạnh phúc)
- Metadata giống như ai đã tạo ra ảnh, ở đâu, khi nào?

Mỗi kiểu truy vấn được liệt kê bên dưới miêu tả mức trừu tượng cao hơn mức trước đó. Và mỗi mức rất khó để trả lời mà không tham khảo thêm tri thức bên ngoài. Điều này dẫn đến kiểu truy vấn được phân làm ba mức tăng dần theo độ phức tạp.

Mức 1: Gồm tra cứu bởi những đặc điểm nguyên thủy như màu sắc, kết cấu, hình dạng hoặc những vị trí đặc biệt của những phần tử ảnh. Ví dụ "Tìm một bức tranh với một đối tượng dài , màu xám ở trên đỉnh góc trái", " Tìm ảnh chứa ngôi sao màu vàng được xếp thành một dãy" hoặc "Tìm bức tranh giống như thế này"... Mức tra cứu này sử dụng các đặc điểm từ chính những ảnh đó mà không cần tham khảo bất kỳ tri thức bên ngoài nào. Nó thường được ứng dụng trong lĩnh vực chuyên gia như việc đăng kí thương hiệu, nhận dạng các bộ sưu tập thiết kế.

Mức 2: Gồm những tra cứu bằng những đặc điểm biến đổi liên quan đến một số kết luận logic về sự đồng nhất của các đối tượng được mô tả trong ảnh. Nó có thể được chia thành:

a) Khôi phục các đối tượng theo kiểu nhất định(ví dụ tìm ảnh của chiếc xe buýt 2 tầng

b) Tra cứu những đối tượng đặc biệt hoặc người (ví dụ tìm bức ảnh của tháp Eiffel)

Để trả lời truy vấn ở mức này cần phải tham khảo một số tri thức bên ngoài, đặc biệt là truy vấn ở mức 2b. Trong ví dụ đầu tiên ở trên hiểu biết trước tiên cần thiết để xác định đối tượng là một chiếc xe buýt hơn là một chiếc xe tải. Trong ví dụ thứ 2 cần một tri thức về một cấu trúc có tên là "tháp Eiffel". Truy vấn mức này thường gặp hơn so với mức 1.

Mức 3: Gồm tra cứu bởi những thuộc tính trừu tượng liên quan đến một số lượng đáng kể suy luận ở mức cao về ý nghĩa và mục đích của đối tượng. Mức này có thể được chia làm:

a) Tra cứu tên gọi của những sự kiện hoặc kiểu của hành động (ví dụ Tìm bức tranh về điệu nhảy dân gian Scottish)

b) Tra cứu ảnh với những cảm xúc ("Tìm bức tranh mô tả sự đau khổ")

Những thành công trong trả lời truy vấn ở mức này đòi hỏi một vài sự tinh tế của công cụ dò tìm. Để tạo ra sự kết nối giữa nội dung ảnh và những khái niệm trừu tượng thì cần phải có những lập luận phức hợp và những ý kiến chủ quan để minh họa. Nhưng truy vấn ở mức độ này ít phổ biến hơn mức độ 2 và thường gặp ở báo chí và những thư viện nghệ thuật.

Chúng ta nhận thấy rằng sự phân lớp của các kiểu truy vấn này có thể có lợi cho việc minh họa điểm mạnh cũng như những hạn chế của các công nghệ tra cứu ảnh khác nhau. Khoảng cách đáng kể hiện nay nằm giữa mức 1 và mức 2. Một số tác giả đề cập

tới mức 2 và mức 3 như là tra cứu ảnh dựa trên ngữ nghĩa, và vì vậy khoảng cách giữa mức 1 và mức 2 là khoảng cách ngữ nghĩa.

1.4 Những ứng dụng cơ bản của tra cứu ảnh

Tra cứu ảnh được ứng dụng trong rất nhiều lĩnh vực, những lĩnh vực thành công bao gồm:

- Ngăn chặn tội phạm
- Quân sự
- Quản lý tài sản trí tuệ
- Thiết kế kiến trúc máy móc
- Thiết kế thời trang và nội thất
- Báo chí quảng cáo
- Chuẩn đoán y học
- Hệ thống thông tin địa lý
- Di sản văn hóa
- Giáo dục và đào tạo
- Giải trí
- Tìm kiếm trang web

1.5 Tra cứu ảnh dựa trên nội dung

Thuật ngữ tra cứu ảnh dựa trên nội dung đã được Kato sử dụng đầu tiên để mô tả những thí nghiệm của ông về lĩnh vực tra cứu tự động những hình ảnh từ một cơ sở dữ liệu dựa trên đặc điểm hình dạng và màu sắc. Từ đó, nó được sử dụng rộng rãi để mô tả quá trình tra cứu những hình ảnh mong muốn từ một tập hợp lớn hình ảnh dựa trên những đặc điểm về màu sắc, kết cấu và hình dạng, và những đặc điểm đó được trích rút một cách tự động từ chính những hình ảnh đó. Những đặc điểm sử dụng cho việc tra cứu có thể là những đặc điểm nguyên thủy hoặc là những đặc điểm ngữ nghĩa, tuy nhiên quá trình trích chọn chủ yếu phải được tự động . Tra cứu ảnh dựa trên việc gán từ khóa (manually assigned keywords) nhất định không phải là tra cứu ảnh dựa

trên nội dung bởi vì thuật ngữ được hiểu một cách chung chung ngay cả khi những từ khóa mô tả nội dung ảnh.

Một số phương pháp của tra cứu ảnh dựa trên nội dung được đưa ra từ lĩnh vực xử lý ảnh và đồ họa máy tính, và nó được lưu tâm bởi một số phương pháp như là một tập con của lĩnh vực đó. Nó khác với những lĩnh vực này chủ yếu thông qua việc nhấn mạnh vào tra cứu ảnh với những đặc điểm mong muốn từ một tập hình ảnh lớn. Những vấn đề nghiên cứu và phát triển về lĩnh vực tra cứu ảnh dựa trên nội dung bao gồm một số đặc điểm chính: đặc điểm màu sắc, kết cấu, hình dạng, ngữ nghĩa.

1.5.1 Những phương pháp quản lý dữ liệu ảnh truyền thống

Sự cần thiết của việc lưu trữ và tra cứu ảnh một cách có hiệu quả đã được những nhà quản lý tập hợp ảnh lớn như thư viện ảnh, bộ sưu tập thiết kế...quan tâm từ nhiều năm nay. Trong khi việc xác định một ảnh mong muốn từ một tập ảnh nhỏ hoàn toàn có thể thực hiện được một cách đơn giản bằng cách duyệt qua thì với một tập ảnh lớn gồm hàng ngàn các đề mục thì cần phải có một công nghệ hiệu quả hơn. Công nghệ thường được sử dụng là gán mô tả dữ liệu bằng hình thức từ khóa, tiêu đề hoặc là mã phân lớp đối với mỗi ảnh khi nó được đưa vào tập hợp ảnh lần đầu tiên và sau đó dùng những ký hiệu mô tả này như là khóa để tìm kiếm.

Nhiều thư viện ảnh dùng từ khóa như là hình thức tra cứu chính của họ. Sơ đồ chỉ số thường được phát triển trong một nhóm phản ánh nét tự nhiên của tập ảnh. Một ví dụ điển hình là hệ thống được phát triển bởi Getty Image [Bjarnestam,1998]. Từ điển chuyên đề của họ trên 10.000 từ khóa được phân thành chín nhóm nghĩa gồm: Địa lý, con người, hoạt động và khái niệm...Lĩnh vực hay sử dụng sơ đồ chỉ số nhất là nghệ thuật và từ điiển chuyên đề về nghệ thuật và kiến trúc (AAT), nó có nguồn gốc từ viện Rensselaer Polytechnic vào đầu những năm 80, và ngày nay nó được sử dụng trong các thư viện nghệ thuật trên khắp thế giới. AAT gồm 120.000 thuật ngữ cho việc mô tả đối tượng, kết cấu vật liệu hình ảnh, kiến trúc và các di sản văn hóa khác. Các thuật ngữ được sắp xếp thành hệ thống phân cấp khái niệm như thuộc tính vật lý, kiểu, giai đoạn, chất liệu...

Một số sơ đồ chỉ số dùng mã phân lớp nhiều hơn từ khóa để mô tả nội dung ảnh bởi vì chúng có thể đưa ra ngôn ngữ độc lập hơn và chỉ ra khái niệm hệ thống phân cấp rõ ràng hơn ví dụ như: CONCLASS của trường Đại học Leiden [Gordon, 1990]

Công nghệ đánh chỉ số ảnh hiện thời có nhiều điểm mạnh đặc biệt là chỉ số từ khóa, nó có thể được sử dụng để mô tả hầu hết các khía cạnh của nội dung ảnh. Nó có thể mở rộng một cách dễ dàng phù hợp với những khái niệm mới và có thể sử dụng để

mô tả nội dung ảnh ở những mức độ biến đổi phức tạp. Có rất nhiều phần mềm tra cứu văn bản có sẵn để tự động hóa quá trình tìm kiếm nhưng quá trình đánh chỉ số hướng dẫn (manual indexing) hoặc là từ khóa hoặc là mã phân lớp đều gặp phải hai hạn chế:

Thứ nhất: Nó vốn là công việc rất tỉ mỉ, thời gian đánh chỉ số được đưa ra cho ảnh tĩnh là khoảng 7-40 phút/1 ảnh [Eakins and Graham, 1999].

Thứ hai: Nó không xuất hiện một cách đáng tin cậy như là nghĩa của vấn đề tra cứu chẳng hạn những người khác nhau lấy những từ khác nhau để gán cho những ảnh bức ảnh giống nhau.

1.5.2 Các chức năng của hệ thống tra cứu ảnh dựa trên nội dung

Một hệ thống tra cứu ảnh dựa trên nội dung tiêu biểu không chỉ liên quan tới các nguồn thông tin trong những dạng khác nhau (ví dụ như văn bản, ảnh, video) mà còn liên quan đến nhu cầu của người sử dụng. Về cơ bản nó phân tích cả nội dung của nguồn thông tin cũng như truy vấn của người sử dụng và sau đó đối sánh chúng để tìm ra những tiêu chí có liên quan này. Những chức năng chính của một hệ thống bao gồm:

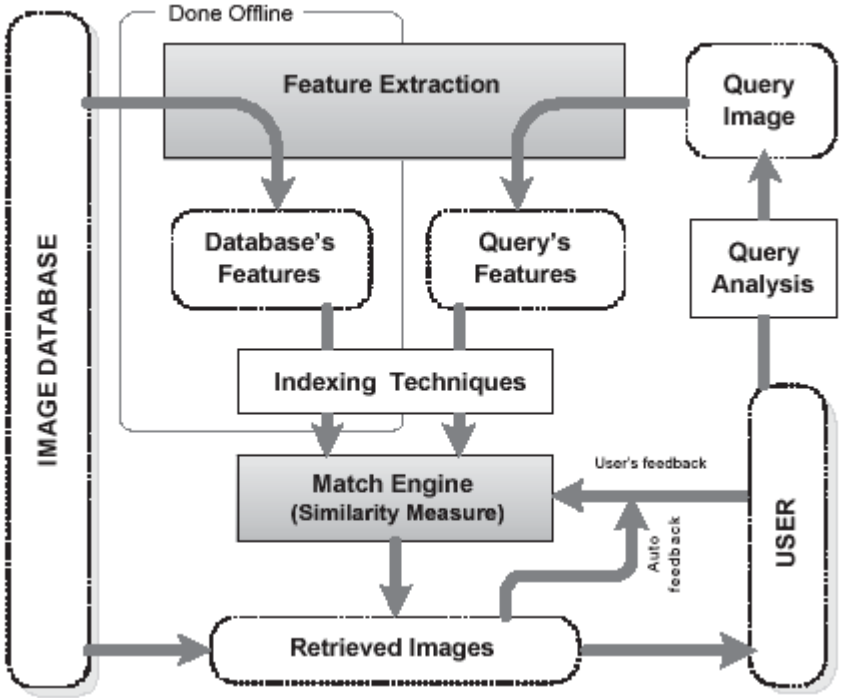
- Phân tích nội dung của nguồn thông tin và biểu diễn nội dung của các nguồn thông tin được phân tích phù hợp với sự đối sánh truy vấn của người sử dụng (không gian của thông tin nguồn được chuyển đổi thành không gian đặc điểm với mục đích đối sánh nhanh trong bước tiếp theo). Bước này thường là mất nhiều thời gian cho việc xử lý tuần tự các thông tin nguồn (ảnh) trong cơ sở dữ liệu. Nó chỉ phải làm một lần và có thể làm độc lập.

- Phân tích các truy vấn của người dùng và biểu diễn chúng thành các dạng phù hợp với việc đố sánh với cơ sở sữ liệu nguồn. Nhiệm vụ của bước này giống với bước trước nhưng chỉ được áp dụng với những ảnh truy vấn.

- Xác định chiến lược để đối sánh tìm kiếm truy vấn với thông tin được lưu trữ trong cơ sở dữ liệu. Bước này có thể thực hiện trực tuyến và thực hiện rất nhanh. Công nghệ đánh chỉ số hiện tại có thể được sử dụng để nhận dạng không gian đặc điểm để tăng tốc độ xử lý đối sánh.

- Tạo ra sự điều chỉnh cần thiết trong hệ thống (thường là bằng cách đối chiếu các tham số trong công nghệ đối sánh) dựa trên phản hồi từ người sử dụng hoặc những hình ảnh được tra cứu.

Rõ ràng là từ sự trình bày ở trên ta thấy một mặt hệ thống tra cứu ảnh dựa trên nội dung có các nguồn thông tin trực quan trong các dạng khác nhau, mặt khác lại có cả các yêu cầu của người sử dụng. Chúng được liên kết với nhau qua một loạt các công việc như được minh hoạ trong hình 1.2.



Hình 1.2: Các chức năng chính của hệ thống tra cứu ảnh dựa trên nội dung

Người sử dụng yêu cầu:

Có rất nhiều cách có thể đưa truy vấn trực quan. Một phương pháp truy vấn tốt là phương pháp tự nhiên với người sử dụng tức là cung cấp đầy đủ thông tin từ người sử dụng để trích chọn những kết quả có ý nghĩa. Những phương pháp dưới đây thường được sử dụng trong kỹ thuật tra cứu ảnh dựa trên nội dung:

Truy vấn bởi ví dụ (QBE-Query By Examble): Trong kiểu truy vấn này người sử dụng chỉ định một ảnh truy vấn gốc dựa trên cơ sở dữ liệu ảnh được tìm kiếm và so sánh. Ảnh truy vấn có thể là một ảnh chuẩn, một ảnh quét với độ phân giải thấp, hoặc người sử dụng vẽ bằng cách sử sụng công cụ vẽ đồ họa. Ưu điểm của kiểu hệ thống này là rất tự nhiên đối với người sử dụng để tra cứu ảnh trong cơ sở dữ liệu ảnh.

Truy vấn bởi đặc điểm (QBF- Query By Feature): Trong hệ thống kiểu này người dùng chỉ định câu hỏi bởi những đặc điểm chỉ định rõ ràng đó là những đặc điểm được quan tâm trong tìm kiếm. Ví dụ người dùng có thể truy vấn cơ sở dữ liệu ảnh bởi việc đưa ra một câu lệnh "Đưa ra tất cả những ảnh có góc bên trên trái chứa 25% điểm màu vàng". Truy vấn này được người dùng chỉ định bởi việc sử dụng công cụ giao diện đồ họa đặc biệt. Những người sử dụng chuyên nghiệp thì có thể tìm kiếm kiểu truy vấn tự nhiên này nhưng những người không chuyên thì rất khó. QBIC là một ví dụ về hệ thống tra cứu ảnh dựa trên nội dung mà người sử dụng truy vấn kiểu này.

Những truy vấn dựa trên thuộc tính (Attribute-based queries): Những truy vấn dựa trên thuộc tính sử dụng những chú giải kết cấu được trích chọn đầu tiên bởi sự nỗ lực của con người như khoá tra cứu. Mô tả kiểu này đòi hỏi phải có mức trìu tượng cao, cái rất khó đạt được mức độ tự động hoá hoàn toàn bởi vì ảnh gồm rất nhiều thông tin và rất khó có thể tổng kết bănngf một ít từ khoá. Trong khi phương pháp này nhìn chung là nhanh hơn và dễ thực thi hơn thì nó vốn có sự chủ quan và mơ hồ ở mức cao như đã giới thiệu phần trước.

Phương pháp truy vấn nào là tự nhiên nhất ? Với người sử dụng nói chung thì chắc chắn là truy vấn dựa trên những thuộc tính. Người sử dụng điển hình chắc chắn thích hỏi hệ thống tra cứu ảnh dựa trên nội dung bởi câu hỏi tự nhiên "Đưa ra cho tôi tất cả những ảnh từ hai năm trước", hoặc là "tìm tất cả các ảnh trên Internet mà có bàn phím của máy tính". Việc ánh xạ câu hỏi bằng ngôn ngữ tự nhiên này thành truy vấn trên cơ sở dữ liệu ảnh là vô cùng khó đối với việc sử dụng những phương pháp được tự động. Khả năng những máy tính thực hiện nhận dạng đối tượng tự động trên những ảnh vẫn đang là vấn đề nghiên cứu mở. Hầu hết những nghiên cứu cũng như các hệ thống mang tính thương mại đều tập trung xây dựng những hệ thống thực hiện tốt với những phương pháp QBE.

1.5.3 Trích chọn những đặc điểm

Trích chọn đặc điểm là cơ sở của tra cứu ảnh dựa trên nội dung. Theo một nghĩa rộng, những đặc điểm có thể gồm cả những đặc điểm dựa trên text (Từ khoá, những chú giải) và những đặc điểm trực quan (màu sắc, kết cấu, hình dạng). Trong phạm vi đặc điểm trực quan, những đặc điểm này lại được phân thành những đặc điểm mức thấp và những đặc điểm mức cao. Những đặc điểm mức thấp bao gồm: màu sắc, kết cấu, hình dạng trong khi đặc điểm mức cao được ứng dụng dựa trên những đặc điểm này ví dụ: mặt người, vân tay. Bởi nhận thức chủ quan, nên không tồn tại cách biểu diễn tốt nhất cho mỗi đặc điểm và vì vậy với mỗi đặc điểm có nhiều cách để biểu diễn mô tả những đặc điểm từ những ngữ cảnh khác nhau.

1.5.3.1 Màu sắc

Màu là đặc điểm trực quan đầu tiên và dễ nhất cho việc đánh chỉ số và tra cứu của ảnh và nó cũng là đặc điểm hay được sử dụng nhất trong lĩnh vực này.

Một ảnh màu điển hình được lấy từ camera số hoặc download từ Internet thường có ba kênh màu (ảnh xám chỉ có một kênh), những giá trị của dữ liệu ba chiều này từ ảnh màu có thể cho ta biết vị trí của những điểm ảnh này trong không gian màu. Những điểm ảnh có giá trị (1, 1, 1) cho những màu khác nhau trong những không gian màu khác nhau. Như vậy mô tả đầy đủ của một ảnh màu điển hình gồm thông tin không gian hai chiều với điểm ảnh trong vùng không gian này và dữ liệu màu ba chiều với điểm ảnh màu trong không gian mà chúng ta đang đề cập. Ở đây giả thiết không gian màu là cố định, bỏ qua thông tin không gian, thông tin màu trong ảnh có thể coi như là tín hiệu ba chiều đơn giản.

Nếu chúng ta coi thông tin màu của ảnh là tín hiệu một, hai, hoặc ba chiều đơn giản thì việc phân tích các tín hiệu sử dụng ước lượng mật độ xác xuất là một cách dễ nhất để mô tả thông tin màu của ảnh. Biểu đồ màu là một công cụ đơn giản nhất, những cách khác mô tả thông tin màu trong tra cứu ảnh dựa trên nội dung gồm những đại diện màu, những moment màu.

1.5.3.2 Kết cấu

Kết cấu được sử dụng rộng rãi và rất trực quan nhưng không có định nghĩa chính xác bởi tính biến thiên rộng của nó. Có rất nhiều cách để mô tả kết cấu: Những phương pháp thống kê thường sử dụng tần số không gian, ma trận biến cố, tần số biên...Từ những đặc điểm đơn giản này như là năng lượng, entropy, độ tương phản , độ thô, tính đồng nhất, tính tương quan, đẳng hướng, pha, độ ráp, đã được nhận ra. Những phương pháp mô tả kết cấu này tính toán các thuộc tính kết cấu khác nhau và hoàn toàn phù hợp nếu cỡ của kết cấu gốc có thể được so sánh với cỡ của điểm ảnh...

1.5.3.3 Hình dạng

Định nghĩa hình dạng của đối tượng thường là rất khó. Hình dạng thường được biểu diễn bằng lời nói hoặc hình vẽ, và mọi người thường sử dụng thuật ngữ như là tròn, méo. Xử lý hình dạng dựa trên máy tính đòi hỏi rất phức tạp, trong khi rất nhiều phương pháp mô tả hình dạng thực tế đang tồn tại nhưng không có một phương pháp chung nào cho mô tả hình dạng. Có hai kiểu đặc điểm hình dạng chính thường được sử dụng: những đặc điểm dựa trên biên và những đặc điểm dựa trên vùng. Đặc điểm dựa trên biên chỉ sử dụng đường bao ngoài của hình dạng trong khi đó đặc điểm vùng sử dụng toàn bộ vùng của hình dạng. Ví dụ những đặc điểm biên bao gồm mã xích, mô tả

fourier, những đường viền hình học đơn giản như uốn cong, chiều dài biên,..., đặc điểm vùng như số chu trình, độ lệch tâm...

1.5.3.4 Những đặc điểm mức cao

Phần lớn những nghiên cứu tra cứu ảnh dựa trên nội dung đều tập trung vào những phương pháp ở mức thấp. Mặc dù vậy, một vài nghiên cứu đã cố gắng làm giảm khoảng cách giữa mức thấp và mức cao, chúng có hướng tập trung vào một trong hai vấn đề sau. Thứ nhất là nhận dạng cảnh, nó thường rất quan trọng để xác định tất cả các kiểu cảnh miêu tả ảnh, nó thường được sử dụng để tìm kiếm và có thể giúp xác định đối tượng một cách rõ ràng. Một trong những hệ thống kiểu này là IRIS (Hermes-1995), chúng sử dụng màu, kết cấu, vùng và thông tin không gian lấy ra từ phần thích hợp nhất của cảnh, tạo ra kí hiệu text để có thể đưa vào bất kỳ hệ thống tra cứu dựa trên text. Những nghiên cứu khác đã đưa ra những kỹ thuật đơn giản cho phân tích cảnh, sử dụng những thành phần tần số thấp của ảnh để huấn luyện mạng neural, hoặc những thông tin màu lân cận được trích chọn từ những ảnh độ phân giải thấp để tạo ra những mẫu do người dùng định nghĩa.

Hướng thứ hai tập trung nghiên cứu nhận dạng đối tượng. Những công nghệ đang được phát triển cho nhận dạng và phân lớp đối tượng với cơ sở dữ liệu trực quan. Kỹ thuật tốt nhất được biết đến trong lĩnh vực này là kỹ thuật cho nhận dạng người trong ảnh. Tất cả những công nghệ này đều dựa trên ý tưởng phát triển mẫu cho mỗi lớp của những đối tượng được nhận dạng, xác định những vùng ảnh chứa đựng những mẫu của những đối tượng và xây dựng lên những mẫu chốt để xác nhận hoặc loại bỏ sự có mặt của đối tượng.

1.5.4 Những khoảng cách tương tự

Khi những đặc điểm của ảnh trong cơ sở dữ liệu được trích chọn và truy vấn của người dùng được thực hiện thì kết quả tìm kiếm được đưa ra bởi việc đo độ tương tự giữa những đặc điểm được trích chọn trong cơ sở dữ liệu và truy vấn của người sử dụng được phân tích. Những thước đo lý tưởng có một số những thuộc tính cơ bản sau:

Độ tương tự trực quan: Đặc điểm khoảng cách giữa hai ảnh là lớn chỉ khi những ảnh không tương tự và ngược lại khoảng cách giữa hai ảnh là nhỏ nếu chúng tương tự. Những ảnh thường được mô tả trong không gian đặc điểm và sự tương tự giữa các ảnh thường được đo bởi những thước đo khoảng cách trong không gian đặc điểm. Số thuộc tính của không gian này cho cảm nhận của con người và hiểu những thuộc tính của những đặc điểm vectơ mô tả ảnh là rất quan trọng trong việc cải thiện thuộc tính độ tương tự trực quan của những thước đo độ tương tự được đề xuất.

Hiệu quả: Sự đo đạc cần phải được tính toán nhanh để nhanh chóng đưa ra kết quả. Những ứng dụng tra cứu ảnh dựa trên nội dung tiêu biểu đòi hỏi phản hồi nhanh. Trong khoảng thời gian ngắn công nghệ tìm kiếm thường phải tính toán hàng ngàn khoảng cách phụ thuộc vào cỡ của cơ sở dữ liệu ảnh, bởi vậy độ phức tạp tính toán là rất quan trọng.

Khả năng biến đổi: Quá trình hệ thống thực hiện không nên bị giảm hiệu quả quá nhiều đối với cơ sở dữ liệu lớn bởi vì một hệ thống có thể tìm kiếm trong cơ sở dữ liệu chứa hàng triệu ảnh. Một sự thi hành đơn giản của một hệ thống tra cứu ảnh dựa trên nội dung là tính toán tất cả khoảng cách giữa ảnh truy vấn và ảnh trong cơ sở dữ liệu ảnh, sau đó những khoảng cách này được sắp xếp để tìm ra những ảnh tương tự nhất với ảnh truy vấn. Độ phức tạp của công nghệ tìm kiếm này tương ứng với cỡ của cơ sở dữ liệu ảnh (hoặc là O (N) với N là số ảnh). Công nghệ đánh chỉ số đa chiều có thể được sử dụng để làm giảm độ phức tạp xuống O (log(N)). Tuy nhiên, theo báo cáo rằng việc thực hiện của những công nghệ đánh chỉ số hiện thời đã giảm bớt được việc quét liên tục khi số chiều cần để đánh chỉ số là lớn hơn 20. Bởi vậy cần phải xem xét nhân tố này khi làm việc với cơ sở dữ liệu lớn.

Hệ thước đo: Vấn đề khoảng cách tương tự là có lên là hệ mét hay không vẫn chưa được quyết định chính thức khi sự nhìn nhận của con người là rất phức tạp và chưa được hiểu một cách đầy đủ. Chúng ta thích khoảng cách tương tự là một hệ đo khi chúng ta xem xét những thuộc tính sau như là những yêu cầu rất tự nhiên:

Sự bất biến của tương tự với chính nó: Khoảng cách giữa một ảnh với chính nó là hằng số độc lập với ảnh.

$$d(A,A)=d(B,B)$$

Sự tối thiểu: Một ảnh giống với nó hơn là với những ảnh khác

$$d(A,A)<d(A,B)$$

Sự đối xứng: Là vô lý nếu chúng ta nói rằng ảnh A giống với ảnh B nhưng ảnh B không giống với ảnh A

$$d(A,B)=d(B,A)$$

Sự bắc cầu: Là vô lý nếu nói rằng ảnh A rất giống với ảnh B, ảnh B không giống vói ảnh C nhứng ảnh C rất giống với ảnh A. Tuy nhiên, thuộc tính bắc cầu này có thể không đúng cho một dãy các ảnh. Thậm chí nếu ảnh I_i là giống với ảnh I_{i+1} với tất cả i=1..N thì điều này không có nghĩa rằng ảnh I_i tương tự với ảnh I_N, ví dụ trong

băng video mỗi khung tương tự với khung kề nó nhưng khung đầu tiên và khung cuối cùng có thể là rất khác nhau.

Sự mạnh mẽ: Hệ thống cần có khả năng để thay đổi những điều kiện ảnh trong cơ sở dữ liệu ảnh, ví dụ nếu ảnh trong cơ sở dữ liệu ảnh được lấy dưới ánh sáng đèn điện (hơi đỏ) thì hệ thống phải có thể tìm được những đối tượng này ngay cả khi đối tượng truy vấn được lấy dưới ánh sáng ban ngày (hơi xanh).

Có rất nhiều thước đo khoảng cách tương tự đã được đưa ra nhưng chúng đều không có đầy đủ các thuộc tính trên. Dưới đây là một vài thước đo chung nhất thường được sử dụng:

Histogram intersection Distanc (Swain and Ballard 1991):

Đây là một trong những thước đo khoảng cách đầu tiên trong tra cứu ảnh dựa trên màu sắc. Khoảng cách được định nghĩa dựa trên cỡ phần chung của hai biểu đồ màu. Cho hai biểu đồ màu h_1, h_2, khoảng cách giữa chúng có thể được định nghĩa như sau:

$$dis_{HI} = 1 - \sum N_{i=1} \min(h_{1i}, h_{2i})$$

Việc đo khoảng cách này rất nhanh bởi nó dựa trên công thức đơn giản. Tuy nhiên thông tin màu không được sử dụng khi nhận được khoảng cách bởi vậy có thể dẫn tới những kết quả không tốt.

L1 Distanc (Stricker and Orengo, 1996):

Khoảng cách dạng Minkowski L_p giữa hai biểu đồ màu được định nghĩa như sau:

$$dis_{Mp} = \left(\sum_i |h_{1i} - h_{2i}|\right)^{1/p}$$

Quadratic form Distanc (Hafner, 1995):

Khoảng cách giữa hai biểu đồ màu N chiều h_1 và h_2 được định nghĩa như sau:

$$dis_{QF} = (h_1 - h_2)A(h_1 - h_2)$$

Với $A=[a_{ij}]$ là ma trận với trọng số biểu thị sự giống nhau giữa bin i và bin j, a_{ij} được tính như sau:

$$a_{ij} = 1-(d_{ij} / d_{max})^k$$

Ở đây d_{ij} là khoảng cách giữa màu i và màu j(thường d_{ij} là khoảng cách Euclidean giữa hai màu trong một vài không gian màu đồng dạng) và $d_{max}=max_{ij}(d_{ij})$. K là hằng số điều khiển trọng số giữa những màu lân cận.

Earth Mover Distance (Rubner, 1998)

Thước đo này dựa trên chi phí tối thiểu để chuyển một phân bố thành phân bố khác. Nếu chi phí của việc di chuyển một đơn vị đặc điểm đơn trong không gian đặc điểm là khoảng cách chung thì khoảng cách giữa hai phân bố sẽ là tổng cực tiểu của giá trị để di chuyển những đặc điểm riêng. Khoảng cách EMD có thể được định nghĩa như sau:

$$dist_{EMD} = \sum_{i\,j} g_{ij} d_{ij} / \sum_{i\,j} g_{ij}$$

Ở đây g_{ij} biểu thị khoảng cách tương tự giữa bin i và bin j và $g_{ij} >= 0$ là sự tối ưu hoá giữa hai phân bố như là tổng giá trị được cực tiểu hoá,

$$\sum_i g_{ij} \leq h_{1i}$$

$$\sum_j g_{ij} \leq h_{2i}$$

$$\sum_{i\,j} g_{ij} = \min (h_{1i}, h_{2i})$$

1.6 Các phương pháp tra cứu ảnh dựa trên nội dung

1.6.1 Tra cứu ảnh dựa trên màu sắc

Tra cứu ảnh dựa trên nền tảng màu sắc tương tự hầu hết là biến đổi dựa trên ý tưởng giống nhau. Mỗi ảnh khi đưa vào tập hợp ảnh đều được phân tích, tính toán một biểu đồ màu đó là tỷ lệ của những điểm ảnh của mỗi màu trong ảnh. Sau đó biểu đồ màu của mỗi ảnh sẽ được lưu trữ trong cơ sở dữ liệu. Khi tìm kiếm người sử dụng có thể xác định tỷ lệ của mỗi màu mong muốn (ví dụ 75% Blue, 25% Red) hoặc đưa ra một ảnh mẫu với biểu đồ màu đã được tính toán. Đồng thời khi đó quá trình đối sánh tra cứu những biểu đồ màu của những hình ảnh này so sánh với biểu đồ màu của truy vấn gần nhất. Kỹ thuật đối sánh được sử dụng phổ biến nhất là biểu đồ màu giao nhau được phát triển đầu tiên bởi Swain and Ballard's[1991]. Những kỹ thuật cải tiến từ kỹ thuật này ngày nay được sử dụng rộng rãi trong các hệ thống tra cứu ảnh hiện thời.

Phương pháp cải tiến dựa trên công nghệ độc đáo của Swain and Ballard's gồm cách sử dụng biểu đồ màu tích lũy [Stricker and Orengo, 1995], kết hợp biểu đồ màu giao nhau với một số thành phần đối sánh không gian [Stricker and Dimai, 1996] và sử dụng vùng truy vấn dựa trên màu sắc [Carrson et al, 1997]. Kết quả của các hệ thống này đã tạo những ấn tượng khá sâu sắc.

1.6.2 Tra cứu ảnh dựa trên kết cấu

Khả năng tra cứu ảnh dựa trên kết cấu tương tự dường như không hiệu quả nhưng khả năng đối sánh dựa trên đặc điểm này thường có lợi cho việc phân biệt các

vùng ảnh với màu tương tự (ví dụ như bầu trời và biển hoặc lá cây và cỏ). Một loạt các kỹ thuật được sử dụng cho việc đo kết cấu tương tự; công nghệ tốt nhất được thiết lập dựa trên sự so sánh những giá trị đã được biết đến như là số liệu thống kê thứ hai được tính toán từ truy vấn và những ảnh được lưu trữ. Từ đó có thể tính toán được khoảng cách của kết cấu ảnh như mức độ tương phản, độ thô, phương hướng và tính cân đối [Tamura et al, 1978] hoặc chu kỳ, phương hướng và tính ngẫu nhiên [Liu and Picard, 1996]. Các phương pháp phân tích kết cấu cho tra cứu bao gồm sử dụng những bộ lọc Gabor [Manjunath and Ma, 1996] và những Fractal [Kaplan et al, 1998]. Các truy vấn kết cấu có thể được trình bày tương tự như truy vấn màu sắc bằng việc lựa chọn những mẫu kết cấu như mong muốn từ bảng màu hoặc bằng việc cung cấp ảnh truy vấn mẫu. Hệ thống sau đó sẽ tra cứu những ảnh với giá trị độ đo kết cấu giống nhau nhất với truy vấn. Gần đây có một sự mở rộng của công nghệ là cuốn từ điển kết cấu được phát triển bởi Ma and Manjunath, nó tra cứu những vùng kết cấu rõ ràng trong ảnh dựa trên nền tảng của sự tương tự để nhận lấy từ mã mô tả các lớp quan trọng của kết cấu trong tập ảnh một cách tự động.

1.6.3 Tra cứu ảnh dựa trên hình dạng

Khả năng tra cứu bởi hình dạng có lẽ là nhu cầu hiển nhiên nhất ở mức độ nguyên thủy. Không như kết cấu, hình dạng là một khái niệm hoàn toàn rõ ràng, và bằng chứng là những vật thể tự nhiên đầu tiên được nhận thấy bởi hình dạng của chúng [Biederman, 1987] . Số lượng những đặc điểm tiêu biểu của hình dạng đối tượng được tính toán cho mỗi đối tượng xác định trong mỗi ảnh được lưu trữ. Sau đó truy vấn được trả lời bởi việc tính toán tập những đặc điểm cho ảnh truy vấn, và việc tra cứu đặc điểm của những hình ảnh được lưu trữ này phải phù hợp với đặc điểm của truy vấn. Hai kiểu chính của đặc điểm hình dạngthường được sử dụng là đặc điểm tổng thể như tỷ lệ bên ngoài, hình tròn [Niblack et al, 1993] và những đặc điểm cục bộ như tập các đoạn biên liên tiếp [Mehrotra and Gary, 1995]. Các phương pháp khác đề cập tới sự đối sánh hình dạng bao gồm sự biến dạng co giãn của các khuân dạng ([Pentland et al, 1996], [delBimbo et al, 1996]), sự so sánh của những biểu đồ định hướng của những biên được trích chọn từ ảnh [jain and Vailaya,1996], khung biểu diễn hình dạng của đối tượng có thể được so sánh bằng việc sử dụng những kỹ thuụât đố sánh đồ thị [Kimia et al,1977], Tirthap et al, 1998]. Những truy vấn đối với hệ thống tra cứu hình dạng thường được biểu diễn bằng cách xác định một hình ảnh mẫu để thực hiện như là hình thức truy vấn hoặc như là một bản phác thảo được vẽ ra bởi người sử dụng [Hirata and Kato, 1992], [Chan and Kung, 1997].

Việc đối sánh hình dạng của các đối tượng 3 chiều là một công việc khó khăn hơn. Trong khi chưa có giải pháp chung cho vấn đề này thì một số cách hữu ích đã được tạo thành cho việc xác định độ đo của đối tượng từ nhiều khía cạnh khác nhau. Một phương pháp đã được sử dụng để xây dựng tập mô hình 3 chiều thích hợp từ ảnh 2 chiều có sẵn và đối sánh chúng với các mẫu khác trong cơ sở dữ liệu [Chen and Stokman, 1996]. Một cách khác nhằm tạo ra một loạt các ảnh 2 chiều khác của mỗi cơ sở dữ liệu đối tượng và mỗi ảnh này được đối sánh với ảnh truy vấn [Dickínon et al, 1998]. Những vấn đề nghiên cứu có liên quan đến lĩnh vực này gồm định nghĩa những độ đo tương tự hình dạng 3 chiều [Shum et al, 1996] và cung cấp phương tiện cho người sử dụng tạo ra những truy vấn hình dạng 3 chiều [Horikoshi and Kasahara,1990].

1.6.4 Tra cứu ảnh bởi các đặc điểm khác

Một trong những phương tiện truy cập dữ liệu có hình ảnh cổ điển nhất là tra cứu bởi vị trí của nó trong ảnh. Truy cập dữ liệu bởi không gian vị trí là một khía cạnh chủ yếu của hệ thống thông tin địa lý, và các phương pháp hiệu quả để thực hiện công việc này đã được áp dụng trong nhiều năm gần đây (ví dụ Chock et al [1984], Roussopoulos et al [1988]. Những công nghệ tương tự cũng đã được áp dụng cho những tập ảnh, cho phép người sử dụng tìm kiếm những ảnh chứa các đối tượng có mối quan hệ không gian xác định với các đối tượng khác (Chang et al[1998], Chang and jungert[1991]). Các thuật toán được cải tiến cho việc tra cứu thuộc lĩnh vực không gian vẫn đang được đề xuất. Việc đánh chỉ số không gian riêng nó thì ít hiệu quả, mặc dù nó chứng tỏ được hiệu quả của nó trong việc kết hợp với các dạng khác như màu sắc và hình dạng.

Một vài kiểu khác của đặc điểm ảnh được đưa ra như là nền tảng cho việc tra cứu ảnh dựa trên nội dung. Hầu hết những kiểu này đều dựa vào sự biến đổi phức tạp của cường độ của điểm ảnh. Đa số các công nghệ đều hướng về việc trích chọn ra những đặc điểm phản ánh một số khía cạnh của hình ảnh tương tự mà đối tượng con người có thể cảm nhận được, ngay cả khi người đó cảm thấy rất khó để mô tả. Kỹ thuật thành công nhất của loại này là sử dụng cách biến đổi wavelet. Kết quả tra cứu đầy hứa hẹn đã được báo cáo bằng việc đối sánh những đặc điểm wavelet được tính toán từ truy vấn và những ảnh được lưu trữ. Một phương pháp khác cũng cho kết quả rất tốt là tra cứu bởi hình thức. Hai phiên bản của phương pháp này đã được phát triển, một cho đối sánh toàn bộ và một cho đối sánh những phần được lựa chọn của ảnh.

CHƯƠNG 2: TRA CỨU ẢNH DỰA TRÊN KẾT CẤU

2.1 Giới thiệu

Kết cấu là một khái niệm trực quan, là một thành phần chủ yếu về nhận thức thuộc tri giác của con người. Giống như màu sắc, kết cấu trở thành đặc điểm cần thiết để xem xét khi truy vấn cơ sở dữ liệu ảnh. Mọi người đều có thể nhận thấy kết cấu tuy nhiên nó rất khó để xác định, điều này được thể hiện bằng một số khái niệm khác nhau về kết cấu. Mặc dù không có một khái niệm chung cho kết cấu nhưng tất cả các nhà nghiên cứu đều tập trung thống nhất trên hai điểm chính:

Trong phạm vi một kết cấu có sự biến đổi đáng kể về mức độ cường độ giữa các điểm ảnh liền kề, đó là giới hạn của độ phân giải, không có sự đồng nhất.

Kết cấu là thuộc tính đồng nhất ở một vài không gian lớn hơn độ phân giải của ảnh, cái hàm ý trong những thuộc tính này của cấu trúc là ảnh có độ phân giải nhất định.

Khác với màu sắc, kết cấu diễn ra trên cả một vùng hơn là tại một điểm, nó thường được định nghĩa bằng những mức xám được hiểu như là màu sắc. Một số nhà nghiên cứu giải quyết bài toán về xác định kết cấu bằng cách mô tả nó trong những thuật ngữ của hệ thống thị giác của con người như hướng, độ thô, độ tương phản…. Một số nhà nghiên cứu khác lại lái định nghĩa kết cấu bởi những ứng dụng. Điều này tạo cho kết cấu những mặt đa dạng và cho nhiều cách để trích chọn kết cấu.

Định nghĩa của kết cấu dựa trên nhận thức của con người là phù hợp cho nghiên cứu và cho bàn luận về nét tự nhiên của kết cấu. Mặc dù vậy một định nghĩa đưa ra những vấn đề khi được sử dụng như là học thuyết cơ bản cho thuật toán phân tích kết cấu.

2.2 Kết cấu theo nhận thức của con người

Julez đã nghiên cứu tổng quát sự nhận thức cấu trúc trong nội dung phân biệt cấu trúc. Câu hỏi được đưa ra là ” Khi nào một cặp kết cấu được phân biệt, nhất định rằng các kết cấu có cùng độ sáng, độ tương phản và màu sắc ?”. Phương pháp của Julez gắn một kết cấu với một kết cấu khác. Nếu phần được gắn vào của kết cấu đứng ngoài kết cấu lân cận thì hai kết cấu được xem như không giống nhau. Nếu hai kết cấu có thể được phân biệt, Julez đã sử dụng hai số liệu thống kê thứ nhất và thứ hai để phân tích.

Số liệu thống kê thứ nhất đo khả năng quan sát một giá trị mức xám tại một vị trí được chọn ngẫu nhiên trong ảnh. Số liệu thống kê này có thể được tính toán từ biểu đồ cường độ điểm ảnh trong ảnh. Điều này chỉ phụ thuộc vào các giá trị điểm ảnh riêng

biệt mà không có sự tương tác hoặc kết hợp giữa các điểm ảnh lân cận. Cường độ trung bình của ảnh là một ví dụ của số liệu thống kê thứ nhất. Số liệu thống kê thứ hai được định nghĩa như là khả năng quan sát một cặp giá trị xám xuất hiện từ một điểm tới một điểm khác với khoảng cách ngẫu nhiên, hướng và vị trí ngẫu nhiên trong ảnh. Đây là những thuộc tính của những cặp giá trị điểm ảnh.

Julez nhận thấy rằng những kết cấu có số liệu thống kê thứ nhất giống nhau nhưng số liệu thống kê thứ hai khác nhau thì dễ phân biệt. Mặc dù vậy Julez không thể tìm thấy những kết cấu có cả hai số liệu thống kê như nhau có thể phân biệt được. Điều này cho phép ông đưa ra phỏng đoán "Chuẩn thứ hai của kết cấu là không thể phân biệt được".

Sau này Caelli đã đưa ra kết luận Chuẩn thứ hai của kết cấu là có thể phân biệt được với nhận thức thị giác của con người trước đây. Hơn nữa một nghiên cứu khác của Julez đã chỉ ra rằng phỏng đoán đầu tiên của ông ấy là sai. Thay vào đó, ông ấy đã nhận thấy rằng kỹ thuật cảm nhận thị giác con người không cần thiết phải sử dụng số liệu thống kê thứ ba để phân biệt những kết cấu thuộc chuẩn hai này, nhưng đúng hơn là sử dụng những số liệu thống kê loại hai của những đặc điểm mà Julez gọi là các Textons. Những cái này được biểu diễn như là nền tảng của kết cấu. Ba lớp của texton được đưa ra là: color, elongated blobs, terminators(endpoint) of elongated blobs. Phỏng đoán ban đầu được xem xét lại rằng " Hệ thống cảm nhận thị giác con người trước đây không thể tính toán những tham số thống kê cao hơn loại hai được". Hơn nữa Julez đã phát biểu rằng hệ thống dựa vào trực giác của con người trước đây thực sự mới chỉ sử dụng số liệu thống kê loại một của những texton này . Từ những nghiên cứu trước về cảm nhận giác quan của con người, nghiên cứu tâm sinh lý học đã tập trung vào việc phát triển những mô hình thích hợp cho sự phân biệt kết cấu, những mô hình này bao gồm những việc xác định những thước đo nào của con người nhạy cảm nhất đối với sự biến đổi của kết cấu. Texton không được xem như là thước đo phân biệt kết cấu hợp lý như đã được vạch ra bởi Julez. Beck đã chứng tỏ rằng cảm nhận của phân đoạn kết cấu trong những kiểu mẫu nhất định là một hàm cơ sở của những phân tích không gian tần số. Nghiên cứu về tâm sinh lý học dã chứng tỏ rằng bộ óc biểu diễn rất nhiều kênh, tần số, hướng trên võng mạc[10,25]. Campbell và Robson đã làm những thí nghiệm tâm sinh lý học sử dụng một loạt những mẫu khác nhau và đã chỉ ra rằng hệ thống trực quan phân tích ảnh thành những bộ lọc ảnh của tần số và hướng khác nhau. De Valois đã nghiên cứu bộ óc của một loài khỉ ở ấn độ được coi là giống với bộ óc của con người nhất về quá trình xử lý hị giác. Họ đã ghi lại phản ứng của các tế bào đơn trong vỏ não của con khỉ thành những lưới hình sin ở những tần số và hướng khác nhau, và

họ đã kết luận rằng những tế bào này thực chất là những phạm vi hẹp về tần số và hướng. Những nghiên cứu này đã trở thành động lực thúc đẩy các nhà nghiên cứu thị giác áp dụng những phương pháp lọc đa kênh vào việc phân tích kết cấu. Tamura đã chỉ rõ những thuộc tính sau đây đóng vai trò quan trọng trong việc mô tả kết cấu: Không đồng dạng, mật độ, độ thô, độ gồ ghề, tính đều đặn, hướng, tần số. Một số tính chất được cảm nhận này là không độc lập, ví dụ tần số không độc lập với mật độ và thuộc tính phương hướng chỉ áp dụng vào những kết cấu phương hướng. Thực tế cảm nhận về kết cấu có quá nhiều mức độ khác nhau, và đây chính là lý do quan trọng dẫn đến tại sao không có một phương pháp duy nhất để biểu diễn kết cấu thích hợp với những kết cấu khác nhau.

2.3 Phương pháp cho phân tích kết cấu

2.3.1 Tiêu chuẩn kết cấu thống kê

Một tập các đặc điểm được sử dụng để biểu diễn những đặc điểm của một kết cấu ảnh, những đặc điểm này đo những thuộc tính như độ tương phản, mối tương quan và entropy. Chúng thường được lấy ra từ những loạt dài giá trị mức xám, giá trị mức xám khác nhau hoặc ma trận kết hợp. những đặc điểm được lựa chọn và ảnh không thể tái tạo lại từ tập các đặc điểm được đánh giá.

2.3.2 Mô hình kết cấu ước lượng (Stochastic)

Một đặc điểm được coi là sự thực hiện của quá trình ước lượng bị ảnh hưởng bởi một số tham số. Việc phân tích được thực hiện bằng việc xác định một mô hình và những tham số ước lượng, bởi vậy xử lý ước lượng có thể được tái tạo từ những mô hình và tham số kết hợp. Những tham số ước lượng có thể đáp ứng như là những đặc điểm cho những bài toán phân đoạn và phân lớp kết cấu. Một khó khăn đối với mô hình kết cấu này là một số kết cấu tự nhiên không phù hợp với sự hạn chế của mô hình đặc biệt.

2.3.3 Tiêu chuẩn kết cấu cấu trúc

Một số kết cấu có thể xem như là những mô hình hai chiều gồm một tập các đặc điểm gốc hoặc các mẫu con được sắp xếp dựa trên luật nhất định. Các đặc điểm gốc này có thể là những hình dạng thay đổi hoặc xác định như hình tròn, hình lục giác hoặc thậm chí là mô hình dấu chấm. Những kết cấu lớn có đặc điểm gốc lớn trong khi những kết cấu nhỏ được tạo ra từ những đặc điểm nguyên thuỷ nhỏ, những thuật ngữ này liên quan tới độ phân giải của ảnh. Ảnh có kết cấu được hình thành từ những đặc điểm nguyên thuỷ bởi các luật cả trên phạm vi ảnh và mối quan hệ giữa các ảnh với nhau. Ví

dụ những kết cấu này gồm những cấu trúc có ô hình mạng như miếng vải mỏng và ảnh của một bức tường gạch, việc xác định những đặc điểm này là rất khó khăn.

2.3.4 Những đặc điểm kết cấu

Mục đích của mô hình trong phân tích ảnh là đạt được những đặc tính bên trong của ảnh với một số tham số cũng như để hiểu nét tự nhiên của hiện tượng tạo ra ảnh. Những mô hình ảnh cũng có lợi cho việc chỉ rõ qui định và những giả định chung về thế giới vật lý và quy trình vẽ hình. Nghiên cứu về mô hình kết cấu với mục đích tìm ra một sự thích hợp và nếu có thể là hoàn hảo cho việc biểu diễn kết cấu thường gặp. Mục đích là sử dụng những mẫu này cho những công việc như phân lớp, phân đoạn các phần của ảnh với những kết cấu khác nhau hoặc khám phá ra những thiếu xót hay sự dị thường trong kết cấu.

Mô hình kết cấu có thể được chia làm ba nhóm chính: Mô hình hàm mật độ xác suất (PDF- Probability Density Function), mô hình hình dạng chung (GS-Gross Shape) và mô hình bộ phận (Partial).

Mô hình hàm mật độ xác suất thống kê PDF là phù hợp với sự phân bố không gian của cường độ trong kết cấu. Điển hình những phương pháp này đo sự tương tác của một số lượng nhỏ các điểm ảnh ví dụ như mô hình ngẫu nhiên Gauss-Markov(MGRF) và những phương pháp đồng sự kiện mức xám (GLC Gray-Level-Co-occurence) đo sự tương tác của các cặp điểm ảnh.

Những phương pháp GS làm mẫu một kết cấu như là một bề mặt, chúng đo những đặc điểm mà con người có thể cảm nhận được như là các biên, đường cường độ cực trị, dạng sóng và hướng. Những phương pháp mày đo sự tương tác của một số lượng lớn các điểm ảnh trên một phạm vi rộng hơn các phương pháp PDF. Những phương pháp điều hoà đo tính chu kỳ trong kết cấu, nó tìm kiếm những đặc điểm trực quan xảy ra liên tiếp với những khoảng thời gian đều đặn. Những phương pháp nguyên thuỷ khám phá ra một tập những đặc điểm trực quan cô đọng như đường biên, cường độ cực trị và cho ra một vec tơ đặc điểm bao gồm cường độ của những đặc điểm trực quan này trong kết cấu.

Những phương pháp cục bộ (Partial) tập trung vào một số khía cạnh đặc biệt của những đặc tính kết cấu không có lợi cho các khía cạnh khác.

2.4 Những phương pháp phân tích kết cấu

2.4.1 Phương pháp Gause Markov Random Field (GMRF)

Đây là phương pháp PDF có tham số, nó có giả định cơ bản là những kết cấu một phần là cấu trúc, một phần là ước lượng. Trong thực tế những phương pháp này thừa nhận rằng cấu trúc trong kết cấu có thể được mô tả một cách cục bộ.

Phương pháp này làm mẫu cường độ của điểm ảnh như là một hàm ước lượng cường độ của những điểm ảnh gần kề nó, đặc biệt là nó sử dụng hàm mật độ xác suất Gauss để làm mẫu cường độ điểm ảnh. Ý nghĩa của phân bố Gauss là hàm tuyến tính của cường độ các điểm ảnh gần kề. Tiêu biểu, phương pháp bình phương tối thiểu được sử dụng để ước lượng những hệ số tuyến tính và sự biến đổi của sự phân bố Gauss. Như là sự lựa chọn phân bố nhị thức được sử dụng hơn là phân bố Gauss, mặc dù vậy, trong các loại tham số được sử dụng thì sự phân bố nhị thức tương đương với sự phân bố Gauss.

Chellappa và Chaterjee đưa ra công thức tiêu biểu sau:

$$I(x,y) = \sum_{(\sigma x,\, \sigma y)\epsilon Ns} \Theta_{(\sigma x,\, \sigma y)} (I(x + \sigma x, y + \sigma y) + I(x - \sigma x, y - \sigma y)) + e(x,y)$$

Ở đây I(x,y) cường độ của điểm ảnh tại toạ độ (x,y) trong ảnh, N là điểm ảnh láng giềng đối xứng (bao gồm cả điểm ảnh đó), N_s là một nửa của N, $\theta(x,y)$ là tham số ước lượng sử dụng phương pháp bình phương tối thiểu và e(x,y) là trung bình nhiễu Gauss không đổi (zero mean stationary Gaussian noise sequence) với những thuộc tính sau:

$$E(e(x,y)e(\sigma x, \sigma y)) = \begin{cases} -\Theta_{(x - \sigma x, y - \sigma y)}^{v} & IF(x - \sigma x, y - \sigma y) \epsilon N, IF(x,y) = (\sigma x, \sigma y) \\ \text{otherwise} \end{cases}$$

Với V là sai số bình phương trung bình của ước lượng bình phương tối thiểu.

2.4.2 Phương pháp Gray-Level Co-occurrence Matrices

Đây là phương pháp PDF không có tham số. Sự khác biệt giữa phương pháp này với các phương pháp có tham số phản ánh sự phân biệt được tạo bởi các con số thống kê giữa hai kỹ thuật tạo mẫu PDF có tham số và không tham số. Không gian mức xám đồng nhất ước lượng những thuộc tính của ảnh có liên quan đến những số liệu thống kê thứ hai. Haralick gợi ý sử dụng ma trận mức xám đồng nhất (GLCM) cái mà đã trở thành một trong những phương pháp nổi tiếng nhất và được sử dụng rộng rãi những đặc điểm kết cấu. Ma trận đồng nhất mức xám Pd(G*G) với vectơ thay thế d=(dx,dy) được định nghĩa như sau:

$P_d(i,j)$=|| {((r, s),(t, v)):(t, v)=(r+dx, s+dy), I(r, s)=i, I(t, v)=j}||

Ở đây (r, s), (t,v) ∈N x N, ||.|| là lực lượng trong tập hợp

1	1	2	2	2
1	1	2	2	2
1	3	3	3	3
3	3	4	4	4
3	3	4	4	4

a)

	1	2	3	4
1	2	2	1	0
2	0	4	0	0
3	0	0	5	2
4	0	0	0	4

b)

Hình 2.1. Tính toán ma trận đồng nhất mức xám

a)Ảnh được lượng tử hoá thành 4 mức cường độ

b)Tính toán ma trận GLC tương ứng với offset(dx,dy)=(0,1)

Một ví dụ đưa ra ở hình 2.1. Trong hình 2.1a có hai cặp điểm được đóng khung có I(x,y)=1 và I(x+dx,y+dy)=2. Bin tương ứng được nhấn mạnh trong hình 2.1b. Lưu ý rằng ma trận Co-occurrent được định nghĩa trong cách này là không cân đối. Một biến thể cân đối có thể được tính toán bằng công thức $P=P_d+P_{-d}$. Ma trận này khám phá ra những thuộc tính nhất định về sự phân bố không gian cấp xám trong kết cấu ảnh. Ví dụ nếu đầu vào trong ma trận đều được tập trung theo đường chéo thì kết cấu sẽ thô đối với véc tơ thay thế d. Haralick đã đưa ra một số các đặc điểm kết cấu có thể được tính toán từ ma trận này, chúng được liệt kê trong bảng 2.1.1. Ở đây μx, μy là phương sai, σx, σy là độ lệch tiêu chuẩn của $P_d(x) = \Sigma_j P_d(x,j)$ và $P_d(y) = \Sigma_i P_d(i,y)$.

Những đặc điểm ma trận Co-occurrent có một số khó khăn đó là: Không có một phương pháp được thiết lập hoàn hảo về sự lựa chọn véc tơ thay thế d và việc tính toán ma trận đối với một số giá trị khác nhau của d là không thể thực hiện được. Hơn nữa, với một giá trị của d có một số lượng lớn các đặc điểm có thể được tính toán. Điều này có nghĩa là một số phương pháp lựa chọn đặc điểm cần phải được sử dụng để lựa chọn những đặc điểm có liên quan nhất.

Bảng 2.1: Một số trích chọn đặc điểm kết cấu từ ma trận đồng nhất mức xám

Đặc điểm kết cấu	Công thức
Energy	$\sum_i \sum_j P_d^2(i,j)$
Entropy	$-\sum_i \sum_j P_d(i,j) \log \sum_i \sum_j P_d(i,j)$
Contrast	$\sum_i \sum_j (i-j)^2 P_d(i,j)$
Homogeneity	$\sum_i \sum_j P_d^2(i,j) / i+\|i-j\|$
Correlation	$\sum_i \sum_j (i - \mu x)(j - \mu y) P_d(i,j) / \sigma x, \sigma y$

2.4.3 Phương pháp Gray-Level Difference (GLD)

Như đã được giải thích ở bởi Weszka, Những phương pháp GLD rõ ràng là tương tự với các phương pháp GLC. Tuy nhiên điểm khác biệt chính giữa chúng là, trong khi phương pháp GLC tính toán ma trận của các cặp cường độ thì phương pháp GLD lại tính toán một véc tơ của những chênh lệch cường độ. Điều này tương đương với việc tổng kết ma trận GLC với những đường chéo của nó.

Cụ thể, cho bất kỳ khoảng cách thay thế d=(dx,dy) thì:

$$I_d(x,y) = \| I(x,y) - I(x+dx,y+dy) \|$$

Cho P_d là mật độ xác suất của $I_d(x,y)$. Nếu có m mức xám thì sẽ tạo thành một véc tơ m chiều trong đó thành phần thứ i chính là xác suất mà $I_d(x,y)$ sẽ có giá trị i. Nếu ảnh I là rời rạc thì dễ dàng tính toán P_d bằng việc đếm số lần mỗi giá trị $I_d(x,y)$ xảy ra. Những đặc điểm tương tự trong bảng 2.1.1 có thể được tính toán.

2.4.4 Phương pháp phân bố kết cấu (Texture spectrum)

Tất cả các phương pháp được mô tả trên lấy mẫu kết cấu như một lĩnh vực ngẫu nhiên. Chúng lấy mẫu cường độ của điểm ảnh như là những hàm ước lượng về cường độ của những điểm ảnh lân cận. Mặc dù vậy, không gian của tất cả các mẫu cường độ trong vùng lân cận là rất lớn. Ví dụ nếu một vùng lân cận 5 x 5 được tính đến (loại trừ điểm trung tâm) thì PDF là một hàm trong không gian 24 chiều. Những phương pháp GMRF và GLC dựa vào những giả định là làm giảm độ phức tạp tính toán của mô hình PDF. Phương pháp GMRF ước lượng cường độ như là một hàm của tất cả các điểm ảnh lân cận, nhưng lại cho rằng sự phân bố là Gaussian và được tập trung trong một

hàm tuyến tính của những cường độ lân cận. Phương pháp GLC sử dụng mô hình biểu đồ; đòi hỏi không gian cường độ phải được chia thành các "bin" biểu đồ. Sự phân chia chỉ nhạy cảm với những tác động loại hai nhưng không nhạy cảm với những tác động loại cao hơn.

Phương pháp phân bố kết cấu sử dụng mô hình PDF cái mà nhạy cảm với những tác động loại cao. Đặc biệt, những phương pháp này sử dụng một mô hình biểu đồ, trong đó sự phân chia không gian cường độ là nhạy cảm với những tác động cao giữa các điểm ảnh. Sự nhạy cảm này có thể được tạo thành bằng việc lượng tử hóa những giá trị cường độ với một số mức nhỏ, cái mà làm giảm đáng kể không gian kích thước. Số lượng lớn nhất của các mức được sử dụng là bốn nhưng hai mức hoặc ngưỡng là phổ biến hơn.

Ojala đã đưa ra một đơn vị kết cấu được biểu diễn bằng tám thành phần, mỗi thành phần có thể có hai giá trị {0,1} đạt được từ những điểm ảnh lân cận 3 x 3. Những đơn vị kết cấu này được gọi là mẫu nhị phân cục bộ (LBP-Local Binary Pattern) và sự phân bố của chúng trên cả một vùng tạo nên sự phân bố kết cấu. LBP được tính toán bởi ngưỡng mỗi điểm ảnh không thuộc trung tâm vói điểm ảnh trung tâm, kết quả là 256 kiểu nhị phân. Phương pháp LBP là độ xám bất biến (gray-scale) và có thể kết hợp dễ dàng với một thước đo tương phản đơn giản bởi việc tính toán mức xám trung bình khác nhau của mỗi điểm ảnh lân cận sau khi lấy ngưỡng có giá trị 0,1 tương ứng. Thuật toán được mô tả chi tiết như sau:

Với mỗi điểm lân cận 3 x 3, xét cường độ P_i của các điểm ảnh hợp thành với P_0 là cường độ của điểm ảnh trung tâm thì:

1. Lấy ngưỡng P_i với giá trị của điểm trung tâm:

$$P' = \{ 0 \text{ if } Pi < Po \text{ , } 1 \text{ otherwise}$$

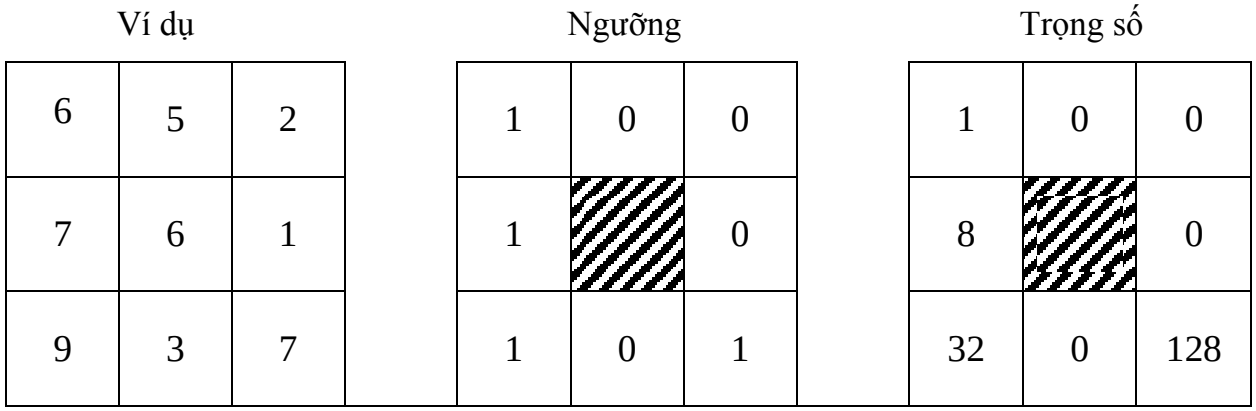
2. Đếm số giá trị điểm khác 0 : $n = \sum_{i-1} P_i r$

3. Tính toán mẫu nhị phân cục bộ : $LBP = \sum_{i-1} P_i' * 2^{i-1}$

4. Tính toán độ tương phản cục bộ

$$C = \{ 0 \text{ if } n = 0 \text{ or } n = 8 \text{ ; } 1/n \sum_{i-1} P_i' P_i - 1/8-n \sum_{i-1} (1-P_i') P_i \text{ otherwise}$$

Phương pháp LBP tương tự với các phương pháp được mô tả bởi Wang và He và Read nhưng những đơn vị kết cấu phân biệt được tạo ra ít hơn. Wang lượng tử hoá ở ba mức cường độ, cho ra 3^8 hoặc 6561 đơn vị kết cấu phân biệt. Read lượng tử hoá ảnh thành bốn mức cường độ nhưng chỉ sử dụng những điểm lân cận 3 x 2, nó tạo ra 4^6 hoặc 4096 đơn vị kết cấu phân biệt.



Hình 2.2: Tính toán thước đo mẫu nhị phân cục bộ và tương phản

Một lớp khác của phương pháp phân bố kết cấu gồm những phương pháp N-tuple. Trong khi các đơn vị kết cấu trong những phương pháp được mô tả ở trên sử dụng tất cả các điểm ở một vùng lân cận nhỏ thì những phương pháp này sử dụng một tập con của những điểm từ một vùng lân cận lớn hơn. Điển hình là những tập con của 6 đến 10 điểm được sử dụng từ vùng lân cận 6 x 6 tới vùng 10 x 10. Những tập nhỏ này được lựa chọn ngẫu nhiên; một kiểu bộ nhớ thuật toán N-tuple phải có 30 đơn vị N-tuple, mỗi đơn vị có một tập con ngẫu nhiên phân biệt trong những điểm lân cận. Biểu đồ đơn vị kết cấu và thông tin lớp kết cấu được tính toán độc lập đối với mỗi đơn vị N-tuple. Thông tin lớp kết cấu từ mỗi đơn vị N-tuple được kết hợp để tạo ra thông tin lớp của bộ nhớ N-tuple.

Tóm lại, các phương pháp phân bố kết cấu nhạy cảm với sự tương tác cao giữa các điểm ảnh. Điều này có thể được thực hiện bằng cách giảm kích thước không gian cường độ bằng việc lượng tử hoá. Thậm chí trong không gian được giảm này chỉ một số lượng giới hạn các điểm ảnh, điển hình là ít hơn 10, cùng với đặc điểm véc tơ để tạo lên đặc tính của kết cấu.

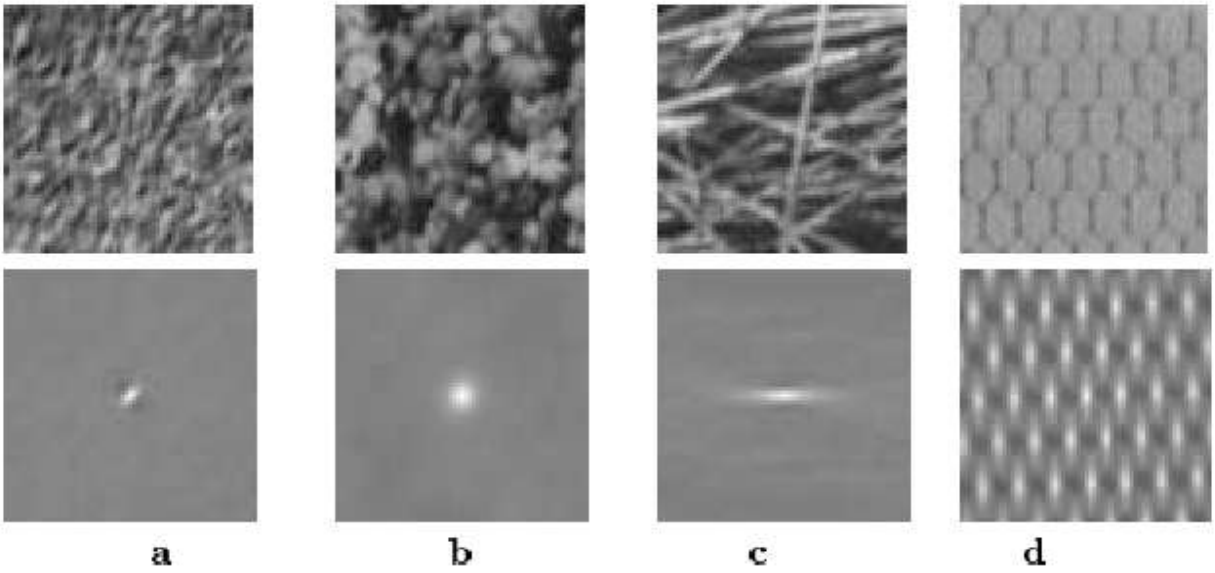
2.5 Mô hình hình dạng chung dùng trong kết cấu (GS-Gross Shape)

2.5.1 Phương pháp Autocorrelation

Một thuộc tính quan trọng của nhiều kết cấu là sự lặp lại tự nhiên của những phần kết cấu trong ảnh. Hàm tương quan tự động của ảnh có thể được sử dụng để truy cập số lượng lớn tính đều đặn cũng như độ mịn, độ thô của kết cấu xuất hiện trong ảnh. Hàm tương quan tự động P của ảnh I được định nghĩa như sau:

$$P(x,y) = \sum^{N}_{u=0}\sum^{N}_{r=0} I(u,v)I(u+x,v+y) / \sum^{N}_{u=0}\sum^{N}_{r=0} I^2(u,v)$$

Ví dụ về hàm tương quan tự động đối với một số kết cấu được chỉ rõ trong hình 2.3. Những hàm tương quan tự động của những kết cấu không tuần hoàn được bao quát bởi một chóp đơn. Bề rộng và bề dài của chóp được xác định bởi độ thô và hướng của kết cấu. Trong kết cấu mịn mặt như hình 3.a, hàm tương quan tự động sẽ giảm nhanh chóng tạo ra một hình nhọn. Mặt khác, trong một kết cấu thô như hình 4.3 hàm tương quan tự động sẽ giảm chậm hơn, tạo ra một đỉnh rộng hơn. Một kết cấu hướng như hình 3.c sẽ tạo ra một đỉnh thon dài, với kết cấu cân đối như hình 3.d hàm tương quan tự động sẽ biểu thị ra các đỉnh và các rãnh.



Hình 2.3: Hàm tương quan tự động tính toán cho 4 kết cấu

Sự phân biệt khả năng của các phương pháp tương quan tự động đã được so sánh với các phương pháp khác, theo thí nghiệm của Wesszka và về mặt lý thuyết của Harlow thì cả hai công trình nghiên cứu đã nhận thấy rằng các phương pháp tương quan tự động phân biệt yếu hơn những phương pháp GLC. Họ giải thích rằng điều này là do sự không thích hợp của mô hình kết cấu.

2.5.2 Phương pháp Tamura

Tamura đề xướng ra phương pháp tiếp cận với những đặc điểm kết cấu dựa trên sự nhận thức tri giác của con người, và đã xác định sáu đặc điểm kết cấu đó là: (độ thô, độ tương phản, hướng, đường nét(line-likeness), trạng thái đều đặn và độ ráp) và so sánh chúng với các thước đo tâm sinh lý của con người. Ba đặc điểm đầu tiên thu được những kết quả thành công và được sử dụng rộng rãi.

Độ thô có quan hệ trực tiếp tới phạm vi và những tỷ lệ lặp và được coi như là đặc điểm kết cấu cơ bản. Một ảnh sẽ chứa các kết cấu ở nhiều phạm vi, độ thô nhằm xác định kích thước lớn nhất mà tại đây tồn tại một kết cấu, thậm chí là một kết cấu rất nhỏ. Đầu tiên là tính trung bình tại mỗi điểm với lân cận có kích thước là luỹ thừa của 2.

Sau đó tại mỗi toạ độ lấy sự khác nhau giữa các cặp trung bình tương ứng trong vùng lân cận không trùng nhau

$$E_{k,h}(x,y) = |A_k(x+2^{k-1},y) - A_k(x - 2^{k-1},y)|$$

Với mỗi điểm chọn kích thước tốt nhất có thể với mỗi k cho E lớn nhất cả về hai hướng. Thước đo độ thô là trung bình $S_{opt}(x,y) = 2^{opt}$ trên khắp ảnh.

Độ tương phản nhằm đạt được vùng động của mức xám trong ảnh như là sự khác biệt của phân bố giữa màu đen và trắng. Trước tiên được đo bằng độ lệch tiêu chuẩn của mức xám, sau đó kurtosis $\propto_4$. Thước đo tương phản vì vậy được định nghĩa như sau:

$F_{con} = \sigma/(\alpha_4)^n$

Ở đây: $\alpha_4 = \mu_4/\sigma^4$

μ_4 moment thứ tư về giá trị trung bình và σ là giá trị biến thiên. Qua thực nghiệm n=1/4 cho sự thoả thuận chặt chẽ nhất với thước đo của con người.

Hướng là một thuộc tính toàn bộ trên cả một vùng. Đặc điểm được mô tả không nhằm mục đích phân biệt các hướng hoặc các mẫu khác nhau nhưng tổng độ của hướng. Hai mặt nạ đơn giản được sử dụng để phát hiện biên ảnh, với mỗi điểm góc và độ lớn được tính toán. Một biểu đồ H_d của những xác suất biên sau đó được tính toán bằng cách tính tất cả các toạ độ với đọ lớn lớn hơn ngưỡng và lượng tử hoá bằng góc biên. Biểu đồ sẽ phản ánh góc độ của hướng. Để rút ra một thước đo từ H_d, đỉnh của các cao độ được tính toán từ moment thứ hai của chúng.

Ảnh tamura là một khái niệm mà ở đó ta tính toán giá trị cho ba đặc điểm tại mỗi điểm ảnh và xử lý chúng như là sự phân bố không gian khớp nối giữa độ thô, độ tương phản và hướng, và như vậy những ảnh có thể được xem như sự phân bố không gian RGB và những đặc điểm kiểu biểu đồ màu được sử dụng. Nét độc đáo của kết cấu là giá trị tại mỗi điểm được tính toán thông qua một cửa sổ.

2.6 Những phương pháp Primitive

Phần này đề cập tới những phương pháp gồm kết cấu biên và các phương pháp hình thái học. Một số kết cấu ban đầu được sử dụng trong các phương pháp này có

phạm vi và hướng đặc trưng. Ví dụ những đường thẳng và những biên có hướng xác định tốt và phạm vi của đường thẳng được xác định bởi chiều rộng của nó. Như chúng ta đã thấy, các phương pháp điều hoà đồng thời đo phạm vi và những đặc điểm đặc trưng hướng, đặc biệt là có một mối quan hệ chặt chẽ giữa các phương pháp đầu tiên như Gabor và biến đổi Fourie. Về cơ bản, Gabor là một phần của biến đổi Fourie. Mặc dù vậy, sự phân biệt giữa các phương pháp này là rõ ràng. Những phương pháp Primitive đo những đặc điểm cục bộ, trong khi những phương pháp điều hoà lại đo những đặc điểm rời rạc thuộc không gian.

Những phương pháp Primitive cũng liên quan tới những phương pháp kết cấu cấu trúc nhưng những phương pháp kết cấu cấu trúc có khuynh hướng tạo nên độ phức tạp trong khi những phương pháp Primitive lấy mẫu kết cấu đơn giản.

2.6.1 Phương pháp Primitive đầu tiên (Early primitive)

Những bộ lọc không gian là cách trực tiếp nhất để đạt được những thuộc tính kết cấu ảnh. Những cố gắng trước đây định nghĩa những phương pháp này tập trung vào đo mật đọ biên trên một đơn vị diện tích. Những kết cấu mịn có xu hướng có mật độ biên trên một đơn vị diện tích cao hơn những kết cấu thô. Thước đo biên thường được tính toán bởi những mặt nạ biên đơn giản như Robert hoặc Lplace. Thước đo biên có thể tính toán trên khắp vùng của ảnh bởi việc tính toán độ lớn từ đáp ứng của mặt nạ Robert hoặc Laplace. Hsu đưa ra phương pháp khác, đo cường độ điểm khác nhau giữa các điểm lân cận với cường độ không đổi, khoảng cách này được sử dụng như là thước đo mật độ biên.

Malik và Peroma đưa ra bộ lọc không gian để làm mẫu trước cảm nhận kết cấu trong hệ thống thị giác của con người. Những bộ lọc cân xứng thường được sử dụng bao gồm những khác nhau của độ lệch các hàm Gauss. Những phương pháp không tuyến tính là rất cần thiết để phân biệt những cặp kết cấu với độ sáng trung bình và số liệu thống kê thứ hai giống hệt nhau.Sự khám phá ra đường ranh giới kết cấu được thực hiện bằng những phương pháp phát hiện biên đơn giản. Những phương pháp này làm việc trên những mẫu kết cấu khác nhau và có thể phân biệt nét tự nhiên tốt như những kết cấu nhân tạo.

2.6.2 Phương pháp Gabor

Một trong số những phương pháp dựa trên xử lý tín hiệu số cho việc trích chọn những đặc điểm kết cấu đã trở thành tác dụng cho bộ lọc Gabor. Chúng có thể lọc trong miền tần số và không gian. Bộ lọc Gabor có thể được sử dụng để tạo mẫu cho những câu trả lời về hệ thống thuộc tri giác của con người. Turner trước tiên sử dụng

một danh sách các bộ lọc Gabor để phân tích kết cấu. Các bộ lọc ở các phạm vi khác nhau và hướng khác nhau cho phép lọc đa kênh của một ảnh để trích chọn thông tin tần số và hướng. Sau đó bộ lọc có thể được sử dụng để phân tích những hình ảnh trong những đặc điểm kết cấu.

Đặc điểm được tính toán bằng cách lọc ảnh với một dãy các bộ lọc hướng và tính toán độ lệch chuẩn độ lệch trung bình của đầu ra trong phạm vi tần số. Việc lọc ảnh I(x,y) với bộ lọc Gabor được phác hoạ như sau:

$$W_{mn}(x,y) = \int I(x,y) g_{mn} * (x - x_1, y - y_1) dx_1 dy_1$$

Độ lệch chuẩn và độ lệch trung bình của đại lượng $|W_{mn}|$ được sử dụng cho đặc điểm véc tơ. Đầu ra của bộ lọc ở những phạm vi khác nhau sẽ cho các vùng khác nhau. Vì lý do này mà mỗi thành phần của đặc điểm véc tơ được chuẩn hoá bằng cách sử dụng độ lệch chuẩn.

CHƯƠNG 3: PHƯƠNG PHÁP PHÂN TÍCH KẾT CẤU MẦU

3.1 Phương pháp Color auto-corrlegram

3.1.1 Giới thiệu:

Như đã trình bày phần trên, biểu đồ màu là một trong các kỹ thuật quan trọng trong việc tra cứu ảnh dựa trên nội dung. Biểu đồ màu cũng đại diện cho xác suất của một điểm ảnh bất kỳ trong ảnh thuộc màu C_i và nó được tính như sau:

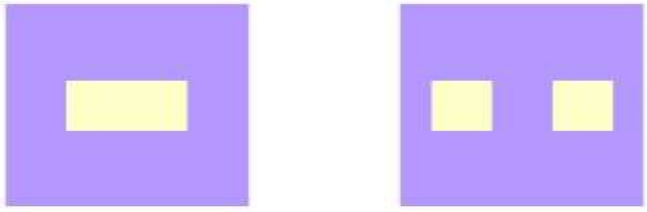
$$Pr(P \epsilon Ci) = h_i /m*n$$

Biểu đồ màu dễ tính toán, nó chỉ cần duyệt qua ảnh một lần và vì vậy độ phức tạp của nó là $O(n^2)$. Màu sắc là một trong số những đặc điểm trực quan nhất, bởi vậy trong một số trường hợp hiệu quả của việc sử dụng biểu đồ màu để tìm kiếm và tra cứu là khá tốt. Tuy nhiên điểm hạn chế của phương pháp biểu đồ màu truyền thống là không có bất kỳ thông tin về không gian, cho dù biểu đồ màu cục bộ đã cải tiến một phần. Có một số kỹ thuật đã được đưa ra để tích hợp thông tin không gian với biểu đồ màu, color auto-Correlogram là một trong những kỹ thuật này. Chúng ta xem xét vấn đề sau: Lấy một điểm ảnh bất kỳ P_1có màu C_i trong ảnh, với khoảng cách K tính từ P_1 lấy điểm ảnh P_2, xác suất để P_2 cũng có màu C_i là gì?

Auto-Correlogram của ảnh I cho mà C_i với khoảng cách k được định nghĩa:

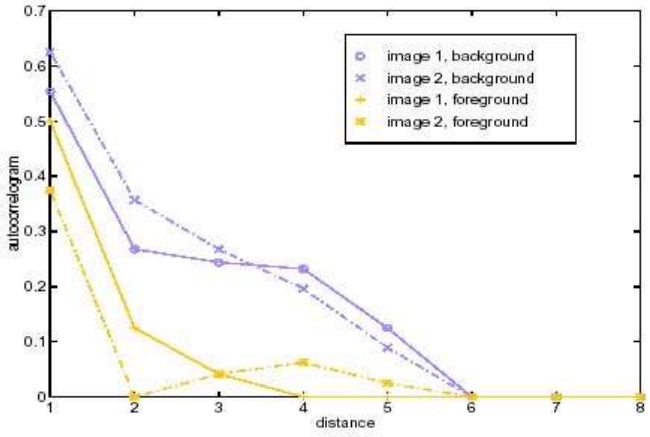
$$Y^{(k)}_c (I) \equiv Pr[| p_1 - p_2| = k, p_2 \epsilon Ic_i \mid p_1 \epsilon Ic_i]$$

Vì vậy, auto-Correlogram chỉ ra mối tương quan tự động giữa không gian của màu thay đổi so với khoảng cách như thế nào. Ví dụ ta xét 2 ảnh trong hình 5.1:



Hình 3.1: Hai ảnh tương tự

Dễ nhận thấy rằng biểu đồ của 2 ảnh này là giống nhau hoàn toàn, nhưng auto-correlogram của chúng sẽ khác nhau như hình 3.2.



Hình 3.2: Auto-Correlogram của hai ảnh trong hình 3.1

Auto-Correlogram tích hợp thông tin màu và thông tin không gian. Đối với mỗi điểm ảnh, phương pháp này cần phải duyệt qua tất cả các láng giềng của điểm ảnh đó. Vì vậy độ phức tạp tính toán của nó là $O(k*n^2)$ với k là số điểm ảnh láng giềng, nó phụ thuộc vào sự lựa chọn khoảng cách. Độ phức tạp tính toán sẽ tăng nhanh khi k lớn (k<= cỡ của ảnh).

3.1.2 Thước đo khoảng cách điểm ảnh

Như đã trình bày, độ phức tạp tính toán của phương pháp auto-Correlogram liên quan trực tiếp tới sự lựa chọn khoảng cách điểm ảnh. Nếu khoảng cách lớn thì hiệu quả của việc tìm kiếm sẽ tốt hơn bởi tập trung nhiều thông tin nhưng điều đó dẫn tới độ phức tạp tăng. Khoảng cách giữa hai điểm sẽ được tính như sau:

$$D_k(p,q) = \max(|p_x - q_x|,|p_y - q_y|)$$

Đây là khoảng cách lớn nhất theo trục x và trục y**.**

3.1.3 Những đặc điểm thước đo khoảng cách

Cho hai ảnh, sự khác nhau giữa hai đặc điểm khoảng cách là độ tương tự của hai ảnh này. Những đặc điểm thường được coi như là những vector vì vậy sự khác nhau chính là khoảng cách giữa hai vector này. Định nghĩa khoảng cách theo thuật ngữ của Euclidean nhưng khoảng cách tuyệt đối là không phù hợp tính toán này. ví dụ có hai điểm ảnh thuộc hai ảnh (a, a') và (b, b'), đặc điểm của hai ảnh này là f(a)=1000, f(a')=1050, f(b)=100, f(b')=150 khi đó khoảng cách tuyệt đối trong hai trường hợp là như nhau nhưng sự khác nhau trong trường hợp thứ hai là đáng kể. Vì vậy, thước đo khoảng cách được sử dụng ở đây sẽ là:

$$d_{(r,s)} = r\text{-}s \,/\, 1 + r + s$$

Số 1 trong mẫu số được thêm vào để tránh phép chia cho 0.

Đối với đặc điểm của biểu đồ màu thì khoảng cách là:

$$|I\text{-}I'|_h = \sum_{i\epsilon[m]} |h_{ci}(I) - h_{ci}(I')| / 1 + h_{ci}(I) + h_{ci}(I')|$$

Đố với đặc điểm auto-Correlogram khoảng cách sẽ là:

$$|I\text{-}I'|_y = \sum_{i\epsilon[m],\, k\epsilon[d]} |Y^{(k)}_{Ci}(I) - Y^{(k)}_{Ci}(I')| / 1 + Y^{(k)}_{Ci}(I) + Y^{(k)}_{Ci}(I')$$

3.2 Phương pháp ma trận đồng mức xám Co-occurrence Matrix

3.2.1 Mô tả những đặc điểm

Như đã được mô tả trong 3.2 những đặc điểm có thể được trích chọn từ ma trận Co-occurrence để giảm kích thước của không gian đặc điểm. Điều này rất quan trọng cho công nghệ tra cứu ảnh dựa trên nội dung bởi vì nó có thể giảm đáng kể độ phức tạp tính toán. Trong phần này định nghĩa hình thức về các đặc điểm từ ma trận này được cung cấp

Đặc điểm kết cấu	Công thức
Energy	$\sum_i \sum_j P_d^{\,2}(i,j)$
Entropy	$-\sum_i \sum_j P_d(i,j) \log \sum_i \sum_j P_d(i,j)$
Contrast	$\sum_i \sum_j (i-j)^2 P_d(i,j)$
Inverse Difference Moment	$\sum_{i,j} 1/1+(i-j)^2 P_d(i,j)$
Cluster Shade	$\sum_{i,j}((i-\mu_i) + (j-\mu_j))^3 P_d(i,j)$
Cluster Prominence	$\sum_{i,j}((i-\mu_i) + (j-\mu_j))^4 P_d(i,j)$
Correlation	$\sum_i, \sum_j (i-\mu_i)(j-\mu_j) P_d(i,j) / \boldsymbol{\delta_i \delta_j}$
Haralick's Correlation	$\sum_i, \sum_j (ij) P_d(i,j) - \mu_x \mu_y / \boldsymbol{\delta} x \boldsymbol{\delta} y$

Trong đó:

$P_d(i,j)$ Là phần tử thứ (i,j) của ma trận co-occurrence P_d

$\sum_i$ Nghĩa là : $\sum_{i=1}$ với M là số hàng

$\sum_j$ Nghĩa là: $\sum_{j=1}$ với N là số cột

$\sum_{i,j}$ Nghĩa là : $\sum_i, \sum_j$

μ_i Được định nghĩa : $\mu_i = \sum_i i \sum P_d(i,j)$

μ_j Được định nghĩa : $\mu_j = \sum_j j \sum_i P_d(i,j)$

σ_i Được định nghĩa : $\boldsymbol{\delta_i} = \sum_i (i - \mu_i)^2 \sum_j P_d(i,j)$

σ_j Được định nghĩa : $\boldsymbol{\delta_j} = \sum_j (j - \mu_j)^2 \sum_i P_d(i,j)$

μ_x, μ_y Là tổng hàng và cột tương ứng

σ_x, σ_i Là độ lệch tiêu chuẩn của hàng và cột tương ứng.

Thứ nhất energry của kết cấu mô tả sự tương tự của kết cấu. Trong ảnh đồng nhất có rất ít chuyển đổi mức xám trội, bởi vậy ma trận co-occurrence sẽ có ít vùng có cường độ lớn. Như vậy energry của ảnh là cao khi ảnh là đồng nhất.

Mô tả thứ hai entropy đo sự ngẫu nhiên của những phần tử trong ma trận khi tất cả những phần tử của ma trận là ngẫu nhiên tối đa thì entropy có giá trị cao nhất. Bởi vậy một ảnh đồng nhất có entropy thấp hơn ảnh không đồng nhất.

Đặc điểm thứ ba có giá trị cao tương đối khi những giá trị cao của ma trận gần với đường chéo chính. Điều này bởi vì $(i-j)^2$ sẽ nhỏ dần khi càng gần đường chéo chính và làm tăng giá trị của : $1 / 1+(i-j)^2$

Đặc điểm này cho hiệu quả đối nhau, khi các giá trị cao của ma trận càng xa đường chéo chính thì giá trị của inertia trở lên cao hơn. Như vậy inertia và inverse difference moment là đo sự phân bố của mức xám trong ảnh.

Cluster shade và cluster prominence là đo đối xứng lệch của ma trận, khi những giá trị này cao thì ảnh không đối xứng, khi giá trị này thấp thì có một đỉnh nhọn xung quanh giá trị trung bình, nghĩa là có sự biến thiên thấp về cấp độ xám.

Đặc điểm correlation đo tương quan giữa các phần tử của ma trận. Khi giá trị này cao thì ảnh phức tạp hơn.

Đặc điểm cuối cùng là Haralik's corrlation. Đặc điểm này đo sự phụ thuộc tuyến tính giữa những điểm ở những vị trí được xác định liên quan với mỗi điểm khác. So sánh với correlation chuẩn đặc điểm này tác động mạnh hơn với độ phức tạp của ảnh.

3.2.2 Thực hiện cải tiến việc tính toán ma trận Co-occerrence

Một cách đơn giản để tính toán ma trận co-occerrence là duyệt mỗi điểm ảnh P_i của mức xám i và đếm tất cả các điểm P_j của mức xám j tại khoảng cách bất kỳ d'=(d,∞) với $|P_i-P_j|$=d' cho mọi i và j có thể. Độ phức tạp tính toán là O(nmb[d']) với n x m là cỡ của ảnh, b là số mức xám, d' là số khoảng cách được sử dụng.

Thuật toán của chúng tôi đưa ra sử dụng định nghĩa của ma trận thay vì nó duyệt tuần tự mỗi mức xám, nó đếm đồng sự kiện của mỗi điểm với mức xám i và j tại khoảng cách d'. Ảnh chỉ được duyệt qua một lần, với thuật toán cũ mất b[d'] lượt. Với mỗi điểm trong ảnh, những điểm tại bốn khoảng cách d' được xét và chúng được lưu trữ trong ma trận co-occurrence trong mỗi lần lặp.

Cho I là một ảnh xám và (δi_α,δj_α) là véc tơ thay thế trong ảnh I để có được điểm ở khoảng cách d'=(1,α) từ điểm I[i,j]. Gọi C là ma trận co-occurrence cho tất cả bốn góc (0^0, 45^0, 90^0, 135^0) và các phần tử được đặt bằng 0. Sau đó thuật toán của chúng tôi thực hiện như sau:

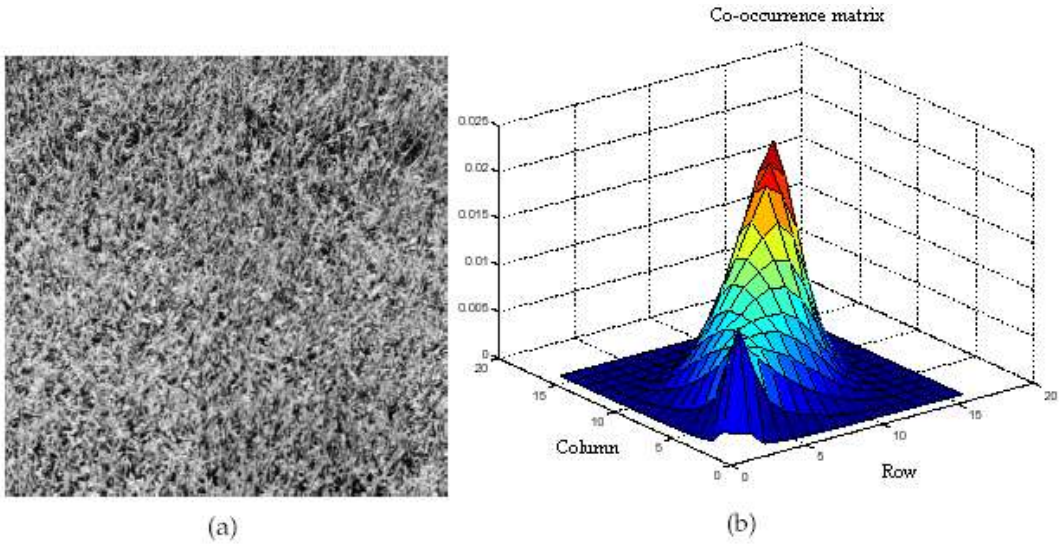
For(i=0; i< imagewidth; i^{++})

For (j=0; j< imagehight; j^{++})

Foreach $\alpha \in \{0^0, 45^0, 90^0, 135^0\}$

C[I[i,j], I[i+δi_α ,j+δj_α]]+=1

Thuật toán này có độ phức tạp tính toán là O(nm). Hình 3.3 là ảnh gốc và kết quả tính toán ma trận co-occurrence của ảnh này



Hình 3.3: a) Ảnh gốc,

b) Biểu đồ của đặc điểm ma trận co-occurrence được tính toán

CHƯƠNG 4: CÀI ĐẶT CHƯƠNG TRÌNH THỬ NGHIỆM

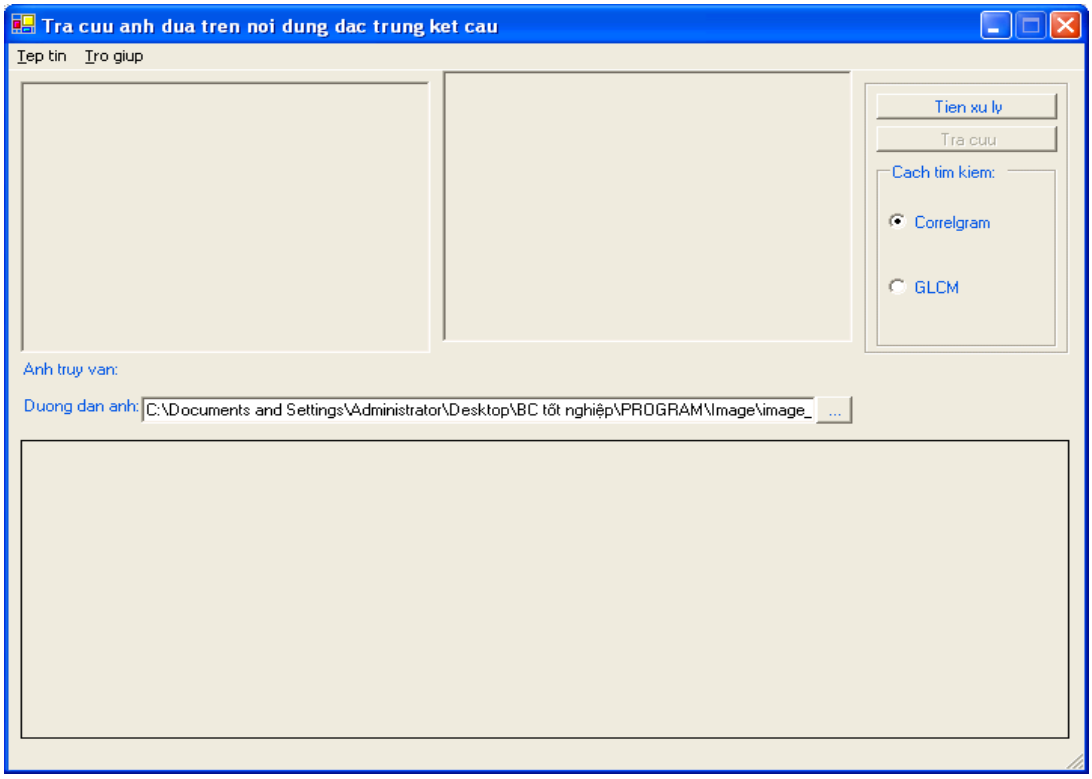
4.1 Môi trường thực nghiệm

Chương trình được lập trình trên Visual studio 2003 C# , một thư viện liên kết tĩnh. DLL của Intel hỗ trợ cho việc truy cập ảnh jpeg.

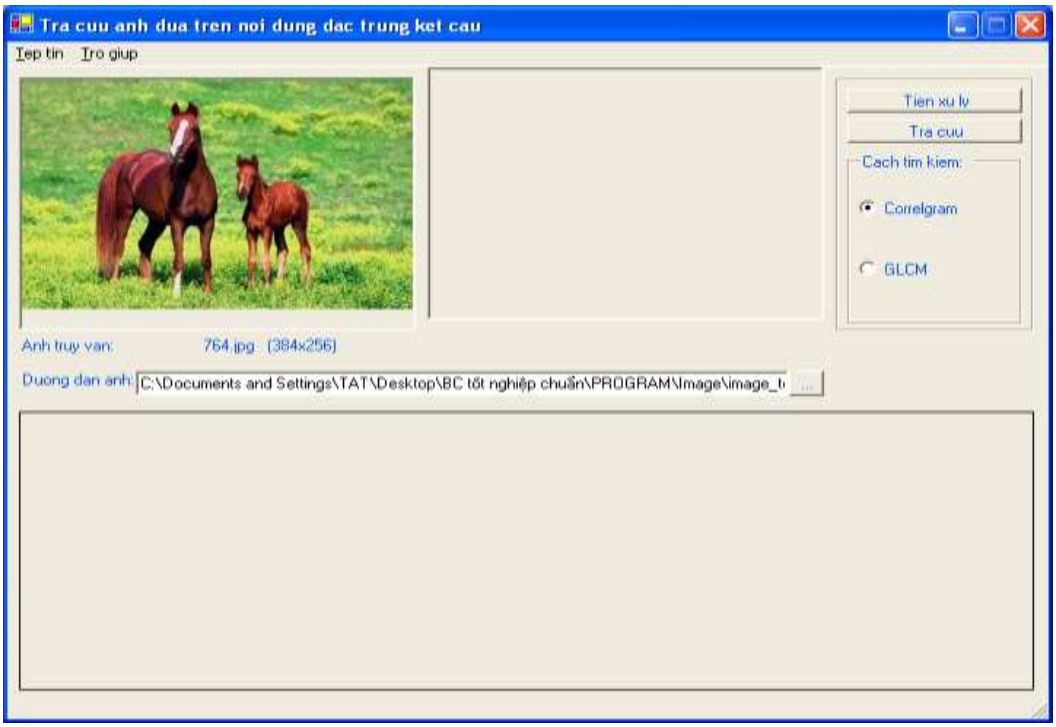
Chương trình tập chung vào nội dung sử dụng đặc trưng kết cấu là nhiều, vì vậy là phần khả thi và có thể ứng dụng được vào trong thực tế với tốc độ và kết quả.

4.2 Kết quả thử nghiệm

4.2.1 Giao diện chương trình



4.2.2 Chọn ảnh cần tìm kiếm



4.2.3 Kết quả tìm kiếm ảnh hoàn thiện

KẾT LUẬN

Chương trình chạy tìm kiếm ảnh theo nội dung kết cấu. Việc tìm kiếm ảnh với tiêu chí kết cấu màu có kết quả tốt, phù hợp với thị giác con người. Một hệ thống tính toán lược đồ màu đã làm cho việc tìm kiếm ảnh dựa vào kết cấu màu đạt kết quả tốt.

- Tìm hiểu được phương pháp tra cứu ảnh theo nội dung. .
- Tìm hiểu được một số phương pháp tra cứu ảnh theo nội dung.
- Xây dựng được chương trình thử nghiệm

Bên cạnh còn có những mặt hạn chế:

- Chưa áp dụng được các kỹ thuật trích chọn đặc điểm khác để so sánh.
- Chưa áp dụng được các kỹ thuật đánh chỉ số để tăng hiệu năng của chương trình.

Có thể phát triển thêm để ứng dụng được vào thực tế:

- Giúp tìm hàng hóa trong siêu thị dựa vào hình dáng
- Tìm ảnh trên mạng
- Xác định ảnh cần tìm

TÀI LIỆU THAM KHẢO

[1.] Nâng cao hiệu quả của các thuật toán nhận dạng ảnh, luận văn PTS khoa học Ngô Quốc Tạo, 1996.

[2.] Nhận dạng các phương pháp và ứng dụng, Hoàng Kiếm, Nguyễn Ngọc Ký và các tác giả, Nhà xuất bản thông kê 7/1992

[3.] Tra cứu ảnh tương tự dựa vào đặc điểm kết cấu(ma trận GLCM) (Tiểu án K7), trường ĐHDL Hải Phòng.

[4.] Tài liệu lý thuyết xử lý ảnh trên trang web http://cuasoit.com

[5.] Selim Akshoy and Robert M.Haralick "Graph-Theoretical Clustering for Image grouping and Retrieval"

[6.] Shengjiu Wang (2001),"A Robust CBIR Approach Using Local Color Histogram", Technique Repost TR 01-13, Edmonton, Alberta, Canada.