

LỜI CẢM ƠN

Trong lời đầu tiên của báo cáo Đồ án Tốt Nghiệp “Tìm hiểu phương pháp trích chọn dấu hiệu của ảnh dựa vào đặc trưng hình dạng” này, em muốn gửi những lời cảm ơn và biết ơn chân thành nhất của mình tới tất cả những người đã hỗ trợ, giúp đỡ em về kiến thức, và tinh thần trong quá trình thực hiện Đồ án.

Trước hết, em chân thành cảm ơn Thầy Giáo. Ths. Ngô Trường Giang, Giảng viên Khoa Công Nghệ Thông Tin, Trường ĐHDL Hải Phòng, người đã trực tiếp hướng dẫn, nhận xét, giúp đỡ em trong suốt quá trình thực hiện Đồ án.

Em chân thành cảm ơn các Thầy Cô trong Khoa Công Nghệ Thông Tin và toàn Thầy Cô trong Trường Đại Học Dân Lập Hải Phòng đã giúp đỡ em trong suốt quá trình học tập.

Em cũng xin bày tỏ lòng biết ơn đến gia đình và những người bạn đã luôn giúp đỡ động viên em rất nhiều trong quá trình học tập và làm Đồ án Tốt Nghiệp.

Do thời gian thực hiện có hạn, kiến thức còn nhiều hạn chế nên Đồ án thực hiện chắc chắn không tránh khỏi những thiếu sót nhất định. Em rất mong nhận được ý kiến đóng góp của Thầy Cô giáo và các bạn để em có thêm kinh nghiệm và tiếp tục hoàn thiện bài báo cáo của mình.

Em xin chân thành Cảm ơn!

Hải Phòng, tháng 11/2011

Sinh viên

Phùng Thị Lệ

MỤC LỤC

MỤC LỤC	2
MỞ ĐẦU	4
DANH MỤC HÌNH VẼ	5
CHƯƠNG 1: TỔNG QUAN VỀ TRA CỨU ẢNH DỰA TRÊN NỘI DUNG.	6
1.1 Tra cứu thông tin trực quan.....	6
1.2 Các thành phần chính của một hệ thống CBIR.....	7
1.3 Các chức năng cơ bản của hệ thống CBIR.	8
1.4 Một số phương pháp tra cứu ảnh dựa trên nội dung.	9
1.5 Một số hệ thống tra cứu ảnh dựa trên nội dung.	10
1.5.1 Hệ thống QBIC (Query By Image Content).	11
1.5.2 Hệ thống Photobook.....	11
1.5.3 Hệ thống VisualSEEK và WebSEEK.	11
1.5.4 Hệ thống RetrievalWare.....	12
1.5.5 Hệ thống Imatch.....	12
CHƯƠNG 2: CÁC PHƯƠNG PHÁP BIỂU DIỄN HÌNH DẠNG.....	13
2.1 Giới thiệu.	13
2.2 Kỹ thuật biểu diễn hình dạng dựa trên biên.	14
2.2.1 Phương pháp toàn cục.	14
2.2.2 Phương pháp cấu trúc.....	16
2.3 Kỹ thuật biểu diễn hình dạng dựa trên vùng.	18
2.3.1 Phương pháp toàn cục.	19
2.3.2 Phương pháp cấu trúc.....	20
CHƯƠNG 3: MỘT SỐ PHƯƠNG PHÁP TRÍCH CHỌN DẤU HIỆU CỦA ẢNH DỰA VÀO ĐẶC TRƯNG HÌNH DẠNG.....	22
3.1 Giới thiệu.	22
3.2 Phương pháp trích chọn đặc trưng dựa trên đường biên.	22
3.2.1 Mã xích (chain code).....	22
3.2.2 Shape number.....	24
3.2.3 Đối sánh các shape number.....	25

3.3 Phương pháp trích chọn đặc trưng dựa trên vùng.....	27
3.3.1 Đồ thị xương.	29
3.3.2 Đối sánh đồ thị xương.....	29
3.3.3 Nhận xét	36
CHƯƠNG 4: THỰC NGHIỆM.	37
4.1 Môi trường thử nghiệm.	37
4.2 Một số kết quả thu được.	37
4.2.1 Giao diện chương trình:	37
4.2.2 Kết quả trên một số đối tượng khác nhau.	38
4.2.3 Một số nhận xét về chương trình.	40
KẾT LUẬN.....	41
TÀI LIỆU THAM KHẢO	42

MỞ ĐẦU

Trong thực tế, bài toán tra cứu ảnh có rất nhiều ứng dụng quan trọng, ví dụ như so sánh mẫu vân tay, tìm kiếm ảnh tội phạm... được áp dụng trong ngành khoa học hình sự, cơ sở dữ liệu ảnh về địa lý, y học.... làm cho lĩnh vực nghiên cứu này phát triển nhanh trong công nghệ thông tin. Tuy nhiên khi số lượng ảnh được lưu trữ trở nên rất lớn thì vấn đề là phải có những phương pháp tổ chức cơ sở dữ liệu ảnh tốt cùng với những kỹ thuật tìm kiếm, tra cứu ảnh hiệu quả, có độ chính xác cao và có hiệu năng tốt. Việc tìm kiếm được một bức ảnh mong muốn trong hàng triệu bức ảnh là rất khó khăn, đòi hỏi phải có những phương pháp hiệu quả và chính xác. Một trong những phương pháp được nhiều người quan tâm nghiên cứu hiện nay là phương pháp “Tra cứu ảnh dựa theo nội dung” (Content-Based Image Retrieval). Ý tưởng phương pháp này là trích chọn các đặc điểm dựa vào nội dung trực quan của ảnh như màu sắc, kết cấu, hình dạng và bố cục không gian của ảnh để làm cơ sở cho việc tra cứu, sắp xếp, tổ chức cơ sở dữ liệu ảnh.

Tra cứu ảnh dựa trên hình dạng sử dụng các đặc trưng hình dạng của các ảnh mục tiêu để tìm kiếm, nó là một khía cạnh rất quan trọng của tra cứu ảnh dựa trên nội dung. Một phần quan trọng của tra cứu ảnh dựa trên hình dạng là nghiên cứu trích chọn dấu hiệu đặc trưng. Cho đến nay vẫn chưa có, định nghĩa toán học chính xác, bao gồm cả hình học, thống kê, hình thái học đo lường về hình dạng, do đó việc mô tả hình dạng là vấn đề hết sức phức tạp. Đề tài này sẽ tìm hiểu một số phương pháp mô tả và trích chọn dấu hiệu của ảnh dựa vào đặc trưng hình dạng ứng dụng trong tra cứu ảnh dựa trên nội dung. Đồ án bao gồm 4 chương:

- Chương 1: Trình bày tổng quan về tra cứu ảnh dựa trên nội dung và giới thiệu một số hệ thống tra cứu ảnh dựa trên nội dung.
- Chương 2: Trình bày một số phương pháp biểu diễn hình dạng và cách phân loại các phương pháp biểu diễn hình dạng.
- Chương 3: Tìm hiểu hai phương pháp trích chọn dấu hiệu của ảnh dựa vào đặc trưng hình dạng: Mã xích và xương.
- Chương 4: Trình bày thực nghiệm và một số kết quả đạt được.

DANH MỤC HÌNH VẼ

Hình 1.1: Các thông tin sử dụng để mô tả hình ảnh.....	7
Hình 1.2: Mô hình hệ thống tra cứu ảnh dựa trên nội dung.....	9
Hình 2.1: Phân loại các kỹ thuật mô tả hình dạng.	14
Hình 2.2: Minh họa độ lệch tâm và tuần hoàn của hình dạng.	15
Hình 2.3: Minh họa phương pháp bề mặt lỗi: (a) Bề mặt lỗi và các thiếu hụt lỗi của nó; (b) Cây lỗi biểu diễn bề mặt lỗi.	20
Hình 2.4: Trục trung vị (xương) của hình chữ nhật.	20
Hình 3.1: Các hướng của đoạn thẳng đơn vị: (a): 4 hướng, (b): 8 hướng.	23
Hình 3.2: Biểu diễn của một chuỗi mã (theo 4 hướng và 8 hướng).....	23
Hình 3.3: Biểu diễn hình dạng sử dụng shape number.	24
Hình 3.4: Các bước tính toán shape number.	25
Hình 3.5: Minh họa tìm kiếm hình dạng tương tự sử dụng shape number: (a) hình dạng; (b) cây tương tự; (c) ma trận tương tự.	26
Hình 3.6: Hình dạng (a) và (b) tương tự nhau nhưng đồ thị khác nhau.....	28
Hình 3.7: Hình dạng (a) và (b) khác nhau nhưng có đồ thị xương (c) giống nhau.	28
Hình 3.8: Minh họa đường dẫn xương: (a) xương của hình con ngựa, (b) đường dẫn ngắn nhất giữa các cặp node cuối.	30
Hình 3.9: Sự tương ứng giữa các node cuối của hai đồ thị xương.....	32
Hình 3.10: Minh họa xương được cắt tỉa bởi DCE.	33
Hình 3.11: Sự tương ứng giữa các yếu tố.	35
Hình 4.1: Giao diện của chương trình.....	37
Hình 4.2: Kết quả thu được với hình con ngựa.....	38
Hình 4.3: Kết quả thu được với hình con ngựa kéo xe.	38
Hình 4.4: Kết quả thu được với hình cá heo.	39
Hình 4.5: Kết quả thu được với hình chữ nhật.....	39

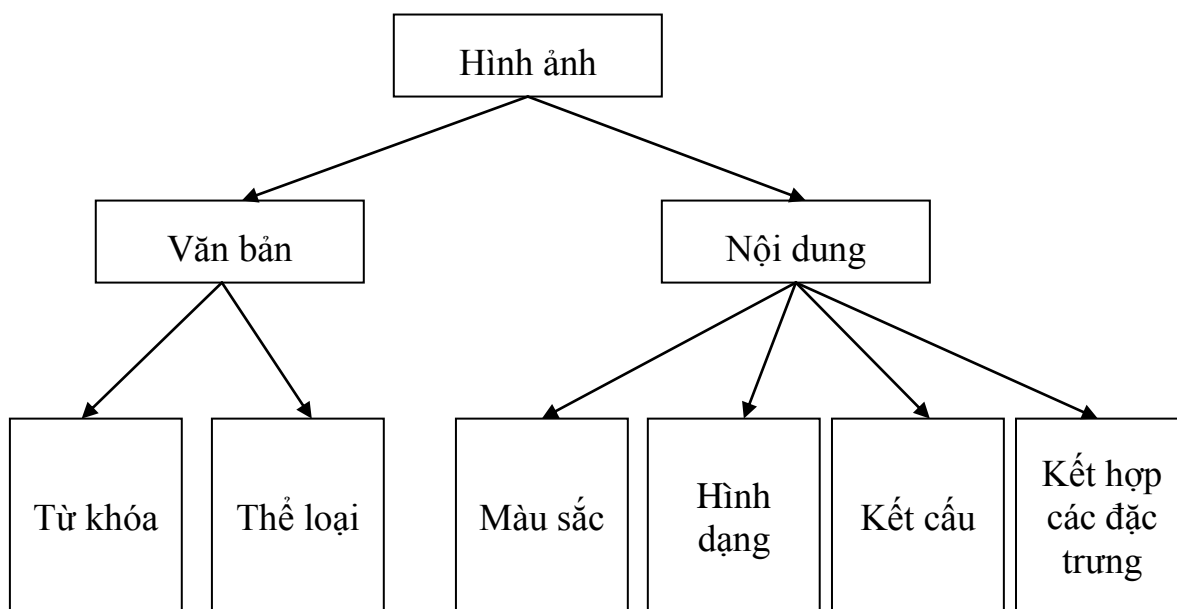
CHƯƠNG 1: TỔNG QUAN VỀ TRA CỨU ẢNH DỰA TRÊN NỘI DUNG.

1.1 Tra cứu thông tin trực quan.

Nội dung của một bức ảnh gồm nội dung trực quan và nội dung ngữ nghĩa. Nội dung trực quan gồm có nội dung tổng quan và nội dung đặc tả. Nội dung ngữ nghĩa được phát hiện thông qua chú thích hoặc suy luận từ nội dung trực quan.

Có hai phương pháp chung để giải bài toán tra cứu thông tin trực quan dựa trên những thông tin trực quan: Phương pháp dựa trên những thuộc tính (tổng quan) và phương pháp dựa trên những đặc điểm (đặc tả). 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 sự can thiệp của con người để trích chọn metadata về đối tượng trực quan và sự chú thích kết cấu. Nhưng việc phân tích kết cấu 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 từ khóa 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 khóa 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 trực quan này.

Thuật ngữ tra cứu ảnh dựa trên nội dung (Content Based Image Retrieval - CBIR) 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 các hình ảnh. 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 [5]. Tra cứu ảnh dựa trên việc gán từ khóa (manually assigned keywords) 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.



Hình 1.1: Các thông tin sử dụng để mô tả hình ảnh.

1.2 Các thành phần chính của một hệ thống CBIR.

Hiện nay, trọng tâm chính của CBIR là nghiên cứu chủ yếu trên 3 chủ đề chính: Trích chọn đặc trưng, đánh chỉ số hiệu quả, giao diện người dùng.

– *Trích chọn đặc trưng*: Các đặc trưng của hình ảnh bao gồm các đặc trưng nguyên thủy và các đặc trưng ngữ nghĩa /đặc trưng logic. Đặc trưng nguyên thủy như màu sắc, hình dạng, kết cấu và các mối quan hệ không gian được định lượng trong tự nhiên, chúng có thể được trích xuất tự động hoặc bán tự động. Đặc trưng logic cung cấp mô tả trừu tượng của dữ liệu hình ảnh ở các cấp độ khác nhau. Thông thường, các đặc trưng logic được chiết xuất bằng tay hoặc bán tự động. Một hoặc nhiều đặc trưng có thể được sử dụng trong một ứng dụng cụ thể. Ví dụ, trong một hệ thống thông tin vệ tinh, các đặc trưng kết cấu là quan trọng, trong khi hình dạng và màu sắc là các đặc trưng quan trọng hơn trong các hệ thống đăng ký nhãn hiệu hàng hoá. Một hoặc nhiều đặc trưng đã được chiết xuất, tra cứu trở thành một nhiệm vụ đo độ giống nhau giữa các đặc trưng hình ảnh.

– *Lập chỉ số hiệu quả*: Để tạo điều kiện truy vấn hiệu quả và xử lý tìm kiếm, các chỉ số hình ảnh cần thiết được tổ chức thành các cấu trúc dữ liệu hiệu quả. Đặc trưng hình ảnh là sự hình dung miêu tả, chúng không thể nhúng chỉ thị và chúng có thể có các thuộc tính liên quan đến nhau. Vì vậy, cấu trúc dữ liệu linh hoạt nên được

sử dụng để tạo điều kiện thuận lợi cho lưu trữ và truy xuất trong hệ thống phục hồi hình ảnh. Các cấu trúc như k-d-tree, R-tree family, R*-tree, quad-tree, và grid file (tập lưới) thường được sử dụng. Các cấu trúc trên có các ưu điểm và nhược điểm riêng vì vậy một số cấu trúc dữ liệu chỉ có thể sử dụng trong những lĩnh vực nhất định nhưng một số cấu trúc dữ liệu có được sử dụng đồng thời trong nhiều lĩnh vực khác nhau.

– *Giao diện người dùng*: Trong các hệ thống thông tin trực quan, tương tác người dùng đóng một vai trò quan trọng trong hầu như tất cả các chức năng. Giao diện người dùng bao gồm một bộ xử lý truy vấn và trình duyệt để cung cấp các công cụ đồ họa tương tác, cơ chế truy vấn và truy cập cơ sở dữ liệu, theo thứ tự định sẵn. Các cơ chế truy vấn thông thường được cung cấp bởi giao diện người dùng là: truy vấn bằng từ khóa, truy vấn bằng cách phác thảo, truy vấn bằng các ví dụ, duyệt theo thể loại, chọn lọc đặc trưng và trọng số thu hồi, tinh chỉnh và phù hợp thông tin phản hồi.

Ba tác vụ trên là ba thành phần chính của một hệ thống CBIR. Trong số ba tác vụ trên, trích chọn đặc trưng (bao gồm cả đo độ giống nhau) là nhiệm vụ quan trọng và khó khăn nhất. Phần lớn các nghiên cứu CBIR đi vào nhiệm vụ đầy thách thức này. Nghiên cứu này tập trung vào khai thác các đặc trưng hình dạng.

1.3 Các chức năng cơ bản của hệ thống CBIR.

Một hệ thống tra cứu ảnh dựa trên nội dung không chỉ xử lý các nguồn thông tin trong những định dạng khác nhau (ví dụ như văn bản, ảnh, video) mà còn xử lý các nhu cầu của người sử dụng. Về cơ bản, hệ thống phân tích các nội dung của nguồn thông tin cũng như truy vấn của người sử dụng, rồi sau đó đối sánh chúng để tìm ra những tiêu chí có liên quan [5]. Những chức năng chính của một hệ thống bao gồm các nội dung sau:

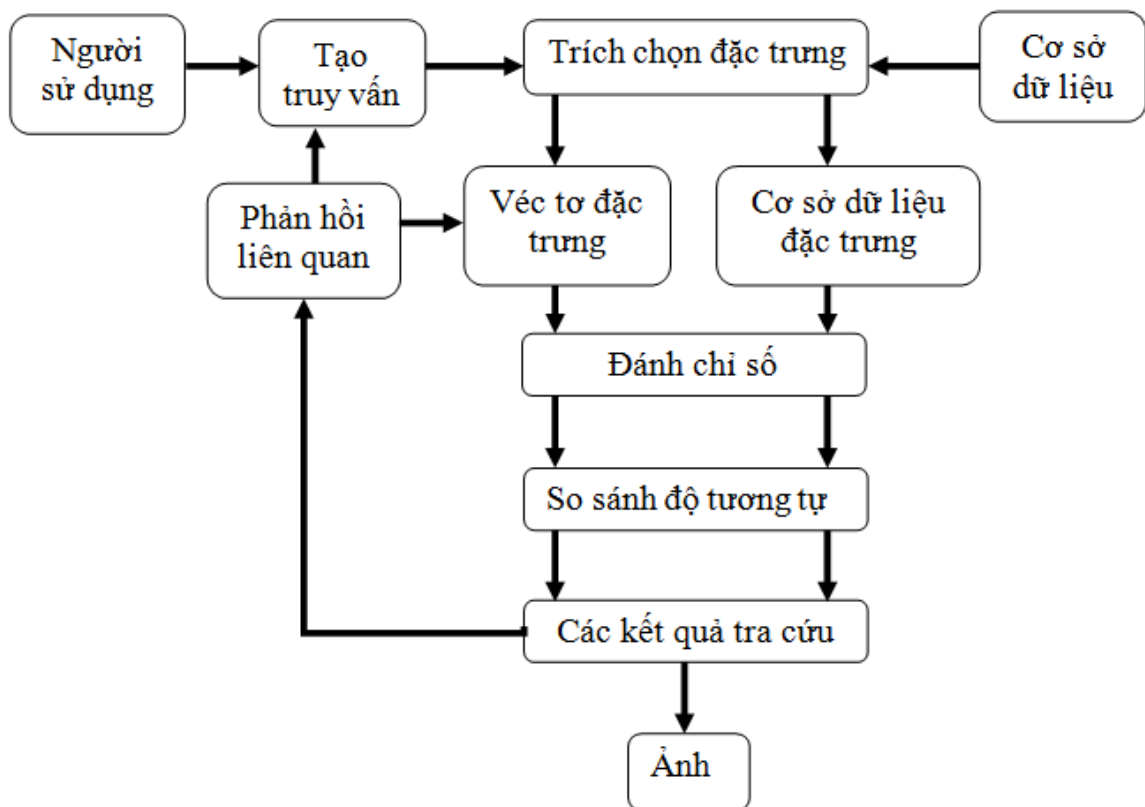
– Phân tích nội dung của nguồn thông tin và biểu diễn nội dung của các nguồn thông tin sao cho phân tích phù hợp với sự so sánh các truy vấn của người sử dụng (không gian của thông tin nguồn được chuyển đổi thành không gian đặc điểm với mục đích đối sánh nhanh trong bước tiếp theo). Bước này thường là mất nhiều thời gian cho việc xử lý tuần tự các thông tin nguồn (hình ảnh) trong cơ sở dữ liệu. Tuy nhiên, bước này 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 đối sánh với cơ sở dữ 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 đại có thể được sử dụng để tổ chức lại không gian đặc điểm để tăng tốc độ xử lý đối sánh.

– Thực hiện các điều chỉnh cần thiết trong hệ thống (thường là bằng cách đổi chiều các tham số trong công nghệ đối sánh) dựa trên phản hồi từ người sử dụng hoặc những hình ảnh được tra cứu.

Từ sự trình bày ở trên ta thấy rằng một mặt hệ thống tra cứu ảnh dựa trên nội dung có các nguồn thông tin trực quan ở các dạng khác nhau, bên cạnh đó lại có cả các yêu cầu của người sử dụng. Mô hình một hệ thống tra cứu ảnh dựa trên nội dung [2] như hình 1.2:



Hình 1.2: Mô hình hệ thống tra cứu ảnh dựa trên nội dung.

1.4 Một số phương pháp tra cứu ảnh dựa trên nội dung.

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. Các phương pháp này chủ yếu thông qua việc nhấn mạnh vào tra cứu ảnh với những đặc điểm mong muốn từ một tập hình ảnh lớn. Việc lựa chọn các đặc trưng và độ đo thích hợp sẽ giúp tăng tốc độ thực hiện và

mức độ chính xác của hệ thống tra cứu ảnh. Vì vậy phải lựa chọn được một tập các đặc trưng tốt nhất cho đầu vào của các hệ thống tìm kiếm ảnh. Nếu số lượng các đặc trưng quá nhiều sẽ che khuất các tín hiệu, mặt khác, nếu số lượng các đặc trưng quá ít sẽ khó nhận dạng được ảnh [1]. Sau đây, một số phương pháp tra cứu ảnh dựa trên nội dung được giới thiệu:

– *Tra cứu ảnh dựa trên màu sắc*: màu sắc là một đặc trưng nổi bật và được sử dụng phổ biến nhất trong tìm kiếm ảnh theo nội dung. Mỗi một điểm ảnh (thông tin màu sắc) có thể được biểu diễn như một điểm trong không gian màu sắc ba chiều. Các không gian màu thường dùng là: RGB, Munsell, CIE, HSV. Tìm kiếm ảnh theo màu sắc tiến hành tính toán biểu đồ màu cho mỗi ảnh để xác định tỉ trọng các điểm ảnh chứa các giá trị đặc biệt (màu sắc). Các nghiên cứu gần đây đang tập trung vào phân vùng ảnh theo các màu khác nhau và tìm mối quan hệ giữa các vùng này.

– *Tra cứu ảnh dựa trên kết cấu*: Trích xuất nội dung ảnh theo kết cấu nhằm tìm ra mô hình trực quan của ảnh và cách thức chúng được xác định trong không gian. Kết cấu cung cấp thông tin về sự sắp xếp về mặt không gian của màu sắc và cường độ của ảnh. Kết cấu được biểu diễn bởi các texel mà sau đó được đặt vào một số các tập phụ thuộc vào số kết cấu được phát hiện trong ảnh. Các tập này không chỉ xác định các kết cấu mà còn chỉ rõ vị trí các kết cấu trong ảnh. Việc xác định các kết cấu đặc biệt trong ảnh đạt được chủ yếu bằng cách mô hình các kết cấu như những biến thể xám hai chiều.

– *Tra cứu ảnh dựa trên hình dạng*: hình dạng của một ảnh hay một vùng là một đặc trưng quan trọng trong việc xác định và phân biệt ảnh trong nhận dạng mẫu. Mục tiêu chính của biểu diễn hình dạng trong nhận dạng mẫu là đo thuộc tính hình học của một đối tượng được dùng trong phân lớp, so sánh và nhận dạng đối tượng. Hình dạng là đặc trưng hình ảnh quan trọng và nó là một trong những đặc trưng nguyên thủy để mô tả nội dung hình ảnh. Tuy nhiên, mô tả nội dung hình dạng là một nhiệm vụ rất khó khăn. Bởi vì, rất khó để định nghĩa các nhận thức về đặc trưng hình dạng và đo lường sự giống nhau giữa các hình dạng.

1.5 Một số hệ thống tra cứu ảnh dựa trên nội dung.

Những năm gần đây, có nhiều hệ thống tra cứu ảnh đã được xây dựng và phát triển rất nhanh. Một số hệ thống của CBIR được biết tới:

1.5.1 Hệ thống QBIC (Query By Image Content).

Hệ thống QBIC của hãng IBM là một hệ thống tra cứu ảnh thương mại đầu tiên và nổi tiếng nhất trong số các hệ thống tra cứu ảnh dựa trên nội dung. Nó cho phép người sử dụng tra cứu ảnh dựa vào màu sắc, hình dạng và kết cấu. QBIC cung cấp một số phương pháp: Simple, Multi-feature, và Multi-pass. Trong phương pháp truy vấn Simple chỉ sử dụng một đặc điểm. Truy vấn Multi-feature bao gồm nhiều hơn một đặc điểm và mọi đặc điểm đều có trọng số như nhau trong suốt quá trình tìm kiếm. Truy vấn Multi-pass sử dụng đầu ra của các truy vấn trước làm cơ sở cho bước tiếp theo. Người sử dụng có thể vẽ ra và chỉ định màu, kết cấu mẫu của hình ảnh yêu cầu. Trong hệ thống QBIC màu tương tự được tính toán bằng thước đo bình phương sử dụng biểu đồ màu k phần tử (k-element) và màu trung bình được sử dụng như là bộ lọc để cải tiến hiệu quả của truy vấn. Bản demo của QBIC tại địa chỉ www.qbic.almaden.ibm.com.

1.5.2 Hệ thống Photobook.

Hệ thống này được phát triển ở Massachusetts Institute of Technology cho phép người sử dụng tra cứu ảnh dựa trên màu sắc, hình dạng và kết cấu. Hệ thống này cung cấp một tập các thuật toán đối sánh gồm: Euclidean, mahalanobis, vector space angle, histogram, Fourier peak, và wavelet tree distance như là những đơn vị đo khoảng cách. Trong hầu hết các phiên bản đã có thể định nghĩa những thuật toán đối sánh của họ. Hệ thống như là một công cụ bán tự động và có thể sinh ra một mẫu truy vấn dựa vào những ảnh mẫu được cung cấp bởi người sử dụng. Điều này cho phép người sử dụng trực tiếp đưa những yêu cầu truy vấn của họ với những lĩnh vực khác nhau, và mỗi lĩnh vực họ có thể thu được những mẫu truy vấn tối ưu.

1.5.3 Hệ thống VisualSEEK và WebSEEK.

Cả hai hệ thống này đều được phát triển tại Trường Đại học Colombia. VisualSEEK là hệ thống cơ sở dữ liệu ảnh. Nó cho phép người sử dụng tra cứu ảnh dựa trên màu sắc, không gian miền và đặc điểm kết cấu. Tập màu và chuyển đổi wavelet dựa trên kết cấu được sử dụng để thực hiện những đặc điểm này. Thêm vào đó VisualSEEK còn cho phép người sử dụng tạo truy vấn bằng việc chỉ định vùng màu và những không gian vị trí của chúng. WebSEEK là một catalog ảnh và là công cụ tìm kiếm cho web. Hệ thống này cung cấp mẫu cho danh sách ảnh và video trên trang web sử dụng kết hợp xử lý dựa trên text và phân tích dựa trên nội dung.

1.5.4 Hệ thống RetrievalWare.

Hệ thống này được phát triển bởi tập đoàn công nghệ Excalibur cho phép người sử dụng tra cứu ảnh bởi nội dung màu, hình dạng, kết cấu, độ sáng, kết cấu màu và hệ số co. Người sử dụng có thể điều chỉnh tỷ trọng của những đặc điểm này trong suốt quá trình tìm kiếm.

1.5.5 Hệ thống Imatch.

Hệ thống này cho phép người sử dụng tra cứu ảnh bởi nội dung màu, hình dạng, và kết cấu. Nó cung cấp một số phương pháp để tra cứu ảnh tương tự: Màu tương tự, màu và hình dạng (Quick), màu và hình dạng (Fuzzy) và sự phân bố màu. Màu tương tự truy vấn những ảnh tương tự với ảnh mẫu dựa trên sự phân bố màu toàn cục. Màu và hình dạng (Quick) tìm hình ảnh tương tự bởi việc kết hợp cả hình dạng, kết cấu và màu. Màu và hình dạng (Fuzzy) thực hiện thêm những bước xác định đối tượng trong ảnh mẫu. Phân bố màu cho phép người sử dụng vẽ ra sự phân bố màu hoặc xác định tỷ lệ phần trăm của một màu trong hình ảnh mong muốn. Imatch cũng cung cấp những đặc điểm khác nội dung để xác định ảnh: ảnh nhị phân, ảnh co kích thước, lưu trữ trong những định dạng khác và những ảnh có tên tương tự.

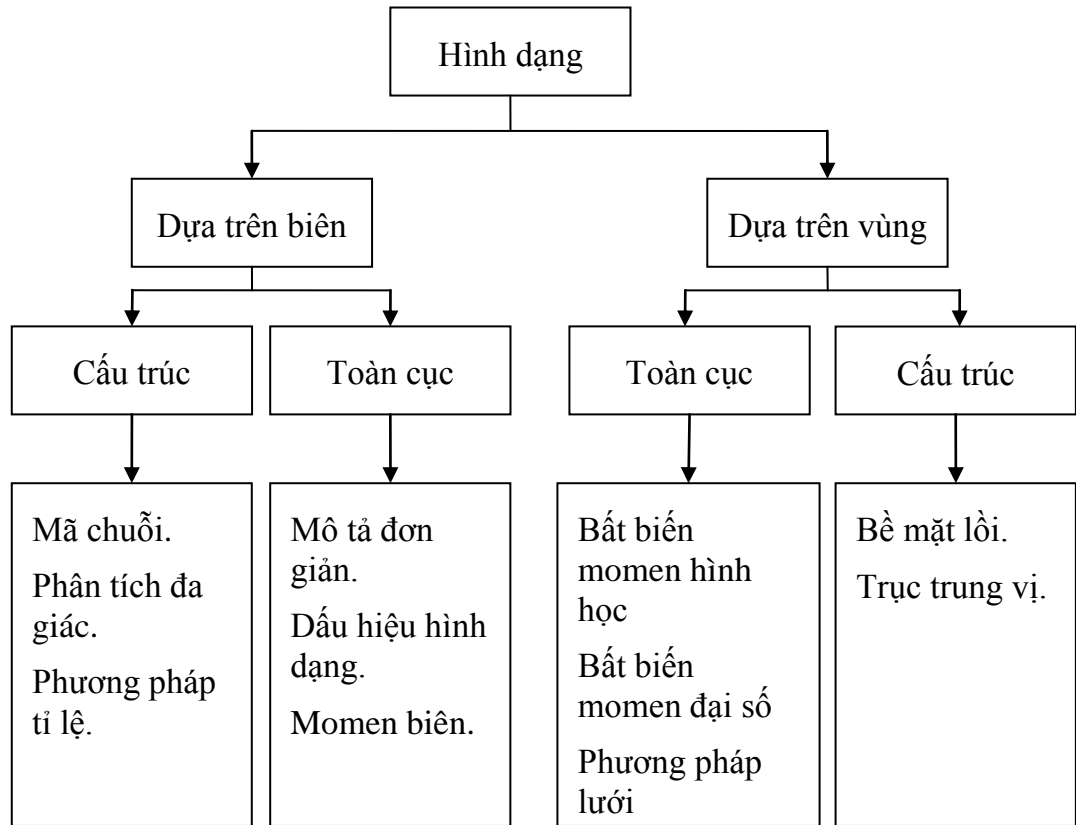
CHƯƠNG 2: CÁC PHƯƠNG PHÁP BIỂU DIỄN HÌNH DẠNG.

2.1 Giới thiệu.

Nghiên cứu về hình dạng đã hoạt động trong hơn 30 năm. Trước đây, nghiên cứu hình dạng được thúc đẩy chủ yếu bởi sự nhận dạng đối tượng, các kỹ thuật mô tả và biểu diễn hình dạng này chủ yếu dựa vào các ứng dụng cụ thể. Trong đó, sự hiệu quả và chính xác là mối quan tâm chính của những kỹ thuật này. Khi các ứng dụng đa phương tiện mới nổi lên, vấn đề quan trọng đặt ra là tra cứu trực tuyến. Trong sự phát triển của MPEG-7, sáu yêu cầu chính được thiết lập để đánh giá một mô tả hình dạng, đó là: trích chọn hiệu quả và chính xác, đặc trưng cô đọng, ứng dụng rộng rãi, độ phức tạp thấp, hiệu suất cao và phân cấp mô tả tốt [5].

Việc phân loại các phương pháp biểu diễn hình dạng phổ biến nhất là dựa trên việc sử dụng các điểm biên hình dạng và điểm vùng. Biểu diễn hình dạng cũng có thể được phân biệt giữa miền không gian và miền đặc trưng. Phương pháp trong miền không gian so sánh các hình dạng dựa trên điểm (hoặc điểm đặc trưng) cơ sở, còn phương pháp miền đặc trưng so sánh các hình dạng dựa trên đặc trưng (vector) cơ sở. Một cách phân loại các kỹ thuật biểu diễn hình dạng khác là dựa trên cơ sở bảo quản thông tin. Phương pháp cho phép xây dựng lại chính xác một hình dạng từ mô tả của nó được gọi là lưu trữ thông tin (Information preserving - IP), còn phương pháp chỉ có khả năng xây dựng lại một phần hoặc mô tả không rõ ràng được gọi là sự không lưu trữ thông tin (Non Information preserving - NIP).

Các phương pháp biểu diễn hình dạng được phân loại theo các cấp bậc, đầu tiên phương pháp phân loại dựa trên đường biên và phương pháp phân loại dựa trên vùng căn cứ vào đặc trưng hình dạng được trích chọn từ đường biên hay toàn bộ các phân vùng hình dạng. Trong mỗi lớp, các phương pháp khác nhau được tiếp tục phân biệt thành cấu trúc và toàn cục dựa vào việc hình dạng được biểu diễn theo toàn bộ hay theo các thành phần con [5]. Sau đó, tiếp tục phân ra các phương pháp cụ thể như mô tả trong sơ đồ (2.1).



Hình 2.1: Phân loại các kỹ thuật mô tả hình dạng.

2.2 Kỹ thuật biểu diễn hình dạng dựa trên biên.

Kỹ thuật mô tả hình dạng dựa trên biên chỉ khai thác thông tin trên biên. Có hai loại phương pháp tiếp cận rất khác nhau cho kỹ thuật dựa trên biên: phương pháp tiếp cận liên tục (toàn cục) và phương pháp tiếp cận rời rạc (cấu trúc). Phương pháp tiếp cận liên tục không phân chia hình dạng thành các phần và một vector đặc trưng có gốc từ đường biên được sử dụng để mô tả hình dạng. Thước đo sự giống nhau về hình dạng là dựa trên sự đối sánh các điểm đặc biệt hoặc dựa trên đặc trưng. Phương pháp tiếp cận rời rạc chia đường biên thành các phân đoạn bằng cách sử dụng một tiêu chuẩn cụ thể. Biểu diễn cuối cùng thường là một chuỗi hoặc một đồ thị (hoặc cây), các biện pháp tương tự được thực hiện bằng cách kết hợp chuỗi hoặc đồ thị một cách phù hợp.

2.2.1 Phương pháp toàn cục.

Kỹ thuật mô tả hình dạng dựa trên đường biên toàn cục thường tính toán vectơ đặc trưng từ thông tin đường biên. Khi đối sánh giữa các hình dạng sử dụng

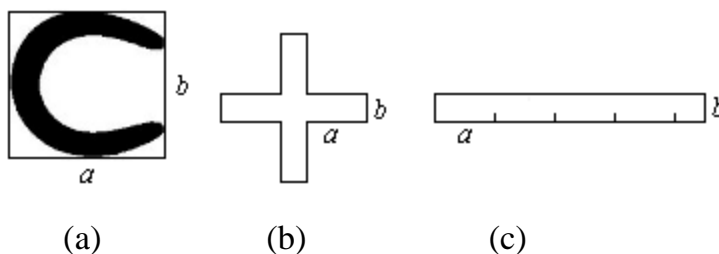
một khoảng cách theo quy tắc, chẳng hạn như khoảng cách Euclide hoặc khoảng cách City block.

Phương pháp mô tả biểu diễn hình dạng bằng đường biên toàn cục là mô tả biểu diễn toàn bộ hình dạng có hệ thống cho một mô tả mẫu. Đối sánh giữa các hình dạng có thể thực hiện được trong miền không gian hoặc trong miền đặc trưng. Đối với một mô tả hình dạng thì luôn cần sự cân bằng giữa độ chính xác và hiệu quả. Một mặt, hình dạng cần được mô tả chính xác, mặt khác, hình dạng mô tả nên càng nhỏ gọn càng tốt để đơn giản hóa quá trình đánh chỉ số và tra cứu. Bởi vậy chúng ta cần phải trích chọn đặc trưng một cách hiệu quả.

Mô tả hình dạng toàn cục đơn giản nhỏ gọn, tuy nhiên mô tả hình dạng không chính xác, nó chỉ có thể được kết hợp với mô tả hình dạng khác để tạo ra các mô tả hình dạng chính xác.

2.2.1.1 Mô tả hình dạng đơn giản.

Mô tả toàn cục đơn giản có thể được biểu diễn thông qua: vùng, tuần hoàn (chu vi²/diện tích), độ lệch tâm (độ dài trục chính/độ dài trục nhỏ), hướng trục chính và khả năng uốn. Những mô tả đơn giản toàn cục thường chỉ có thể phân biệt hình dạng với khác biệt lớn, do đó sẽ thường sử dụng các bộ lọc để loại bỏ các truy cập sai hoặc kết hợp với mô tả hình dạng khác để phân biệt hình dạng. Phương pháp không phù hợp với mô tả hình dạng độc lập. Ví dụ, lệch tâm của hình dạng trong hình 2.2(a) là gần tới 1 ($a=b$), nó không chính xác để mô tả hình dạng. Trong trường hợp này, tuần hoàn là một mô tả tốt hơn. Hai hình dạng trong hình 2.2(b) và 2.2(c) có cùng tuần hoàn ($a=2b$), tuy nhiên, chúng là những hình dạng rất khác nhau. Trong trường hợp này, độ lệch tâm là mô tả tốt hơn.



Hình 2.2: Minh họa độ lệch tâm và tuần hoàn của hình dạng.

2.2.1.2 Dấu hiệu hình dạng (Shape Signature).

Dấu hiệu hình dạng (SS) mô tả hình dạng bởi hàm một chiều thu được từ điểm biên. SS bao gồm: tọa độ phức hợp, tọa độ cực, khoảng cách tâm, góc tiếp tuyến, góc quỹ tích, độ cong, diện tích và chiều dài dây cung.

SS không bị tác động bởi dịch chuyển và co giãn hình dạng. Bên cạnh đó, SS có thể được lượng tử hóa thành một biểu đồ dấu hiệu, biểu đồ này bất biến với phép quay và có thể sử dụng cho đối sánh. SS thường nhạy cảm với nhiễu, những thay đổi nhỏ trên biên có thể gây ra những lỗi lớn trong đối sánh. Vì vậy, SS không thực tế và không hiệu quả trong tra cứu hình dạng.

2.2.1.3 Momen biên (Boundary Moment).

Momen biên (BM) có thể được dùng để giảm kích thước của các biểu diễn biên. Giả sử biên đã được biểu diễn như một SS $Z(i)$, momen thứ r là m_r và momen tâm là μ_r , có công thức ước tính:

$$m_r = \frac{1}{N} \sum_{i=1}^N [z_i]^r$$

Và

$$\mu_r = \frac{1}{N} \sum_{i=1}^N [z_i - m_i]^r$$

Trong đó, N là số các điểm biên.

Chuẩn hóa các momen: $m_r = \frac{m_r}{(M_2)^{r/2}}$

Và $M_r = \frac{M_r}{(M_2)^{r/2}}$

Để mô tả bất biến với các phép dịch chuyển, phép quay và co giãn.

2.2.2 Phương pháp cấu trúc.

Một phương pháp khác trong phân tích hình dạng là biểu diễn hình dạng cấu trúc. Với cách tiếp cận cấu trúc, hình dạng được chia thành các phân đoạn biên được gọi là đối tượng ban đầu. Phương pháp cấu trúc khác nhau trong việc lựa chọn đối tượng ban đầu và tổ chức các đối tượng này cho việc biểu diễn hình dạng. Phân tích biên thường dựa trên xấp xỉ đa giác, phân tích đường cong và đối sánh các đường cong.

Điều đạt được của tiếp cận cấu trúc đó là có khả năng giải quyết sự bế tắc trong chuỗi hoạt động liên tục và cho phép đối sánh từng phần, tuy nhiên, nó vẫn còn một vài điều hạn chế.

2.2.2.1 Biểu diễn bằng mã xích (chain code).

Mã xích mô tả đường biên đối tượng bằng một chuỗi các đoạn thẳng đơn vị với các hướng đã được xác định. Nền tảng này đã được giới thiệu vào năm 1961 bởi Freeman, ông đã mô tả một phương pháp cho phép mã hóa các cấu hình hình học theo ý muốn. Trong phương pháp này, một đường cong bất kỳ được biểu diễn bởi một chuỗi các vector đơn vị chiều dài và thiết lập một giới hạn các hướng cho phép, do đó gọi là phương pháp vector đơn vị. Trong thực hiện, một hình ảnh được đặt chồng lên một lưới, từ đó các điểm biên lấy xấp xỉ với điểm lưới gần nhất, sau đó lấy mẫu của hình ảnh thu được. Từ một điểm khởi đầu được lựa chọn trên biên, một mã xích có thể được tạo ra bằng cách mã hóa các đoạn thẳng biểu diễn biên. Các đoạn thẳng đơn vị có thể định hướng theo 4 hướng, 8 hướng hoặc N hướng (với $N > 8$ và $N = 2k$), mã xích sử dụng đoạn thẳng đơn vị định hướng theo N hướng được gọi là mã xích tổng quát.

Mã xích dùng để biểu diễn hình dạng phải không phụ thuộc vào sự lựa chọn điểm ảnh biên bắt đầu trong chuỗi. Một khả năng để chuẩn hóa chuỗi mã xích là tìm các điểm ảnh trong trình tự biên mà kết quả mô tả là các số nguyên tối thiểu, sau đó chúng được sử dụng như là các điểm ảnh bắt đầu. Ngoài ra, biên có thể được biểu diễn bởi sự khác biệt về các chỉ thị tiếp theo trong chuỗi mã thay vì biểu diễn cho biên theo chỉ số tương đối. Sự chuẩn hóa sự khác biệt chuỗi mã được gọi là shape number, shape number sẽ được sử dụng để biểu diễn hình dạng đối tượng (phần này sẽ được trình bày cụ thể trong mục 3.2.2).

Dùng mã xích biểu diễn hình dạng và đối sánh có nhiều hạn chế, mã xích bị ảnh hưởng nhiều đường biên và biến dạng, thêm vào đó là kích thước của chuỗi mã dài. Mã xích mà thường được sử dụng là đầu vào của những phân tích ở mức độ cao, ví dụ như xấp xỉ đa giác và tìm điểm uốn.

2.2.2.2 Phân tích đa giác (Polygon Decompositon).

Trong phương pháp này, đường biên được chia nhỏ thành các đoạn bởi xấp xỉ đa giác. Các đỉnh đa giác được sử dụng như một đối tượng ban đầu. Đặc trưng của mỗi đối tượng ban đầu được mô tả như một chuỗi bao gồm 4 yếu tố: góc nội tiếp, khoảng cách đến đỉnh tiếp theo, các tọa độ x và y. Các đặc trưng này được tổ chức thành một cây nhị phân hoặc m-arytree. Đối sánh hình dạng có hai bước: Bước đầu tiên đối sánh đặc trưng với đặc trưng, bước thứ hai, đối sánh hình dạng với hình dạng. Trong bước đầu tiên, chúng ta thu được dữ liệu đặc trưng của các hình dạng truy vấn. Các đặc trưng này được tìm kiếm thông qua chỉ số cây, nếu một mẫu đặc

trung cụ thể trong cơ sở dữ liệu được tìm thấy tương tự như dữ liệu đặc trưng thì danh sách các hình dạng liên quan đến mô hình đặc trưng được lấy ra. Trong bước thứ hai, đối sánh giữa hình dạng truy vấn và mẫu thu được, việc đối sánh được thực hiện dựa vào khoảng cách biến đổi giữa hai chuỗi các đối tượng ban đầu.

2.2.2.3 Phương pháp không gian tỉ lệ (Scale Space method).

Dudek và Tsotsos phân tích hình dạng trong không gian tỉ lệ và sử dụng sơ đồ đối sánh mô hình với mô hình. Trong phương pháp này, trước tiên hình dạng gốc (nguyên thủy) thu được từ kỹ thuật làm mịn đường cong. Sau đó, thiết lập một mô tả đoạn bao gồm chiều dài phân đoạn, thứ tự vị trí và giá trị điều chỉnh độ cong được trích chọn từ mỗi hình dạng nguyên thủy. Cuối cùng, một chuỗi các mô tả đoạn được tạo ra để mô tả hình dạng. Ví dụ với hai hình dạng A và B được mô tả bởi hai chuỗi: $A = (S_1^A, S_2^A \dots, S_N^A)$ và $B = (S_1^B, S_2^B \dots, S_M^B)$, đối sánh mô hình với mô hình sử dụng lập trình động để thu được số điểm tương đồng của hai hình dạng. Để làm tăng hiệu quả trong quá trình tính toán đối sánh, chúng ta đưa các đặc trưng hình dạng vào không gian có độ cong tỉ lệ để hình dạng có thể được đối sánh ở các tỉ lệ khác nhau. Tuy nhiên, do trong mô tả đoạn có bao gồm chiều dài phân đoạn nên mô tả này bất biến với co giãn.

2.3 Kỹ thuật biểu diễn hình dạng dựa trên vùng.

Trong phương pháp biểu diễn dựa trên vùng phải kể đến tất cả những pixel trong vùng hình dạng thu được trong biểu diễn hình dạng. Phương pháp biểu diễn vùng thường sử dụng các momen để mô tả hình dạng. Và một số phương pháp khác thường sử dụng gồm có: phương pháp lưới, bề mặt lồi và trục trung vị.

Biểu diễn hình dạng dựa trên vùng xem xét đến toàn bộ vùng hình dạng và sử dụng hiệu quả thông tin của toàn bộ pixel chứa trong vùng. Những phương pháp này đo sự phân phối pixel của vùng hình dạng, chúng ít có khả năng giả tạo bởi nhiễu và biến dạng. Phương pháp vùng phổ biến là những phương pháp moment. Ở mức thấp thứ tự moment hay bất biến moment mang theo những ý nghĩa vật lý kết hợp với sự phân phối pixel. Tuy nhiên nó rất khó khăn để kết hợp thứ tự moment cao hơn với sự giải thích vật lý. Phương pháp lưới là dựa trên khả năng trực quan sát hình dạng, nó không phản ánh sự thống kê phân bố của vùng hình dạng và bị ảnh hưởng bởi nhiễu và không cô đọng như bất biến moment.

2.3.1 Phương pháp toàn cục.

Phương pháp toàn cục xem xét đến toàn bộ hình dạng, kết quả của mô tả là vector số đặc trưng (numeric feature vector), nó được sử dụng để biểu diễn hình dạng.

2.3.1.1 Bất biến momen hình học (Geometric Moment Invariant).

Hu đã công bố bài nghiên cứu đầu tiên về việc sử dụng các momen bất biến cho ứng dụng nhận dạng mẫu hai chiều. Phương pháp tiếp cận của ông dựa trên các nghiên cứu của các nhà toán học thế kỷ 19 và lý thuyết đại số:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \rho(x, y) dx dy \quad p, q = 0, 1, 2 \dots$$

Sử dụng kết hợp phi tuyến các momen có thứ tự thấp, một tổ hợp các bất biến momen (thường được gọi là momen hình học), trong đó các thuộc tính bất biến với co giãn và phép quay được rút ra. Việc sử dụng các momen có thứ tự cao cho phân tích mẫu không được áp dụng. Vấn đề chính với momen hình học là chỉ có một số bất biến được rút ra từ thứ tự thấp của momen, như vậy không đủ để mô tả chính xác hình dạng, nhưng cũng rất khó để lấy được những bất biến thứ tự cao hơn.

2.3.1.2 Bất biến moment đại số (Algebraic Moment Invariant).

Bất biến momen đại số (AMI) được Taubin và Cooper giới thiệu và sử dụng trong QBIC. Các AMI được tính toán cho từ m momen trung tâm đầu tiên và được đặt ra như là giá trị riêng của ma trận định trước $M[j,k]$, trong đó các phần tử tỉ lệ với các yếu tố của các momen trung tâm. Khác với phương pháp bất biến momen hình học của Hu, các bất biến momen đại số có thể được xây dựng từ các thứ tự bất kỳ. AMI có xu hướng làm việc tốt trên các đối tượng có điểm ảnh được phân bố và không phải là hình dạng phức tạp.

2.3.1.3 Phương pháp dựa trên lưới (Grid Based Method).

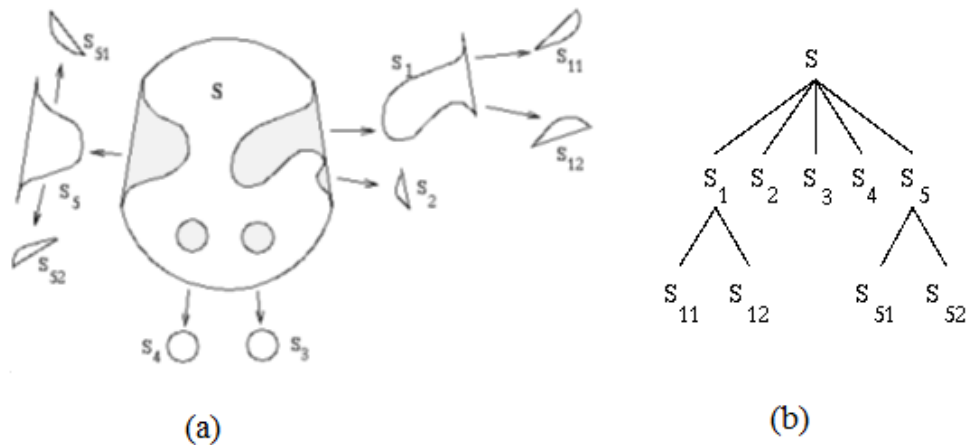
Lưới mô tả hình dạng được đề xuất bởi Lu và Sajjanhar, nó đã được sử dụng trong Mars và một số ứng dụng khác. Về cơ bản, hình dạng sẽ được chiếu lên một lưới có kích thước cố định, một chuỗi nhị phân mô tả hình dạng sẽ được tạo ra bằng cách quét lưới này từ trái sang phải và từ trên xuống và cho kết quả là một bitmap. Các ô bao phủ hình dạng được chỉ định giá trị 1, và những ô không bao phủ hình dạng được chỉ định giá trị 0. Khoảng cách Hamming hoặc khoảng cách cityblock được sử dụng để đo lường sự giống nhau giữa hai hình.

2.3.2 Phương pháp cấu trúc.

Cũng giống như phương pháp cấu trúc biên, cấu trúc vùng phân tích hình dạng vùng theo từng phần rời sử dụng chúng cho mô tả và biểu diễn hình dạng.

2.3.2.1 Bề mặt lồi (Convex Hull).

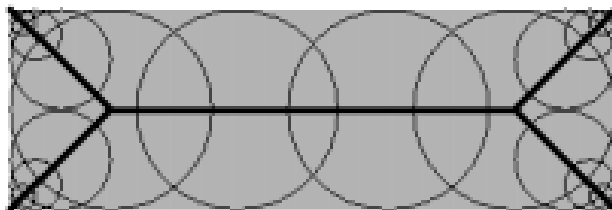
Một vùng R là lồi khi và chỉ khi với 2 điểm bất kỳ $x_1, x_2 \in R$ thì toàn bộ đoạn x_1x_2 nằm bên trong vùng. Bề mặt lồi của một vùng là vùng lồi H nhỏ nhất đáp ứng điều kiện $R \subset H$. Sự khác biệt của $R-H$ được gọi là thiếu hụt lồi của vùng R (convex deficiency). Đầu tiên, bề mặt lồi của một đối tượng thu được với các thiếu hụt lồi của nó, sau đó lại tìm bề mặt lồi và thiếu hụt lồi của các thiếu hụt lồi đã tìm thấy ở bước trước, quá trình tiếp tục cho đến khi các thiếu hụt lồi đều là các vùng lồi. Hình 2.3(a) minh họa quá trình này, và hình dạng được mô tả như một cây lốm trong hình 2.3(b).



Hình 2.3: Minh họa phương pháp bề mặt lồi: (a) Bề mặt lồi và các thiếu hụt lồi của nó; (b) Cây lốm biểu diễn bề mặt lồi.

2.3.2.2 Trục trung vị (Media Axis) (hay còn gọi là xương).

Cũng giống như bề mặt lồi, xương cũng có thể được sử dụng để mô tả và biểu diễn hình dạng. Xương (trục trung vị) là quỹ tích tâm của các đĩa cực đại của hình dạng như trong hình 2.4, đường in đậm là xương của hình chữ nhật.



Hình 2.4: Trục trung vị (xương) của hình chữ nhật.

Ý tưởng cơ bản của việc sử dụng xương là loại bỏ các thông tin dư thừa trong khi vẫn giữ được các thông tin topo có liên quan đến cấu trúc của đối tượng để có thể nhận dạng đối tượng. Xương có thể được phân tách thành các đoạn và được biểu diễn dưới dạng các đồ thị theo một tiêu chí nhất định. Như vậy, việc đối sánh giữa các hình dạng sẽ trở thành việc đối sánh giữa các đồ thị. Tuy nhiên việc tính toán đối với xương khá phức tạp, hơn nữa xương rất nhạy cảm với nhiễu và các biến dạng.

CHƯƠNG 3: MỘT SỐ PHƯƠNG PHÁP TRÍCH CHỌN DẤU HIỆU CỦA ẢNH DỰA VÀO ĐẶC TRƯNG HÌNH DẠNG.

3.1 Giới thiệu.

Về cơ bản, tra cứu hình ảnh dựa trên hình dạng là đo đặc sự tương tự giữa các biểu diễn hình dạng. Vì vậy, hai bước cần thiết trong tra cứu ảnh dựa trên hình dạng là trích chọn đặc trưng và đo sự giống nhau giữa các đặc trưng đó. Trong tìm kiếm ảnh theo nội dung, hình dạng là một cấp cao hơn so với màu sắc và kết cấu. 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: biểu diễn các đường biên bao ngoài.
- Biểu diễn theo vùng: biểu diễn trên từng vùng.

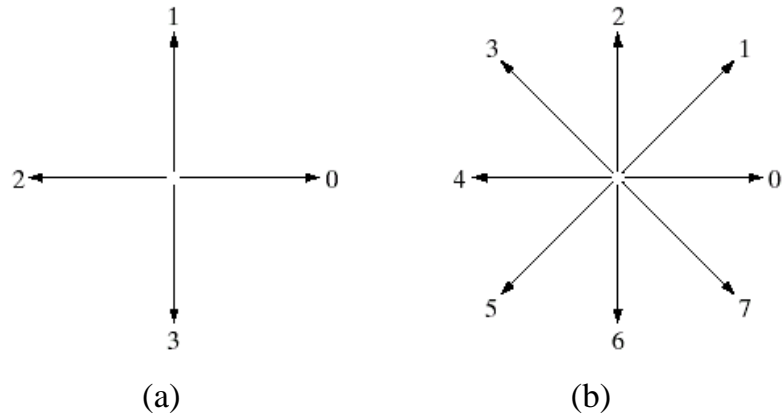
3.2 Phương pháp trích chọn đặc trưng dựa trên đường biên.

Như đã giới thiệu trong mục 2.2.2.1, một phương pháp điển hình trong biểu diễn hình dạng dựa trên đường biên là sử dụng mã xích. Trong phần này, chúng ta sẽ xem xét cụ thể phương pháp sử dụng mã xích cùng với shape number để biểu diễn và nhận dạng đối tượng.

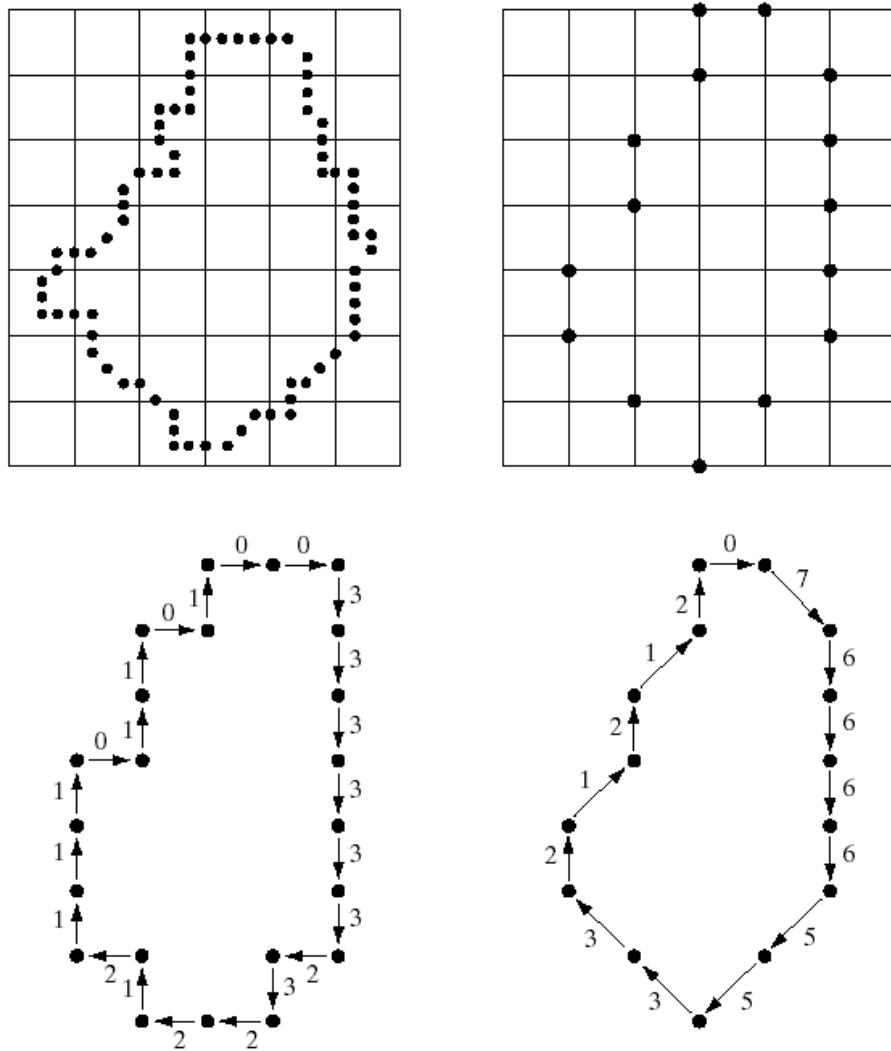
3.2.1 Mã xích (*chain code*).

Mã xích biểu diễn đường biên đối tượng bằng một chuỗi kết nối của các phân đoạn đường thẳng có độ dài quy định và định hướng [8]. Thông thường, biểu diễn này dựa trên 4 hoặc 8 hướng kết nối của các phân đoạn đường thẳng. Hướng của mỗi phân đoạn được mã hóa bằng cách sử dụng một lược đồ số như được hiển thị trong hình 3.1. Những hình ảnh kỹ thuật số thường được xử lý với định dạng lưới với khoảng cách bình đẳng với các hướng x và y . Một chuỗi mã có thể tạo ra bằng cách định hướng các phân đoạn đường thẳng dọc theo biên theo chiều kim đồng hồ như minh họa trong hình 3.2.

Vấn đề đặt ra là một chuỗi mã phụ thuộc vào điểm bắt đầu và giải pháp được đưa ra là coi chuỗi mã như một chuỗi kín và xác định điểm bắt đầu để chuỗi kết quả không phụ thuộc vào sự lựa chọn điểm bắt đầu đó. Chúng ta có thể chuẩn hóa mã xích với phép quay bằng cách sử dụng sự khác biệt đầu tiên (first difference) của mã xích thay vì bản thân mã. Sự khác biệt này thu được bằng cách đếm số lượng các hướng thay đổi (theo hướng ngược chiều kim đồng hồ) giữa 2 yếu tố liền kề.



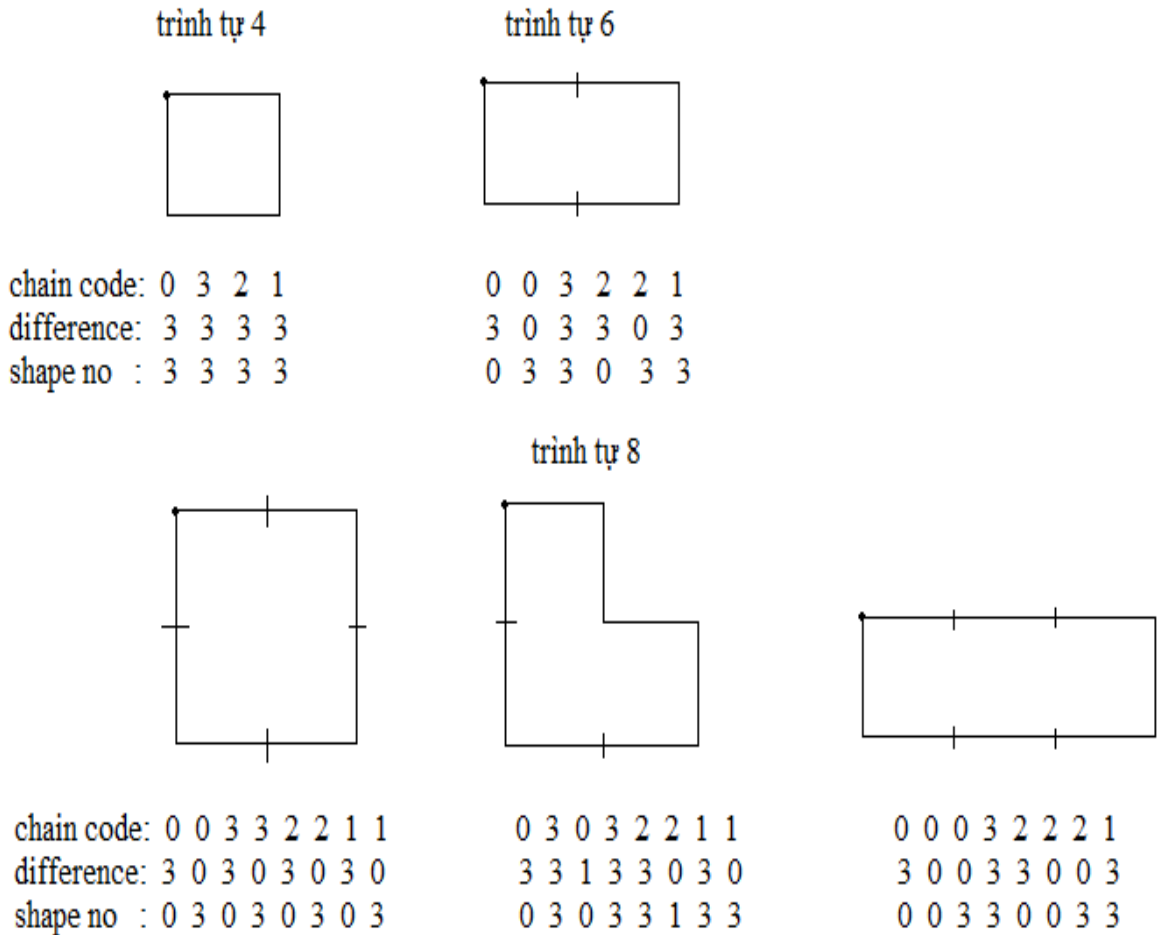
Hình 3.1: Các hướng của đoạn thẳng đơn vị: (a): 4 hướng, (b): 8 hướng.



Hình 3.2: Biểu diễn của một chuỗi mã (theo 4 hướng và 8 hướng)

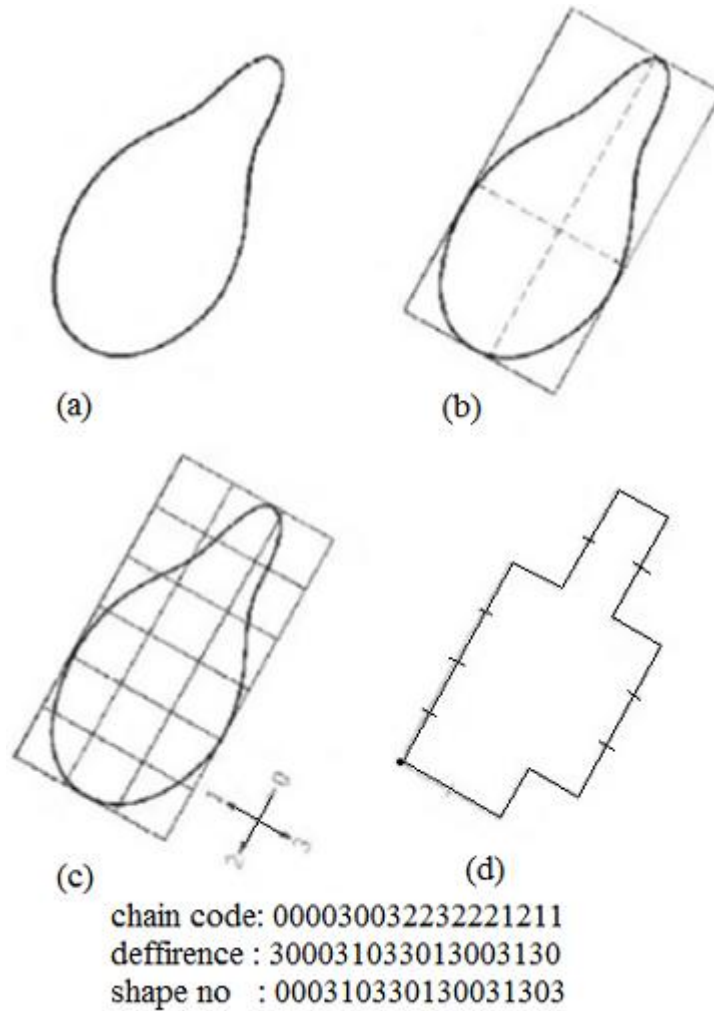
3.2.2 Shape number.

Shape number của một biểu diễn đường biên được định nghĩa là sự khác biệt đầu tiên của cường độ nhỏ nhất [8]. Trình tự n của một shape number là số lượng các chữ số được biểu diễn. Hình 3.3 minh họa hình dạng của trình tự 4,6,8.



Hình 3.3: Biểu diễn hình dạng sử dụng shape number.

Chúng ta xét một ví dụ cụ thể, giả sử n=18 được quy định cụ thể cho biên như hình 3.4(a). Để có được một shape number của trật tự này đòi hỏi phải làm theo các bước sau: Bước đầu tiên là tìm các hình chữ nhật cơ bản như trong hình 3.4(b). Hình chữ nhật gần nhất của trật tự 18 là hình chữ nhật 3x6, yêu cầu phải chia nhỏ hình chữ nhật cơ bản như trong hình 3.4(c). Cuối cùng có được chuỗi mã và sử dụng điểm khác biệt đầu tiên (first difference) để tính toán shape number.



Hình 3.4: Các bước tính toán shape number.

3.2.3 *Đôi sánh các shape number.*

Mức độ tương tự k giữa 2 hình dạng được định nghĩa là thứ tự lớn nhất mà shape number vẫn còn trùng khớp. Ví dụ, cho hai hình dạng a và b được biểu diễn bởi một chuỗi mã 4 hướng, hai hình dạng có độ tương tự k nếu:

$$S_j a = S_j b \text{ với } j=4, 6, 8, \dots k$$

$$S_j a \neq S_j b \text{ với } j=k+2, k+4, \dots$$

Trong đó S cho biết shape number và chỉ số dưới là trình tự.

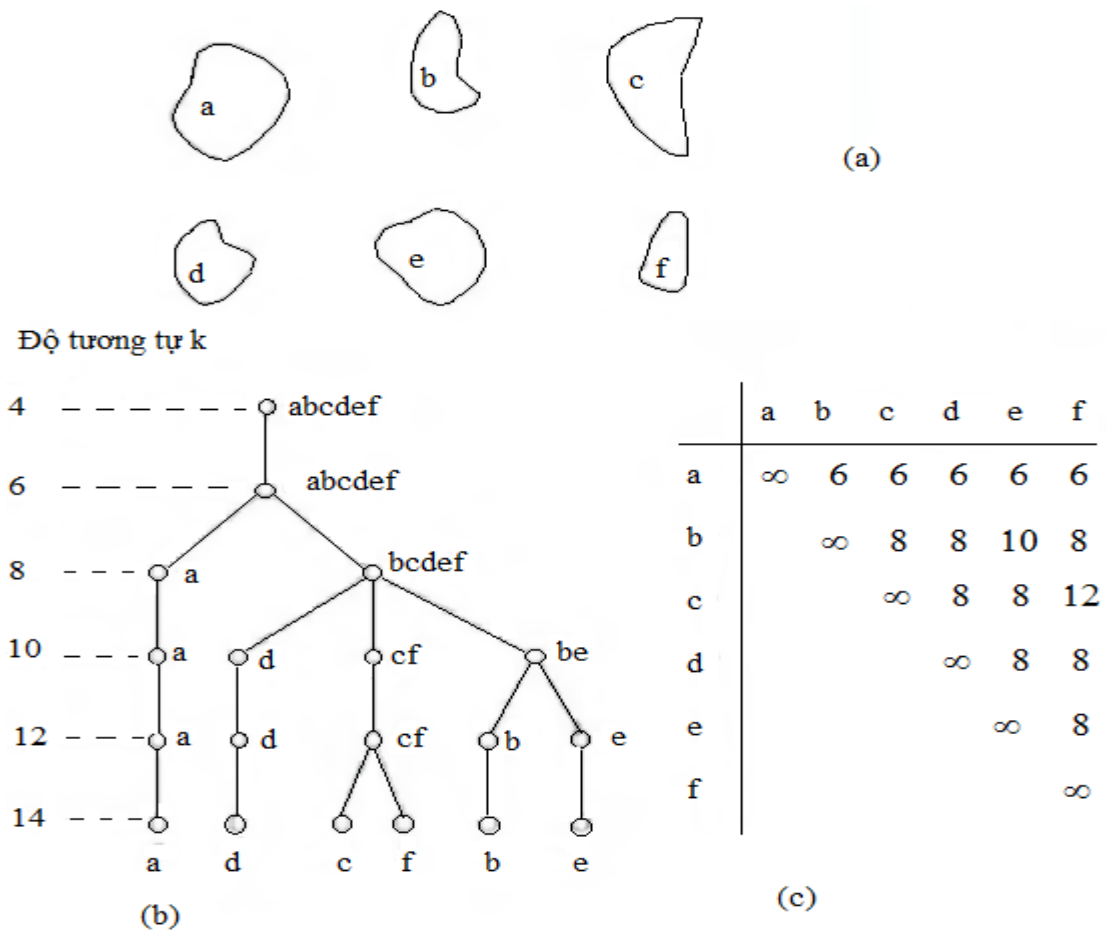
Khoảng cách giữa hai hình a và b được định nghĩa là nghịch đảo của mức độ tương tự:

$$D_{a,b} = \frac{1}{k}$$

Khoảng cách này có các thuộc tính sau:

1. $D(a,b) \geq 0$
2. $D(a,b) = 0$ nếu $a=b$
3. $D(a,b) \leq \max[D(a,b), D(b,c)]$

Giả sử, chúng ta có một hình dạng f và muốn tìm tập thích hợp của nó trong tập các hình dạng khác (a, b, c, d, e) như hiển thị trong hình 3.5(a). Việc tìm kiếm có thể được hình dung như cây tương tự trong hình 3.5(b.) Gốc của cây tương ứng với mức độ tương tự thấp nhất có thể thực hiện được, ở trong ví dụ này là bằng 4. Giả sử các hình dạng giống nhau lên tới cấp 8, ngoại trừ hình dạng a , có mức độ tương tự đối với tất cả các hình dạng khác là 6. Thực hiện từ trên xuống, tìm thấy hình dạng d có mức độ tương tự 8 đối với các hình dạng khác. Hai hình dạng f và c có mức độ tương tự cao hơn so với các hình khác. Ngược lại, nếu có một hình dạng chưa biết, sử dụng phương pháp này là tương tự như 5 hình dạng với mức độ tương tự là 6. Thông tin tương tự có thể được tóm lược trong ma trận tương tự như trong hình 3.5(c).



Hình 3.5: Minh họa tìm kiếm hình dạng tương tự sử dụng shape number: (a) hình dạng; (b) cây tương tự; (c) ma trận tương tự.

3.3 Phương pháp trích chọn đặc trưng dựa trên vùng.

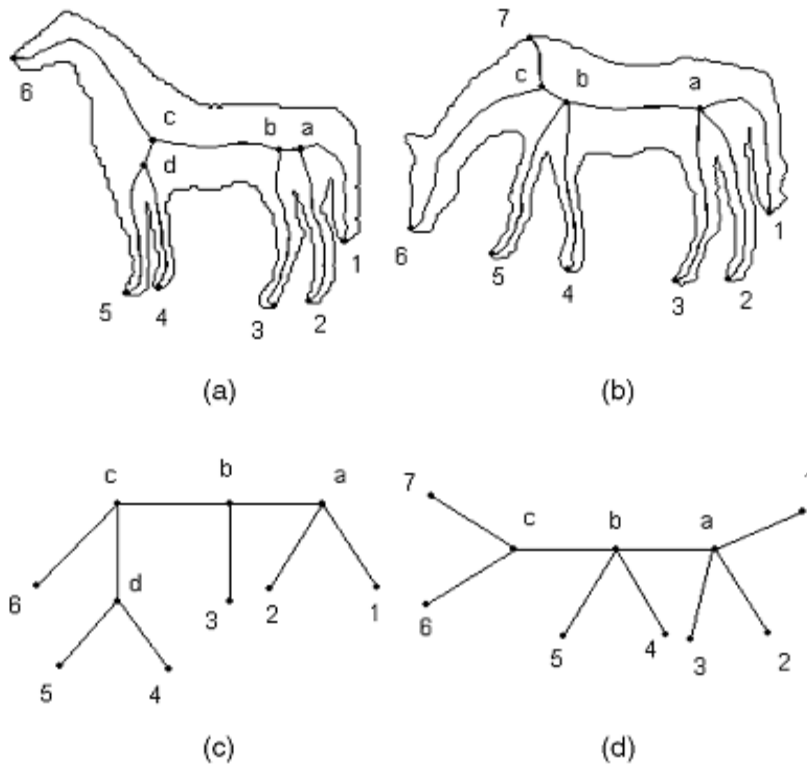
Xương (hay còn gọi là trục trung vị) tích hợp các tính năng hình học và topo của đối tượng, là một mô tả hình dạng quan trọng đối với nhận dạng đối tượng [3]. Sự tương đồng về hình dạng dựa trên đồ thị xương thường thực hiện tốt hơn so với dựa trên đường biên hoặc các mô tả hình dạng khác với sự có mặt của chồng lấp từng phần và khớp nối nhiều phần. Tuy nhiên đó là nhiệm vụ đầy thách thức để sử dụng xương cho việc nhận dạng tự động do độ nhạy cảm của xương đối với biến dạng biên. Hơn nữa, một hạn chế chính của phương pháp nhận dạng dựa trên xương là cấu trúc phức tạp của cây hoặc đồ thị biểu diễn của xương.

Phương pháp đối sánh sự tương đồng của đồ thị xương được đề xuất bởi X.Bai và L.Jan Latecki là một phương pháp thực hiện khá hiệu quả trong việc nhận dạng đối tượng dựa trên xương [4]. Ý tưởng chính của phương pháp là đối sánh đồ thị xương bằng cách so sánh các đường dẫn tới điểm cuối xương. Phương pháp đối sánh này không dựa trên cấu trúc topo hình học, bởi một thực tế trực quan là những bộ xương tương tự có thể có cấu trúc topo hình học khác nhau. Việc so sánh các đường dẫn giữa các điểm cuối của đồ thị xương mang lại kết quả chính xác phù hợp với mọi trường hợp. Thông thường dùng cho nhận dạng là các nhánh xương đã được cắt tía. Các xương được cắt tía bởi phân chia đường biên với DCE (Discrete Curve Evolution) có điểm cuối của nhánh xương tương ứng với phần trực quan của đối tượng. Kết quả thực nghiệm cho thấy rằng phương pháp này có thể tạo ra kết quả chính xác với sự có mặt của sự khớp xương, sự kéo dài xương và biến dạng đường biên.

Chúng ta sử dụng sự tương tự của các đường đi ngắn nhất giữa mỗi cặp điểm cuối xương để thiết lập mối quan hệ tương ứng với điểm cuối trong đồ thị khác. Ví dụ, đỉnh 1 trong hình 3.6(a) tương ứng với đỉnh 1 trong hình 3.6(b) kể từ đường đi ngắn nhất tới các đỉnh 2,3,4,5,6 là tương tự. Cuối cùng giá trị không giống nhau giữa các đồ thị là tính ước lượng khoảng cách giữa các điểm cuối tương ứng. Vì vậy ý tưởng cơ bản của phương pháp này là xác định sự giống nhau của các cấu trúc phức tạp (complex structures) của đồ thị hoặc cây bằng cách kiểm tra đường đi ngắn nhất giữa các điểm cuối của chúng. Phương pháp đề xuất đạt được kết quả chính xác và nhanh hơn so với các phương pháp đối sánh đồ thị và cây hiện có.

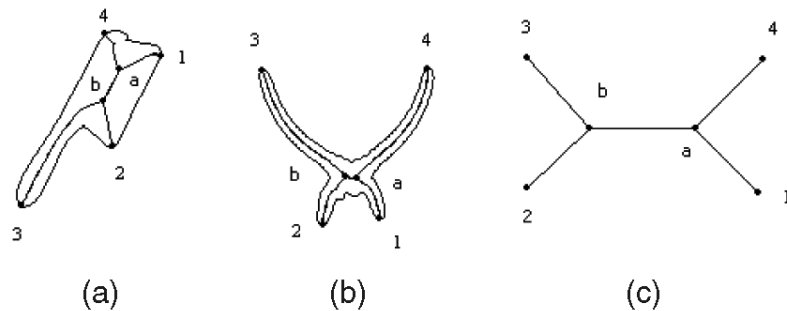
Có lẽ thách thức quan trọng nhất cho đo độ tương tự xương là cấu trúc topo của cây xương hoặc đồ thị của các đối tượng tương tự nhau lại có thể hoàn toàn khác nhau. Thực tế này được minh họa trong hình 3.6, mặc dù bộ xương của 2 con

ngựa (hình 3.6a và hình 3.6b) là tương tự, nhưng đồ thị xương (hình 3.6c và hình 3.6d) lại rất khác nhau. Ví dụ này minh họa những khó khăn phải đối mặt bởi cách tiếp cận dựa trên hoạt động chỉnh sửa đồ thị trong đối sánh xương. Để có một đồ thị xương hay cây xương được hiển thị như trong hình 3.6 phải sử dụng một số hoạt động chỉnh sửa (cắt, trộn...).



Hình 3.6: Hình dạng (a) và (b) tương tự nhau nhưng đồ thị khác nhau.

Mặt khác, đồ thị xương của các đối tượng khác nhau có thể có cấu trúc topo giống nhau, như trong hình 3.7. Các xương của bàn chải trong hình 3.7(a) và kìm trong hình 3.7(b) có cùng topo như thể hiện trong hình 3.7(c).



Hình 3.7: Hình dạng (a) và (b) khác nhau nhưng có đồ thị xương (c) giống nhau.

Đề xuất đối sánh đồ thị xương dựa trên giả định rằng sự tương tự xương có sự tương tự cấu trúc của node cuối (sự giống nhau của đường đi ngắn nhất đến các node cuối khác). Và việc đo đường dẫn xương được mô tả như một chuỗi bán kính của đĩa cứng đại. Mặc dù không xem xét tường minh cấu trúc topo của đồ thị xương nhưng cũng không hoàn toàn bỏ qua cấu trúc này, nó được ngầm hiểu biểu diễn một thực tế là chồng chéo các phần của đường xương tương tự, khi chồng các phần có các chuỗi con có cùng bán kính. Ví dụ trong hình 3.6(a), 3.6(b) đường đi từ 6→1 và từ 5→2 là chồng chéo lên nhau. Thực tế là những phân đoạn chồng chéo nhau này hơi khác nhau trong hình 3.6(a), 3.6(b) không ảnh hưởng đến sự giống nhau của chuỗi bán kính. Vì vậy cách tiếp cận của trên đủ linh hoạt để thực hiện tốt trên những hình dạng khớp nhau nhưng cũng không gây nhầm lẫn với hình dạng khác nhau.

3.3.1 Đồ thị xương.

Các định nghĩa sau áp dụng cho xương liên tục (continuous skeletons) và xương trong ảnh số (digital images).

Định nghĩa 1: Điểm xương chỉ có một điểm lân cận (adjacent point) được gọi là điểm cuối xương (endpoint). Điểm xương có từ 3 lân cận trở lên được gọi là điểm giao nhau (junction point). Điểm xương không phải điểm cuối và không phải điểm giao nhau được gọi là điểm kết nối (connection point).

Định nghĩa 2: Chuỗi các điểm kết nối giữa 2 điểm xương kết nối trực tiếp gọi là một nhánh xương (skeleton branch). Một tiêu chuẩn để xây dựng đồ thị xương đó là: điểm cuối và điểm giao nhau đều được chọn là các node cho đồ thị, và các nhánh xương giữa các node là các cạnh giữa các node. Ví dụ, hình 3.6(c) và 3.6(d) là đồ thị xương trong hình 3.6(a) và 3.6(b) tương ứng.

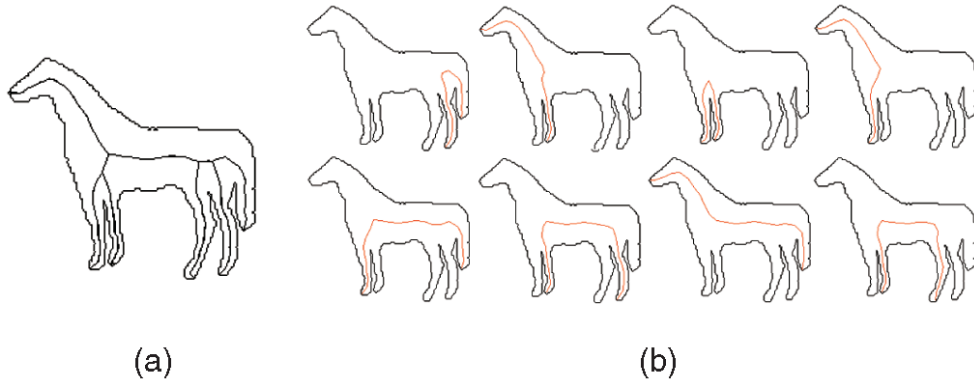
Định nghĩa 3: Điểm cuối trong đồ thị xương gọi là node cuối và điểm giao nhau trong đồ thị xương gọi là node giao nhau.

3.3.2 Đối sánh đồ thị xương.

Đối sánh đồ thị xương bằng cách thiết lập sự tương ứng của các node cuối, các node này là các điểm nổi bật trên đường biên, và tất cả các nhánh xương cuối trên đường biên có thể được xem như một phần trực quan của hình dạng ban đầu. Vì vậy sự đối sánh không liên quan đến node giao nhau.

3.3.2.1 Mô tả đường dẫn

Định nghĩa 4: Đường đi ngắn nhất giữa 1 cặp node cuối trên đồ thị xương được gọi là đường dẫn xương (skeleton path), ví dụ xem hình 3.8b.



Hình 3.8: Minh họa đường dẫn xương: (a) xương của hình con ngựa, (b) đường dẫn ngắn nhất giữa các cặp node cuối.

Giả sử có N node cuối trên đồ thị xương G và cho v_i ($i=1,2,\dots,N$) với i là node cuối dọc theo biên theo chiều kim đồng hồ. Cho $p(v_m, v_n)$ là đường dẫn xương từ v_m đến v_n . Lấy mẫu $p(v_m, v_n)$ với những điểm M cách đều đó là tất cả những điểm xương. Cho $R_{m,n}(t)$ là bán kính đĩa cực đại của điểm xương với chỉ số t trong $p(v_m, v_n)$. Vectơ của bán kính đĩa cực đại có tâm tại điểm M trên $p(v_m, v_n)$ được ký hiệu là:

$$R_{m,n} = (R_{m,n}(t))_{t=1,2,\dots,M} = (r_1, r_2, \dots, r_M) \quad (3.1)$$

Trong bài viết này, bán kính $R_{m,n}(t)$ là xấp xỉ với khoảng cách biến đổi $DT(t)$ (Distance Transform) tại mỗi điểm xương có chỉ số t . Giả sử có N_0 pixel trong hình dạng S ban đầu, để phương pháp thực hiện bất biến với co giãn, chúng ta chuẩn hóa $R_{m,n}(t)$ như sau:

$$\mathbf{R}_{m,n} = \frac{DT(t)}{\frac{1}{N_0} \sum_{i=1}^{N_0} DT S_i} \quad (3.2)$$

Trong đó S_i ($i=1,2,\dots,N$) biến đổi trên tất cả N_0 pixel trong hình dạng.

Định nghĩa 5: Sự khác nhau về hình dạng giữa hai đường xương được gọi là khoảng đường dẫn (path distance). Nếu R và R' là các vectơ của bán kính của 2 đường dẫn $p(u, v)$ và $p(u', v')$ tương ứng, khoảng đường dẫn được định nghĩa là:

$$pd(p(u, v), p(u', v')) = \sum_{i=1}^M \frac{(r_i - r'_i)^2}{r_i + r'_i} + \alpha \frac{(l - l')^2}{l + l'} \quad (3.3)$$

Trong đó l và l' là độ dài của $p(u,v)$ và $p(u',v')$ và α là trọng số (weight factor). Để biểu diễn bất biến với co giãn, chiều dài của đường dẫn là được chuẩn hóa. Bằng cách này, đường dẫn và khoảng cách đường dẫn là bất biến với co giãn.

Để giải quyết sự giống nhau của hình dạng khớp nhau, khoảng đường trong (3.3) không xét đến biến dạng đường đi, không thay đổi các vecto bán kính và độ dài đường đi.

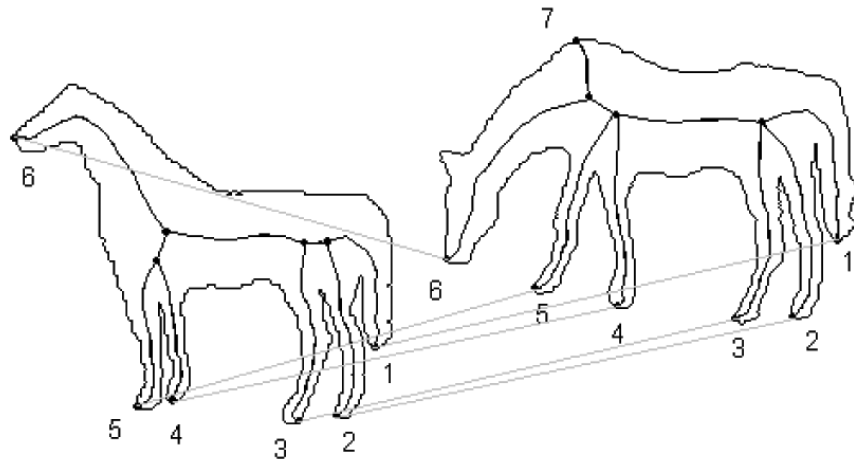
3.3.2.2 Đối sánh node cuối sử dụng đường xương.

Trong đồ thị xương, mỗi node cuối có đường xương cho tất cả các node khác trong đồ thị. Cho G và G' là hai đồ thị được đối sánh, và v_i, v'_j là các node cuối của G và G' . Cho số lượng các node trong G và G' là $K+1$ và $N+1$ ($K \leq N$). Chi phí đối sánh $c(v_i, v'_j)$ giữa v_i và v'_j là ước tính dựa trên đường dẫn từ các đỉnh khác trong G và G' , chúng bắt nguồn từ v_i và v'_j . Đầu tiên, chúng ta để tất cả các node cuối trong G theo chiều kim đồng hồ với node khởi đầu là v_i gọi là v_{i0} (tất cả các điểm cuối xương đều nằm trên biên). Chúng ta có 1 chuỗi các node cuối $v_{i0}, v_{i1} \dots v_{iK}$ trong G và tương tự trong G' là $v'_{j0}, v'_{j1}, \dots, v'_{jN}$. Sau đó, chúng ta tính toán khoảng đường đi giữa 2 chuỗi. Chúng ta thu được ma trận của khoảng đường đi với (3.3) :

$$pd_{v_i, v'_j} = \begin{pmatrix} pd(p_{v_{i0}, v_{i1}}, p_{v'_{j0}, v'_{j1}}) & pd(p_{v_{i0}, v_{i1}}, p_{v'_{j0}, v'_{j2}}) \\ pd(p_{v_{i0}, v_{i2}}, p_{v'_{j0}, v'_{j1}}) & pd(p_{v_{i0}, v_{i2}}, p_{v'_{j0}, v'_{j2}}) \\ \vdots & \vdots \\ pd(p_{v_{i0}, v_{iK}}, p_{v'_{j0}, v'_{j1}}) & pd(p_{v_{i0}, v_{i2}}, p_{v'_{j0}, v'_{j1}}) \\ \dots & pd(p_{v_{i0}, v_{i1}}, p_{v'_{j0}, v'_{jN}}) \\ \dots & pd(p_{v_{i0}, v_{i2}}, p_{v'_{j0}, v'_{jN}}) \\ \vdots & \vdots \\ \dots & pd(p_{v_{i0}, v_{iK}}, p_{v'_{j0}, v'_{jN}}) \end{pmatrix} \quad (3.4)$$

Để tính toán giá trị không giống nhau giữa 2 node cuối v_i và v'_j sử dụng phương pháp mở rộng của tương tự từng phần của chuỗi, phương pháp mở rộng được giới thiệu là OSB (Optimal Subsequence Bijection). Các thuộc tính chính của OSB là nó có thể bỏ qua yếu tố outlier (yếu tố nằm ngoài) trong trường hợp có thể bỏ qua 1 số điểm cuối xương. Bằng cách áp dụng OSB cho ma trận (3.4) chúng ta có được sự khác nhau của 2 node cuối v_i và v'_j :

$$c(v_i, v'_j) = OSB(pd(v_i, v'_j)) \quad (3.5)$$



Hình 3.9: Sự tương ứng giữa các node cuối của hai đồ thị xương.

Đối với 2 đồ thị G và G' với các node cuối v_i và v_j' , chúng ta tính toán tất cả các giá trị không giống nhau giữa các node cuối của chúng và thu được 1 ma trận mới :

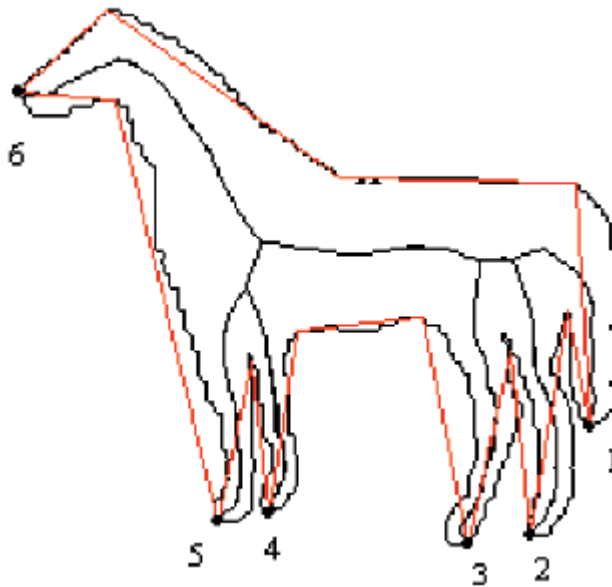
$$C(G, G') = \begin{matrix} c_{v_0, v'_0} & c_{v_0, v'_1} & \cdots & c_{v_0, v'_N} \\ c_{v_1, v'_0} & c_{v_1, v'_1} & \cdots & c_{v_1, v'_N} \\ \vdots & \vdots & \vdots & \vdots \\ c_{v_K, v'_0} & c_{v_K, v'_1} & \cdots & c_{v_K, v'_N} \end{matrix} \quad (3.6)$$

Cuối cùng chúng ta tính tổng các giá trị không giống nhau $c(G, G')$ giữa G và G' với thuật toán Hungary trên $C(G, G')$. Đối với mỗi node cuối v_i trong G , thuật toán Hungary có thể tìm thấy v_j' tương ứng trong G' . Từ G và G' có thể có các số khác nhau của node cuối, tổng giá trị khác nhau nên bao gồm penalty cho các node cuối mà không tìm thấy bất kỳ đối tác nào. Để đạt được điều này, chỉ cần bổ sung thêm các hàng với giá trị không đổi cho (3.6) để $C(G, G')$ trở thành ma trận vuông, hàng giá trị không đổi là trung bình của các giá trị trong $C(G, G')$. Sử dụng thuật toán Hungary để có sự phù hợp một - một (one to one) trong chuyển đổi của các node cuối, với một số node cuối có thể gán giá trị const biểu diễn một node giả.

Trong cách tiếp cận này không yêu cầu bất kỳ sự tương ứng nào của các node giao nhau. Điều này rất quan trọng bởi trong nhiều trường hợp sự tương ứng của các node giao nhau là không thể thiết lập trực tiếp, và do đó cách tiếp cận chỉnh sửa đồ thị hoặc cây là cần thiết nếu có yêu cầu sự tương ứng của các node giao nhau. Một điều quan trọng cần phải tuân thủ là không thể thay đổi cấu trúc của các node giao nhau với việc cắt tỉa xương mà không cần loại bỏ một số node quan trọng.

Mặt khác cắt tia xương có thể làm giảm tập hợp các node cuối liên quan đến cấu trúc bằng cách loại bỏ các node cuối giả.

Phương pháp giới thiệu ở trên đo độ tương tự dựa trên hình dạng xương. Đó là đối sánh đồ thị xương dựa trên tương tự của đường đi ngắn nhất, đường đi ngắn nhất giữa mỗi cặp điểm cuối xương được biểu diễn bằng chuỗi bán kính của đĩa cực đại (radii of the maximal disks) của điểm xương tương ứng. Chúng ta có lợi từ thực tế các điểm cuối của xương kế thừa một trật tự tuần tự từ đường biên, điều này là có thể khi xương được cắt tia dựa trên phân vùng biên với DCE đảm bảo tất cả các điểm cuối của xương đều nằm trên biên. Ví dụ trong hình 3.10, tất cả các điểm cuối (1,2,...6) của bộ xương ngựa đều là đỉnh của đa giác được đơn giản hóa bởi DCE[6] là sự ổn định của nó ở chỗ có thể loại bỏ các nhánh giả trong khi vẫn giữ được cấu trúc của nhánh có liên quan.



Hình 3.10: Minh họa xương được cắt tia bởi DCE.

Như vậy, đồ thị xương đề xuất đối sánh dựa trên giả định rằng những đường xương tương đồng có cấu trúc các node cuối tương đồng và được đo bằng sự giống nhau của những đường đi ngắn nhất đến các node cuối khác.

3.3.2.3 Phương pháp OSB (Optimal Subsequence Bijection).

Thuật toán đối sánh 2 chuỗi có độ dài khác nhau m và n đã được đề xuất trong [7], cụ thể hơn chúng ta xét hai chuỗi hữu hạn của các node cuối xương

$a=(a_1,\dots,a_n)$ và $b=(b_1,\dots,b_n)$. Mục đích để tìm dãy con a' của a và b' của b sao cho a' đối sánh tốt nhất với b' . Bỏ qua không đối sánh 1 số yếu tố của a và b là cần thiết, bởi vì cả hai chuỗi có thể chứa 1 số yếu tố outlier. Tuy nhiên nếu bỏ qua quá nhiều yếu tố của chuỗi làm nguy cơ cho những kết quả không phù hợp. Để ngăn chặn điều này xảy ra, sử dụng 1 penalty bỏ qua yếu tố. Cách thức biểu diễn dưới dạng đối sánh thêm 1 yếu tố b_∞ . Do đó chúng ta mở rộng chuỗi b thêm 1 yếu tố b_∞ .

Mục tiêu là tìm thấy sự tương ứng tốt nhất có thể có của chuỗi a với chuỗi con b' của b . Định nghĩa sự tương ứng $f: \{1,\dots,m\} \rightarrow \{1,\dots,n,\infty\}$ như một ánh xạ đơn điệu đối với phạm vi giới hạn của hàm $f: (1,\dots,m) \rightarrow (1,\dots,n)$, có nghĩa là 1 hàm f có $f(i) < f(i+1) < \infty$ trong đó a_i là ánh xạ bởi $b_{f(i)}$ cho tất cả $i \in \{1,\dots,m\}$, và cho phép ánh xạ nhiều - một đến ∞ . Việc gán $f(i) = \infty$ có nghĩa là bỏ qua yếu tố i trong chuỗi a . Tập hợp các chỉ số i_k và $f(i_k)$ với $f(i_k) < \infty$ cho $i_k \in \{1,\dots,m\}$ xác định dãy con a' của a và b' của b chẳng hạn f bị giới hạn bởi (i_k) là một bijection.

Chúng ta giả định rằng các hàm khoảng cách d (distance function d) có thể tính toán giá trị không giống nhau giữa các yếu tố của a và b , đó là, $d(a_i, b_j)$ được đưa ra cho $(i,j) \in \{1,\dots,m\} \times \{1,\dots,n,\infty\}$. Do không hạn chế hàm tính khoảng cách d nên bất kỳ hàm khoảng cách nào cũng có thể thực hiện được. Trong bài viết này, sử dụng khoảng đường đi pd trong định nghĩa (3.3) là khoảng cách d . Trong hầu hết các ứng dụng, $d(a_i, b_j)$ được đưa ra cho $(i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}$, việc bổ sung yếu tố khoảng cách $d(a_i, b_\infty)$ được lựa chọn cẩn thận. Thông thường $d(a_i, b_\infty)$ là hằng số cho tất cả các $i \in \{1,\dots,m\}$.

Xác định chi phí bỏ qua (cost of skipping) cho yếu tố bất kỳ trong a , gọi hằng số này là $jumpcost$. Trong bài viết này $d(a_i, b_\infty) = jumpcost$ được tính toán như sau:

$$jumpcost = mean_i(\min_j d(a_i, b_j) + std(\min_j(d(a_i, b_j))) \quad (3.7)$$

Như vậy mỗi yếu tố a_i tìm được một yếu tố gần nhất b_j (closest element b_j) và sau đó chúng ta đem mean cộng với độ lệch chuẩn (standard deviation) của khoảng cách đối với yếu tố gần nhất. Ví dụ, nếu chuỗi a và b là tương tự với trường hợp ngoại lệ của yếu tố outlier, gọi nó là a_k , sau đó với mọi a_i với $i \neq k$ tìm 1 yếu tố b_j với khoảng cách $d(a_i, b_j)$ nhỏ. Do đó $jumpcost$ sẽ nhỏ, vì vậy mà khoảng cách đến các yếu tố gần nhất trong b từ a_k lớn hơn $jumpcost$ và yếu tố a_k được loại trừ với một

penalty tương đối nhỏ, đó là chúng ta nhảy qua nó. Với bất kỳ tương ứng nào, chúng ta cần xác định khoảng cách giữa hai trình tự như sau:

$$d(a, b, f) = \frac{1}{m} \sum_{i=1}^m d(a_i, b_{f(i)})^2 \quad (3.8)$$

Mục tiêu là tìm thấy sự tương ứng f để $d(a', b', f)$ là tối thiểu. Chính xác hơn, 1 tương ứng f tối ưu của yếu tố trong chuỗi a đến các yếu tố trong chuỗi b được định nghĩa là giá trị nhỏ nhất của $d(a, b, f)$ trên tất cả các tương ứng có thể:

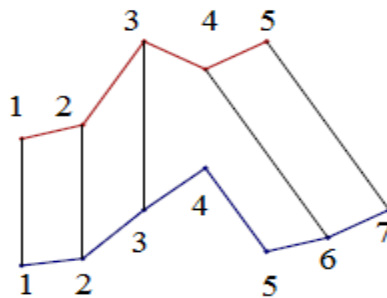
$$f = \arg \min \{d(a, b, f) : f \text{ là một tương ứng}\} \quad (3.9)$$

Cuối cùng, khoảng cách tối ưu cho bởi công thức (3.8) cho $f = f$. Sự tương ứng tối ưu có thể tìm thấy với thuật toán tìm đường đi ngắn nhất trên DAG (Directed Acyclic Graph). Chúng ta biểu thị khoảng cách tối ưu $d(a, b)$ trong (3.10) với $OSB(a, b)$ cho khoảng cách OSB. Node của DAG là tất cả các cặp chỉ số $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$ và trọng số cạnh w được định nghĩa:

$$w((i, j)) = \begin{cases} d(a_i, b_j) & \text{if } i + 1 = k \text{ and } j + 1 \leq l \\ k - i - 1 \cdot \text{jumpcost} & \text{if } i + 1 < k \text{ and } j + 1 \leq l \\ \infty & \text{otherwise} \end{cases} \quad (3.10)$$

Không có penalty rõ ràng để bỏ qua một yếu tố của chuỗi b . Dựa vào khả năng trực giác là chúng ta kỳ vọng tất cả các yếu tố của chuỗi a để tìm ra một tương ứng trong chuỗi b có khả năng bỏ qua một số yếu tố của b .

Minh họa một ví dụ đơn giản như trong hình 3.11 cho thấy sự tương ứng được tìm thấy bởi OSB như một đường dẫn ngắn nhất cho 2 chuỗi $a = \{1, 2, 8, 6, 8\}$ và $b = \{1, 2, 9, 15, 3, 5, 9\}$ với khoảng cách giữa hai yếu tố là bình phương khác biệt. Trình tự các chỉ số được tính toán bởi OSB (không phải các giá trị): (1,1) (2,2) (3,3) (4,6) (5,7). Quan sát thấy rằng yếu tố $b_4 = 15$ và $b_5 = 3$ được bỏ qua, jumcost được tính là $c = 1.15$.



Hình 3.11: Sự tương ứng giữa các yếu tố.

3.3.3 Nhận xét

Phương pháp đối sánh đồ thị xương được mô tả trong phần 3.3.2 khác với phương pháp tiếp cận hiện có, phương pháp này không xem xét tường minh cấu trúc topo của cây hoặc đồ thị xương. Thay vào đó, tập trung vào sự giống nhau của các đường dẫn kết nối tới điểm cuối xương.

Việc sử dụng đo đạc đường đi ngắn nhất không phải là mới, tuy nhiên có sự khác biệt đáng kể trong cách tiếp cận trên. Chúng ta không xem xét đường đi ngắn nhất giữa tất cả các node xương và đường đi ngắn nhất giữa các điểm biên, mà chỉ xem xét tới đường đi ngắn nhất giữa các điểm cuối xương, điều này cho phép chúng ta tránh được những bất ổn của điểm xương giao nhau.

Phương pháp này sử dụng cấu trúc hai lớp. Trong lớp đầu, đường xương ngắn nhất xuất phát từ điểm cuối xương đã định sẵn. Trong lớp thứ hai, chúng ta tính toán sự giống nhau của hai hình dạng bằng cách đối sánh các mô tả hình dạng của điểm cuối xương.

Khi xương tương tự có số lượng điểm cuối khác nhau, chúng ta phải sử dụng sự tương ứng từng phần của các điểm cuối (mục 3.3.2.3). Bằng cách sử dụng phương pháp này chúng ta có thể đối sánh từng phần xương của đối tượng hoàn chỉnh. Và đối sánh từng phần với từng phần là một yêu cầu cần thiết để nhận dạng đối tượng.

CHƯƠNG 4: THỰC NGHIỆM.

4.1 Môi trường thử nghiệm.

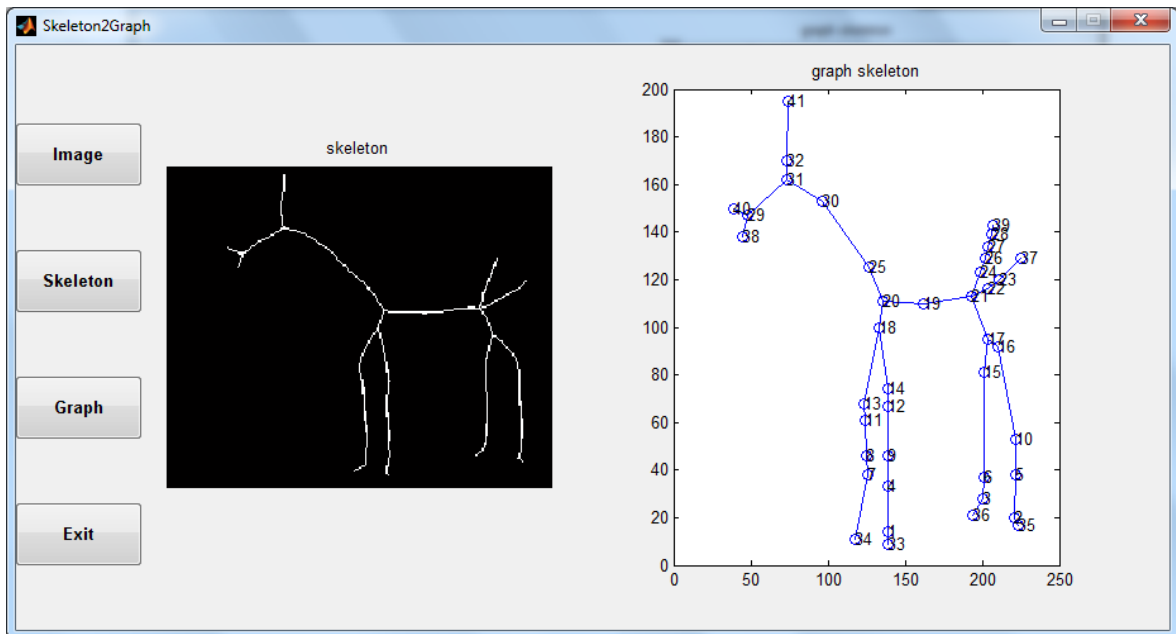
Chương trình được cài đặt trên Môi trường Windows 7, sử dụng ngôn ngữ Matlab với máy tính có cấu hình như sau:

- CPU : Dual core 2.16GHz.
- Ram: 2G.
- HDD: 140G.

Tập dữ liệu được sử dụng trong thử nghiệm là tập dữ liệu thuộc: MPEG-7

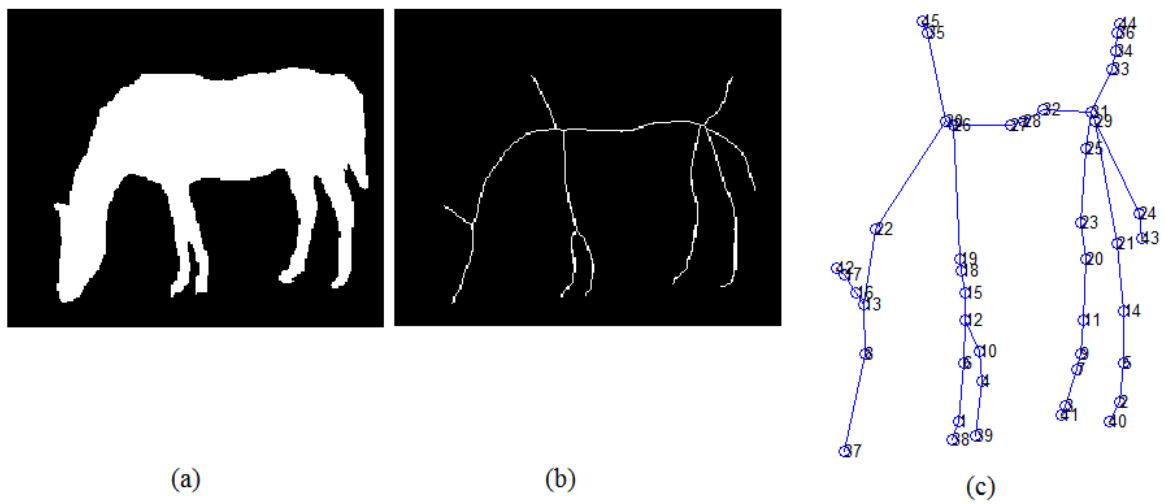
4.2 Một số kết quả thu được.

4.2.1 Giao diện chương trình:



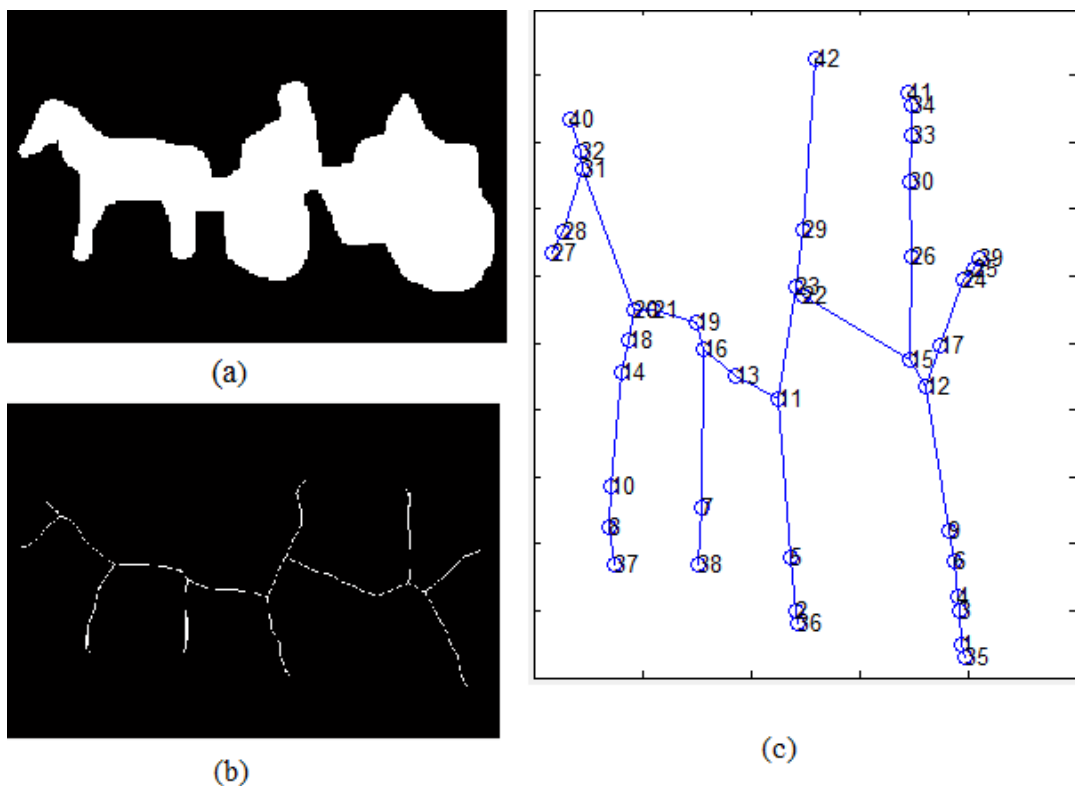
Hình 4.1: Giao diện của chương trình.

4.2.2 Kết quả trên một số đối tượng khác nhau.



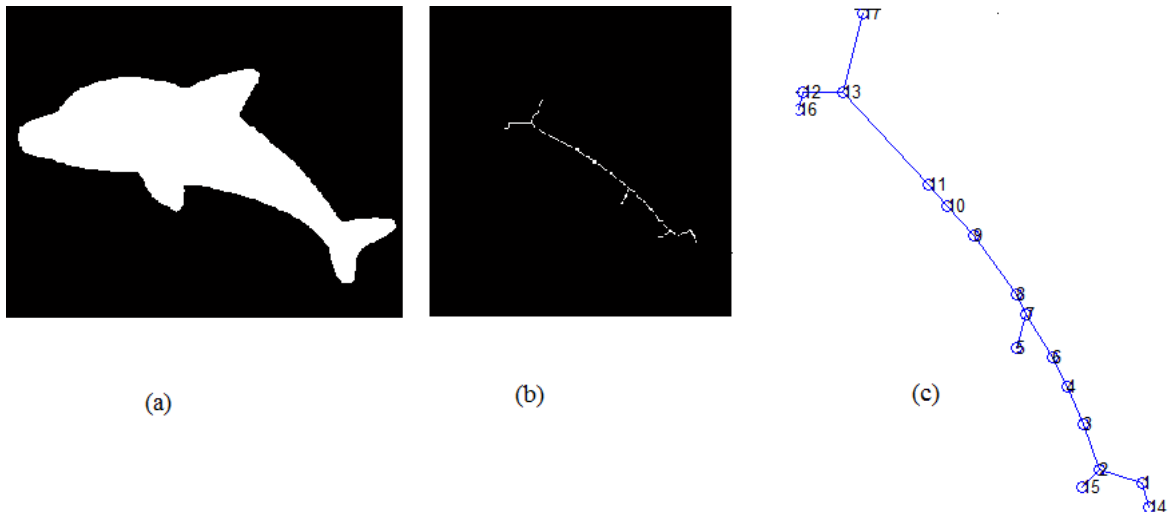
Hình 4.2: Kết quả thu được với hình con ngựa

(a): Hình con ngựa, (b): xương, (c): đồ thị.



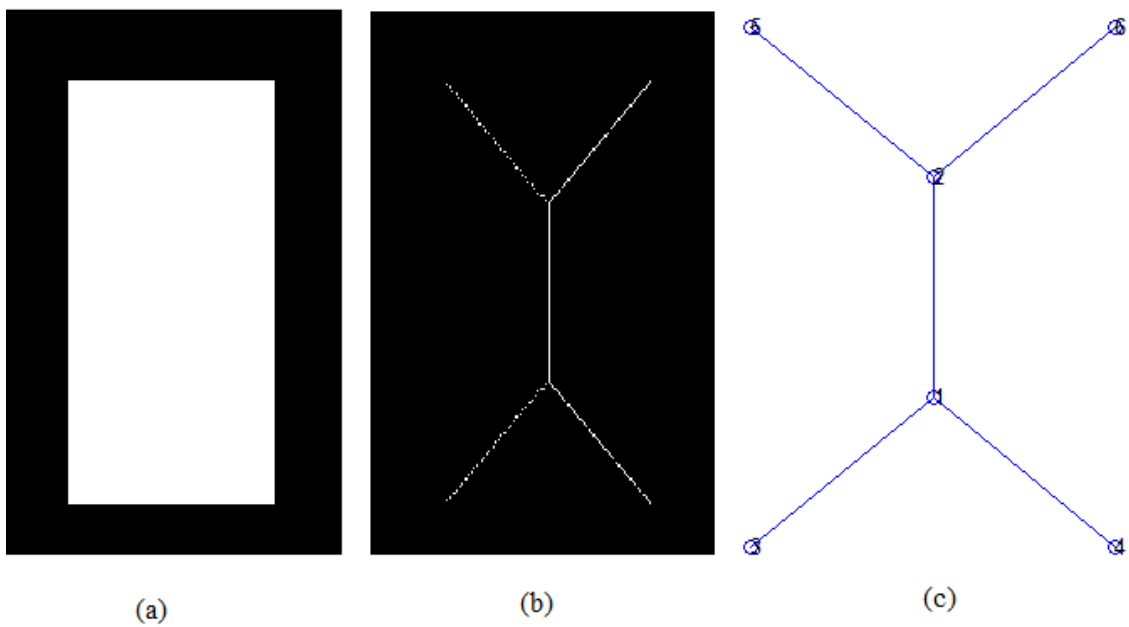
Hình 4.3: Kết quả thu được với hình con ngựa kéo xe.

(a): hình con ngựa kéo xe, (b): xương, (c): đồ thị.



Hình 4.4: Kết quả thu được với hình cá heo.

(a): hình cá heo, (b): xương, (c): đồ thị.



Hình 4.5: Kết quả thu được với hình chữ nhật.

(a): hình chữ nhật, (b): xương, (c): đồ thị.

4.2.3 Một số nhận xét về chương trình.

Chương trình mô tả cho quá trình chuyển đổi từ xương sang đồ thị dựa trên ma trận kề. Sau đây là danh sách một số tập tin và chức năng của nó trong chương trình:

- cListaNodos: mảng của lớp node.
- LimpiarNodosFd.m: hàm tinh chỉnh danh sách node ban đầu.
- Nodo.m: lớp node đồ thị.
- Skel2Graph.m: hàm xây dựng đồ thị.
- vCracV1: hàm xây dựng ma trận kề.
- Skeleton2Graph: hàm chính để chạy chương trình.

Chương trình vẫn còn một số hạn chế, trên lý thuyết, biểu diễn đồ thị từ xương chỉ sử dụng các điểm đặc trưng (điểm cuối và điểm giao nhau) và các liên kết giữa chúng, nhưng việc thực hiện chương trình trên thực tế có một số điểm thuộc nhánh xương bị thêm vào.

Bên cạnh đó, chương trình chưa tính toán được trọng số của đồ thị để phục vụ cho việc so sánh và tìm kiếm ảnh.

KẾT LUẬN

Sau một thời gian tìm hiểu và nghiên cứu đề án đã trình bày được một số vấn đề sau:

Về lý thuyết: Trình bày tổng quan tra cứu ảnh dựa trên nội dung; một số phương pháp trích chọn đặc trưng ảnh dựa trên hình dạng và đặc biệt trình bày cụ thể phương pháp trích chọn đặc trưng ảnh dựa trên xương.

Về thực nghiệm, em đã cài đặt thử nghiệm chương trình chuyển đổi từ xương của ảnh sang biểu diễn bằng đồ thị nhằm mục đích phục vụ cho việc đối sánh và tra cứu ảnh.

Tuy nhiên trong quá trình thực hiện, thời gian không có nhiều, năng lực chuyên môn còn nhiều hạn chế, nên đề tài mới chỉ dừng lại ở mức đọc, dịch hiểu và tìm hiểu tóm lược về phương pháp mà chưa thể cài đặt hoàn thiện chương trình. Nếu có điều kiện, em sẽ tìm đọc tài liệu để nghiên cứu nhằm tổng hợp nhiều phương pháp và đưa ra được những đánh giá kết luận dựa trên những gì đã tìm hiểu được và hoàn thiện chương trình hơn nữa.

Em rất mong nhận được sự đóng góp ý kiến của các Thầy Cô và các bạn để em có thêm kiến thức và kinh nghiệm tiếp tục hoàn thiện nội dung nghiên cứu trong đề tài.

Em xin chân thành Cảm ơn!

TÀI LIỆU THAM KHẢO

Tài liệu tiếng việt:

[1] Nguyễn Thị Hoàn, “*Phương pháp trích chọn đặc trưng ảnh trong thuật toán học máy tìm kiếm ảnh áp dụng vào bài toán tìm kiếm sản phẩm.*” Khóa luận tốt nghiệp, Đại học Công nghệ, 2010.

[2] Bùi Thị Thúy Nga, “*Tìm hiểu một số phương pháp trích chọn đặc trưng và ứng dụng cho tra cứu ảnh theo nội dung.*” Đồ án tốt nghiệp, Đại học Dân lập Hải Phòng, 2011.

Tài liệu tiếng anh:

[3] H.Blum, “*Biological Shape and Visual Science,*” J. Theoretical Biology, vol. 38, pp. 205-287, 1973.

[4] X.Bai and Longin Jan Latecki, “*Path Similarity Skeleton Graph Matching,*” IEEE Transaction on Pattern Analysis and Machine Intelligence, 2008.

[5] Dengsheng Zhang, “*Image Retrieval Based on Shape,*” Monash University, 2002.

[6] L.J Latecki and R.Lakamper, “*Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution,*” Computer Vision and Image Understanding, vol.73, no.3, pp. 441-454, 1999.

[7] L.J Latecki, Qiang Wang, Suzan Koknar-Tezel, and Vasileios Megalooikonomou “*Optimal Subsequence Bijection,*” Temple University.

[8] Rafael C.Gonzalez and Richard E.Woods “*Digital Image Processing,*” Prentice Hall, 2002.