BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

ISO 9001:2015

ĐỒ ÁN TỐT NGHIỆP

NGÀNH: CÔNG NGHỆ THÔNG TIN

Sinh viên : Lê Minh Quý

Giảng viên hướng dẫn: TS. Ngô Trường Giang

HẢI PHÒNG - 2018

BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

KỸ THUẬT ĐỐI SÁNH HÌNH DẠNG SỬ DỤNG ĐẶC TRƯNG DỰA TRÊN ĐƯỜNG BAO ĐỐI TƯỢNG

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

NGÀNH: CÔNG NGHỆ THÔNG TIN

Sinh viên : Lê Minh Quý

Giảng viên hướng dẫn: TS. Ngô Trường Giang

HẢI PHÒNG - 2018

BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

NHIỆM VỤ ĐỀ TÀI TỐT NGHIỆP

Sinh viên: Lê Minh Quý Mã SV: 1412101051

Lớp: CT1802 Ngành: Công nghệ thông tin

Tên đề tài: Kỹ thuật đối sánh hình dạng sử dụng đặc trưng dựa trên đường bao đối tượng

MỤC LỤC

LỜI CẢM ƠN **4**

LỜI MỞ ĐẦU **5**

DANH MỤC HÌNH VẼ **6**

CHƯƠNG 1: TỔNG QUAN VỀ PHÁT HIỆN BIÊN VÀ ĐỐI SÁNH ẢNH **7**

1.1 Biên và phát hiện biên 7

1.1.1 Khái niệm về biên 7

1.1.2 Phân loại các kỹ thuật phát hiện biên 8

1.1.3 Quy trình phát hiện biên trực tiếp 9

1.1.4 Một số phương pháp phát hiện biên 10

1.2 Mô tả hình dạng dựa trên đường bao 15

1.2.1 Mô tả theo tiếp cận toàn cục 16

1.2.2 Mô tả theo tiếp cận cấu trúc 18

1.3 Đối sánh ảnh 24

1.3.1 Giới thiệu về đối sánh ảnh 24

1.3.2 Đối sánh ảnh dựa trên đặc trưng 27

CHƯƠNG 2: ĐỐI SÁNH HÌNH DẠNG SỬ DỤNG NGỮ CẢNH HÌNH DẠNG **30**

2.1 Giới thiệu 30

2.2 Độ đo khoảng cách hình dạng 30

2.2.1 Khoảng cách min-max 30

2.2.2 Khoảng cách Euclid 31

2.2.3 Khoảng cách toàn phương 31

2.2.4 Khoảng cách Chi Squared distance 31

2.2.5 Khoảng cách Hausdorff 31

2.2.6 Độ đo khoảng cách trong 32

2.3 Mô tả ảnh sử dụng ngữ cảnh hình dạng (Shape context) 35

2.4 Đối sánh hình dạng ngữ cảnh 36

2.4.1 Đối sánh shape sử dụng quy hoạch động 36

2.4.2 Đối sánh hình dạng dựa trên đồ thị 37

CHƯƠNG 3: THỰC NGHIỆM **44**

3.1 Môi trường thực nghiệm.....44
3.1.1 Phần cứng.....44
3.1.2 Phần mềm.....44
3.2 Đối sánh ảnh dựa trên ngữ cảnh hình dạng sử dụng opencv.....45
3.2.1 Tìm đường bao và lấy mẫu các điểm trên đường bao.....45
3.2.2 Tìm khoảng cách và đối sánh giữa hai đường bao đã được lấy mẫu.....49
KẾT LUẬN.....54
TÀI LIỆU THAM KHẢO.....55

LỜI CẢM ƠN

Em xin chân thành cảm ơn các thầy cô trong khoa công nghệ thông tin 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.

Em xin gửi lời cảm ơn chân thành tới thầy giáo TS. Ngô Trường Giang, Thầy đã tận tình hướng dẫn và giúp đỡ em trong suốt quá trình làm đồ án, giúp em hoàn thành báo cáo đúng kế hoạch. Với sự chỉ bảo của thầy, em đã có những định hướng tốt trong việc triển khai và thực hiện các yêu cầu trong quá trình làm đồ án tốt nghiệp.

Ngoài ra, em cũng xin gửi lời cảm ơn tới tất cả bạn bè, đặc biệt là các bạn trong lớp CT1802 đã luôn gắn bó, cùng học tập và giúp đỡ em trong những năm qua và trong suốt quá trình thực hiện đồ án này.

Em xin chân thành cảm ơn.

Hải Phòng, ngày 3 tháng 11 năm 2018

Sinh viên

Lê Minh Quý

LỜI MỞ ĐẦU

Phát hiện biên của ảnh là một trong những nhiệm vụ quan trọng trong xử lý ảnh. Nhận dạng ảnh dùng máy tính liên quan tới việc nhận dạng và phân loại các đối tượng trong bức ảnh do đó phát hiện biên là một công cụ quan trọng. Phát hiện biên sẽ làm giảm một cách đáng kể khối lượng dữ liệu cần xử lý và loại bỏ các thông tin không cần thiết trong khi vẫn đảm bảo các thuộc tính quan trọng về cấu trúc của ảnh. Có rất nhiều kỹ thuật phát hiện biên hiện đang được sử dụng, mỗi kỹ thuật này thường làm việc một cách có hiệu quả cao đối với một loại đường biên cụ thể.

Còn 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. Hình dạng (Shape) là một đặc trưng quan trọng của việc phân đoạn vùng của ảnh, và tính hiệu quả và thiết thực của nó đóng vai trò quan trọng trong việc tra cứu ảnh. Phép biểu diễn hình dạng sử dụng đường cong rời rạc để làm đơn giản hóa đường viền giúp cho thuận lợi việc lọc nhiễu đã được hai tác giả Latecki và Lakamper nghiên cứu, ngoài ra việc sử dụng đường cong rời rạc còn loại bỏ được các đặc trưng hình dạng không thích hợp. Một phương pháp mô tả hình dạng để đo độ tương tự đó chính là sử dụng Shape Context để đối sánh hình dạng, phương pháp này đã được đề xuất bởi tác giả Belongie, ưu điểm của phương pháp này là nó khá tối ưu, đơn giản nhưng hiệu quả mang lại chưa cao cho việc liên quan đến biến đổi hình học và tra cứu dựa trên hình dạng.

Trong phạm vi đề tài này, em sẽ tập trung tìm hiểu về các kỹ thuật phát hiện biên, mô tả các điểm đặc trưng sử dụng ngữ cảnh hình dạng và đối sánh tập đặc trưng để ước lượng khoảng cách giữa hai ảnh hình dạng đối tượng.

DANH MỤC HÌNH VẼ

Hình 1-1: Một số kiểu đường biên thông dụng 8

Hình 1-2: Toán tử Sobel 11

Hình 1-3: Toán tử Prewitt 11

Hình 1-4: Toán tử Roberts 12

Hình 1-5: Kỹ thuật Laplace 13

Hình 1-6: Toán tử Laplacian 15

Hình 1-7: Minh họa độ lệch tâm của hình dạng 16

Hình 1-8: Các hướng của đoạn thẳng đơn vị: (a): 4 hướng, (b): 8 hướng 20

Hình 1-9: Biểu diễn của một chuỗi mã (theo 4 hướng và 8 hướng) 20

Hình 1-10: Biểu diễn hình dạng sử dụng shape number. 21

Hình 1-11: Các bước tính toán shape number. 22

Hình 1-12: Phân tích đường cong mịn 23

Hình 1-13: Ảnh gốc 29

Hình 1-14: Phát hiện cạnh 29

Hình 2-1: Ví dụ khoảng cách trong 32

Hình 2-2: Ví dụ về khoảng cách trong của x và y trong hình O 33

Hình 2-3: Quá trình biểu diễn khoảng cách trong của đối tượng 34

Hình 2-4: Tính toán ngữ cảnh hình dạng 36

Hình 3-1: Hình được hiển thị Shape 46

Hình 3-2: Kết quả tìm biên bằng phương pháp Canny từ ảnh đầu vào 47

Hình 3-3: Kết quả tìm đường bao và lấy mẫu. 49

CHƯƠNG 1: TỔNG QUAN VỀ PHÁT HIỆN BIÊN VÀ ĐỐI SÁNH ẢNH

1.1 Biên và phát hiện biên

1.1.1 Khái niệm về biên

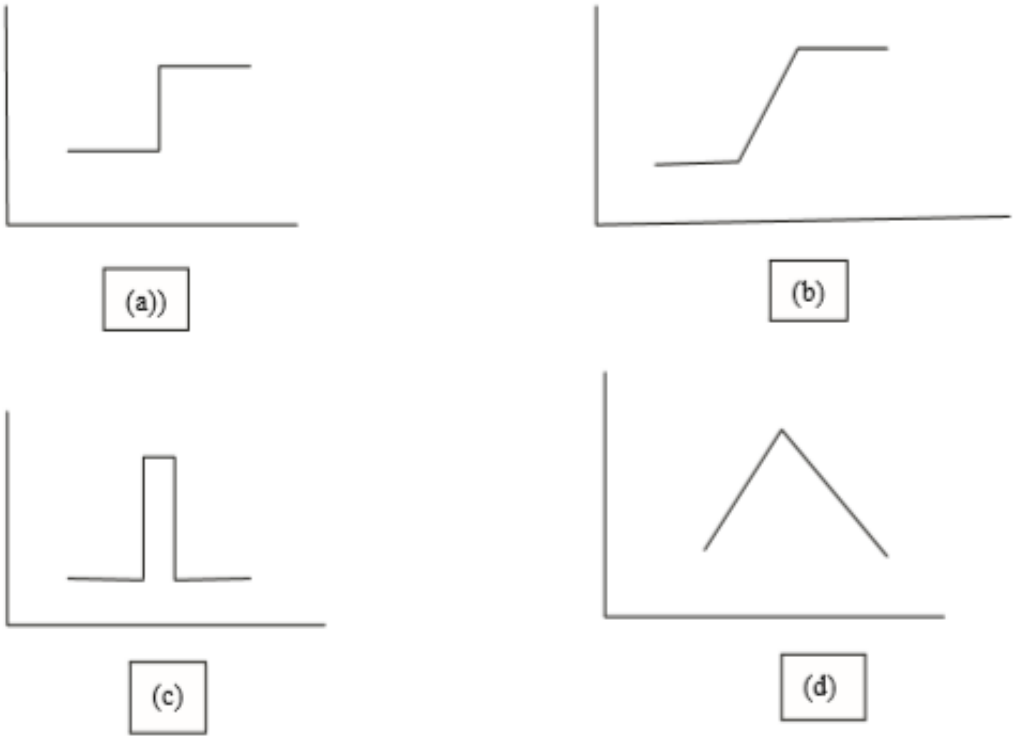
Biên là một phần chủ yếu trong phân tích ảnh vì các kỹ thuật phân đoạn ảnh chủ yếu dựa vào biên. Một điểm ảnh có thể coi là điểm biên nếu ở đó có sự thay đổi đột ngột về mức xám. Tập hợp các điểm biên tạo thành biên hay đường bao ảnh của ảnh. Ví dụ, trong một ảnh nhị phân, một điểm có thể gọi là biên nếu đó là điểm đen và có ít nhất một điểm trắng là lân cận.

Để hình dung tầm quan trọng của biên ta xét ví dụ sau: Khi người hoạ sĩ vẽ một cái bàn gỗ, chỉ cần vài nét phác thảo về hình dáng như cái mặt bàn, chân bàn mà không cần thêm các chi tiết khác, người xem đã có thể nhận ra nó là một cái bàn, nếu ứng dụng của ta là phân lớp nhận diện đối tượng, thì coi như nhiệm vụ đã hoàn thành. Tuy nhiên nếu đòi hỏi thêm về các chi tiết khác như vân gỗ hay màu sắc,v.v. thì với chừng ấy thông tin là chưa đủ.

Nhìn chung về mặt toán học người ta coi điểm biên của ảnh là điểm có sự biến đổi đột ngột về độ xám. Đường biên là tập các điểm biên. Một số kiểu đường biên hay gặp trên thực tế được minh họa trên hình 1-1.

Trong đó:

- a) Biên dạng nhẩy bậc
- b) Biên dốc
- c) Biên dạng xung vuông
- d) Biên dạng hình nón



Hình 1-1: Một số kiểu đường biên thông dụng

Phát hiện biên là một công cụ quan trọng trong xử lý ảnh số. Phương pháp phát hiện biên làm giảm một cách đáng kể khối lượng dữ liệu cần tính toán, chỉ giữ lại một số ít những thông tin cần thiết đồng thời vẫn bảo toàn được những cấu trúc quan trọng trong bức ảnh. Như vậy phát hiện biên một cách lý tưởng là xác định được tất cả các đường bao trong các đối tượng. Định nghĩa toán học của biên ở trên là cơ sở cho các kỹ thuật phát hiện biên. Điều quan trọng là sự biến thiên mức xám giữa các ảnh trong một vùng thường là nhỏ, trong khi đó biến thiên mức xám của điểm vùng giáp ranh (khi qua biên) lại khá lớn.

1.1.2 Phân loại các kỹ thuật phát hiện biên

Xuất phát từ định nghĩa toán học của biên người ta thường sử dụng hai phương pháp phát hiện biên là phương pháp phát hiện biên trực tiếp và phương pháp phát hiện biên gián tiếp. Các phương pháp này sẽ được trình bày trong các phần dưới đây.

1.1.2.1 Phương pháp phát hiện biên trực tiếp

Phương pháp này nhằm làm nổi biên dựa vào sự biến thiên về giá trị độ sáng của điểm ảnh. Kỹ thuật chủ yếu dùng phát hiện biên ở đây là kỹ thuật đạo hàm. Nếu lấy đạo hàm bậc nhất của ảnh ta có phương pháp Gradient; nếu lấy đạo hàm bậc hai ta có kỹ thuật Laplace. Hai phương pháp trên được gọi là phương pháp dò biên cục bộ. Ngoài ra người ta còn sử dụng phương pháp "đi theo đường bao" gọi là phương pháp dò biên tổng thể dựa vào nguyên lý quy hoạch hoạt động.

1.1.2.2 Phương pháp gián tiếp

Nếu bằng cách nào đấy, ta phân được ảnh thành các vùng thì đường phân ranh giữa các vùng đó chính là biên. Việc phân vùng ảnh thường dựa vào kết cấu (texture) bề mặt của ảnh.

Cũng cần lưu ý rằng, kỹ thuật dò biên và phân vùng ảnh là hai bài toán đối ngẫu của nhau. Thực vậy, dò biên để thực hiện phân lớp đối tượng và một khi đã phân lớp xong có nghĩa là đã phân vùng được ảnh. Và ngược lại, khi phân vùng, ảnh đã phân lập được thành các đối tượng, ta có thể phát hiện được biên. Phương pháp dò biên trực tiếp tỏ ra khá hiệu quả vì ít chịu ảnh hưởng của nhiễu song nếu sự biến thiên độ sáng không đột ngột, phương pháp này lại kém hiệu quả. Phương pháp dò biên gián tiếp tuy có khó cài đặt xong lại áp dụng khá tốt khi sự biến thiên độ sáng nhỏ.

1.1.3 Quy trình phát hiện biên trực tiếp

Bước 1: Khử nhiễu ảnh

- Vì ảnh thu nhận thường có nhiễu, nên bước đầu tiên là phải khử nhiễu, việc khử nhiễu được thực hiện bằng các kỹ thuật khử nhiễu khác nhau.

Bước 2: Làm nổi biên

- Tiếp theo là làm nổi biên bởi các toán tử đạo hàm.

Bước 3: Định vị điểm biên

- Vì các kỹ thuật làm nổi biên có hiệu ứng phụ là tăng nhiễu, do vậy sẽ có một số điểm biên giả cần loại bỏ.

Bước 4: Liên kết và trích chọn biên

- Như đã nói, phát hiện biên và phân vùng ảnh là một bài toán đối ngẫu, vì thế cũng có thể phát hiện biên thông qua việc phân vùng ảnh.

1.1.4 Một số phương pháp phát hiện biên

Các phương pháp phát hiện biên truyền thống thường dựa trên kết quả của phép tích chập (convolution) giữa bức ảnh cần nghiên cứu và một bộ lọc 2D (filter) thường được gọi là mặt nạ (mask).

Cấu trúc và giá trị của các toán tử phát hiện biên sẽ xác định hướng đặc trưng mà toán tử nhạy cảm với biên. Có một số toán tử thích hợp cho các đường biên có hướng nằm ngang, một số toán tử lại thích hợp cho việc tìm kiếm biên dạng thẳng đứng hay theo hướng đường chéo.

Hiện nay thì có nhiều phương pháp phát hiện biên đang được sử dụng, tuy nhiên có hai phương pháp phát hiện biên cơ bản đó là: Phương pháp Gradient và phương pháp Laplace.

1.1.4.1 Phương pháp Gradient

Đạo hàm bậc nhất theo hướng ngang và dọc được tính theo công thức sau:

$$\Delta f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \tag{1.1}$$

Biên độ của gradient vector hay độ lớn tổng cộng của giá trị đạo hàm nằm tại biên là kết hợp của cả hai giá trị này theo công thức:

$$\Delta f = |\Delta f| = \sqrt{G_x^2 + G_y^2} \tag{1.2}$$

Hướng của gradient vector được xác định theo:

$$\nabla f = \tan^{-1}\left(\frac{G_x}{G_y}\right) \quad (1.3)$$

Hướng của biên sẽ vuông góc với hướng của gradient vector này.

1.1.4.1.1 Toán tử sobel

Trên thực tế Sobel sử dụng hai mặt nạ có kích thước [3 x 3] trong đó một mặt nạ chỉ đơn giản là sự quay của mặt nạ kia đi một góc 900 như ở hình1- 2. Các mặt nạ này được thiết kế để tìm ra các đường biên theo chiều đứng và chiều ngang một cách tốt nhất. Khi thực hiện phép convolution giữa ảnh và các mặt nạ này ta nhận được các gradient theo chiều đứng và chiều ngang Gx, Gy. Toán tử Sobel có dạng như hình 1-2.

–1	–2	–1
0	0	0
1	2	1

–1	0	1
–2	0	2
–1	0	1

Hình 1-2: Toán tử Sobel

1.1.4.1.2 Toán tử Prewitt

Phương pháp Prewitt gần giống với Sobel. Đây là phương pháp lâu đời nhất, cổ điển nhất. Toán tử Prewitt được mô tả trên hình 1-3.

–1	–1	–1
0	0	0
1	1	1

–1	0	1
–1	0	1
–1	0	1

Hình 1-3: Toán tử Prewitt

1.1.4.1.3 Toán tử Roberts

Tương tự như Sobel, ta tính đường biên ngang và dọc một cách riêng rẽ dùng hai mặt nạ như hình 1-4, sau đó tổng hợp lại để cho đường biên thực của ảnh. Tuy nhiên do mặt nạ của Robert khá nhỏ nên kết quả là bị ảnh hưởng khá nhiều của nhiễu.

0	0	0
0	–1	0
0	0	1

0	0	0
0	0	–1
0	1	0

Hình 1-4: Toán tử Roberts

1.1.4.1.4 Phương pháp Canny

Phương pháp này sử dụng hai mức ngưỡng cao và thấp. Ban đầu ta dùng mức ngưỡng cao để tìm điểm bắt đầu của biên, sau đó chúng ta xác định hướng phát triển của biên dựa vào các điểm ảnh liên tiếp có giá trị lớn hơn mức ngưỡng thấp. Ta chỉ loại bỏ các điểm có giá trị nhỏ hơn mức ngưỡng thấp. Các đường biên yếu sẽ được chọn nếu chúng được liên kết với các đường biên khỏe.

Phương pháp Canny bao gồm các bước sau:

- Bước 1. Trước hết dùng bộ lọc Gaussian để làm mịn ảnh.

$$G'(x) = \left(-\frac{x}{\sigma^2}\right) e^{-\left(\frac{x^2}{2\sigma^2}\right)} \tag{1.4}$$

- Bước 2. Sau đó tính toán gradient (1.5) và (1.6) của đường biên của ảnh đã được làm mịn.

$$C_x[x, y] = -\left(\frac{j}{\sigma^2}\right) e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \tag{1.5}$$

$$C_y[x, y] = -\left(\frac{i}{\sigma^2}\right) e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \tag{1.6}$$

- Bước 3. Tiếp theo là loại bỏ những điểm không phải là cực đại.
- Bước 4. Bước cuối cùng là loại bỏ những giá trị nhỏ hơn mức ngưỡng.

Phương pháp này hơn hẳn các phương pháp khác do ít bị tác động của nhiễu và cho khả năng phát hiện các biên yếu. Nhược điểm của phương pháp này là nếu chọn ngưỡng quá thấp sẽ tạo ra biên không đúng, ngược lại nếu

chọn ngưỡng quá cao thì nhiều thông tin quan trọng của biên sẽ bị loại bỏ. Căn cứ vào mức ngưỡng đã xác định trước, ta sẽ quyết định những điểm thuộc biên thực hoặc không thuộc biên. Nếu mức ngưỡng càng thấp, số đường biên được phát hiện càng nhiều (nhưng kèm theo là nhiễu và số các đường biên giả cũng xuất hiện càng nhiều). Ngược lại nếu ta đặt mức ngưỡng càng cao, ta có thể bị mất những đường biên mờ hoặc các đường biên sẽ bị đứt đoạn.

Các ưu điểm của phương pháp Canny là:

- Cực đại hóa tỷ số tín hiệu trên nhiễu làm cho việc phát hiện các biên thực càng chính xác.
- Đạt được độ chính xác cao của đường biên thực.
- Làm giảm đến mức tối thiểu số các điểm nằm trên đường biên nhằm tạo ra các đường biên mỏng, rõ.

1.1.4.2 Phương pháp Laplace

Các phương pháp đánh giá Gradient ở trên làm việc khá tốt khi mà độ sáng thay đổi rõ nét. Khi mức sáng thay đổi chậm, miền chuyển tiếp trải rộng, phương pháp cho hiệu quả hơn đó là sử dụng phương pháp đạo hàm bậc hai gọi là phương pháp Laplace. Kết quả nghiên cứu cho thấy phương pháp Gradient rất nhạy cảm với nhiễu và thường tạo nên biên kép. Toán tử Laplace dùng nhiều kiểu mặt nạ khác nhau để xấp xỉ đạo hàm bậc hai. Dưới đây là 3 kiểu mặt nạ hay dùng.

$$H1=\begin{bmatrix}0 & -1 & 0\\ -1 & 4 & -1\\ 0 & -1 & 0\end{bmatrix} \qquad H2=\begin{bmatrix}-1 & -1 & -1\\ -1 & 8 & -1\\ -1 & -1 & -1\end{bmatrix} \qquad H2=\begin{bmatrix}1 & -2 & 1\\ -2 & 8 & -2\\ 1 & -2 & 1\end{bmatrix}$$

Hình 1-5: Kỹ thuật Laplace

Kỹ thuật Laplace cho đường biên mảnh, tức là đường biên có độ rộng bằng một điểm ảnh. Tuy nhiên, kỹ thuật này rất nhạy cảm với nhiễu vì đạo hàm bậc hai thường không ổn định.

1.1.4.3 Phương pháp Laplacian

Dùng phương pháp Gradient sẽ cho kết quả là ảnh nhận được có cấu trúc không rõ nét do tạo nên những đường biên dày, không sắc nét. Để nhận được các đường biên mỏng và rõ nét, ta phải tiến hành các bước xử lý tiếp theo như loại bỏ những điểm không phải là cực trị đồng thời áp dụng kỹ thuật liên kết biên. Ngoài ra ta còn gặp phải vấn đề là làm thế nào để xác định được mức ngưỡng một cách chính xác. Việc chọn đúng giá trị ngưỡng phụ thuộc rất nhiều vào nội dung của từng bức ảnh. Nếu ta tăng gấp đôi kích thước của một bức ảnh mà không thay đổi giá trị cường độ của các điểm ảnh, ta sẽ nhận được gradients bị suy giảm đi một nửa. Mặt khác kích thước của mặt nạ (masks) cũng ảnh hưởng nhiều đến giá trị của Gradient trong ảnh.

Phương pháp Gradient chỉ thích hợp cho các vùng ảnh độ tương phản thay đổi có tính nhảy bậc, điều này gây khó khăn cho phát hiện các đường thẳng. Để khắc phục nhược điểm này ta thường dùng đạo hàm bậc hai. Phương pháp Laplacian cho phép xác định đường biên dựa vào giá trị 0 của đạo hàm bậc hai của ảnh. Laplacian của một ảnh tại điểm I(x, y) được tính theo (1.7):

$$L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{1.7}$$

Laplacian được kết hợp với bộ lọc làm mịn ảnh để tìm biên.Xét công thức sau:

$$h(r) = -e^{-\frac{r^2}{2\sigma^2}} \tag{1.8}$$

Ở đây $r^2=x^2+y^2$và σ là độ lệch chuẩn. Nếu thực hiện phép tích chập của hàm này với ảnh cần tìm biên, kết quả là ảnh sẽ bị mờ đi, mức độ mờ phụ thuộc vào giá trị của σ. Laplacian của *h* tức đạo hàm bậc hai của *h* theo *r* là:

$$\nabla^2 h(r) = -\left[\frac{r^2 - \sigma^2}{\sigma^4}\right] e^{-\frac{r^2}{2\sigma^2}} \tag{1.9}$$

Hàm này thường được gọi là Laplacian of a Gaussian (LoG) do (1.8) có dạng Gaussian. Trong phương pháp này, bộ lọc Gaussian được kết hợp với Laplacian cho phép hiển thị những vùng ảnh có cường độ thay đổi nhanh do đó làm tăng hiệu quả phát hiện biên. Nó cho phép làm việc với một diện tích rộng hơn xung quanh điểm ảnh đang được nghiên cứu nhằm phát hiện chính xác hơn vị trí của đường biên. Nhược điểm của phương pháp này là không xác định được hướng của biên do sử dụng hai bộ lọc Laplacian quá khác nhau có dạng như trên hình 1-6.

0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

Hình 1-6: Toán tử Laplacian

1.2 Mô tả hình dạng dựa trên đường bao

Mô tả hình dạng dựa trên biên sẽ chỉ khai thác thông tin biên của đường bao đối tượng được mô tả. Có hai kiểu phương pháp tiếp cận để mô tả đường bao hình dạng đó là phương pháp tiếp cận toàn cục và phương pháp tiếp cận cấu trúc.

Phương pháp tiếp cận toàn cục không phân chia hình dạng thành các phần mà dùng một vector xác định đường bao để mô tả hình dạng đặc trưng từ đường biên được sử dụng để mô tả hình dạng. Độ đo khoảng cách giữa các vector đặc trưng thường được sử dụng để đo độ tương tự hình dạng.

Phương pháp tiếp cận cấu trúc thì phân tách các đường bao của hình dạng thành các đoạn dựa trên các điều kiện phân tách. Biểu diễn cuối cùng của nó thường sử dụng các xâu, một chuỗi hoặc một đồ thị, các biện pháp tương tự được thực hiện bằng cách kết hợp chuỗi hoặc đồ thị một cách phù hợp. Theo hướng tiếp cận này thì các chuỗi, đồ thị hoặc cây sẽ được biểu diễn để đạt được những kết quả cuối cùng. Các thuật toán đối sánh chuỗi hoặc đối sánh đồ thị sẽ được dùng để đo độ tương tự hình dạng.

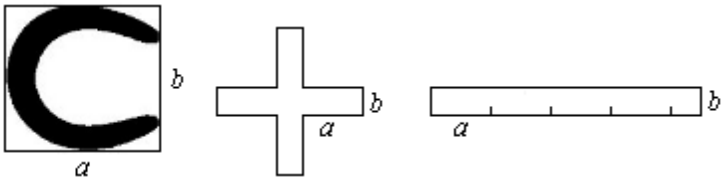
1.2.1 Mô tả theo tiếp cận toàn cục

Kỹ thuật biểu diễn đường bao hình dạng toàn cục nó thường tính toán vector đặc trưng đa chiều từ các thông tin đường bao của hình dạng. Việc đối sánh giữa hai hình dạng với nhau là một quá trình đơn giản thường xây dựng bằng cách sử dụng một độ đo khoảng cách, chẳng hạn như khoảng cách Euclide, hoặc khoảng cách cityblock, và nó cũng thường được sử dụng trong các ứng dụng thực tế.

Mô tả hình dạng toàn cục đơn giản nhỏ gọn, tuy nhiên mô tả hình dạng không được 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.

1.2.1.1 Mô tả hình dạng đơn giản(Simple shape descriptors)

Mô tả hình dạng đơn giản nhất có thể được biểu diễn bằng các yếu tố như: diện tích, vùng, hướng trục chính, độ tròn (perimeter^2/area), độ uốn, độ lệch tâm. Những mô tả toàn cục thường chỉ có thể phân biệt hình dạng có sự khác biệt lớn, do đó chúng thường được sử dụng để lọc, để loại bỏ những cái sai hoặc kết hợp với các mô tả hình dạng khác để phân biệt hình dạng. Chúng không phù hợp với các mô tả hình dạng độc lập. Ví dụ, độ lệch tâm của hình dạng trong hình 1-7(a) là gần tới 1 vì (a=b), do đó nó không mô tả đúng về hình dạng bởi vì theo quan sát thì nó một hình thon dài. Trong trường hợp này, độ tròn sẽ mô tả tốt hơn. Hai hình dạng trong hình 1-7(b) và 1-7(c) có độ tròn tương tự nhau vì (a=2b), tuy nhiên, chúng là những hình dạng rất khác nhau. Trong trường hợp này, độ lệch tâm là mô tả tốt hơn**.**



(a) (b) (c)

Hình 1-7: Minh họa độ lệch tâm của hình dạng

1.2.1.2 Dấu hiệu đặc trưng hình dạng

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

Dấu hiệu đặc trưng hình dạng không bị tác động bởi dịch chuyển và co dãn hình dạng. Dấu hiệu hình dạng có thể được lượng tử hóa thành một biểu đồ dấu hiệu, biểu đồ này có thể sử dụng cho đối sánh và bất biến với phép quay. Dấu hiệu hình dạng thường nhạy cảm với nhiễu và những thay đổi trên đường bao, do vậy nó có thể gây ra những lỗi trong việc đối sánh hình dạng. Vậy nên, dấu hiệu đặc trưng hình dạng thường không sử dụng trực tiếp để mô tả hình dạng.

1.2.1.3 Momen đường bao

Momen biên có thể được dùng để giảm kích thước của các biểu diễn trên đường bao. Giả sử biên hình dạng đã được biểu diễn bởi một dấu hiệu hình dạng Z(i), khi đó momen thứ r là m_r và momen tâm là μ_r, có công thức ước tính như sau:

$$m_r = \frac{1}{N}\sum_{i=1}^{N}[z(i)]^r \tag{1.10}$$

Và

$$\mu_r = \frac{1}{N}\sum_{i=1}^{N}[z(i) - m_i]^r \tag{1.11}$$

Trong đó, N là số các điểm biên

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

Và $\overline{M}_r = M_r / (M_2)^{r/2}$

Để mô tả bất biến với các phép dịch chuyển, phép quay và co dãn của hình dạng.

Ưu điểm của mô tả momen đường bao chính là nó dễ dàng được thực hiện tuy nhiên rất khó để gán những momen bậc cao hơn với các giải thích liên quan tới vật lý.

1.2.2 Mô tả theo tiếp cận 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 đoạn đường bao và sau đó được mã hóa thành các chuỗi tổng quát: $S=S_1, S_2, \ldots S_n$.

Ở đây S_i là các phần tử của mã xích, một cạnh của đa giác, hình vuông hoặc là một mặt spline. S_i có thể chứa một số thuộc tính ví dụ như chiều dài, độ cong trung bình, độ cong lớn nhất, khả năng uốn,v.v. Các chuỗi có thể sử dụng trực tiếp để mô tả hoặc có thể sử dụng như là một đầu vào. Dưới đây là một vài mô tả biểu diễn dưới dạng cấu trúc.

1.2.2.1 Biểu diễn mã xích

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.

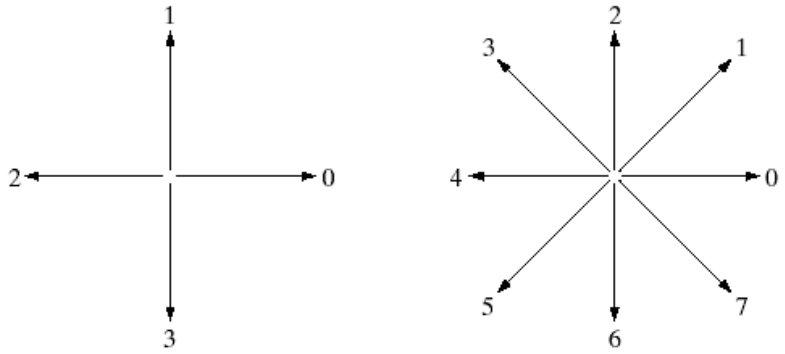
Mã xích dùng để biểu diễn hình dạng phải không phụ thuộc vào sự lựa chọn điểm ảnh biên bắt đầu trong chuỗi. Một khả năng để chuẩn hóa chuỗi mã

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.

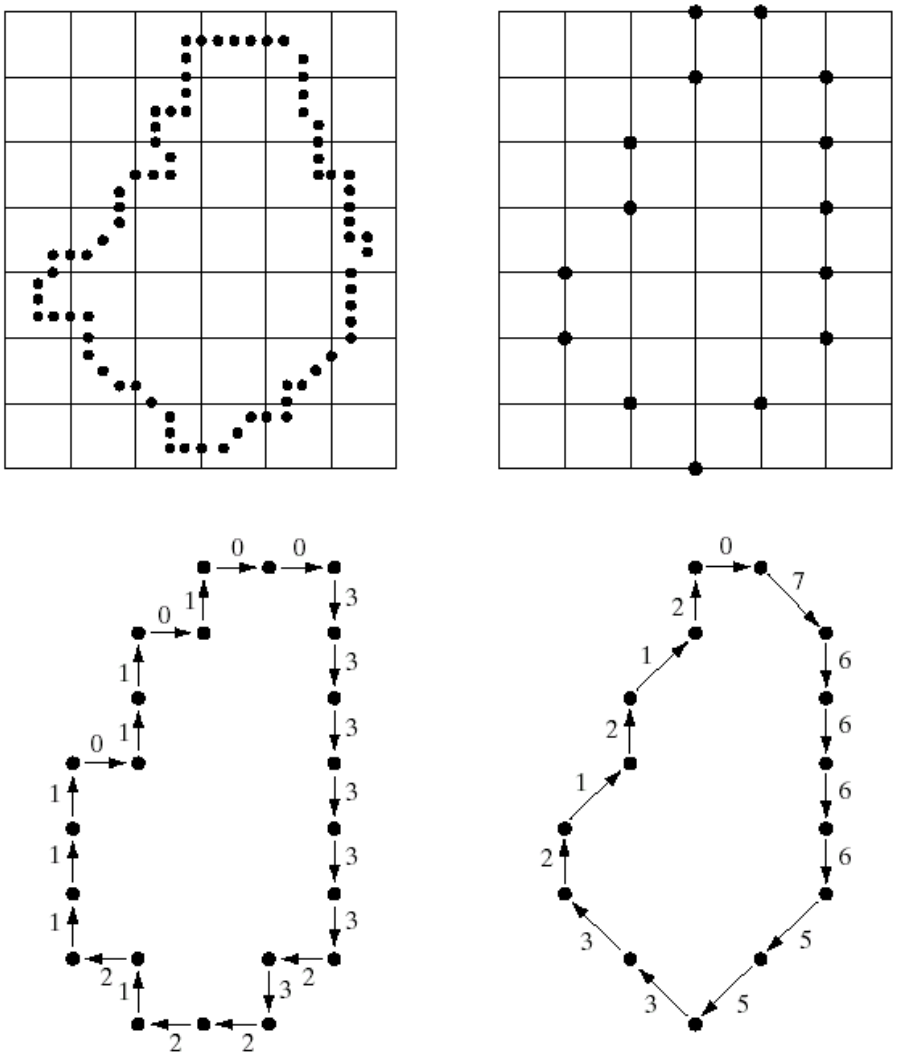
Dùng mã xích biểu diễn hình dạng và đối sánh có nhiều hạn chế, mã xích bị ảnh hưởng bởi nhiễu đường biên và biến dạng, thêm vào đó là kích thước của chuỗi mã dài. Mã xích mà thường được sử dụng là đầu vào của những phân tích ở mức độ cao, ví dụ như xấp xỉ đa giác và tìm điểm uốn.

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. 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 1-8. 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 1-8.

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 giữa 2 yếu tố liền kề.



Hình 1-8: Các hướng của đoạn thẳng đơn vị: (a): 4 hướng, (b): 8 hướng

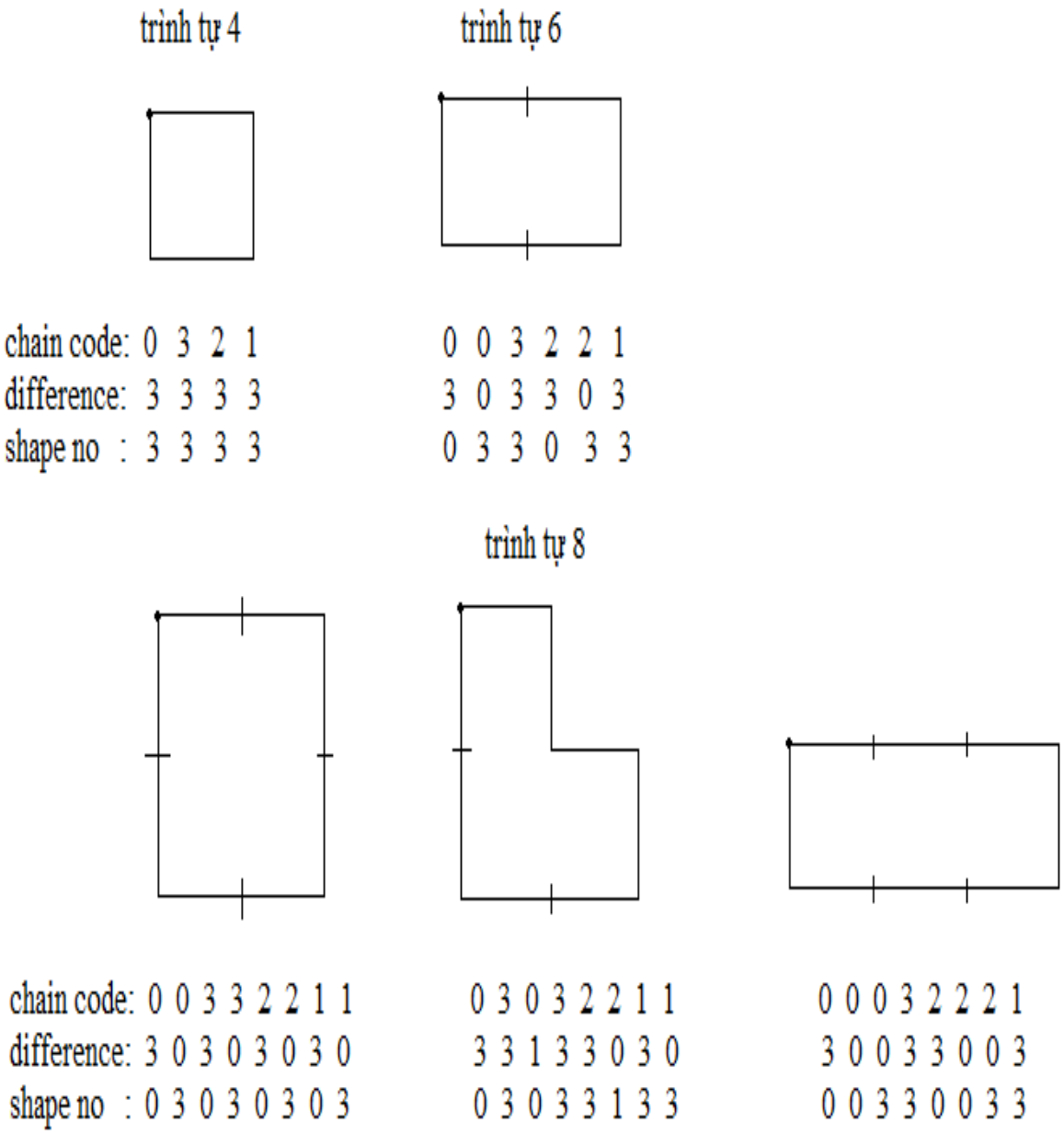


Hình 1-9: Biểu diễn của một chuỗi mã (theo 4 hướng và 8 hướng)

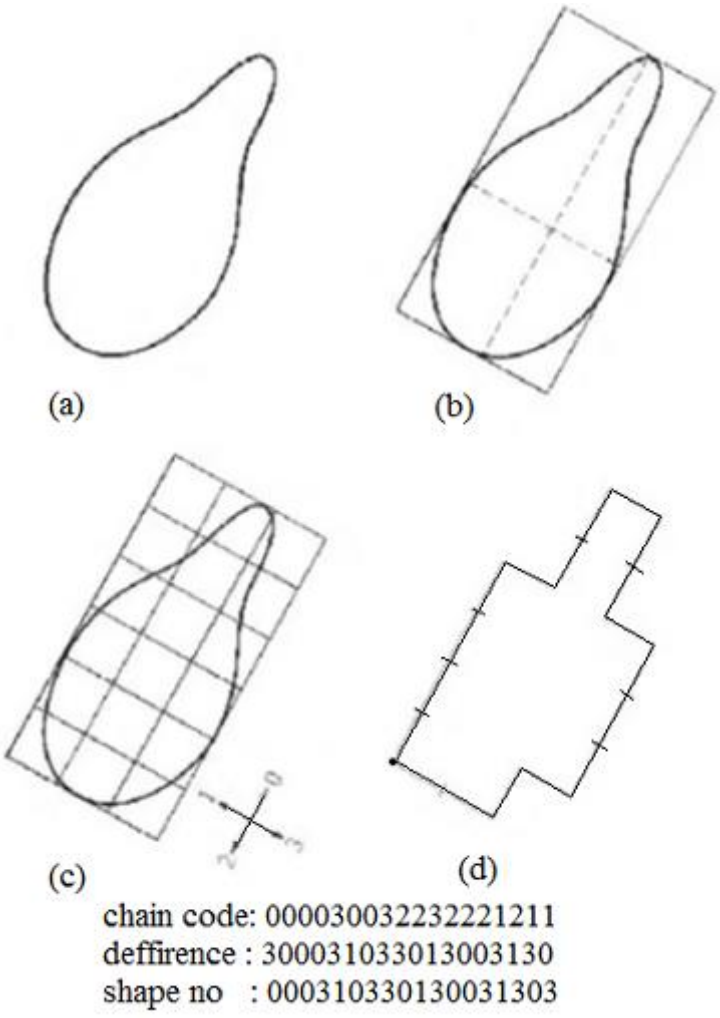
1.2.2.2 Shape number

Shape number của một biểu diễn đường biên được định nghĩa là sự khác biệt đầu tiên của cường độ nhỏ nhất. Trình tự *n* của một Shape number

là số lượng các chữ số được biểu diễn. Hình 1-10 minh họa hình dạng của trình tự 4, 6, 8.



Hình 1-10: Biểu diễn hình dạng sử dụng Shape number



Hình 1-11: Các bước tính toán shape number

Chúng ta xét một ví dụ cụ thể, giả sử n=18 được quy định cụ thể cho biên như hình 1-11(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 1-11(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 1-11(c). Cuối cùng có được chuỗi mã và sử dụng điểm khác biệt đầu tiên (first difference) để tính toán shape number.

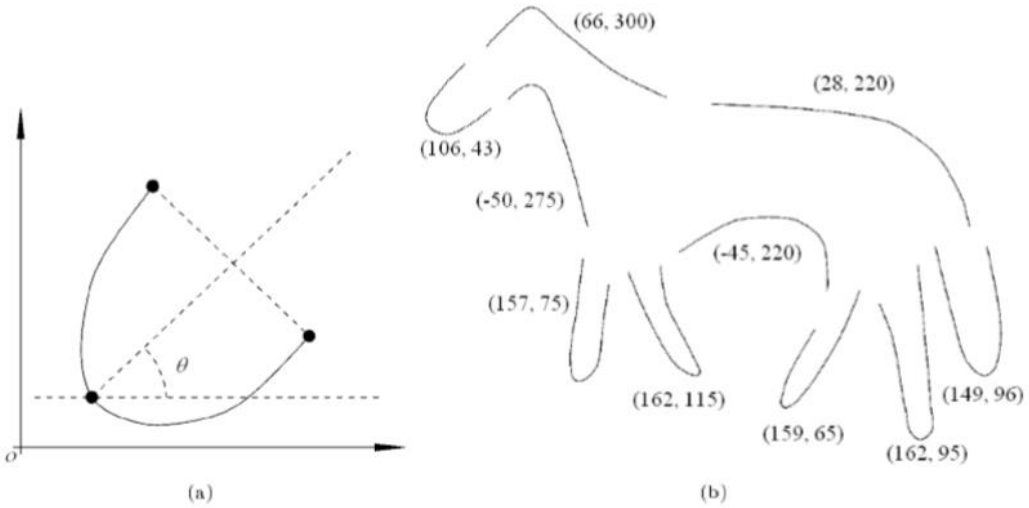
1.2.2.3 Phân tích đa giác

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 bốn 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 trưng 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.

1.2.2.4 Kỹ thuật làm mịn đường cong

Phân tích đường cong mịn như mô tả hình dạng. Phân đoạn giữa các điểm uốn độ cong từ một điểm biên được làm phẳng, được gọi là các mã thông báo. Tính năng cho mỗi mã thông báo là độ cong tối đa và hướng của nó. Trong hình 1-12, số đầu tiên trong ngoặc đơn là độ cong tối đa của nó và số thứ hai là hướng của nó.



Hình 1-12: Phân tích đường cong mịn

Trong đó: (a) θ là định hướng của mã thông báo này; (b) một ví dụ về phân hủy đường cong mịn.

Sự tương tự giữa hai hình (a) và (b) được đo bằng khoảng cách Euclide có trọng số. Độ tương tự của hình dạng được đo theo khoảng cách không trọng số. Việc tìm kiếm hình dạng dựa trên biểu diễn mã thông báo đã thể hiện rõ ràng khi có các đối tượng bị dịch, mở rộng và xoay một phần.

1.2.2.5 Phương pháp không gian tỷ lệ

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: $A = (S_1^A, S_2^A, ..., S_M^B)$ và $B = (S_1^B, S_2^B, ..., S_M^B)$, 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.

1.3 Đối sánh ảnh

1.3.1 Giới thiệu về đối sánh ảnh

Đối sánh ảnh là một bài toán đã và đang thu hút được sự quan tâm của các nhà nghiên cứu và phát triển. Mỗi khi bài toán này được giải quyết, nó mở ra rất nhiều các ứng dụng hữu ích như: tìm kiếm ảnh, nhận dạng, theo dõi và phát hiện đối tượng, ghép ảnh, v.v.

Đối sánh hai ảnh là tìm ra những vùng giống nhau trên hai ảnh. Thông thường, để đối sánh ảnh cần so sánh các phần tử cơ bản cấu thành nên nó. Giải pháp đầu tiên cho vấn đề đối sánh ảnh được đề xuất bởi Hobrough vào cuối những năm 1950. Hệ thống tự động tìm kiếm các điểm tương quan được

giới thiệu lần đầu bởi công ty Wild Heerbrugg năm 1964 nhưng lại không được sử dụng phổ biến. Tuy nhiên, ý tưởng áp dụng mối tương quan chéo của Hobrough lại được nhiều người sử dụng. Từ những năm 1970, việc tập trung phát triển đối sánh ảnh và đối sánh tương quan gặt hái được nhiều thành công và được áp dụng trong hệ thống đo độ tương tự cho ảnh (Helava, 1978). Ngày nay, công nghệ đối sánh ảnh được tính hợp trong nhiều phần mềm xử lý ảnh được sử dụng như là một công cụ tính toán. Có rất nhiều nghiên cứu được thực hiện với mong muốn tìm cặp điểm tương đồng trên hai bức ảnh. Thuật toán tìm kiếm điểm tương đồng có thể thực hiện được trên ảnh 2D.

Vấn đề chính của việc đối sánh ảnh là việc chọn thực thể đối sánh (Một thực thể trong ảnh này được so sánh với một thực thể trong ảnh khác) và lựa chọn độ đo tương tự (Một độ đo định lượng đánh giá đối sánh của toàn bộ các thực thể). Đối sánh từng pixel sẽ không khả thi với những ảnh có kích thước lớn vì nó sẽ cần tính toán nhiều hơn, mất nhiều thời gian hơn, hoặc muốn rút ngắn thời gian thì cần có phần cứng xử lý mạnh hơn. Hơn nữa, nó thường dẫn đến sự nhập nhằng do các giá trị mức xám của ảnh xuất hiện lặp đi lặp lại và do nhiễu của ảnh. Do đó bài toán đối sánh ảnh thuộc về nhóm bài toán giả định yếu (ill-posed problems). Để chuyển đổi bài toán đối sánh ảnh thành bài toán giả định chặt (Well-posed problem) thì các thực thể đối sánh, các độ đo tương tự, các ràng buộc hình học và các giả thiết phải được định nghĩa trong một giới hạn phạm vi nhất định, nghĩa là không gian của tất cả các giải pháp sẽ bị hạn chế. Hai phương pháp cơ bản của đối sánh ảnh đã được phát triển và sử dụng trong quan trắc và thị giác máy là phương pháp dựa trên vùng và phương pháp dựa trên đặc trưng. Tổng quan của các phương pháp này được chỉ ra trong Bảng 1.1.

Bảng 1.1: Phương pháp đối sánh hình ảnh

Phương pháp đối sánh	Độ tương tự	Đối tượng đối sánh
Dựa trên vùng	Tương quan, đối sánh hình vuông nhỏ nhất.	Giá trị mức xám
Dựa theo đặc trưng	Hàm chi phí	Điểm quan tâm, cạnh, vùng

Các giá trị mức xám là những thực thể trong đối sánh dựa trên vùng. Đối sánh từng điểm ảnh dễ gặp phải vấn đề nhập nhằng, do vậy, các giá trị mức xám của một vài điểm ảnh lân cận sẽ được sử dụng. Một phần ảnh được cắt từ ảnh được gọi là mẫu, được sử dụng để tìm kiếm trong ảnh thứ hai. Mẫu gồm m*n điểm ảnh (thông thường là m=n). Vị trí của mẫu là điểm ảnh trung tâm, do vậy, m và n thường là lẻ. Mẫu sẽ được so sánh với phần ảnh có kích thước tương tự trong ảnh thứ hai. Việc so sánh được hạn chế với vùng được gọi là tìm kiếm dựa trên vùng hoặc là tìm kiếm cửa sổ. Giá trị độ đo tương tự được tính toán tại mỗi vị trí của mẫu trong vùng tìm kiếm. Dựa trên đặc tính của độ đo tương tự, mà các điểm tương ứng với tâm của mẫu sẽ là những điểm có độ đo tương tự lớn nhất hoặc nhỏ nhất. Trong phép quan trắc thì tương quan chéo và đối sánh bình phương nhỏ nhất là những công nghệ được sử dụng nhiều cho đối sánh dựa trên vùng. Bên cạnh đó thông tin tương hỗ và khoảng cách ảnh cũng có thể được áp dụng.

Ngược lại với đối sánh dựa trên vùng sử dụng các toán tử trực tiếp trên các giá trị mức xám, các phương pháp dựa trên đặc trưng sẽ dựa trên việc đối sánh các đặc trưng được trích chọn như điểm, cạnh hoặc vùng. Các thủ tục đối sánh dựa trên đặc trưng bao gồm ba bước (được điều chỉnh từ Forstner, 1986):

- Chọn các đặc trưng riêng biệt (điểm,cạnh, góc) trong các ảnh riêng biệt.

- Xây dựng danh sách sơ bộ các cặp ứng viên của các đặc trưng tương ứng dựa trên độ đo tương tự được lựa chọn.
- Lấy danh sách cuối cùng các cặp đặc trưng phù hợp với mô hình đối tượng.

1.3.2 Đối sánh ảnh dựa trên đặc trưng

1.3.2.1 Điểm đặc trưng

Những điểm có độ chênh lệch cao về giá trị mức xám hoặc hàm tương quan tự động cao và có độ dốc gradient lớn thì gọi là những điểm đặc trưng. Các điểm đặc trưng nên có sự khác biệt, bất biến đối với các biến đổi hình học và biến đổi phổ, ổn định (một điểm nên xuất hiện trong tất cả các ảnh được đối sánh) (Forstner, 1986). Thủ tục để tìm tìm kiếm những điểm đặc trưng trong ảnh đối sánh được thực hiện qua hai bước:

- Tính toán các tham số đặc trưng ở mỗi cửa sổ của ảnh được chọn.
- So sánh giá trị của các tham số với một ngưỡng cho trước.

Các tham số đặc trưng là khác nhau cho mỗi toán tử nhưng về cơ bản dựa trên các giá trị mức xám (cấu trúc và kết cấu) bên trong cửa sổ được đánh giá. Chỉ có những cửa sổ mà có giá trị tham số lớn hơn hoặc nhỏ hơn ngưỡng mới được chấp nhận là điểm đặc trưng. Một danh sách các điểm đặc trưng của mỗi ảnh được đối sánh với tọa độ của nó (điểm trung tâm của mỗi cửa sổ trượt) và mô tả của nó là kết quả của quá trình này. Trong (Luhmann và Altrogge, 1986) đã đề xuất 3 toán tử điểm đặc trưng tên là Moravec, Forstner và Dreschler. Tuy nhiên Moravec và Forstner hoạt động tốt hơn với việc tìm kiếm các điểm đặc trưng theo các điều kiện hình học khác nhau.

1.3.2.2 Cạnh và vùng

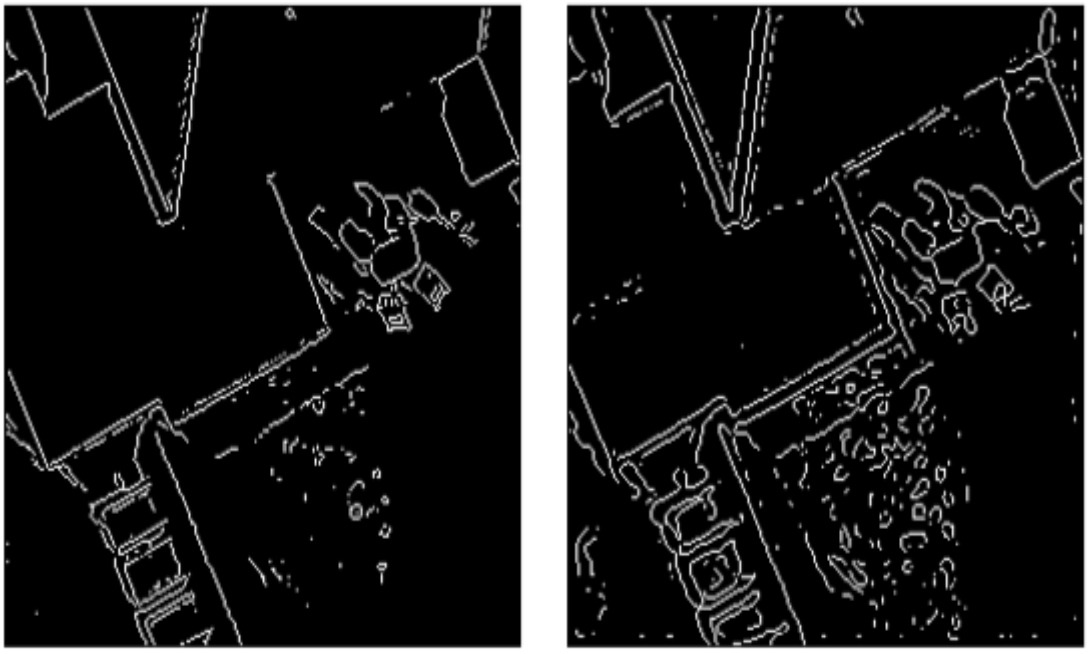
Các cạnh có thể được mô tả như các điểm gián đoạn trong hàm mức xám, ví dụ: các giá trị mức xám thay đổi nhanh chóng trong một khu vực nhỏ. Các cạnh thường tương ứng với đường bao của các đối tượng được hiển thị trong hình ảnh. Quá trình trích chọn cạnh khá phức tạp và bao gồm các bước sau (Schenk, 1999):

- Xác định các điểm ảnh nằm trên cạnh, giá trị mức xám không liên tục được phát hiện bởi một phép đo được gọi là toán tử cạnh. Một ngưỡng cho các giá trị mức xám khác nhau được thiết lập để quyết định các điểm là điểm cạnh.
- Liên kết các điểm cạnh thành các cạnh.
- Nhóm các cạnh: tức là xác định phân đoạn đường thẳng, đường đa giác, đường gấp khúc, đường song song, v.v.

Toán tử cạnh sẽ phát hiện ra sự thay đổi của giá trị mức xám trong ảnh, một trong số chúng dựa trên đạo hàm bậc nhất của các hàm mức xám để tìm ra cực trị và định vị điểm cạnh. Một số toán tử cạnh có thể dùng như toán tử Robert (Robert Cross), toán tử Sobel (Sobel Operator). Cả 2 toán tử đều dựa trên hướng, phát hiện theo chiều ngang, phát hiện theo chiều dọc. Toán tử Sobel sẽ ít bị ảnh hưởng bởi nhiễu của ảnh vì bao gồm cả những điểm ảnh lân cận.

Toán tử Laplacian thuộc về nhóm các toán tử đạo hàm bậc hai, nó độc lập về hướng để giảm ảnh hưởng của nhiễu trong ảnh. Nó kết hợp với toán tử Gaussian làm mịn ảnh. Sau khi áp dụng kết quả Laplacian của toán tử Gaussian trên ảnh gốc thì các điểm cạnh tương ứng với giá trị zero. Toán tử LoG giống toán tử đạo hàm bậc nhất được mô tả chi tiết ở Hình 1-13 và 1-14:

Hình 1-13: Ảnh gốc



Hình 1-14: Phát hiện cạnh

CHƯƠNG 2: ĐỐI SÁNH HÌNH DẠNG SỬ DỤNG NGỮ CẢNH HÌNH DẠNG

2.1 Giới thiệu

Hình dạng là một đặc trưng thị giác quan trọng và nó là một trong những đặc trưng cơ bản được sử dụng để mô tả nội dung ảnh. Tuy nhiên việc biểu diễn hình dạng và mô tả chúng là một việc khó khăn. Điều này là do khi một đối tượng thế giới thực dạng 3-D được chiếu lên một mặt phẳng 2-D, một chiều của thông tin đối tượng sẽ bị mất đi. Do vậy là hình dạng được trích chọn từ hình ảnh chỉ biểu diễn một phần đối tượng được chiếu. Bên cạnh đó vấn đề phức tạp hơn khi hình dạng thường bị hỏng, biến dạng, cắt bỏ nhiễu, trùng lặp.

Biểu diễn hình dạng thường đi tìm kiếm những đặc trưng hình dạng quan trọng và hiệu quả dựa trên thông tin đường bao hoặc đường bao với nội dung bên trong của nó. Nhiều kỹ thuật mô tả hình dạng đã được phát triển trong thời gian qua như: dấu hiệu hình dạng, ngữ cảnh hình dạng, ma trận hình dạng, phổ, v.v. Trong phần này chủ yếu trình bày phương pháp mô tả hình dạng sử dụng ngữ cảnh hình dạng.

2.2 Độ đo khoảng cách hình dạng

2.2.1 Khoảng cách min-max

Được thực hiện dựa trên ý tưởng lấy phần giao của hai lược đồ cần so sánh, ta sẽ được một lược đồ, tính tổng các giá trị có được từ lược đồ này cho ta được độ đo min – max. Đối với độ đo min: ta tính dựa vào giá trị min tại mỗi K bin.

$$\text{Intersection: (h(I), h(M)} = \sum_{j=1}^{K} \min\{h(I)[j], \text{h(M) j}]\} \qquad (2.1)$$

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

$$\text{Intersection (h(I), h(M)} = \sum_{j=1}^{K} \max\{h(I)[j], h(M)[j]\} \qquad (2.2)$$

$$\text{Matching(h(I), h(M))} = \frac{Intersection(h(I), h(M))}{\max\left(\sum_i h(I)[i], \sum_i h(M)[i]\right)} \quad (2.3)$$

2.2.2 Khoảng cách Euclid

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

$$\text{Intersection (h(I), h(M)} = \sum_{j=1}^{K} \left| h(I) - h(M)^2 \right| \quad (2.4)$$

Hoặc có thể là:

$$\text{Intersection (h(I), h(M)} = \sum_{j=1}^{K} \left| h(I) - h(M) \right| \quad (2.5)$$

2.2.3 Khoảng cách toàn phương

$$\text{Intersection (h(I), h(M)} = \sum_{j=1}^{k} \sum_{j=1}^{K} [h(i) - h(j) a_{ij} [h(i) - h(j)] \quad (2.6)$$

2.2.4 Khoảng cách Chi Squared distance

Khoảng cách Chi Squared là số liệu khoảng cách giữa các bin được sử dụng để so sánh các biểu đồ. Biểu thức tính toán của khoảng cách Chi Squared giữa hai biểu đồ được tính như sau:

$$\chi^2(a,b) = \frac{1}{2} \sum_1 \frac{(a_1 - b_1)^2}{(a_1 + b_1)^2} \quad (2.7)$$

2.2.5 Khoảng cách Hausdorff

Khoảng cách Hausdorff là một trong những phương pháp đối sánh hình dạng dựa trên tương quan cổ điển. Khoảng cách Hausdorff thường được sử dụng để xác định vị trí trong một ảnh và đo độ tương tự trong hình dạng. Với hai hình dạng được biểu diễn bằng tập những điểm $A=\{a_1, a_2, ...a_p\}$ và $B=\{b_1, b_2, ...b_q\}$ thì khoảng cách giữa A và B được biểu diễn:

H(a, B) = max(h(A, B), h(B, A)}

Trong đó:

$$h(A,B) = \max_{a \in b} \min_{b \in B} \|a - b\| \quad (2.8)$$

Tuy nhiên độ đo khoảng cách này nhạy cảm với nhiễu và ngoại lai. Điểm đơn ở trong *A*, các điểm trong *B* sẽ tạo ra *h(A,B)* rất lớn.Do đó, khoảng cách Hausdorff đã được cải tiến bởi Rucklidge:

$$h^f(A,B) = f^{th}_{a\in A} \min_{b\in B} \|a-b\| \tag{2.9}$$

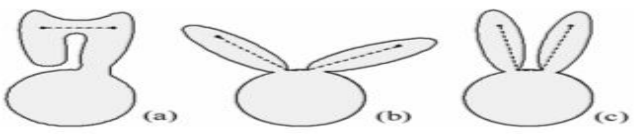
$f^{th}_{x\in X} g(x)$ là giá trị vi phân thứ f^{th} của g(x) trên tập x với một vài giá trị của f là 0 và 1.Ví dụ, giá trị vi phân thứ nhất chính là lớn nhất và giá trị vi phân 1/2 là trung bình. Trong thực tế f thường đặt là 1/2. Ưu điểm của đối sánh hình dạng sử dụng khoảng cách Hausdroff chính là hình dạng có thể được đối sánh cục bộ. Tuy nhiên khoảng cách này là không bất biến với các phép tịnh tiến, phép co dãn và phép quay.

2.2.6 Độ đo khoảng cách trong

2.2.6.1 Giới thiệu

Cấu trúc thành phần đóng vai trò quan trọng trong việc phân loại những hình dạng phức tạp. Tuy nhiên, việc thu lại được những cấu trúc thành phần chưa bao giờ là một công việc đơn giản, nhất là khi xét đến cấu trúc hình dạng có khớp nối. Những kiểu hình dạng này là sự biến đổi phi tuyến giữa các hình dạng, hơn nữa, một vài hình dạng có thể có cấu trúc nhập nhằng. Để giải quyết cho những vấn đề này, một kĩ thuật biểu diễn hình dạng được gọi là khoảng cách trong đã được đề xuất.

Khoảng cách trong được định nghĩa là khoảng cách ngắn nhất của đường dẫn bên trong đường biên hình dạng nhằm xây dựng sự nhận diện hình dạng ảnh. Có thể dễ dàng thấy được, khoảng cách trong không nhạy cảm với các hình dạng khớp nối. Ví dụ trong hình 2-1.



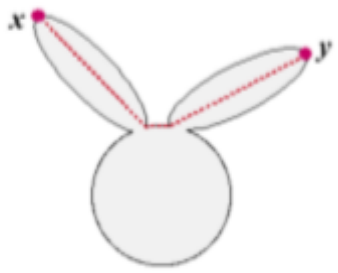
Hình 2-1: Ví dụ khoảng cách trong

Ta có thể thấy, mặc dù trong hình (a) và hình (c) đều có sự phân bố không gian tương tự nhau, nhưng chúng lại hoàn toàn khác nhau về cấu trúc thành phần của chúng. Mặt khác, hình (c) và hình (b) lại xuất hiện từ cùng một loại hình dạng chỉ khác nhau ở các khớp nối. Khoảng cách trong giữa hai điểm được đánh dấu trong hình (a) và hình (b) là hoàn toàn khác nhau trong khi, phần lớn sự giống nhau lại nằm ở hình (b) và hình (c). Bằng trực giác,ví dụ này cho ta thấy rằng, khoảng cách trong là không nhạy cảm đối với cấu trúc khớp nối, và nhạy cảm đối với cấu trúc thành phần, khoảng cách trong đáng để hướng tới cho việc đối sánh các hình dạng phức tạp. Trong khi đó khoảng cách Euclidean không có những thuộc tính đó đối với ví dụ trên. Bằng chứng cho việc này chính là khoảng cách trong được định nghĩa như là độ dài của những đoạn nét đứt giữa các điểm được đánh dấu, còn khoảng cách Euclidean thì không xem xét đến có những đoạn nét đứt chồng chéo lên nhau.

Việc sử dụng khoảng cách trong như là một giải pháp để thay thế cho những độ đo tương tự khác nhằm xây dựng một mô tả hình dạng mới mà có khả năng bất biến (không nhạy cảm) đối với hình dạng có cấu trúc khớp nối.

2.2.6.2 Khoảng cách trong

Trước tiên, cho hình *O* là một tập đóng và có kết nối của R^2, hai điểm *x* và *y* thuộc *O*, khoảng cách trong giữa *x* và *y* được ký hiệu là: *d(x, y; O)* và được định nghĩa là độ dài của đường dẫn ngắn nhất kết nối hai điểm x và y ở trong hình *O*. Ví dụ hình 2-2.



Hình 2-2: Ví dụ về khoảng cách trong của x và y trong hình O

Vấn đề đặt ra:

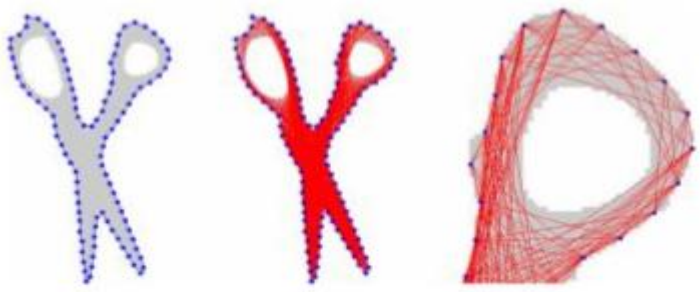
- Trong một vài trường hợp hiếm gặp, có thể tồn tại nhiều đường dẫn ngắn nhất giữa các điểm cho trước, khi đó, ta tùy ý chọn một đường dẫn ngắn nhất trong số đó.
- Chúng ta đã quen với việc định nghĩa Shape bởi những đường biên của chúng, do đó, chỉ những điểm biên được sử dụng như là những điểm đánh dấu. Hơn nữa hình dạng được xấp xỉ bởi một hình đa giác, đa giác này được hình thành nên bởi những điểm được đánh dấu của chúng.

Cách đơn giản nhất để tính toán khoảng cách trong là sử dụng thuật toán tìm đường dẫn ngắn nhất, thuật toán này được chia làm hai bước:

- Bước một: Xây dựng một đồ thị với các điểm mẫu. Trước tiên, mỗi điểm mẫu được coi như là một nút ở trong đồ thị, sau đó đối với mỗi cặp điểm mẫu p_1 và p_2, nếu đoạn nối liền p_1 và p_2 nằm hoàn toàn trong đối tượng thì một cạnh giữa p_1 và p_2 được thêm vào đồ thị cùng với trọng số của nó là khoảng cách Euclidean $||p_1 – p_2||$. Ví dụ: hình 2-3

Một vài chú ý được đề cập tới đó là:

- Thứ nhất: các điểm biên láng giềng thì luôn luôn liên thông với nhau.
- Thứ hai: Khoảng cách trong không sử dụng những điểm mẫu của đường biên lỗ hổng.



Hình 2-3: Quá trình biểu diễn khoảng cách trong của đối tượng

- Bước thứ hai: Áp dụng thuật toán tìm đường đi ngắn nhất cho đồ thị. Nhiều thuật toán đã được áp dụng, trong đó có thuật toán FloydWarshall's có độ phức tạp là $O(n^3)$ với n là số điểm lấy mẫu. Thuật toán khoảng cách trong đã được tác giả chỉ ra có độ phức tạp thuật toán là $O(n^3)$. Trước tiên, mất một khoảng thời gian O(n) để kiểm tra xem đoạn nối giữa hai điểm nằm trong hình dạng. Tiếp theo, việc xây dựng đồ thị thì có độ phức tạp là $O(n^3)$. Khi đồ thị đã được tính toán xong, thuật toán dùng để tính toán tất cả các cặp có đường dẫn ngắn nhất có độ phức tạp thuật toán là $O(n^3)$. Do vậy, độ phức tạp tính toán toàn bộ là $O(n^3)$.

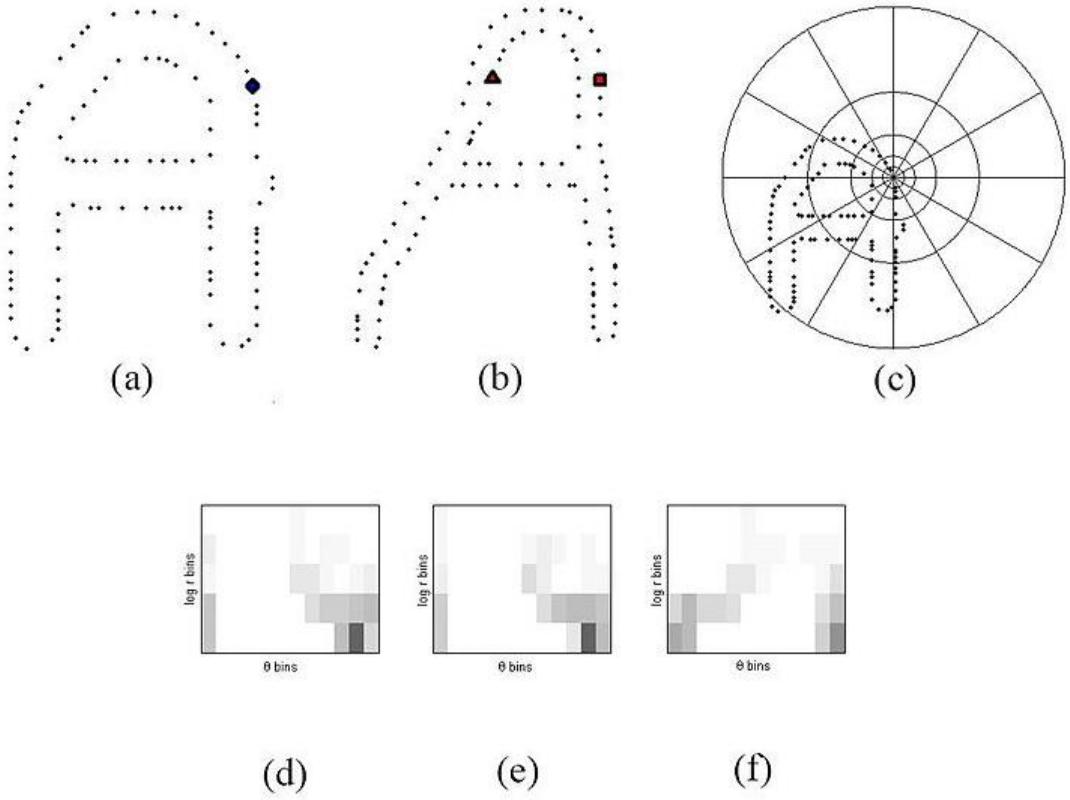
2.3 Mô tả ảnh sử dụng ngữ cảnh hình dạng (Shape context)

Ngữ cảnh hình dạng là một bộ mô tả tính năng được sử dụng trong nhận dạng đối tượng. Ngữ cảnh hình dạng được giới thiệu bởi Belongie. Nó mô tả mối quan hệ phân bố không gian của các điểm đại diện được đánh dấu xung quanh những điểm đặc trưng: với n điểm mẫu x_1, x_2, …, x_n trên một hình dạng. Ngữ cảnh hình dạng tại điểm x_i được định nghĩa như là biểu đồ h_i của mối quan hệ tọa độ của n - 1 điểm còn lại.

Ta có công thức:

$$h_i(k) = \#\{x_j : j \neq i, x_j - x_i \in bin(k)\} \tag{2.10}$$

Trong đó: Với các bin được chia đều trong không gian log-polar. Hình 2-4 là ví dụ tính toán biểu đồ (a) và (b) là các điểm được lấy mẫu trên hai hình dạng. (c) là một biểu đồ log-polar được sử dụng để tính toán với 5 bin bán kính và 12 bin góc. (d) là ngữ cảnh hình dạng với các điểm được đánh dấu với hình tròn ở trong (a). (e) là các điểm được đánh dấu ở trong (b) và (f) là cho hình tam giác. Trên hình vẽ ta thấy (d) và e là ngữ cảnh hình dạng cho hai điểm liên quan gần giống nhau nên chúng tương tự nhau. Trong khi đó ngữ cảnh hình dạng trong (f) là rất khác.



Hình 2-4: Tính toán ngữ cảnh hình dạng

2.4 Đối sánh hình dạng ngữ cảnh

2.4.1 Đối sánh shape sử dụng quy hoạch động

Bài toán đối sánh đường bao được phát biểu như sau: cho hai hình *A* và hình *B*, ta mô tả chúng bằng các dãy điểm trên đường bao của chúng. Ta có: p_1, p_2, ..., p_n là *n* điểm thuộc hình *A* và m điểm q_1, q_2, ..., q_m thuộc hình *B*. Với giả sử n >= m, một đối sánh π từ *A* đến *B* là một ánh xạ từ 1, 2, …, n đến 0, 1, 2, …, m trong đó p_i được đối sánh với q_π (i) nếu p(i) khác 0 và ngược lại thì không đối sánh. Hàm chi phí đối sánh $C(\pi)$ được định nghĩa như sau:

$$C(\pi) = \sum_{1<=i<=n} c(i, \pi()) \tag{2.12}$$

Trong đó c(i, 0) = τ là giá trị phạt cho việc điểm p_i không được đối sánh, và với 1<= j <=m, c(i, j) là chi phí đối sánh giữa p_i với q_j. Độ đo này được đo bằng cách sử dụng hàm thống kê χ_2 như công thức sau:

$$c(i,j) = \frac{1}{2}\sum_{1 \subset k \subset K} \frac{[hA,i(k) - hB,j(k)]^2}{hA,i(k) + hB,j(k)} \quad (2.13)$$

Ở đây, $h_{A,i}$ và $h_{B,j}$ là những biểu đồ ngữ cảnh hình dạng của p_i và q_j, và *K* là số bin của biểu đồ.

Do các đường bao được cung cấp các thứ tự cho các chuỗi điểm $p_1,p_2,...p_n$ và $q_1,q_2,...q_m$, là tự nhiên nên nó vô tình làm hạn chế sự đối sánh. Để giải quyết vấn đề này kỹ thuật Quy hoạch động (DP) đã được sử dụng để giải quyết cho vấn đề đối sánh này. DP được sử dụng rộng rãi cho vấn đề đối sánh đường bao.

2.4.2 Đối sánh hình dạng dựa trên đồ thị

Bài toán đối sánh đường bao được phát biểu như sau: cho hai hình *A* và hình *B*, ta mô tả chúng bằng các dãy điểm trên đường bao của chúng. Ta có: $p_1, p_2, ..., p_n$ là *n* điểm thuộc hình *A* và m điểm $q_1,q_2, ...,q_m$ thuộc hình B. Giả sử n >= m, sự đối sánh π từ A đến B là một ánh xạ từ 1, 2, …, n đến 0, 1, 2, …, m trong đó p_i được đối sánh với $q_{\pi(i)}$ nếu $\pi(i)$ khác 0 và ngược lại thì không đối sánh nên được cực tiểu hóa chi phí đối sánh và được định nghĩa là $H(\pi)$:

$$C(\pi) = \sum_{1 \subset i \subset n} c(i, \pi()) \quad (2.14)$$

Trong đó c(i, 0) = τ là hình phạt cho việc bỏ qua p_i không đối sánh, và cho 1<= j <=m, c(i, j) là chi phí của đối sánh p_i với q_j. Chúng ta có thể áp dụng thuật toán ghép cặp trên đồ thị với trọng số nhỏ nhất để giải quyết bài toán này.

2.4.2.1 Bài toán ghép cạnh với trọng số nhỏ nhất

Input: Đồ thị hai phần đầy đủ G = (X∪Y, E), X = {X_1, X_2, ..., X_n}, Y = {Y_1, Y_2, ..., Y_n} được cho bởi ma trận vuông C cỡ n×n, c[i, j] là trọng số cạnh nối đỉnh X_i với Y_j. Giả thiết c[i, j] ≥ 0 với mọi i, j.

Output: Ghép cặp hoàn hảo với trọng số nhỏ nhất.

2.4.2.2 Thuật toán

Để cho gọn, ta gọi những cạnh trọng số 0 của G là những 0_cạnh. Xét một ghép cặp M chỉ gồm những 0_cạnh.

- Những 0_cạnh thuộc M gọi là những 0_cạnh đã ghép, những 0_cạnh còn lại là những 0_cạnh chưa ghép. Nếu ta định hướng lại các 0_cạnh như sau: Những 0_cạnh chưa ghép cho hướng từ tập X sang tập Y, những 0_cạnh đã ghép cho hướng từ tập Y về tập X. Khi đó:
- Đường xen kẽ là một đường đi đơn xuất phát từ một X_đỉnh chưa ghép đi theo các 0_cạnh đã định hướng ở trên. Như vậy dọc trên đường xen kẽ, các 0_cạnh chưa ghép và những 0_cạnh đã ghép xen kẽ nhau. Vì đường xen kẽ chỉ là đường đi đơn trên đồ thị định hướng nên việc xác định những đỉnh nào có thể đến được từ $x \in X$ bằng một đường xen kẽ có thể sử dụng các thuật toán tìm kiếm trên đồ thị (BFS hoặc DFS). Những đỉnh và những cạnh được duyệt qua tạo thành một cây pha gốc x.

  Một đường mở là một đường xen kẽ đi từ một X_đỉnh chưa ghép tới một Y−đỉnh chưa ghép. Như vậy:

- Đường đi trực tiếp từ một X_ đỉnh chưa ghép tới một Y_đỉnh chưa ghép qua một 0_cạnh chưa ghép cũng là một đường mở.
- Dọc trên đường mở, số 0_cạnh chưa ghép nhiều hơn số 0_cạnh đã ghép đúng 1 cạnh.

2.4.2.2.1 Thuật toán Hungari

Bước 1:

- Khởi tạo một ghép cặp $M = \emptyset$.

Bước 2:

Với mọi đỉnh $x* \in X$, ta tìm cách ghép $x*$ như sau: Bắt đầu từ đỉnh $x*$ chưa ghép, thử tìm đường mở bắt đầu ở $x*$ bằng thuật toán tìm kiếm trên đồ

thị (BFS hoặc DFS thông thường nên dùng BFS để tìm đường qua ít cạnh nhất) có hai khả năng xảy ra:

- Hoặc tìm được đường mở, dọc theo đường mở, ta loại bỏ những cạnh đã ghép khỏi M và thêm vào M những cạnhchưa ghép, ta được một ghép cặp mới nhiều hơn ghép cặp cũ 1 cạnh và đỉnh x∗ trở thành đã ghép.
- Hoặc không tìm được đường mở, do ta dùng thuật toán tìm kiếm trên đồ thị nên có thể xác định được hai tập:
- VisitedX = Tập những X_ đỉnh có thể đến được từ x∗ bằng một đường xen kẽ;
- VisitedY = Tập những Y_đỉnh có thể đến được từ x∗ bằng một đường xen kẽ;

Gọi Δ là trọng số nhỏ nhất của các cạnh nối giữa một đỉnh thuộc VisitedX với một đỉnh không thuộc VisitedY. Dễ thấy Δ > 0 bởi nếu Δ = 0 thì tồn tại một 0_cạnh (x, y) với x thuộc VisitedX và y không thuộc VisitedY. Vì x∗ đến được x bằng một đường xen kẽ và (x, y) là một 0_cạnh nên x∗ cũng đến được y bằng một đường xen kẽ, dẫn tới y ∈ VisitedY, điều này vô lý.

Biến đổi đồ thị G như sau: Với mọi x ∈VisitedX, trừ Δ vào trọng số những cạnh liên thuộc với x, với mọi y ∈ VisitedY, cộng Δ vào trọng số những cạnh liên thuộc với y.

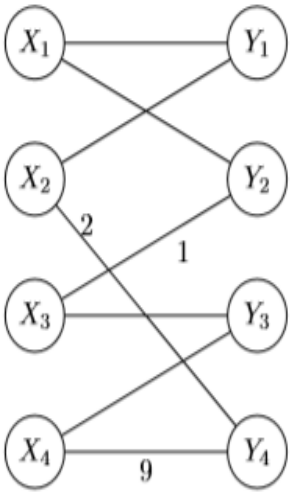
Lặp lại thủ tục tìm kiếm trên đồ thị thử tìm đường mở xuất phát ở x∗ cho tới khi tìm ra đường mở.

Bước 3:

- Sau bước 2 thì mọi X_đỉnh đều được ghép, in kết quả về ghép cặp tìm được.

Ví dụ:

Tìm ghép cặp lớn nhất có trọng số nhỏ nhất trong đồ thị sau. Để không bị rối hình, ta hiểu những cạnh không ghi trọng số là những 0_cạnh, những cạnh không vẽ mang trọng số rất lớn trong trường hợp này không cần thiết phải tính đến. Những cạnh nét đậm là những cạnh đã ghép, những cạnh nét thanh là những cạnh chưa ghép.

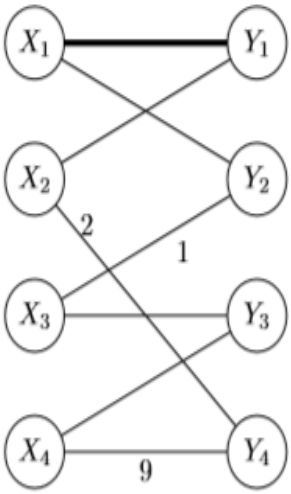


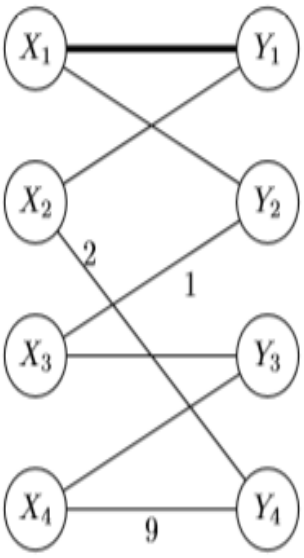
$x^* = X_1$

tìm được đường mở

$X_1 \to Y_1$

tăng cặp



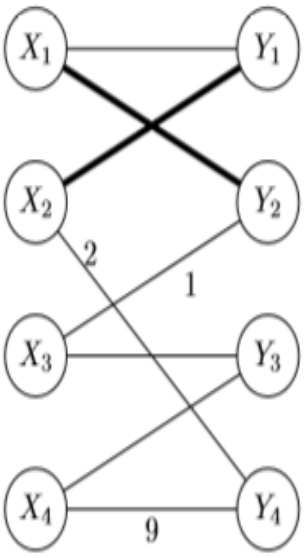


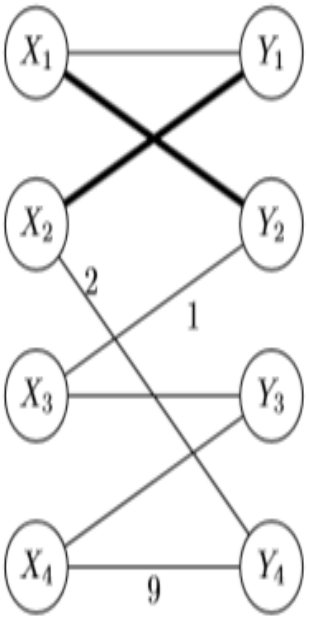
$x^* = X_2$

tìm được đường mở

$X_2 \to Y_1 \to X_1 \to Y_2$

tăng cặp



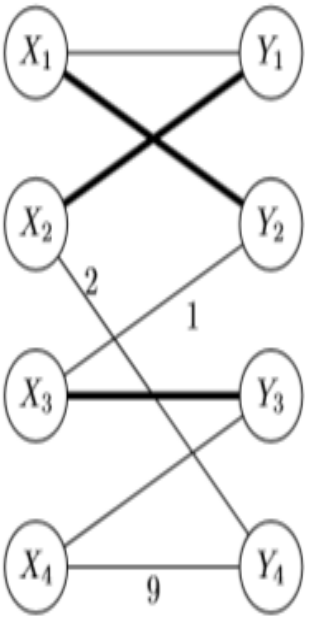


$x^* = X_3$

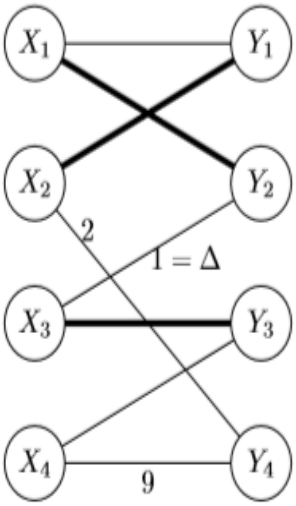
tìm được đường mở

$X_3 \to Y_3$

tăng cặp



$x^* = X_4$



không tìm được đường mở

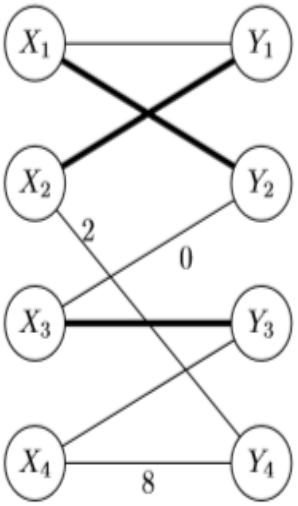
Tập X_đỉnh đến được từ X_4 bằng một đường xen kẽ $\{X_3, X_4\}$

Tập Y_đỉnh đến được từ X_4 bằng một đường xen kẽ $\{Y_3\}$

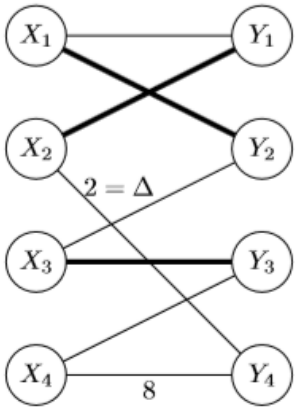
Giá trị xoay $\Delta = 1$ (cạnh $X_3 - Y_2$)

Trừ tất cả trọng số những cạnh liên thuộc $\{X_3, X_4\}$ đi 1.

Cộng tất cả trọng số những cạnh liên thuộc Y_3 lên 1.



$x^* = X_4$



không tìm được đường mở

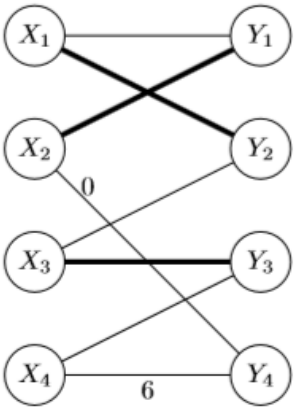
Tập X_đỉnh đến được từ X_4 bằng một đường xen kẽ $\{X_1, X_2, X_3, X_4\}$

Tập Y_đỉnh đến được từ X_4 bằng một đường xen kẽ $\{Y_1, Y_2, Y_3\}$

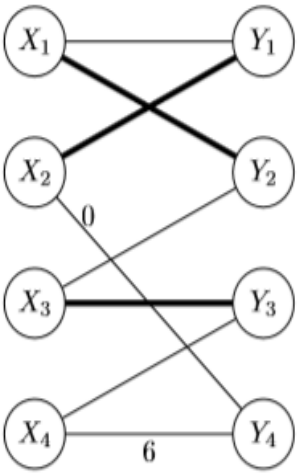
Giá trị xoay $\Delta = 2$ (cạnh $X_2 - Y_4$)

Trừ tất cả trọng số những cạnh liên thuộc với $\{X_1, X_2, X_3, X_4\}$ đi 2.

Cộng tất cả trọng số những cạnh liên thuộc $\{Y_1, Y_2, Y_3\}$ lên 2.



$x^* = X_4$



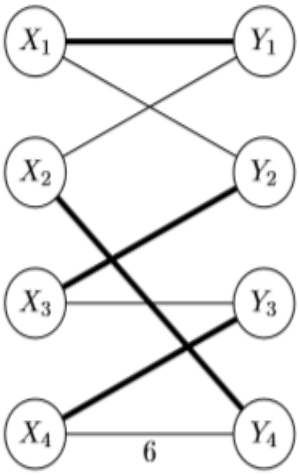
tìm được đường mở:

$X_4 \to Y_3 \to X_3 \to Y_2 \to$

$\to X_1 \to Y_1 \to X_2 \to Y4$

Tăng cặp

Kết thúc



2.4.2.3 Bài toán ghép cặp với trọng số lớn nhất

Input: Đồ thị hai phần đầy đủ G = (X∪Y, E), X = {X_1, X_2, ..., X_n}, Y = {Y_1, Y_2, ..., Y_n} được cho bởi ma trận vuông C cỡ n×n, c[i, j] là trọng số cạnh nối đỉnh X_i với Y_j. Giả thiết c[i, j] ≥ 0 với mọi i, j.

Output: Ghép cặp hoàn hảo với trọng số lớn nhất.

2.4.2.3.1 Thuật toán

Bước 1: Khởi tạo:

- Một ghép cặp M = ∅
- Khởi tạo hai dãy Fx và Fy thỏa mãn Fx[i]+Fy[j] ≥ c[i, j] với mọi i, j; chẳng hạn có thể gán Fx[i] với giá trị lớn nhất trên dòng i của ma trận C và các Fy[j] = 0.

Bước 2: Với mọi đỉnh x∗ ∈ X, ta tìm cách ghép x∗ như sau: Với cách hiểu 0−cạnh là cạnh thỏa mãn c[i, j] = Fx[i]+Fy[j], bắt đầu từ đỉnh X∗, thử tìm đưởng mở bắt đầu từ x∗.Có hai khả năng xảy ra:

- Hoặc tìm được đường mở thì dọc theo đường mở, ta loại bỏ những cạnh đã ghép khỏi M và thêm vào M những cạnh chưa ghép, ta được một ghép cặp mới nhiều hơn bộ ghép cũ 1 cạnh và đỉnh x∗ trở thành đã ghép.
- Hoặc không tìm được đường mở thì ta xác định được hai tập:
- VisitedX = Tập những X_đỉnh có thể đến được từ x∗ bằng một đường xen kẽ;
- VisitedY = Tập những Y_đỉnh có thể đến được từ x∗ bằng một đường xen kẽ;

Đặt Δ = min x∈[a;b]{Fx[i]+Fy[j]−c[i, j], X[i] thuộc V isitedX, Y [j] không thuộc V isitedY}

Với mọi X[i] thuộc tập VisitedX, gán Fx[i]:= Fx[i]−Δ;

Với mọi Y [j] thuộc tập VisitedY, gán Fx[j]:= Fx[j] + Δ;

Lặp lại thủ tục tìm đường mở xuất phát từ x∗ cho tới khi tìm ra đường mở.

Bước 3: Sau bước 2 thì mọi X_đỉnh đều được ghép, ta được một ghép cặp hoàn hảo với trọng số lớn nhất.

CHƯƠNG 3: THỰC NGHIỆM

3.1 Môi trường thực nghiệm

3.1.1 Phần cứng

Thành phần	Thông số
CPU	Core i5-2540M 2.60GHz
RAM	4GB
Hệ điều hành	Windows 10 Pro 64-bit
Bộ nhớ	250GB

3.1.2 Phần mềm

Chương trình cài đặt trên môi trường visual studio 2017 với thư viện mã nguồn mở OpenCV (Open Source Computer Vision Library).

OpenCV là một thư viện mã nguồn mở về thị giác máy tính và học máy. Thư viện được xây dựng để cung cấp nền tảng cho các ứng dụng thị giác máy tính nhằm đẩy mạnh sự phát triển về hàm lượng tri thức máy tính trong các sản phẩm thương mại. Nhờ giấy phép bản quyền BSD và được nhiều công ty lớn hàng đầu thế giới như Google, Yahoo, Microsoft, Intel, IBM, … cùng đóng góp xây dựng thư viện, OpenCV là một trong những công cụ mạnh và được sử dụng rộng rãi trong trường học cũng như các công ty khởi nghiệp.

Thư viện OpenCV bao gồm nhiều giao diện dành cho C++, C, Python, Java, MATLAB và hỗ trợ cho các hệ điều hành khác nhau như Windows, Linux, Android, MacOS. Trong phiên bản OpenCV 3.1, giao diện sử dụng cho CUDA và OpenCL cũng đã được phát triển hoàn thiện. OpenCV được viết nguyên bản bằng ngôn ngữ C++. OpenCV có rất nhiều chức năng. Sau đây là những tóm tắt cơ bản về hệ thống các nhóm hàm trong OpenCV.

- Image and Video I/O. Là các nhóm hàm cho phép đọc dữ liệu ảnh từ file hoặc trực tiếp từ video.

- Các thuật toán xử lý ảnh và thị giác máy (General computer-vision and image-processing algorithms (mid – and low level APIs)): Bao gồm các hàm thực hiện các thuật toán xử lý ảnh từ video.
- Graphic: Bao gồm các hàm cho phép viết chữ và vẽ trên hình ảnh. Thêm vào đó những hàm này còn được sử dụng để ghi nhãn và đánh dấu. Ví dụ nếu ta viết một chương trình cần nhận dạng nhiều đối tượng thì nó sẽ rất có ích cho tạo nhãn ảnh (label image) với kích thước và vị trí.

OpenCV cho Linux và MacOSX được đóng gói như một gói mã nguồn lưu trữ. Ta sẽ phải xây dựng cả thư viện tĩnh và cả những đối tượng chia sẻ (shared-object). Ta cần xây dựng RPM trước tiên và cài đặt từ nó, hoặc biên dịch và cài đặt nó ngay. Cấu trúc cho cả hai được nằm trong INSTALL. Với Window ta khi cài đặt OpenCV, nó sẽ copy file OpenCV vào thư mục mà ta chọn. Mặc định nó cài đặt đến C:/Program Files/OpenCV/<version>.

Trong thư mục OpenCV có chứ một vài thư mục khác. Thư mục docs chứa file văn bản html cho toàn bộ các hàm OpenCV và kiểu dữ liệu. Từ file văn bản này ta có thể làm các ví dụ, ta cũng có thể muốn xem thư mục "samples". Những file header sẽ cần thiết khi ta dịch chương trình sử dụng OpenCV.

3.2 Đối sánh ảnh dựa trên ngữ cảnh hình dạng sử dụng opencv

3.2.1 Tìm đường bao và lấy mẫu các điểm trên đường bao

Bước 1: Đọc và hiển thị ảnh

Hàm đọc ảnh trong opencv: imread

Mat imread(const string&filename, int flags)

Filename: Tên ảnh đầu vào

Flags:

- CV_LOAD_IMAGE_ANYDEPTH: Nếu ảnh đầu vào có chiều sâu tương ứng thì ảnh trả về 16-bit/32-bit, còn ngược lại thì nó trả về ảnh 8-bit.
- CV_LOAD_IMAGE_COLOR: Nếu dùng hàm này, thì ảnh chuyển đổi thành ảnh một màu.
- CV_LOAD_IMAGE_GRAYSCALE: Nếu dùng hàm này, sẽ trả về hình ảnh đa mức xám.

Hàm hiển thị ảnh trong opencv: imshow

Void imshow(const string & winname, InputArray mat)

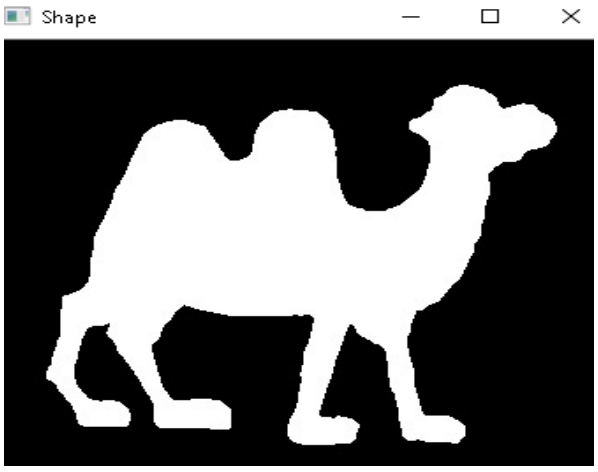
Winname: Tên cửa sổ hiển thị ảnh

Image: Hình ảnh được hiển thị

Hàm imshow hiển thị một hình ảnh trong cửa sổ được chỉ định. Nếu cửa sổ được tạo bằng CV_WINDOW_AUTOSIZE, hình ảnh được hiển thị với kích thước ban đầu của nó, tuy nhiên nó vẫn bị giới hạn bởi độ phân giải màn hình. Nếu không, hình ảnh được thu nhỏ để vừa với cửa sổ. Hàm có thể chia tỷ lệ hình ảnh, tùy thuộc vào độ sâu của nó.

Ví dụ đọc và hiển thị ảnh:

```
Mat img = imread("D:/Anh/Shape1.jpg");
imshow("Shape", img);
```



Hình 3-1: Hình được hiển thị Shape

Bước 2: Tìm đường bao

- Sử dụng hàm findContours(currentQuery, _contoursQuery, RETR_LIST, CHAIN_APPROX_NONE) để tìm đường bao.

```
vector<vector<Point> > _contoursQuery;
vector <Point> contoursQuery;
findContours(currentQuery, _contoursQuery, RETR_LIST, CHAIN_APPROX_NONE);
```

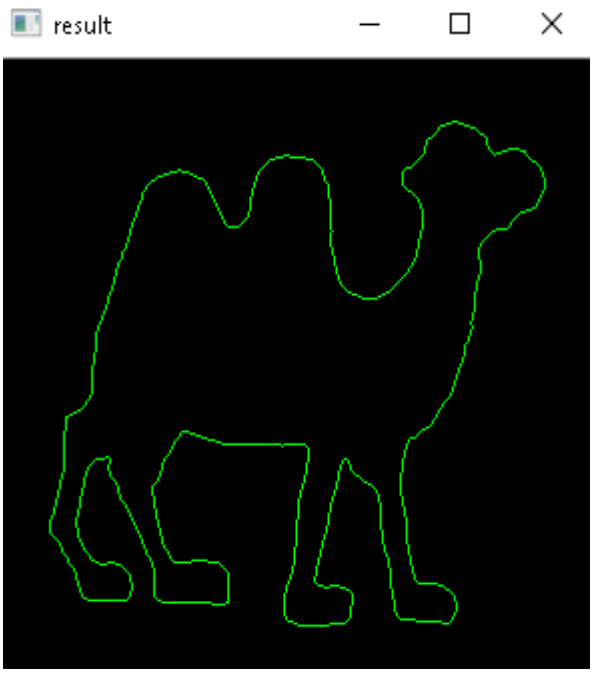
currentQuery: Ảnh đầu vào.

contoursQuery: Các đường bao được tìm.Mỗi đường bao được lưu trữ trong một vector của các điểm.

RETR_LIST: Trả về tất cả các đường bao mà không thiết lập bất kỳ mối quan hệ thứ bậc nào.

CHAIN_APPROX_NONE: Lưu trữ tất cả các điểm đường bao.Nghĩa là 2 điểm bất kỳ liền nhau x_1, y_1 và x_2, y_2 của đường bao sẽ thuộc các láng giềng ngang, dọc và chéo, nghĩa là $\max(abs(x_1-x_2), abs(y_2-y_1))=1$.

- Kết quả tìm đường bao theo thủ tục trên được minh họa như hình 3-2:



Hình 3-2: Kết quả tìm biên bằng phương pháp Canny từ ảnh đầu vào

Bước 3: Lấy mẫu *n* điểm trên đường bao

- Trong trường hợp số điểm thực tế trên đường bao nhỏ hơn *n* thì thêm các điểm vào cho đủ:

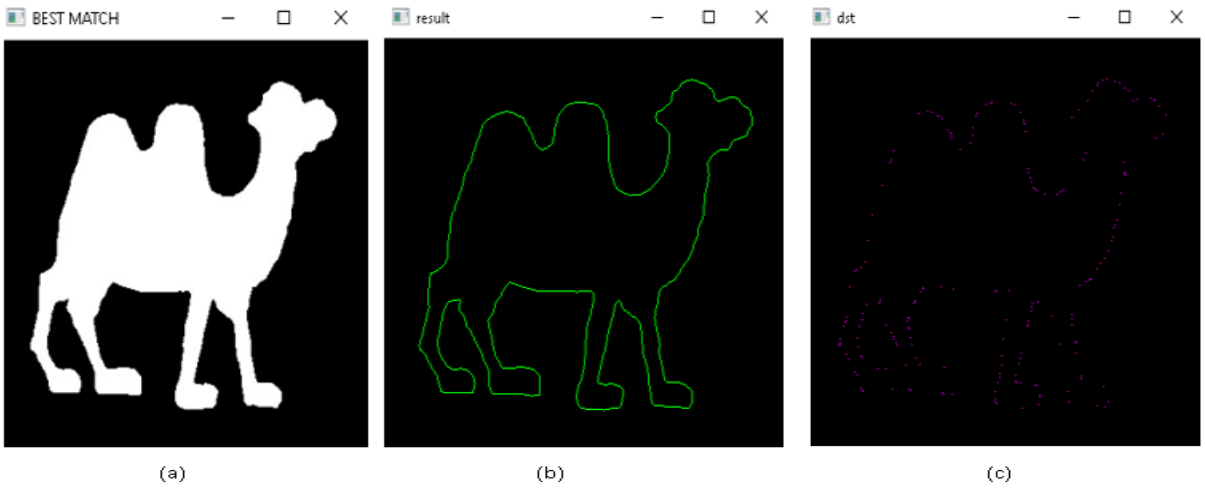
```
int dummy = 0;
for (int add = (int)_contoursQuery[largest_contour_index].size() - 1; add < n; add++)
{
    contoursQuery.push_back(contoursQuery[dummy++]);
}
```

- Lấy mẫu *n* điểm ngẫu nhiên trên đường bao sử dụng phương thức:

```
random_shuffle(contoursQuery.begin(), contoursQuery.end());
vector<Point> contQuery;
for (int i = 0; i < n; i++)
{
      contQuery.push_back(contoursQuery[i]);
}
```

- Việc tìm đường bao và lấy mẫu được thực hiện theo thủ tục sau:

```
int k = _contoursQuery.size();
std::vector<cv::Vec3b> colors;
for (int i = 0; i <= 1; i++)
{
      int b = 255*i;
      int g = 255*i;
      int r = 255*i;
     colors.push_back(cv::Vec3b((uchar)b,(uchar)g,(uchar)r));
}
cv::Mat dst = cv::Mat::zeros(currentQuery.size(),  CV_8UC3);
for (int i = 0; i < n; i++)
{
      for (int j = 0; j < n; j++)
      {
dst.at<cv::Vec3b>(contQuery[i].y,contQuery[i].x)=colors[1];
      }
}
cv::imshow("dst",  dst);
Mat src1 = Mat::zeros(currentQuery.size(), CV_8UC3);
imshow("result", src);
return contQuery;
```



Hình 3-3: Kết quả tìm đường bao và lấy mẫu.

Trong đó:

- (a): Ảnh đầu vào
- (b): Ảnh biên được phát hiện theo phương pháp Canny
- (c): Ảnh biên được lấy mẫu 200 điểm ngẫu nhiên

3.2.2 Tìm khoảng cách và đối sánh giữa hai đường bao đã được lấy mẫu

Trong thủ tục này, các điểm biên của hai ảnh đã được lấy mẫu conQuery1, conQuery2 sẽ được sử dụng để tính toán ngữ cảnh hình dạng, sau đó khoảng cách giữa chúng được tính toán.

Thủ tục tính toán khoảng cách và đối sánh giữa hai hình dạng được thực hiện như sau:

```
static float getShapeContextDistance(string fileName1, vector<Point> contQuery1,  string fileName2,  vector<Point> contQuery2)
{
	cv::Ptr <cv::ShapeContextDistanceExtractor> mysc;
	mysc = cv::createShapeContextDistanceExtractor();
	float dis = mysc->computeDistance(contQuery1,contQuery2);
	int bestMatch = 0;
	float bestDis = FLT_MAX;
	moveWindow("TEST",  0,  0);
	std::cout << " distance between " << fileName1 << " and " << fileName2 << " is: " << dis << std::endl;
	return dis;
}
```

Bảng sau chỉ ra kết quả tính toán khoảng cách giữa hai hình ảnh với 200 số điểm lấy mẫu:

Bảng 3.1: Ước lượng khoảng cách giữa các cặp ảnh với 200 điểm mẫu

<table>
<tr><td colspan="2">Cặp ảnh đối sánh</td><td>Khoảng cách</td></tr>
<tr><td>

Ảnh đầu vào Shape 1</td><td>

Ảnh đã được lấy mẫu Shape 1</td><td rowspan="2">0</td></tr>
<tr><td>

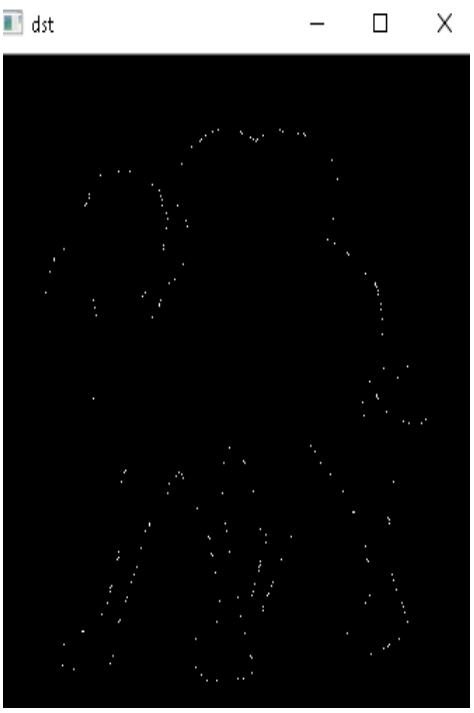
Ảnh đầu vào Shape 1</td><td>

Ảnh đã được lấy mẫu Shape 1</td></tr>
<tr><td colspan="3">Khoảng cách giữa ảnh được lấy mẫu Shape1 và Shape1 là: 0</td></tr>
</table>

<table>
<tr>
<td>

Ảnh đầu vào Shape 1</td>
<td>

Ảnh đã được lấy mẫu Shape 1</td>
<td rowspan="2">2.39</td>
</tr>
<tr>
<td>

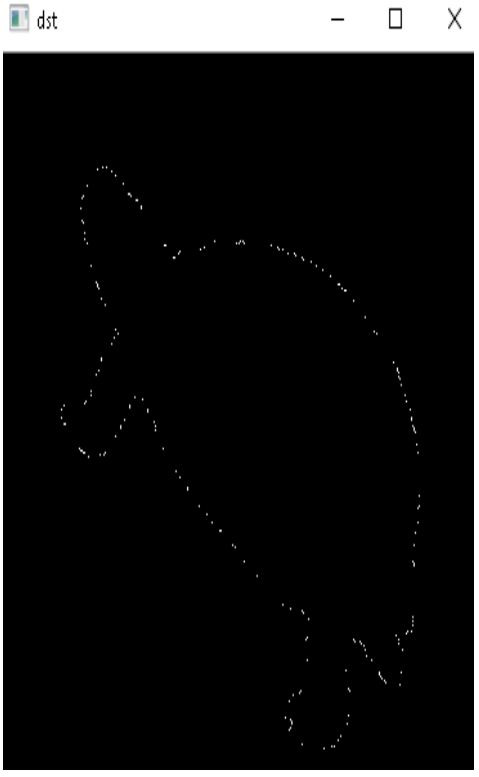
Ảnh đầu vào Shape2</td>
<td> 

Ảnh đã được lấy mẫu Shape2</td>
</tr>
<tr>
<td colspan="3">Khoảng cách giữa ảnh được lấy mẫu Shape1 và Shape2 là: 2.39</td>
</tr>
</table>

<table>
<tr><td>

Ảnh đầu vào Shape 1</td><td>

Ảnh đã được lấy mẫu Shape1</td><td rowspan="2">4.53</td></tr>
<tr><td>

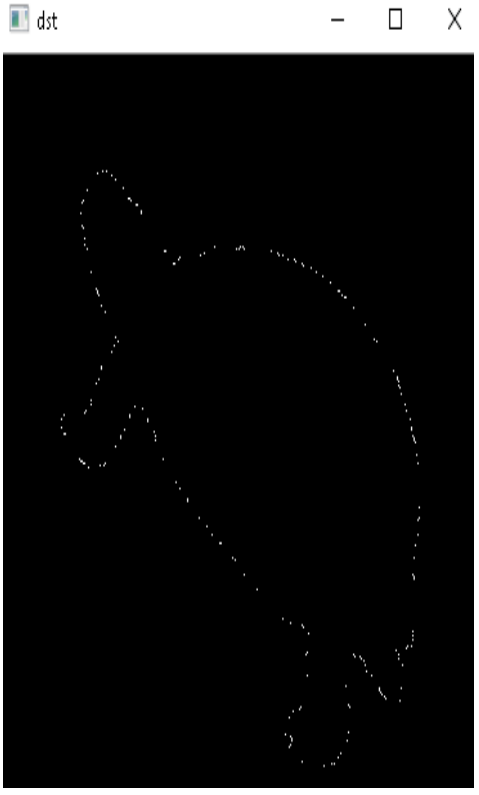
Ảnh đầu vào Shape 3</td><td> 

Ảnh đã được lấy mẫu Shape3</td></tr>
<tr><td colspan="3">Khoảng cách giữa ảnhđược lấy mẫu Shape1 và Shape3 là: 4.53</td></tr>
</table>

<table>
<tr><td>

Ảnh đầu vào Shape 2</td><td>

Ảnh đã được lấy mẫu Shape2</td><td rowspan="2">4.22</td></tr>
<tr><td>

Ảnh đầu vào Shape 3</td><td> 

Ảnh đã được lấy mẫu Shape3</td></tr>
<tr><td colspan="3">Khoảng cách giữa ảnhđược lấy mẫu Shape2 và Shape3 là: 4.22</td></tr>
</table>

KẾT LUẬN

Sau một thời gian thực hiện khóa luận này, đồ án đã đạt được kết quả như sau:

- Nắm được một số phương pháp phát hiện biên.
- Mô tả và giải thích các vấn đề cơ bản trong việc tìm đường bao đối tượng.
- Hiểu được kỹ thuật đối sánh ảnh dựa trên ngữ cảnh hình dạng.
- Nắm bắt được việc tính toán khoảng cách ngữ cảnh hình dạng và đối sánh giữa hai đường bao đã được lấy mẫu.

Tuy nhiên, do điều kiện thời gian và khả năng có hạn, nên em chưa đi sâu tìm hiểu được thêm về ứng dụng, vì vậy việc xử lý code trên OpenCV còn chưa tốt, chương trình vẫn còn nhiều thiết sót.

Trong thời gian tới, em sẽ cố gắng tìm hiểu thêm các vấn đề về đối sánh ảnh dựa trên đường bao đối tượng và xây dựng ứng dụng hoàn chỉnh bao gồm cả giao diện người dùng và CSDL đặc trưng một cách đầy đủ và hệ thống.

TÀI LIỆU THAM KHẢO

[1] Markéta Potůčková**, "**Image matching and its applications in photogrammetry**",** February 2006.

[2] Beant Kaur, Anil Garg, Comparative study of different edge detection techniques, International journal of Engineering Science and Technology (IJEST), vol. 3, No. 3 March 2011.

[3] Raman Maini and Dr. Himanshu Aggarwai, Study and Comparison of various Image Edge Detection Techniques, International journal of Image Processing, Volume 3, Issue 1.

[4] H. Ling and D. Jacobs, "Shape classification using the innerdistance, "IEEE Transactions on Patter n Analysis and Machine Intelligence, . 29 (2007).

[5] Xie, J., Heng, P.A., Shah, M.: Shape matching and modeling using skeletal context. Pattern Recognition. 41 (2008).

[6] S. Belongie, J. Malik and J. Puzicha. "Shape Matching and Object Recognition Using Shape Context, " IEEE Trans. Pattern Anal. Mach. Intell., 24(24):509-522, 2002.

[7] Dengsheng Zhang, "Image Retrieval Based on Shape, " Monash University, 2002.

[8] 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.

[9] A Survey of Shape Feature Extraction Techniques Mingqiang Yang, Kidiyo Kpalma, Joseph Ronsi.

[10] D. Zhang and G. Lu, "Review of shape representation and description techniques, " Pattern Recognition, vol. 37, pp. 1-19, 2004.