

ĐẠI HỌC THÁI NGUYÊN
ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG


Ngô Huy Cường

**TRA CỨU IC MÁY TÍNH DỰA VÀO HÌNH DẠNG ĐỐI
TƯỢNG TRONG ẢNH**

Chuyên ngành: Khoa học máy tính
Mã số: 60.48.01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên - 2014

MỞ ĐẦU

1. Đặt vấn đề

Trong xã hội hiện nay, ảnh số đóng một vai trò quan trọng đối với đời sống con người. Ảnh số không chỉ được sử dụng trong cuộc sống hàng ngày mà nó còn góp phần quan trọng trong việc cung cấp thông tin về vật thể, sự kiện, ... trong công tác khoa học.

Càng ngày con người càng phải đối mặt với một lượng lớn hình ảnh phải xử lý. Bài toán tra cứu hình ảnh nhằm tìm ra các ảnh tương tự trong cơ sở dữ liệu như ảnh mẫu tra cứu được nhiều nhóm quan tâm. Xuất phát trong hoàn cảnh đó tôi chọn đề tài “Tra cứu IC máy tính dựa vào hình dạng đối tượng trong ảnh” nhằm nghiên cứu một số kỹ thuật tra cứu ảnh dựa vào hình dạng. Trên cơ sở kiến thức được hệ thống hóa, áp dụng cho bài toán tra cứu IC máy tính.

Luận văn gồm có bố cục như sau:

Chương I: Khái quát về biểu diễn hình dạng và bài toán tra cứu IC.

Trong chương này trình bày khái quát về xử lý ảnh, sơ lược về IC và bài toán tra cứu IC.

Chương II: Một số kỹ thuật tra cứu ảnh dựa vào hình dạng.

Chương này trình bày một số kỹ thuật tra cứu ảnh dựa vào hình dạng, từ đó trích rút ra các đặc trưng áp dụng vào bài toán tra cứu IC.

Chương III: Chương trình thực nghiệm.

Giới thiệu bài toán IC, phân tích, xây dựng, tra cứu ảnh dựa vào hình dạng đối tượng trong ảnh.

CHƯƠNG I

KHÁI QUÁT VỀ BIỂU DIỄN HÌNH DẠNG VÀ BÀI TOÁN TRA CỨU IC

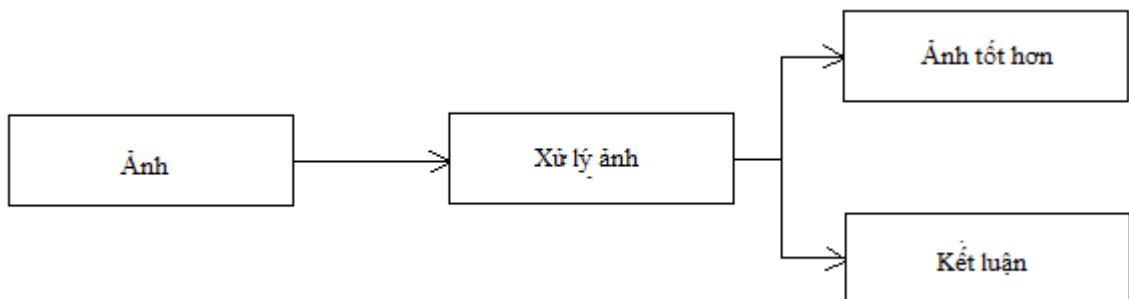
1.1. Khái quát về biểu diễn hình dạng

1.1.1. Xử lý ảnh, các vấn đề cơ bản trong xử lý ảnh

1.1.1.1. Xử lý ảnh là gì?

Con người thu nhận thông tin qua các giác quan, trong đó thị giác đóng vai trò quan trọng nhất. Những năm trở lại đây với sự phát triển của phần cứng máy tính, xử lý ảnh và đồ họa phát triển một cách mạnh mẽ và có nhiều ứng dụng trong cuộc sống. Xử lý ảnh và đồ họa đóng một vai trò quan trọng trong tương tác người máy.

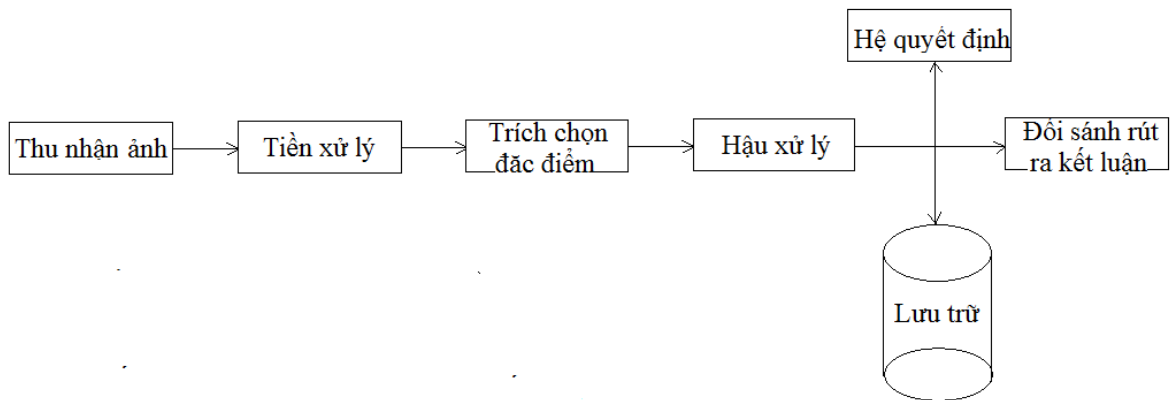
Quá trình xử lý ảnh được xem như là quá trình thao tác ảnh đầu vào nhằm cho ra kết quả mong muốn. Kết quả đầu ra của một quá trình xử lý ảnh có thể là một ảnh “tốt hơn” hoặc một kết luận.



Hình 1.1. Quá trình xử lý ảnh

Ảnh có thể xem là tập hợp các điểm ảnh và mỗi điểm ảnh được xem như là đặc trưng cường độ sáng hay một dấu hiệu nào đó tại một vị trí nào đó của đối tượng trong không gian và nó có thể xem như một hàm n biến $P(c_1, c_2, \dots, c_n)$. Do đó, ảnh trong xử lý ảnh có thể xem như ảnh n chiều.

Sơ đồ tổng quát của một hệ thống xử lý ảnh:



Hình 1.2: Các bước cơ bản trong một hệ thống xử lý ảnh

1.1.1.2. Các vấn đề cơ bản trong xử lý ảnh

a. Một số khái niệm cơ bản

* Ảnh và điểm ảnh:

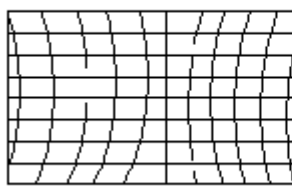
Điểm ảnh được xem như là dấu hiệu hay cường độ sáng tại 1 tọa độ trong không gian của đối tượng và ảnh được xem như là 1 tập hợp các điểm ảnh.

* Mức xám, màu

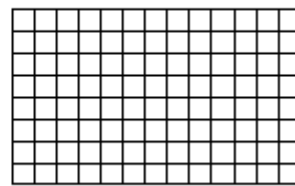
Là số các giá trị có thể có của các điểm ảnh của ảnh

b. Nắn chỉnh biến dạng

Ảnh thu nhận thường bị biến dạng do các thiết bị quang học và điện tử.



Ảnh thu nhận



Ảnh mong muốn

Hình 1.3. Ảnh thu nhận và ảnh mong muốn

Để khắc phục người ta sử dụng các phép chiếu, các phép chiếu thường được xây dựng trên tập các điểm điều khiển.

Giả sử (P_i, P_i') $i = 1, n$ có n các tập điều khiển. Giả sử (P_i, P_i') $i = 1, n$ có n

các tập điều khiển Tìm hàm $f: P_i \in f(P_i)$ sao cho $\sum_{i=1}^n \|f(P_i) - P_i'\|^2 \rightarrow \text{Min}$

Giả sử ảnh bị biến đổi chỉ bao gồm: Tịnh tiến, quay, tỷ lệ, biến dạng, bậc nhất tuyến tính. Khi đó hàm f có dạng:

$$f(x, y) = (a_1x + b_1y + c_1, a_2x + b_2y + c_2)$$

Ta có:

$$\varphi = \sum_{i=1}^n (f(P_i) - P'_i)^2 = \sum_{i=1}^n [(a_1x_i + b_1y_i + c_1 - x'_i)^2 + (a_2x_i + b_2y_i + c_2 - y'_i)^2]$$

Đề $\varphi \rightarrow \text{Min}$

$$\begin{cases} \frac{\partial \varphi}{\partial a_1} = 0 \\ \frac{\partial \varphi}{\partial b_1} = 0 \\ \frac{\partial \varphi}{\partial c_1} = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{i=1}^n a_1 x_i^2 + \sum_{i=1}^n b_1 x_i y_i + \sum_{i=1}^n c_1 x_i = \sum_{i=1}^n x_i x'_i \\ \sum_{i=1}^n a_1 x_i y_i + \sum_{i=1}^n b_1 y_i^2 + \sum_{i=1}^n c_1 y_i = \sum_{i=1}^n y_i x'_i \\ \sum_{i=1}^n a_1 x_i + \sum_{i=1}^n b_1 y_i + n c_1 = \sum_{i=1}^n x'_i \end{cases}$$

Giải hệ phương trình tuyến tính tìm được a_1, b_1, c_1

Tương tự tìm được a_2, b_2, c_2

Xác định hàm f .

c. Khử nhiễu

Có 2 loại nhiễu cơ bản trong quá trình thu nhận ảnh:

- Nhiễu hệ thống: Là nhiễu có quy luật có thể khử bằng các phép biến đổi.
- Nhiễu ngẫu nhiên: Vết bản không rõ nguyên nhân \rightarrow khắc phục bằng các phép lọc.

d. Chỉnh mức xám

Nhằm khắc phục tính không đồng đều của hệ thống gây ra. Thông thường có 2 hướng tiếp cận:

- Giảm số mức xám: Thực hiện bằng cách nhóm các mức xám gần nhau thành một bó. Trường hợp chỉ có 2 mức xám thì chính là chuyển về ảnh đen trắng. Ứng dụng: In ảnh màu ra máy in đen trắng.

- Tăng số mức xám: Thực hiện nội suy ra các mức xám trung gian bằng kỹ thuật nội suy. Kỹ thuật này nhằm tăng cường độ mịn cho ảnh.

e. Trích chọn đặc điểm

Các đặc điểm của đối tượng được trích chọn tùy theo mục đích nhận dạng trong quá trình xử lý ảnh. Có thể nêu ra một số đặc điểm của ảnh sau đây:

- Đặc điểm không gian: Phân bố mức xám, phân bố xác suất, biên độ, điểm uốn v.v..

- Đặc điểm biên đổi: Các đặc điểm loại này được trích chọn bằng việc thực hiện lọc vùng (Zonal filtering). Các bộ vùng được gọi là “mặt nạ đặc điểm” (Feature mask) thường là các khe hẹp với hình dạng khác nhau (chữ nhật, tam giác, cung tròn v.v..).

- Đặc điểm biên và đường biên: Đặc trưng cho đường biên của đối tượng và do vậy rất hữu ích trong việc trích chọn các thuộc tính bất biến được dùng khi nhận dạng đối tượng. Các đặc điểm này có thể được trích chọn nhờ toán tử gradient, toán tử la bàn, toán tử Laplace, toán tử “chéo không” (Zero crossing) v.v..

Việc trích chọn hiệu quả các đặc điểm giúp cho việc nhận dạng các đối tượng ảnh chính xác, với tốc độ tính toán cao và dung lượng nhớ lưu trữ giảm xuống.

f. Nhận dạng

Nhận dạng tự động (Automatic Recognition), mô tả đối tượng, phân loại và phân nhóm các mẫu là những vấn đề quan trọng trong thị giác máy, được ứng dụng trong nhiều ngành khoa học khác nhau. Tuy nhiên, một câu hỏi đặt ra là: Mẫu (Pattern) là gì?

Watanabe, một trong những người đi đầu trong lĩnh vực này đã định nghĩa: “Ngược lại với hỗn loạn (Chaos), mẫu là một thực thể (Entity), được xác định một cách ang áng (Vaguely defined) và có thể gán cho nó một tên gọi nào đó”.

Ví dụ mẫu có thể là ảnh của vân tay, ảnh của một vật nào đó được chụp, một chữ viết, khuôn mặt người hoặc một ký đồ tín hiệu tiếng nói.

Khi biết một mẫu nào đó, để nhận dạng hoặc phân loại mẫu đó có thể: Phân

loại có mẫu (Supervised classification), chẳng hạn phân tích phân biệt (Discriminant analysis), trong đó mẫu đầu vào được định danh như một thành phần của một lớp đã xác định.

Phân loại không có mẫu (Unsupervised classification hay Clustering) trong đó các mẫu được gán vào các lớp khác nhau dựa trên một tiêu chuẩn đồng dạng nào đó. Các lớp này cho đến thời điểm phân loại vẫn chưa biết hay chưa được định danh.

Hệ thống nhận dạng tự động bao gồm ba khâu tương ứng với ba giai đoạn chủ yếu sau đây:

- Thu nhận dữ liệu và tiền xử lý.
- Biểu diễn dữ liệu.
- Nhận dạng, ra quyết định.

Bốn cách tiếp cận khác nhau trong lý thuyết nhận dạng là:

- Đối sánh mẫu dựa trên các đặc trưng được trích chọn.
- Phân loại thống kê.
- Đối sánh cấu trúc.
- Phân loại dựa trên mạng nơ-ron nhân tạo.

Trong các ứng dụng rõ ràng là không thể chỉ dùng có một cách tiếp cận đơn lẻ để phân loại “tối ưu” do vậy cần sử dụng cùng một lúc nhiều phương pháp và cách tiếp cận khác nhau. Do vậy, các phương thức phân loại tổ hợp hay được sử dụng khi nhận dạng và nay đã có những kết quả có triển vọng dựa trên thiết kế các hệ thống lai (Hybrid System) bao gồm nhiều mô hình kết hợp.

Việc giải quyết bài toán nhận dạng trong những ứng dụng mới nảy sinh trong cuộc sống không chỉ tạo ra những thách thức về thuật giải, mà còn đặt ra những yêu cầu về tốc độ tính toán. Đặc điểm chung của tất cả những ứng dụng đó là những đặc điểm đặc trưng cần thiết thường là nhiều, không thể do chuyên gia đề xuất, mà phải được trích chọn dựa trên các thủ tục phân tích dữ liệu.

g. Nén ảnh

Nhằm giảm thiểu không gian lưu trữ. Thường được tiến hành theo cả hai khuynh hướng là nén có bảo toàn và không bảo toàn thông tin. Nén không bảo toàn thì thường có khả năng nén cao hơn nhưng khả năng phục hồi thì kém hơn. Trên cơ sở hai khuynh hướng, có 4 cách tiếp cận cơ bản trong nén ảnh:

- Nén ảnh thống kê: Kỹ thuật nén này dựa vào việc thống kê tần xuất xuất hiện của giá trị các điểm ảnh, trên cơ sở đó mà có chiến lược mã hóa thích hợp. Một ví dụ điển hình cho kỹ thuật mã hóa này là *.TIF

- Nén ảnh không gian: Kỹ thuật này dựa vào vị trí không gian của các điểm ảnh để tiến hành mã hóa. Kỹ thuật lợi dụng sự giống nhau của các điểm ảnh trong các vùng gần nhau.

Ví dụ: Cho kỹ thuật này là mã nén *.PCX

- Nén ảnh sử dụng phép biến đổi: Đây là kỹ thuật tiếp cận theo hướng nén không bảo toàn và do vậy, kỹ thuật hướng nén hiệu quả hơn. *.JPG chính là tiếp cận theo kỹ thuật nén này.

- Nén ảnh Fractal: Sử dụng tính chất Fractal của các đối tượng ảnh, thể hiện sự lặp lại của các chi tiết. Kỹ thuật nén sẽ tính toán để chỉ cần lưu trữ phần gốc ảnh và quy luật sinh ra ảnh theo nguyên lý Fractal.

1.1.2. Thu nhận và biểu diễn ảnh

1.1.2.1. Thu nhận, các thiết bị thu nhận ảnh

Các thiết bị thu nhận ảnh bao gồm Camera, Scanner các thiết bị thu nhận này có thể cho ảnh đen trắng.

Các thiết bị thu nhận ảnh có 2 loại chính ứng với 2 loại ảnh thông dụng Raster, Vector.

Các thiết bị thu nhận ảnh thông thường Raster là Camera các thiết bị thu nhận ảnh thông thường Vector là Sensor hoặc bằng số hoá Digitalizer hoặc được chuyển đổi từ ảnh Raster. Nhìn chung các hệ thống thu nhận ảnh thực hiện 1 quá trình.

- Cảm biến: Biến đổi năng lượng quang học thành năng lượng điện
- Tổng hợp năng lượng điện thành ảnh

1.1.2.2. Biểu diễn ảnh

Ảnh trên máy tính là kết quả thu nhận theo các phương pháp số hoá được nhúng trong các thiết bị kỹ thuật khác nhau. Quá trình lưu trữ ảnh nhằm 2 mục đích:

- Tiết kiệm bộ nhớ.
- Giảm thời gian xử lý.

Việc lưu trữ thông tin trong bộ nhớ có ảnh hưởng rất lớn đến việc hiển thị, in ấn và xử lý ảnh được xem như là một tập hợp các điểm với cùng kích thước nếu sử dụng càng nhiều điểm ảnh thì bức ảnh càng đẹp, càng mịn và càng thể hiện rõ hơn chi tiết của ảnh người ta gọi đặc điểm này là độ phân giải.

Việc lựa chọn độ phân giải thích hợp tùy thuộc vào nhu cầu sử dụng và đặc trưng của mỗi ảnh cụ thể, trên cơ sở đó các ảnh thường được biểu diễn theo 2 mô hình cơ bản.

1.1.3. Mô hình Raster

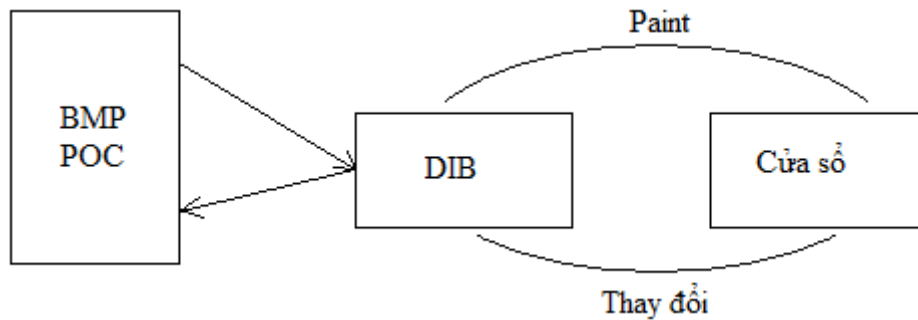
Đây là cách biểu diễn ảnh thông dụng nhất hiện nay, ảnh được biểu diễn dưới dạng ma trận các điểm (điểm ảnh). Thường thu nhận qua các thiết bị như Camera, Scanner. Tùy theo yêu cầu thực tế mà mỗi điểm ảnh được biểu diễn qua một hay nhiều bit.

Mô hình Raster thuận lợi cho hiển thị và in ấn. Ngày nay công nghệ phần cứng cung cấp những thiết bị thu nhận ảnh Raster phù hợp với tốc độ nhanh và chất lượng cao cho cả đầu vào và đầu ra. Một thuận lợi cho việc hiển thị trong môi trường Windows là Microsoft đưa ra khuôn dạng ảnh DIB (Device Independent Bitmap) làm trung gian. Hình 1.4 thể hiện quy trình chung để hiển thị ảnh Raster thông qua DIB.

Một trong những hướng nghiên cứu cơ bản trên mô hình biểu diễn này là kỹ thuật nén ảnh các kỹ thuật nén ảnh lại chia ra theo hai khuynh hướng là nén bảo toàn và không bảo toàn thông tin. Nén bảo toàn có khả năng phục hồi hoàn toàn dữ liệu ban đầu còn nếu không bảo toàn chỉ có khả năng phục hồi độ sai số cho phép

nào đó. Theo cách tiếp cận này người ta đã đề ra nhiều quy cách khác nhau như BMP, TIF, GIF, PCX...

Hiện nay trên thế giới có trên 50 khuôn dạng ảnh thông dụng bao gồm cả trong đó các kỹ thuật nén có khả năng phục hồi dữ liệu 100% và nén có khả năng phục hồi với độ sai số nhận được.



Hình 1.4. Quá trình hiển thị và chỉnh sửa, lưu trữ ảnh thông qua DIB

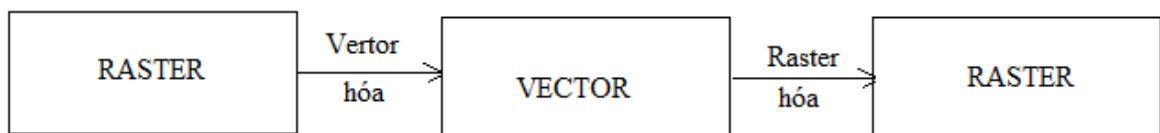
1.1.4. Mô hình Vector

Biểu diễn ảnh ngoài mục đích tiết kiệm không gian lưu trữ dễ dàng cho hiển thị và in ấn còn đảm bảo dễ dàng trong lựa chọn sao chép di chuyển tìm kiếm... Theo những yêu cầu này kỹ thuật biểu diễn Vector tỏ ra ưu việt hơn.

Trong mô hình Vector người ta sử dụng hướng giữa các Vector của điểm ảnh lân cận để mã hoá và tái tạo hình ảnh ban đầu ảnh Vector được thu nhận trực tiếp từ các thiết bị số hoá như Digital hoặc được chuyển đổi từ ảnh Raster thông qua các chương trình số hoá

Công nghệ phần cứng cung cấp những thiết bị xử lý với tốc độ nhanh và chất lượng cho cả đầu vào và ra nhưng lại chỉ hỗ trợ cho ảnh Raster.

Do vậy, những nghiên cứu về biểu diễn vectơ đều tập trung từ chuyển đổi từ ảnh Raster.



Hình 1.5: Sự chuyển đổi giữa các mô hình biểu diễn ảnh