

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



ĐỒ ÁN TỐT NGHIỆP

NGÀNH CÔNG NGHỆ THÔNG TIN

HẢI PHÒNG 2009

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

-----o0o-----

TÌM HIỂU BÀI TOÁN NHẬN DẠNG BIẾN SỐ XE

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công nghệ thông tin

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

TÌM HIỂU BÀI TOÁN NHẬN DẠNG BIỂU SỐ XE

ĐỒ ÁN TỐT NGHIỆP HỆ CHÍNH QUY
Ngành: Công nghệ thông tin

Sinh viên thực hiện: Phạm Thị Thanh Thủy
Giáo viên hướng dẫn: PGS.TS. ĐỖ NĂNG TOÀN
Mã số sinh viên: 090125

BỘ GIÁO DỤC VÀ ĐÀO TẠO TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG	CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM Độc lập - Tự do - Hạnh phúc -----o0o-----
--	---

NHIỆM VỤ THIẾT KẾ TỐT NGHIỆP

Sinh viên: Phạm Thị Thanh Thủy.

Mã số: 090125.

Lớp: CT902.

Ngành: Công nghệ thông tin.

Tên đề tài:

TÌM HIỂU BÀI TOÁN NHẬN DẠNG BIẾN SỐ XE

NHIỆM VỤ ĐỀ TÀI

1. Nội dung và các yêu cầu cần giải quyết trong nhiệm vụ đề tài tốt nghiệp

a. Nội dung:

- Tìm hiểu về biển số xe và hệ thống nhận dạng biển số xe
- Phát biểu bài toán và hướng giải quyết
- Nghiên cứu một số thuật toán ứng dụng trong việc nhận dạng biển số xe

b. Các yêu cầu cần giải quyết:

- Tìm hiểu khái quát về xử lý ảnh và bài toán nhận dạng biển số xe
- Tìm hiểu thông tin về biển số xe và phân loại biển số xe của Việt Nam
- Tìm hiểu các công đoạn chính của bài toán nhận dạng biển số xe gồm 2 khâu chính:
 - Phát hiện biển số xe
 - Nhận dạng biển số xe
- Cài đặt thử nghiệm

2 . Địa điểm thực tập:

VIỆN CNTT-VIỆN KH&CN Việt Nam
Đc: Số 8 HOÀNG QUỐC VIỆT-HÀ NỘI

CÁN BỘ HƯỚNG DẪN ĐỀ TÀI TỐT NGHIỆP

Người hướng dẫn thứ nhất:.....

Họ và tên: ĐỖ NĂNG TOÀN

Học hàm, học vị: PGS.TS.....

Cơ quan công tác: VIỆN CNTT-VIỆN KH&CN Việt Nam

Nội dung hướng dẫn:

.....

.....

.....

Người hướng dẫn thứ hai:.....

Họ và tên: ĐỖ NĂNG TOÀN

Học hàm, học vị: PGS.TS.....

Cơ quan công tác: VIỆN CNTT-VIỆN KH&CN Việt Nam

Nội dung hướng dẫn:

.....

.....

.....

Đề tài tốt nghiệp được giao ngày 06 tháng 04 năm 2009

Yêu cầu phải được hoàn thành trước ngày 11 tháng 07 năm 2009

Đã nhận nhiệm vụ: Đ.T.T.N
Sinh viên

Đã nhận nhiệm vụ: Đ.T.T.N
Cán bộ hướng dẫn Đ.T.T.N

Hải Phòng, ngày..... tháng..... năm 2009

HIỆU TR- ỜNG

GS.TS.NGŨT Trần Hữu Nghị

PHẦN NHẬN XÉT ĐÁNH GIÁ CỦA CÁN BỘ CHẤM PHẢN BIỆN
ĐỀ TÀI TỐT NGHIỆP

1. Đánh giá chất lượng đề tài tốt nghiệp (về các mặt như cơ sở lí luận, thuyết minh chương trình, giá trị thực tế...)

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

2. Cho điểm của cán bộ phản biện
(Điểm ghi bằng số và chữ)

.....

.....

Ngày.....tháng.....năm 2009
Cán bộ chấm phản biện
(Ký, ghi rõ họ tên)

LỜI CẢM ƠN

Em xin chân thành cảm ơn PGS.TS ĐỖ NĂNG TOÀN - VIỆN CNTT - VIỆN KH&CN Việt Nam, người đã trực tiếp hướng dẫn tận tình và tạo mọi điều kiện thuận lợi cho em để em hoàn thành khoá luận của mình.

Em cũng xin chân thành cảm ơn tất cả các anh chị trong Viện CNTT - Viện KH&CN Việt Nam đã nhiệt tình chỉ dạy và cung cấp những kiến thức quý báu để em có thể hoàn thành tốt báo cáo tốt nghiệp này.

Vì thời gian có hạn, kinh nghiệm bản thân còn ít. Cho nên trong đề tài của em không tránh khỏi những thiếu sót, em rất mong được sự góp ý quý báu của tất cả các thầy cô giáo cũng như các bạn để đề tài của em được hoàn thiện hơn.

Em xin chân thành cảm ơn!

Hải Phòng, ngày 4 tháng 07 năm 2009.

Sinh viên

Phạm Thị Thanh Thủy

Mục lục

Phần giới thiệu

Chương 1: TỔNG QUAN VỀ BÀI TOÁN NHẬN DẠNG BIÊN SỐ XE. 21

1.1. Khái quát về xử lý ảnh:	21
1.2. Khái niệm về nhận dạng biên số xe:	21
1.2.1 Khái niệm & ứng dụng:	21
1.2.2 Phân loại biên số xe:	24
1.3. Một số hướng giải quyết bài toán nhận dạng biên số xe:.....	27
1.3.1. Hướng tiếp cận phát triển vùng:	27
1.3.2. Hướng tiếp cận dò biên và biến đổi Hough:	27
1.4 Hướng giải quyết:	28

Chương 2: PHÁT HIỆN VÙNG CHỨA BIÊN SỐ XE..... 31

2.1 Một số khái niệm cơ bản:	31
2.1.1 Tổng quan về ảnh.....	31
2.1.2 Phương pháp tách dò ngưỡng tự động.....	32
2.2 Biên và các phương pháp phát hiện biên.	33
2.2.1 Phương pháp gradient	33
2.2.2. Kỹ thuật Laplace:.....	35
2.3 Phát hiện vùng chứa biên số xe.....	37
2.3.1. Nhị phân hóa ảnh	37
2.3.2 Tách biên:.....	38
2.3.3 Biến đổi HOUGH	39
2.3.4 Trích chọn đoạn thẳng và tính giao điểm	42
2.3. Xác định chính xác vùng chứa biên số xe.....	43
2.3.1. Bước ban đầu:	44
2.4.2 Tiêu chí tỷ lệ chiều dài/rộng.	45
2.4.3 Tiêu chí số ký tự trong vùng biên số xe.....	46

Chương 3: BÀI TOÁN NHẬN DẠNG KÝ TỰ..... 47

3.1 Tổng quan về nhận dạng	47
----------------------------------	----

3.1.1 Không gian biểu diễn đối tượng, không gian diễn dịch	47
3.1.2 Mô hình và bản chất của quá trình nhận dạng.....	48
3.1.2.1 Mô hình	48
3.1.2.2 Bản chất của quá trình nhận dạng	50
3.2 Mô hình mạng nơron nhân tạo	51
3.2.1 Mô hình nơron nhân tạo.....	51
3.2.2 Mạng Nơron.....	52
3.2.2.1 Phân loại các mạng nơron	53
3.2.2.2 Hai chức năng của mạng nơron.....	54
3.2.3 Mạng Kohonen	56
3.2.3.1 Cấu trúc mạng.....	56
3.2.3.3 Sử dụng mạng.....	59
3.2.4 Mạng nơron nhiều lớp lan truyền ngược sai số	61
3.2.4.1 Kiến trúc mạng	61
3.2.4.2 Huấn luyện mạng.....	61
3.2.4.3 Sử dụng mạng.....	63
3.3 Sử dụng mạng nơron lan truyền ngược hướng cho nhận dạng ký tự.....	63
3.3.1 Nhận dạng bằng mạng nơron lan truyền ngược hướng (kn chung). 63	
3.3.2 Cài đặt mạng lan truyền ngược hướng cho nhận dạng ký tự	64
3.3.3 Nhận dạng các ký tự sử dụng mạng lan truyền ngược hướng	66
Kết luận	67
PHẦN KẾT LUẬN.....	68
TÀI LIỆU THAM KHẢO	69

PHẦN GIỚI THIỆU

Ngày nay trên thế giới bên cạnh việc tăng trưởng kinh tế là sự phát triển của các ngành khoa học kỹ thuật nói chung, mà trong đó ngành công nghiệp sản xuất các phương tiện giao thông lại là một trong những ngành có tốc độ phát triển cực nhanh. Sự phát triển ấy, được thể hiện rõ ràng nhất thông qua hình ảnh các phương tiện giao thông trên thế giới ngày một tăng cao và đa dạng. Tuy nhiên, điều đó lại gây ra một áp lực đối với những người và cơ quan các cấp quản lý, làm cho công tác quản lý và giám sát sẽ khó khăn hơn,...

Và đây cũng là một trong những vấn nạn ở Việt Nam. Công tác quản lý phương tiện giao thông nói chung và quản lý ô tô, xe máy là vô cùng phức tạp... cũng như công tác phát hiện, xử phạt các hành vi vi phạm giao thông, chống trộm,... sẽ tốn nhiều thời gian và công sức hơn...

Để làm giảm lượng nhân lực trong việc công tác quản lý, kiểm soát phương tiện giao thông, trên thế giới đã nhanh chóng xây dựng hệ thống giám sát tự động đối với các phương tiện giao thông. Và các hệ thống giám sát đều lấy biển số xe là mục tiêu giám sát. Hệ thống này đã được sử dụng rộng rãi tuy nhiên ở Việt Nam đây vẫn là một lĩnh vực mới mẻ.

Do đó em chọn làm đề tài “Tìm hiểu hệ thống nhận dạng biển số xe” với mục đích để tìm hiểu nhằm trợ giúp cho công tác giám sát, quản lý các phương tiện giao thông một cách hiệu quả, dễ dàng và nhanh chóng hơn

Em tin ở Việt Nam mình trong tương lai gần hệ thống này sẽ được sử dụng rộng rãi.

Bố cục trình bày trong báo cáo của em gồm 3 phần:

- **Chương 1:** Tổng quan về bài toán nhận dạng biển số xe
- **Chương 2:** Phát hiện vùng chứa biển số xe
- **Chương 3:** Nhận dạng ký tự

Chương 1:

TỔNG QUAN VỀ BÀI TOÁN NHẬN DẠNG BIỂN SỐ XE

1.1. Khái quát về xử lý ảnh:

Xử lý ảnh là một trong những mảng quan trọng nhất trong kỹ thuật thị giác máy tính, là tiền đề cho nhiều nghiên cứu thuộc lĩnh vực này. Hai nhiệm vụ cơ bản của quá trình xử lý ảnh là nâng cao chất lượng thông tin hình ảnh và xử lý số liệu cung cấp cho các quá trình khác trong đó có việc ứng dụng thị giác vào điều khiển.

Quá trình bắt đầu từ việc thu nhận ảnh nguồn (từ các thiết bị thu nhận ảnh dạng số hoặc tương tự) gửi đến máy tính. Dữ liệu ảnh được lưu trữ ở định dạng phù hợp với quá trình xử lý. Người lập trình sẽ tác động các thuật toán tương ứng lên dữ liệu ảnh nhằm thay đổi cấu trúc ảnh phù hợp với các ứng dụng khác nhau.

- Chuyển ảnh màu thành ảnh xám
- Lược đồ xám của ảnh (Histogram)
- Các bộ lọc không gian
 - + Lọc tuyến tính
 - + Lọc phi tuyến
- Tách biên đối tượng

1.2. Khái niệm về nhận dạng biển số xe:

1.2.1 Khái niệm & ứng dụng:

a) *Khái niệm:*

Hệ thống nhận dạng biển số xe là hệ thống có khả năng phân tích hình ảnh và xác định biển số trên xe, thông qua video, thiết bị ghi hình và hình ảnh. Sau cùng là xác định các thông tin như: chủ sở hữu xe, theo dõi xe với tốc độ chậm,

* *Phân loại ứng dụng nhận dạng biển số xe:*

Ứng dụng nhận dạng biển số xe là ứng dụng có khả năng phân tích hình ảnh và xác định biển số xe từ các hình ảnh chụp được từ các thiết bị thu hình.

Nguồn hình ảnh cho ứng dụng có rất nhiều. Và phát triển, hình ảnh được trực tiếp thu nhận từ camera. Trong báo cáo tốt nghiệp của em chỉ dừng lại ở mức xác định được biển số xe (xác định các chữ) từ các bức ảnh.

Có nhiều cách thức khác nhau để phân loại các ứng dụng nhận dạng biển số xe. Một trong những cách đơn giản là phân loại ứng dụng nhận dạng biển số xe thông qua mục đích sử dụng. Có thể chia ứng dụng nhận dạng biển số xe thành hai loại sau:

Loại 1: Giới hạn vùng nhìn

- Đầu vào: Ảnh thu trực tiếp từ các thiết bị ghi nhận ảnh kỹ thuật số. Ảnh được ghi nhận thường chỉ giới hạn trong vùng có biển số xe.

- Nguyên lý hoạt động: Các phương tiện giao thông phải chạy với một tốc độ đủ chậm để máy ghi nhận hình ảnh có thể thu được ảnh vùng biển số xe.

- Ứng dụng: Những ứng dụng nhận dạng biển số xe loại này thường được dung tại các trạm kiểm soát, các trạm thu phí, các bãi gửi xe tự động, các trạm gác cổng.

Loại 2: Không giới hạn vùng nhìn

- Đầu vào: Ảnh đầu vào thu được từ các thiết bị ghi hình tự động, không phụ thuộc vào góc độ, các đối tượng xung quanh, ảnh không cần bắt buộc chỉ chụp vùng chứa biển số xe, mà có thể ảnh tổng hợp như chứa thêm các đối tượng như người, cây, đường phố..., miễn là vùng biển số xe phải đủ rõ để có thể thực hiện nhận dạng được các ký tự trong vùng đó.

- Nguyên lý hoạt động: Do đặc tính không giới hạn vùng nhìn mà ảnh đầu vào có thể thu được từ một thiết bị ghi hình (camara, máy ảnh...). Và do đó, công việc đầu tiên là dò tìm trong ảnh, để xác định đúng vùng nào là biển số xe. Sau đó, thực hiện tách vùng và nhận dạng. Cuối cùng tùy thuộc vào mục đích sử dụng mà kết quả nhận dạng được truyền đi hay lưu trữ để phục vụ nhu cầu của người dùng cuối.

- Ứng dụng: Vì không phụ thuộc vào hình ảnh thu được nên có thể dùng ứng dụng tại nhiều nơi như tại những nơi điều tiết giao thông, tại các vị trí nhảy

cảm của giao thông như ngã ba, ngã tư đường giao nhau. Kiểm soát, phát hiện những hành vi vi phạm an toàn giao thông.

- Yêu cầu về thiết bị:
 - Máy quét
 - Đèn chiếu
 - Bộ phận thu – tách hình ảnh
 - Máy tính
 - Các thiết bị khác tùy mục đích sử dụng

Trong quá trình tìm hiểu, xây dựng ứng dụng của mình. Ứng dụng mà em hướng tới trong quá trình xây dựng là ứng dụng loại 2. Vì vậy, trong toàn bộ báo cáo này, chỉ nêu cách thức giải quyết là làm sao nhận dạng (lọc ra) được các ký tự số và chữ.

b) Ứng dụng của hệ thống nhận dạng biển số xe:

Hệ thống nhận dạng biển số xe được xây dựng nhằm mục đích giám sát, kiểm soát các phương tiện. Dưới đây chúng ta đề cập đến một số ứng dụng phổ biến đối với hệ thống nhận dạng biển số xe:

+) Thu phí giao thông: Lắp đặt hệ thống “Nhận dạng biển số xe” tại các trạm thu phí nhằm hỗ trợ hoặc tự động hóa công tác thu phí.

+) Kiểm soát xe tại các đường biên giới: Mỗi quốc gia đều có những quy định riêng về biển số xe, để phục vụ cho công tác quản lý và phát hiện những phương tiện giao thông (xe) vượt biên giới bất hợp pháp. Việc lắp hệ thống “Nhận dạng biển số xe” tại các trạm kiểm soát sẽ góp phần hỗ trợ công tác kiểm tra và an ninh quốc gia.

+) Các trạm gác cổng: Việc lắp đặt hệ thống “Nhận dạng biển số xe” sẽ hỗ trợ hoặc tự động hóa công tác mở cổng cho xe vào.

Ngoài ra, hệ thống còn được ứng dụng vào công tác chống trộm xe, các bãi giữ xe tự động, điều tiết giao thông (chẳng hạn như Thành phố Dublin đã ứng dụng công nghệ “Nhận dạng biển số xe tự động” trong việc điều tiết giao thông theo dạng biển số chẵn/lẻ)

1.2.2 Phân loại biển số xe:

Trước tiên là quy định biển số của 64 tỉnh thành (Biển trắng chữ đen):

11 - Cao Bằng	43 - Đà Nẵng	77 - Bình Định
12 - Lạng Sơn	47 - Đắk Lắk	78 - Phú Yên
14 - Quảng Ninh	48 - Đắk Nông	79 - Khánh Hòa
15,16 - Hải Phòng	49 - Lâm Đồng	80 - Các đơn vị kinh tế thuộc TW (hàng không)
17 - Thái Bình	50 đến 59 - TP. Hồ Chí Minh	81 - Gia Lai
18 - Nam Định	60 - Đồng Nai	82 - KonTum
19 - Phú Thọ	61 - Bình Dương	83 - Sóc Trăng
20 - Thái Nguyên	62 - Long An	84 - Trà Vinh
21 - Yên Bái	63 - Tiền Giang	85 - Ninh Thuận
22 - Tuyên Quang	64 - Vĩnh Long	86 - Bình Thuận
23 - Hà Giang	65 - Cần Thơ	88 - Vĩnh Phúc
24 - Lào Cai	66 - Đồng Tháp	89 - Hưng Yên
25 - Lai Châu	67 - An Giang	90 - Hà Nam
26 - Sơn La	68 - Kiên Giang	92 - Quảng Nam
27 - Điện Biên	69 - Cà Mau	93 - Bình Phước
28 - Hòa Bình	70 - Tây Ninh	94 - Bạc Liêu
29,30,31,32 - Hà Nội	71 - Bến Tre	95 - Hậu Giang
33 - Hà Tây	72 - Bà Rịa - Vũng Tàu	97 - Bắc Cạn
34 - Hải Dương	73 - Quảng Bình	98 - Bắc Giang
35 - Ninh Bình	74 - Quảng Trị	99 - Bắc Ninh
36 - Thanh Hóa	75 - Huế	
37 - Nghệ An	76 - Quảng Ngãi	
38 - Hà Tĩnh		

**) Những quy định về màu sắc và chữ số đặc biệt:*

1. Màu xanh chữ trắng là biển xe của các cơ quan hành chính sự nghiệp:

- Trục thuộc chính phủ thì là biển xanh 80

- Các tỉnh thành thì theo số tương ứng

2. Màu đỏ chữ trắng là biển xe trong quân đội:

AT: Binh đoàn 12

AD: Quân Đoàn 4 , Binh đoàn cứu long

BB: bộ binh

BC: Binh chủng Công Binh

BH: Binh chủng hoá học

BS: Binh đoàn Trường Sơn

BT: Binh chủng thông tin liên lạc

BP: Bộ tư lệnh biên phòng

HB: Học viện lục quân

HH: Học viện quân y

KA: Quân khu 1

KB: Quân khu 2

KC: Quân khu 3

KD: Quân khu 4

KV: Quân khu 5

KP: Quân khu 7

KK: Quân khu 9

PP: Các quân y viện

QH: Quân chủng hải quân

QK, QP: Quân chủng phòng không không quân

TC: Tổng cục chính trị

TH: Tổng cục hậu cần

TK: Tổng cục công nghiệp quốc phòng

TT: Tổng cục kỹ thuật

TM: Bộ tổng tham mưu

VT: Viettel

3. Màu trắng 2 chữ, 5 số là biển dành cho người nước ngoài:

- NG là xe ngoại giao

- NN là xe của các tổ chức, cá nhân nước ngoài: Trong đó 3 số ở giữa là mã quốc gia, 2 số tiếp theo là số thứ tự.

* Xe số 80 NG xxx-yy là biển cấp cho các đại sứ quán, thêm gạch đỏ ở giữa và 2 số cuối là 01 là biển xe của tổng lãnh sự

4. Những xe mang biển 80 gồm có :

- Các Ban của Trung ương Đảng
- Văn phòng Chủ tịch nước
- Văn phòng Quốc hội
- Văn phòng Chính phủ
- Bộ Công an
- Xe phục vụ các đồng chí uỷ viên Trung ương Đảng công tác tại Hà Nội và các thành viên Chính phủ
- Bộ ngoại giao
- Viện kiểm soát nhân dân tối cao
- Toà án nhân dân tối cao
- Đài truyền hình Việt Nam
- Đài tiếng nói Việt Nam
- Thông tấn xã Việt Nam
- Báo nhân dân
- Thanh tra Nhà nước
- Học viện Chính trị quốc gia
- Ban quản lý Lăng, Bảo tàng, khu Di tích lịch sử Hồ Chí Minh;
- Trung tâm lưu trữ quốc gia
- Uỷ ban Dân số kế hoạch hoá gia đình
- Tổng công ty Dầu khí Việt Nam
- Các đại sứ quán, tổ chức quốc tế và nhân viên người nước ngoài
- Uỷ ban Chứng khoán Nhà nước
- Cục Hàng không dân dụng Việt Nam
- Kiểm toán nhà nước

5. Các biến A :

Xe của Công An - Cảnh Sát tương ứng với các tỉnh

ví dụ: 31A = xe của Công An - Cảnh Sát thành phố Hà Nội

1.3. Một số hướng giải quyết bài toán nhận dạng biển số xe:

Có rất nhiều phương pháp tiếp cận. Trong đó có hai cách tiếp cận phổ biến dưới đây:

1.3.1. Hướng tiếp cận phát triển vùng:

- Nhóm tác giả Nigel Whyte and Adrien Kiernan được đại diện cho cách tiếp cận này

- Ý tưởng của phương pháp này: đó là biển số xe thường chứa một màu đồng nhất, chẳng hạn màu trắng, và có diện tích tương đối nhất định. Vì vậy có thể dùng phương pháp phát triển vùng, hoặc sử dụng khung chữ nhật di chuyển trong để tìm ra vùng có tính chất thỏa mãn biển số xe và tiến hành nhận dạng.

- Ưu điểm: rất đơn giản, và xử lý rất nhanh đối với những ảnh chỉ chứa vùng biển số xe.

- Nhược điểm: khi ảnh có thêm nhiều đối tượng không phải là vùng biển số xe, chẳng hạn là ảnh chụp tổng quát gồm cả cảnh vật bên ngoài thì cách tiếp cận này trở nên không hiệu quả. Vì vậy phương pháp này rất hiệu quả đối với hệ thống trạm thu phí, trạm gác cổng, gửi xe tự động

1.3.2. Hướng tiếp cận dò biên và biến đổi Hough:

- Nhóm tác giả Michael Lidenbaum, Rosen Alexander, Vichik Sergey, Sandler Roman được đại diện cho cách tiếp cận này.

- Ý tưởng của cách tiếp cận này là: Biển số xe được bao bọc bởi đường viền. Do đó, có thể dùng phương pháp phát hiện biên, sau đó dùng phép biến đổi Hough để trích những đoạn thẳng dọc, ngang tồn tại trong ảnh. Giao điểm của những đoạn thẳng này chính là vùng bao chứa biển số xe. Và cuối cùng là tiến hành nhận dạng các ký tự ở trên mỗi vùng con.

- Ưu điểm: độ chính xác cao. Và các hệ thống nhận dạng đa phần đều phát triển theo hướng tiếp cận này.

- Nhược điểm: Độ phức tạp tính toán khá cao. Khi ảnh có thêm nhiều đối tượng khác thì khối lượng tính toán tăng lên rất nhiều. Do mục đích là phải xác định được vùng con nào chứa biển số xe.

Ngoài hai cách tiếp cận trên, còn có nhiều cách tiếp cận khác để xác định chính xác vùng nào chứa biển số xe và bước cuối cùng là tiến hành nhận dạng ký tự. Mỗi cách tiếp cận có một ưu và nhược điểm. Đa số các ứng dụng đều sử dụng cách tiếp cận biến đổi Hough. Trong báo cáo đề tài của em, em xin trình bày cách tiếp cận Hough.

1.4 Hướng giải quyết:

Ở phần 1.3 chúng ta đã tìm hiểu 2 hướng giải quyết cho việc xác định vùng chứa biển số xe. Mỗi cách giải quyết có những ưu điểm và hạn chế riêng của nó.

**) Một số đặc điểm về biển số xe ở Việt Nam:*



- Tiêu chuẩn về kích thước: Ở mỗi nước thường có tiêu chuẩn về kích thước nhất định. Đối với nước ta, biển số xe qui định khá đồng đều cho mỗi loại xe, tỷ lệ chiều dài, rộng cho mỗi loại xe là như nhau. Đối với loại xe có một hàng ký tự thì tỉ lệ dài/ rộng là: $3.5 \leq W/H \leq 4.5$. Đối với loại xe có hai hàng ký tự thì tỷ lệ đó là: $0.8 \leq W/H \leq 1.4$. Từ các đặc tính này, ta có thể xác định được các vùng con thỏa mãn các tiêu chí về ngưỡng tỷ lệ dài/rộng. Và chỉ những vùng con thỏa mãn thì khả năng chứa biển số xe là cao

- Số lượng ký tự trong biển số xe. Mỗi ký tự thường có tỷ lệ kích thước về chiều rộng, chiều cao tương ứng với chiều dài và rộng của biển số xe. Ví dụ, chiều cao của mỗi ký tự luôn nhỏ hơn 85% chiều cao của biển số xe và luôn lớn hơn 33% chiều cao của biển xe. Còn chiều rộng của ký tự không lớn hơn 20% chiều dài của biển số xe. Mỗi ký tự của biển số xe được xem như là một vùng liên thông con. Do đó, chúng ta có thể đếm vùng liên thông con thỏa mãn tính chất đó là ký tự. Chú ý số ký tự trên biển số xe là từ 6 đến 10 ký tự. Ở nước ta chỉ có số ký tự trên mỗi biển số xe nằm trong khoảng 6 đến 8 ký tự. Vậy ta có thể dùng ngưỡng [6.8] để nhận dạng vùng biển số xe.

Từ những nhận xét trên, chúng ta có thể đưa ra giải pháp cho bài toán nhận dạng: sử dụng phương pháp phát hiện biên và biến đổi Hough. Sau đó, sử dụng hai tính chất trên biển số xe để xác định chính xác vùng con chứa biển số xe. Khi đã xác định chính xác vùng con chứa biển số xe thì tiến hành nhận dạng các ký tự.

Để giải quyết bài toán nhận dạng biển số xe, trong báo cáo em xin trình bày 3 bước như sau:

- Bước 1: Ảnh vào ảnh mức xám $I(x,y)$ thực hiện theo phương pháp dò biên và biến đổi Hough để tìm ra các vùng con có khả năng chứa biển số xe. Gọi tập con này là I_c .

- Bước 2: Xác định chính xác vùng con nào chứa biển số xe bằng hai thao tác được miêu tả ở trên đó là tiêu chí tỷ lệ chiều dài với chiều rộng và số ký tự trong biển số xe. Kết quả của bước 2 là cho ra một tập ảnh con chứa biển số xe. Gọi tập con này là I'_c .

- Bước 3: Giải quyết bài toán nhận dạng ký tự cho tập I'_c . Bằng cách áp dụng phương pháp và kỹ thuật nhận dạng ký tự

Qua ba bước như trên ta có thể nhận dạng được biển số xe I'_c . Trong bước 3: nhận dạng ký tự em sử dụng phương pháp mạng nơron truyền ngược cho việc nhận dạng ký tự.

Trong phần tiếp theo đó là chi tiết từng bước xử lý bài toán nhận dạng biên số xe, và một số khái niệm cơ bản quen thuộc mà có liên quan đến nhận dạng biên số xe.

Chương 2:

PHÁT HIỆN VÙNG CHỨA BIẾN SỐ XE

2.1 Một số khái niệm cơ bản:

2.1.1 Tổng quan về ảnh

a. Ảnh và điểm ảnh:

Ảnh là mảng số thực hai chiều $\mathbf{I}_{m,n}$, có kích thước (MxN), trong đó mỗi giá trị $\mathbf{I}_{m,n}$ (tại một điểm ảnh), biểu thị mức xám của ảnh tại vị trí $\mathbf{I}_{m,n}$ tương ứng

Một ảnh là ảnh nhị phân nếu giá trị $\mathbf{I}_{m,n}$ bằng 0 hoặc 1.

b. Mức xám:

Mức xám là kết quả sự mã hóa tương ứng một cường độ sáng của mỗi điểm ảnh với một giá trị số- kết quả của quá trình lượng hóa. Cách mã hóa kinh điển thường dùng 16, 32, 64. Mã hóa 256 mức là phổ dụng nhất do lý do kỹ thuật. Vì $2^8 = 256$, nên với 256 mức, mỗi pixel được mã hóa 8bit.

c. Đối tượng ảnh:

Trong phần này ta chỉ xét với ảnh nhị phân, vì mọi ảnh nhị phân đều có thể đưa về ảnh nhị phân bằng các kỹ thuật phân ngưỡng. Ta ký hiệu E là tập các điểm vùng (điểm đen) và \bar{E} là tập các điểm nền (điểm trắng). Hai điểm \mathbf{I}_s và $\mathbf{I}_e \in \mathbf{I}_{m,n}$ nằm trong E (hoặc \bar{E}) được gọi là 4 liên thông (8 liên thông) nếu tồn tại một dãy các điểm gọi là đường đi:

- $\mathbf{I}_{j_0} = \mathbf{I}_s$ và $\mathbf{I}_{j_n} = \mathbf{I}_e$
- $\mathbf{I}_{j_0} \rightarrow \mathbf{I}_{j_1} \rightarrow \dots \rightarrow \mathbf{I}_{j_n}$ mà $\mathbf{I}_{j_k} \in E$ với mọi $k=0,1,\dots,n$
- \mathbf{I}_{j_k} là 4 láng giềng (8 láng giềng) của $\mathbf{I}_{j_{k-1}}$ với mọi $k=1,2,\dots,n$

d. 4- Láng giềng và 8- láng giềng:

Nếu $\mathbf{I}_{n,n}$ là một điểm ảnh, thì 4 láng giềng của nó là các điểm ở ngay bên trên, dưới, phải, và trái. Ta ký hiệu N_4 là tập 4 láng giềng của điểm $\mathbf{I}_{n,n}$.

$$\Rightarrow N_4 = \{\mathbf{I}_{n-1,n}, \mathbf{I}_{n+1,n}, \mathbf{I}_{n,n-1}, \mathbf{I}_{n,n+1}\}$$

Tương tự ta có tập 8- láng giềng N_8

$$\Rightarrow N_8 = N_4 \cup \langle n+1, n+1 \rangle, \langle n-1, n-1 \rangle, \langle n+1, n-1 \rangle, \langle n-1, n+1 \rangle$$

e. Chu tuyến của ảnh:

Định nghĩa chu tuyến:

Chu tuyến của một đối tượng ảnh $\langle m, n \rangle$ là dãy các điểm của đối tượng: $p_0 \rightarrow p_1 \rightarrow \dots \rightarrow p_n$. Sao cho p_{i-1}, p_{i+1} là 8 láng giềng của p_i , $\exists p' \notin I$ và p' là 4 láng giềng của p_i , và $p_0 \equiv p_n$. Khi đó ta gọi n là độ dài hay chu vi của chu tuyến.

Chu tuyến đối ngẫu:

Hai chu tuyến $C = \langle P_1, P_2, \dots, P_n \rangle$ và $C' = \langle Q_1, Q_2, \dots, Q_n \rangle$ được gọi là hai chu tuyến đối ngẫu của nhau nếu và chỉ nếu:

- $\forall i \exists j$ sao cho P_i và Q_j là 8 láng giềng của nhau
- Các điểm P_i là ảnh thì Q_j là nền và ngược lại.

Chu tuyến trong:

Chu tuyến C được gọi là chu tuyến trong nếu và chỉ nếu:

- Chu tuyến đối ngẫu C' của nó là chu tuyến của các điểm nền.
- Độ dài của chu tuyến C' nhỏ hơn độ dài của chu tuyến C

Chu tuyến ngoài:

Chu tuyến C được gọi là chu tuyến ngoài nếu và chỉ nếu:

- Chu tuyến đối ngẫu C' của C là chu tuyến các điểm nền
- Độ dài của chu tuyến C' lớn hơn độ dài của chu tuyến C

Từ định nghĩa, ta thấy chu tuyến ngoài của một đối tượng là một đa giác có độ dày bằng một bao quanh đối tượng.

2.1.2 Phương pháp tách dò ngưỡng tự động

h_g : là tổng số mức xám g

$$\text{Gọi: } t_g = \sum_{i \leq g} h_i$$

Trong đó: P – Số điểm ảnh được xét = $m \cdot n$

G – Số mức xám được xét

Gọi $m_g = \frac{\left[\sum_{i=0}^g ih \right]}{t} = \sum_{i=0}^g iP$ là giá trị trung bình cấp xám g

$$f_g = \frac{t}{P-t} * h_g = m_{G-1} - 1$$

$$\theta = \arg \max f_g \quad \text{với } 0 < g < G - 1$$

Vậy suy ra θ là ngưỡng của ảnh

2.2 Biên và các phương pháp phát hiện biên.

**) Khái niệm về biên:*

Biên là một vấn đề 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 có sự thay đổi đột ngột và mức xám hay biên là điểm có cấp xám có giá trị khác hẳn các điểm xung quanh. Tập hợp các điểm biên tạo thành biên hay đường bao của ảnh

**) Các phương pháp phát hiện biên:*

**) Phương pháp tiếp cận theo kiểu cổ điển*

Đây là phương pháp 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.

2.2.1 Phương pháp gradient

Dựa vào cực đại hóa của đạo hàm. Theo định nghĩa, gradient là một vecto có các thành phần biểu thị tốc độ thay đổi giá trị của điểm ảnh theo 2 hướng x và y. Các thành phần của Gradient được tính bởi:

$$\frac{\partial f(x, y)}{\partial x} = f_x \approx \frac{f(x + dx, y) - f(x, y)}{dx}$$

$$\frac{\partial f(x, y)}{\partial y} = f_y \approx \frac{f(x, y + dy) - f(x, y)}{dy}$$

Đổi sang tọa độ cực

$$x = r \cos \theta \quad y = r \sin \theta$$

Suy ra:

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \cos\theta + \frac{\partial f}{\partial y} \sin\theta$$

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} (-r \sin\theta) + \frac{\partial f}{\partial y} \cos\theta$$

Với dx là khoảng cách giữa các điểm theo hướng x (khoảng cách tính bằng số điểm) và tương tự với dy . Trên thực tế người ta hay dùng với $dx = dy = 1$

Với một ảnh liên tục $f(x, y)$, các đạo hàm riêng của nó cho phép xác định vị trí cục bộ theo hướng của biên. Thực vậy, gradient của một ảnh liên tục, được biểu diễn bởi một hàm $f(x, y)$, dọc theo r với góc θ , được định nghĩa bởi:

$$\frac{df}{dr} = \frac{\partial f}{\partial x} \frac{dx}{dr} + \frac{\partial f}{\partial y} \frac{dy}{dr} = f_x \cos\theta + f_y \sin\theta$$

Chú ý: khi ta nói lấy đạo hàm của ảnh nhưng thực ra chỉ là mô phỏng và xấp xỉ đạo hàm bằng các kỹ thuật nhân chập (phép cuộn). Do ảnh số là tín hiệu rời rạc nên đạo hàm không tồn tại

Kỹ thuật Gradient sử dụng một cặp mặt nạ H_1 và H_2 trực giao (theo 2 hướng vuông góc). Nếu định nghĩa g_1, g_2 là gradient tương ứng theo 2 hướng x và y , thì biên độ của gradient, ký hiệu là g tại điểm (m, n) được tính theo công thức:

$$A_0 = g(m, n) = \sqrt{g_1^2(m, n) + g_2^2(m, n)} \quad (1)$$

$$\theta(m, n) = \tan^{-1}(g_2(m, n) / g_1(m, n)) \quad (2)$$

Chú ý: để giảm tính toán, công thức (1) được tính gần đúng bởi:

$$A_0 = |g_1(m, n)| + |g_2(m, n)|$$

Các toán tử đạo hàm được áp dụng là khá nhiều, ở đây, ta chỉ xét một số toán tử tiêu biểu: toán tử Robert, Solbel...

***) Kỹ thuật Robert**

Với mỗi điểm ảnh $I(x, y)$ của I , đạo hàm theo x , theo y được ký hiệu tương ứng bởi g_x, g_y được tính:

$$\begin{cases} g_x = I(x+1, y) - I(x, y) \\ g_y = I(x, y+1) - I(x, y) \end{cases}$$

Điều này tương đương với việc chập ảnh với 2 mặt nạ H_1 và H_2 :

$$H_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad H_2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Quá trình tính toán được thực hiện qua các bước sau:

- Bước 1: Tính $I_x = I \otimes H_x$ và $I_y = I \otimes H_y$
- Bước 2: Tính $\sqrt{I_x^2 + I_y^2}$

Từ ma trận $\sqrt{I_x^2 + I_y^2}$ chọn ra các điểm cao thứ 2, hoặc thứ 3.. chiếm đa số

***)Kỹ thuật PreWitt:**

Kỹ thuật này sử dụng hai mặt nạ H_1 và H_2 :

$$H_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad H_2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Quá trình tính toán được thực hiện qua 2 bước:

- Bước 1: Tính $I \otimes H_x$ và $I \otimes H_y$
- Bước 2: Tính $|I \otimes H_x| + |I \otimes H_y|$

***)Kỹ thuật Sobel:**

Tương tự như kỹ thuật PreWitt, kỹ thuật Sobel sử dụng 2 ma trận mặt nạ nhân chập là:

$$H_1 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad H_2 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

2.2.2. Kỹ thuật Laplace:

Các phương pháp đánh giá Gradient ở trên làm việc khá tốt khi mà độ sáng thay đổi rõ nét. Khi mức xám (giá trị tại một điểm của ảnh) thay đổi chậm, miền chuyển tiếp trải rộng, thì ta có phương pháp Laplace (đạo hàm bậc hai) có hiệu quả hơn. Toán tử Laplace được định nghĩa như sau:

$$\nabla^2 f = \frac{\partial^2 f}{dx^2} + \frac{\partial^2 f}{dy^2}$$

Vậy suy ra ta có:

$$\begin{aligned}\frac{\nabla^2 f}{\nabla x^2} &= \frac{\nabla}{\nabla x} \left(\frac{\nabla f}{\nabla x} \right) \approx \frac{\nabla [f(x+1, y) - f(x, y)]}{\nabla x} \approx [f(x+1, y) - f(x, y)] - [f(x, y) - f(x-1, y)] \\ &= f(x+1, y) - 2f(x, y) + f(x-1, y)\end{aligned}$$

Tương tự ta có:

$$\frac{\nabla^2 f}{\nabla y^2} = f(x, y+1) - 2f(x, y) + f(x, y-1)$$

Toán tử Laplace dùng nhiều kiểu mặt nạ khác nhau để xấp xỉ rời rạc đạo hàm bậc 2. Dưới đây là 3 kiểu mặt nạ hay dùng:

$$H_1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad H_2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad H_3 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

Quá trình tính toán được thực hiện qua các bước sau:

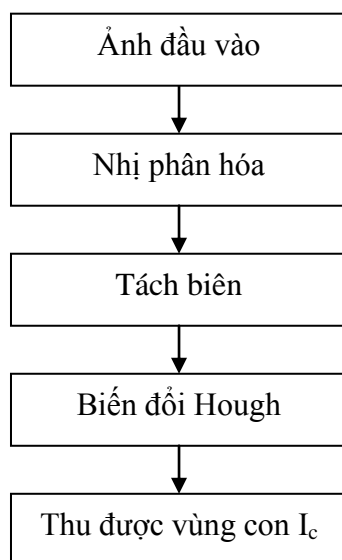
- Bước 1: $H \otimes I$
- Bước 2: $[H \otimes I(x, y)]$
- Bước 3: Tách ngưỡng

Ý nghĩa hình học:

$$\begin{aligned}H \otimes I(x, y) &= 4I(x+1, y+1) - I(x+1, y) - I(x, y+1) - I(x+2, y+1) - I(x+1, y+2) \\ &= I(x+1, y+1) - I(x+1, y) \\ &\quad + I(x+1, y+1) - I(x, y+1) \\ &\quad + I(x+1, y+1) - I(x+2, y+1) \\ &\quad + I(x+1, y+1) - I(x+1, y+2) \\ &= \frac{\partial I(x+1, y)}{\partial y} + \frac{\partial I(x, y+1)}{\partial x} - \frac{\partial I(x+1, y+1)}{\partial y} - \frac{\partial I(x+1, y+1)}{\partial x} \\ &= -\frac{\partial^2 I(x+1, y+1)}{\partial x^2} - \frac{\partial^2 I(x+1, y+1)}{\partial y^2} = -\nabla^2 I\end{aligned}$$

2.3 Phát hiện vùng chứa biển số xe

Sơ đồ các bước được mô tả trong hình dưới



Hình 2.1: Sơ đồ giải quyết

Ảnh đầu vào: là một ảnh có 256 mức xám, được nhị phân hóa thành ảnh nhị phân. Mục đích của giai đoạn nhị phân hóa ảnh là nhằm làm nổi bật vùng biển số xe. Khi ta tách biên, vùng bao của biển số xe sẽ hiện lên rõ ràng. Sau đó dùng phương pháp phát hiện biên để có được biên dọc vào ngang của ảnh. Kết quả của công đoạn này, ảnh thu được là ảnh nhị phân chỉ chứa các cạnh dọc và ngang. Thực hiện biến đổi Hough cho các đoạn biên vừa lấy được và xác định các đoạn thẳng đi qua tập các điểm biên của mỗi biên, kết quả sẽ là các đoạn thẳng ngang và dọc. Giao của những đoạn thẳng này sẽ cho ra vùng con I_c

2.3.1. Nhị phân hóa ảnh

Ảnh ban đầu được sử dụng là ảnh 256 mức xám. Việc sử dụng ảnh 256 mức xám không làm giảm đi tính đa năng của ứng dụng. Trên thực tế, ảnh 256 mức xám vẫn được sử dụng nhiều, và nhiều thiết bị ghi hình cũng có khả năng tự chuyển ảnh màu thành ảnh 256 mức xám. Tuy nhiên, nếu để ảnh 256 mức xám thì việc phát hiện biên không hiệu quả, vì sự thay đổi liên tục của các mức xám làm cho việc xác định biên không phải dễ dàng, và việc tìm ra các vùng liên tục của biên khá hạn chế. Vì vậy, chúng ta thực hiện chuyển ảnh về dạng nhị phân để thực hiện việc lấy biên nhanh hơn.

void Binarize// Nhị phân hóa ảnh

{

// Ảnh đầu vào: ảnh 256 mức xám

// Đầu ra là ảnh nhị phân

P: là tổng số điểm ảnh được xét (m,n)

g(i,j) tương ứng là mức xám của điểm ảnh I(i,j)

θ : là ngưỡng của ảnh được xác định theo phương pháp ở trên.

for(int i= 0; i< m; i++)

for(int j= 0; j< n; j++)

if(g(i, j)<= θ)

{

I_c(I, j)= 0

}else I_c(I, j)= 1

}

Vậy ta thu được ảnh nhị phân I_c, ảnh nhị phân thu được vẫn đảm bảo tách biệt giữa vùng chứa biển số xe với vùng xung quanh. Đồng thời loại bỏ những vùng đồng nhất và ít biến thiên.

2.3.2 Tách biên:

Vì biển số xe có viền bao quanh, nên chúng ta cần làm nổi bật đường biên (boundary). Các đường biên có thể được xem là các cạnh dọc và ngang. Mục đích của giai đoạn này là tách ra các cạnh dọc và ngang để tìm ra vùng con chứa biển số xe nhờ tính giao điểm của các cạnh dọc và ngang. Ở đây, ảnh đầu vào là ảnh nhị phân, nên thích hợp với phương pháp đạo hàm bậc nhất. Dùng hai ma trận Sobel theo hai hướng x(dọc) và y(ngang) để tách các cạnh của ảnh

$$H_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Void BoundaryDetach()

{// Tách biên của ảnh

// Ảnh đầu vào: ảnh 256 mức xám

// Đầu ra là I'(i, j)

P: là tổng số điểm ảnh được xét (m,n)

g(j,j) tương ứng là mức xám của điểm ảnh I(i,j)

θ : là ngưỡng của ảnh được xác định theo phương pháp ở trên.

// Trước hết tính hai ma trận ảnh theo trục dọc x và ngang y

$I_x = H_1 * I, I_y = H_2 * I$

for(int i= 0; i< m; i++)

for(int j= 0; j< n; j++)

I'(I,j)= I_x(I, j)+ I_y(I, j);

}

Kết quả thu được , một ảnh cạnh dọc I_x và một cạnh ngang I_y. Có thể xem ảnh ở dưới với phương pháp Sobel.

2.3.3 Biến đổi HOUGH

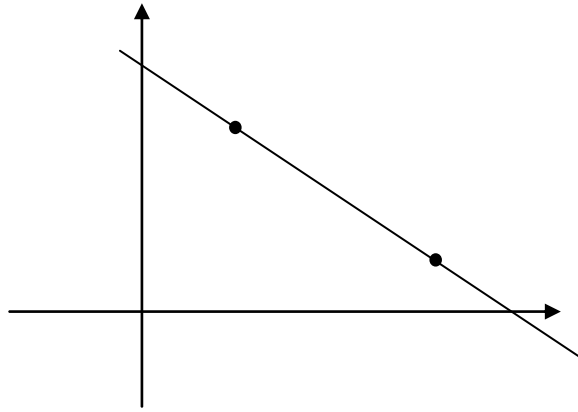
Biến đổi Hough là phương pháp dùng để xác định đường thẳng (đường tròn elip) gần đúng đi qua một tập hợp điểm.

Với (x,y) là một điểm $y = mx + c \Rightarrow c = -mx + y$

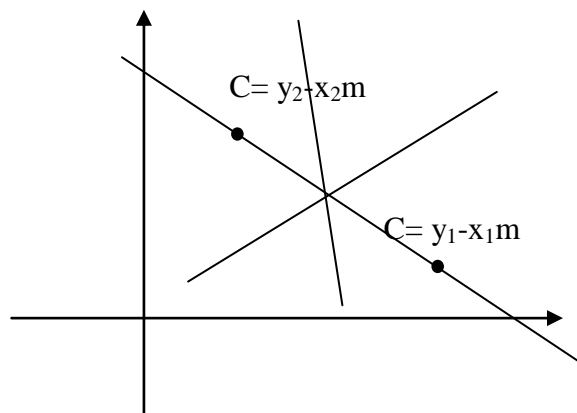
Như vậy nếu có N điểm nằm trên 1 đường thẳng

$$\begin{cases} mx_i + c = y_i \\ i = 1, N \end{cases} \Rightarrow \begin{cases} c = y_i - x_i m \\ i = 1, N \end{cases}$$

Thay vì tìm N điểm trên đường thẳng, người ta xét tất cả các điểm, xem điểm nào có nhiều đường thẳng đi qua nhất



Hình 2.2 Trục tọa độ đề các đi qua 2 điểm



Hình 2.3. Trục tọa độ đề các

Thực chất biến đổi Hough là biến điểm thành đường thẳng

$$\begin{cases} a \sin(\theta) = 0 & \forall m, c \\ \forall x_i, y_i \Rightarrow \forall m, c & \text{thi} \\ c \approx y_i - x_i m & \text{thi} \end{cases}$$

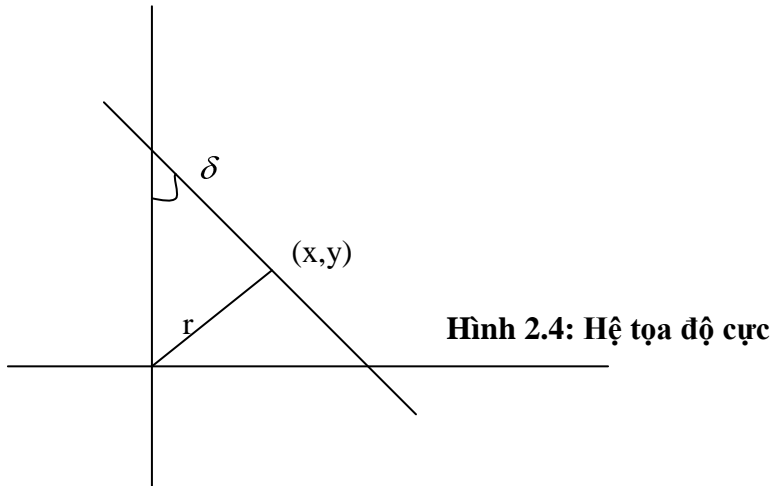
Sau đó đếm trên ma trận

Hạn chế: hệ số $0 < m < \infty$

$$|m| > 1 \text{ thì lưu } a \left[\frac{1}{m} \right]$$

$$|m| \leq 1 \text{ thì lưu } a$$

Biến đổi Hough theo tọa độ cực

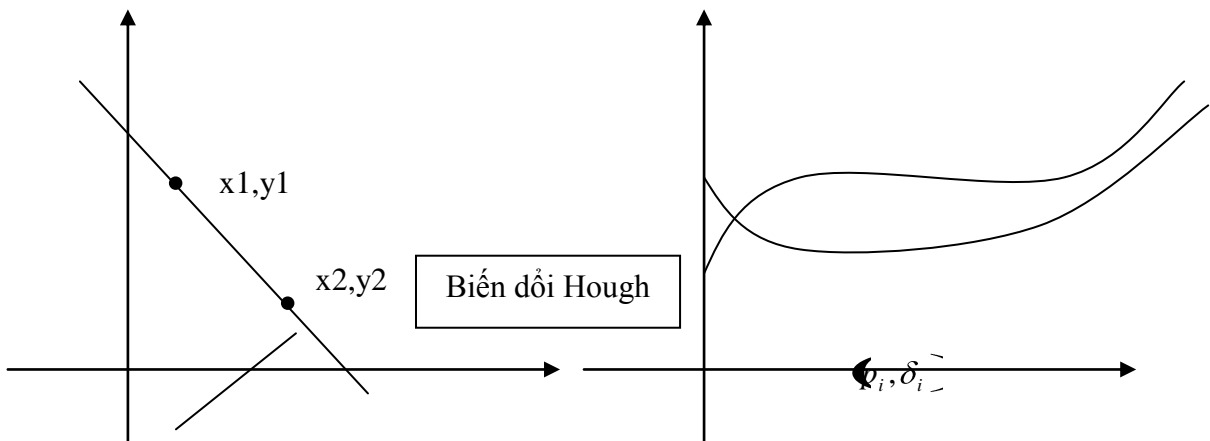


Hình 2.4: Hệ tọa độ cực

Các điểm trên đường thẳng có tọa độ cực t/m $r = x \cdot \cos \delta + y \cdot \sin \delta$ với $0 \leq \delta \leq 360$ và $r \leq \frac{\sqrt{M^2 + N^2}}{2}$ với M và N là chiều cao và chiều rộng của ảnh

Lấy tại tâm ảnh

$$\begin{cases} \forall r, \delta & a[\delta] = 0 \\ \forall r, \delta \\ \forall (x, y) \text{ neu } & |r - x \cos \delta - y \sin \delta| \leq \epsilon \\ \text{thi } & a[\delta] = + \end{cases}$$



Hình 2.5. Đường thẳng Hough trong tọa độ cực

Biến đổi Hough ánh xạ N điểm này thành N đường sin trong tọa độ cực mà các đường này đều đi qua điểm (r_i, δ_i) . Giao điểm (r_i, δ_i) của N đường sin sẽ xác định một đường thẳng trong hệ tọa độ các. Như vậy, những đường thẳng đi

qua điểm (x_i, y_i) sẽ cho duy nhất một cặp (r_i, δ_i) và có bao nhiêu đường qua (x_i, y_i) sẽ có bấy nhiêu cặp giá trị (r_i, δ_i) . Mục đích là tìm ra cặp (r, δ) sao cho số đường hình sin đi qua nhiều nhất, và cặp đó chính là cặp tham số cho đường thẳng

2.3.4 Trích chọn đoạn thẳng và tính giao điểm

Sau khi xây dựng các đường thẳng Hough, chúng ta thu được hai tập đường thẳng: tập đường thẳng dọc và tập đường ngang các ảnh nhị phân cạnh dọc và ngang. Tiếp theo, chúng ta xác định giao điểm của các đường này để tạo thành các vùng con là các tứ giác có khả năng chứa biển số xe. Tuy nhiên, lượng vùng con có được là rất nhiều. Do vậy, chúng ta thay vì tính giao điểm của các đường thẳng, mà chúng ta sẽ thực hiện việc tính giao điểm của các đoạn thẳng. Các đoạn thẳng chính là các đoạn đi qua tập hợp điểm được xác định thông qua đường thẳng Hough. Việc xác định các đoạn thẳng này là đơn giản thông qua hai đầu mút của đoạn thẳng. Một cách đơn giản, ta có thể xem đầu mút trên (hoặc bên trái) là điểm trong tập hợp điểm nằm trên đường thẳng Hough có tổng tọa độ theo trục x và y là nhỏ nhất. Điểm đầu mút dưới (hoặc bên phải) là điểm nằm trong tập hợp điểm nằm trên đường thẳng Hough có tổng tọa độ theo x và y là lớn nhất. Để đảm bảo rằng các đường thẳng dọc, ngang (từ ảnh cạnh dọc, ngang) có thể cắt nhau như trong thực tế, chúng ta cần mở rộng các đoạn thẳng về hai hướng mỗi đoạn 5 điểm. Như vậy đoạn thẳng mà chúng ta sử dụng so với đoạn thẳng thực tế sẽ dài hơn nhiều hơn khoảng 10 điểm. Kết quả của phép biến đổi Hough và trích chọn đoạn thẳng được mô tả trong hình dưới

Việc tính giao điểm của các đoạn thẳng là khá đơn giản. Chúng ta chỉ cần tính giao điểm của các đường thẳng Hough và kiểm tra xem giao điểm đó có nằm trên đoạn thẳng được trích chọn ra hay không. Giao điểm của các đoạn thẳng sẽ là các vùng con I_c có khả năng chứa biển số xe.

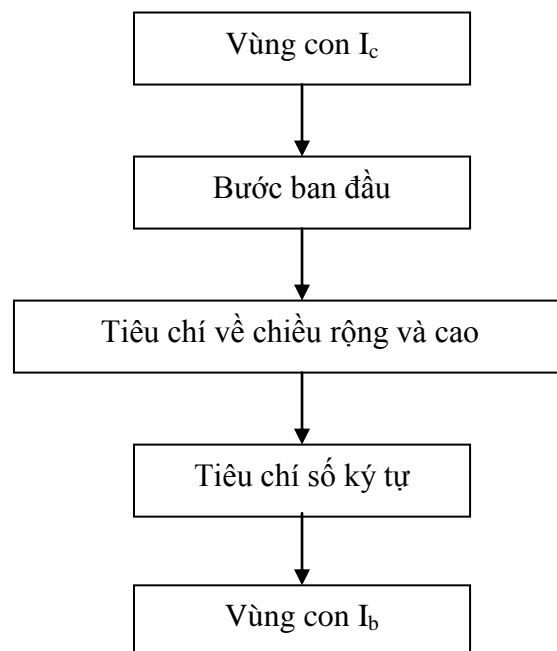
Kết quả của phần trên cho chúng ta một tập các vùng con I_c là các tứ giác. Đến đây, chúng ta có thể khẳng định bài toán trên chính là thực hiện theo phương pháp biến đổi Hough. Tuy vậy, điểm khác biệt (cũng là điểm tiến bộ)

của phương pháp trình bày trong tiểu luận này là: phương pháp biến đổi Hough chỉ dừng lại ở bước này, và sau đó thực hiện tiến hành nhận dạng các ký tự trong các vùng con ngay. Việc nhận dạng có thể có nhiều phương pháp khác nhau, nhưng với mỗi vùng con đang còn một lượng khá lớn. Vậy trong bài tiểu luận này em không chỉ dừng lại trong việc tìm ra các vùng con I_c , mà tìm tiếp những vùng con có xác suất chứa biển số xe, loại đi những vùng mà khả năng tồn tại của biển số xe là rất ít. Vậy có, tập các vùng con I_c được thu hẹp, làm cho không gian bài toán nhận dạng thu hẹp lại. Vì vậy, cách giải quyết này trở nên nhanh hơn, hiệu quả hơn cách tiếp cận biến đổi Hough, không phụ thuộc nhiều vào không gian ảnh đầu vào.

2.3. Xác định chính xác vùng chứa biển số xe

Kết quả của bài toán trên đưa ra tập các vùng con I_c có khả năng chứa biển số xe. Các vùng con này là các tứ giác. Tuy nhiên, số lượng các vùng con I_c là khác nhiều, chưa thể đảm bảo chính xác vùng nào chứa biển số xe để thực hiện việc cuối cùng là nhận dạng ảnh. Vì vậy, phải loại bỏ đi những vùng con trong I_c không có khả năng chứa biển số xe.

Sơ đồ thực hiện bài toán này:



Hình 2.6: Sơ đồ thực hiện bài toán này

2.3.1. Bước ban đầu:

Ta biết: Biển số xe trên thực tế có hình dạng là hình chữ nhật. Vì vậy, khi chụp ảnh của biển số xe sẽ có dạng tựa hình bình hành. Trường hợp tối ưu là ảnh hình chữ nhật. Do đó, trong quá trình xét duyệt các vùng con I_c , nếu hình nào không có hình dạng tựa hình bình hành thì có thể loại bỏ ngay mà không cần tính đến. Ta có thể xem hình tứ giác tựa hình bình hành có những đặc điểm sau:

- Các góc không có nhỏ. Có thể lấy ngưỡng là 45^0
- Hai góc đối không chênh lệch quá lớn. Lấy ngưỡng 30^0

Biển số xe phải có một diện tích nào đó, và đủ lớn để có thể nhận diện ra ký tự tồn tại trên đó. Vì vậy, những vùng con có diện tích nhỏ hơn một ngưỡng nào đó, thì loại bỏ ngay và chú ý kích thước chiều dài và rộng của vùng con I_{c_i} . Vậy thì chu vi của vùng con nhỏ hơn ngưỡng nào đó thì có thể loại bỏ ngay

Void Filter

```
{
    // Đầu vào là tập vùng con  $I_c$ 
    // Đầu ra: tập vùng con  $I'_c$ 
    //  $\theta$  là ngưỡng về chu vi
    // Perimeter( $I_{c_i}$ ) là chu vi của mỗi vùng con  $I_c$ 
    // N là số vùng con
    For(int i=0; i< N; i++){
        If(Perimeter( $I_{c_i}$ ) >  $\theta$ )
            // Caclulate goc A,B,C,D của mỗi vùng con  $I_{c_i}$ 
            If( A && B && C && D thỏa mãn ngưỡng)
                Copy( $I'_{c_i}$ ,  $I_{c_i}$ )
    }
}
```

Sau khi loại bỏ đi những vùng con theo hai tiêu chí trên, chúng ta thu được tập con I'_c . Vì thực tế biển số xe có hình chữ nhật, nên ta có thể dùng phép tịnh tiến, phép quay, phép tỷ lệ để đưa các vùng con I'_c thành các hình chữ nhật.

Lý do để đưa các hình tứ giác thành hình chữ nhật vì biên số xe có dạng hình chữ nhật, các ký tự nằm trong vùng biên số xe vuông góc với cạnh dài của hình chữ nhật. Khi thu được ảnh, có nhiều nguyên nhân làm cho ảnh biên số xe bị nghiêng. Kéo theo đó, các ký tự cũng bị nghiêng theo, làm cho việc nhận dạng trở nên không chính xác. Việc nắn tứ giác trở lại thành hình chữ nhật và cũng nắn các ký tự trở nên thành đứng. Và khi trở thành hình dạng chữ nhật, thì biên số xe mới thể hiện rõ tính tỷ lệ chiều dài/rộng. Và ta có các tiêu chí dưới đây.

2.4.2 Tiêu chí tỷ lệ chiều dài/rộng.

Với mỗi quốc gia, thì biên số xe có kích thước nhất định. Và thể hiện thông qua tỷ lệ giữa các cạnh. Ví dụ với biên số xe ở nước ta: với biên số có một hàng thì tỷ lệ nằm trong khoảng $3.5 \leq W/H \leq 4.5$ và với biên số xe có hai hàng thì tỷ lệ là $0.8 \leq W/H \leq 1.4$. Và kết quả của tiêu chí tỷ lệ chiều dài/rộng là thu được một tập con của I'_c chứa biên số xe. Vậy ta có giải thuật

Void RatioWH

```
{
    // Đầu vào là tập con  $I'_c$ 
    // Đầu ra là tập con  $I''_c$  của  $I'_c$ 
    // Gọi edge_ratio= tỷ lệ chiều dài/rộng
    //  $\theta$  là ngưỡng chiều dài/rộng [0.8,4.5]
    For(int i=0;i< N;i++)
    {
        Int m= edge_ratio( $I'_c$ )
        If(m  $\in$  [0.8,4.5] )
            Copy( $I''_c$ ,  $I'_c$ )
    }
}
```

Kết quả: tập các vùng con I''_c có khả năng chứa biên số xe. Với số vùng con nhỏ hơn hẳn số vùng con mà ta thu được ban đầu trong biến đổi Hough.

2.4.3 Tiêu chí số ký tự trong vùng biển số xe

Với mỗi nước thì số ký tự trong biển là khác nhau. Ở nước ta, số ký tự trong biển số xe thường là 6,7,8 tương ứng đối với các xe quân đội, xe máy cũ và xe ô tô, đối với xe máy bây giờ.

Mỗi ký tự có các đặc trưng sau:

- $0.33 \leq Height \leq 0.85$
- $Width \leq 0.22$

Từ đó ta có ngưỡng sử dụng là [6,8]

Void Character

```
{  
    // Đầu vào tập vùng con  $I_{c_i}''$   
    // Đầu ra tập vùng con  $I_{c_i}'''$   
    // N là tổng số vùng con của  $I_{c_i}''$   
    For(int i= 0; i< N; i++)  
    {  
        Với mỗi vùng con  $I_{c_i}''$   
        - Tìm vùng liên thông của mỗi  $I_{c_i}''$   
        - Lưu các thông số về chiều rộng, cao của mỗi vùng liên thông  
        If(thỏa mãn ngưỡng ) thì tiến hành nhận dạng  
    }  
}
```

Kết quả nếu tìm được biển số xe đầu tiên thỏa mãn, chúng ta có thể dừng thuật toán ngay và chuyển sang bước 3 là nhận dạng ký tự. Nếu tìm tất cả các biển số xe tồn tại trong ảnh, thì bắt buộc phải duyệt qua toàn bộ vùng ảnh.

Chương 3:

BÀI TOÁN NHẬN DẠNG KÝ TỰ

3.1 Tổng quan về nhận dạng

Nhận dạng là quá trình phân loại các đối tượng được biểu diễn theo một mô hình nào đó và gán cho chúng vào một lớp (gán cho đối tượng một tên gọi) dựa theo những quy luật và các mẫu chuẩn. Quá trình nhận dạng dựa vào những mẫu học biết trước gọi là nhận dạng có thầy hay *học có thầy* (supervised learning); trong trường hợp ngược lại gọi là *học không có thầy* (non supervised learning). Chúng ta sẽ lần lượt giới thiệu các khái niệm này.

3.1.1 Không gian biểu diễn đối tượng, không gian diễn dịch

**)Không gian biểu diễn đối tượng*

Các đối tượng khi quan sát hay thu thập được, thường được biểu diễn bởi tập các đặc trưng hay đặc tính. Người ta thường phân các đặc trưng theo các loại như: đặc trưng tô pô, đặc trưng hình học và đặc trưng chức năng. Việc biểu diễn ảnh theo đặc trưng nào là phụ thuộc vào ứng dụng tiếp theo.

Ở đây ta đưa ra một cách hình thức việc biểu diễn các đối tượng. Giả sử đối tượng X (ảnh, chữ viết, dấu vân tay, v...,v) được biểu diễn bởi n thành phần (n đặc trưng): $X = \{x_1, x_2, \dots, x_n\}$; mỗi x_i biểu diễn một đặc tính. Không gian biểu diễn đối tượng thường gọi tắt là không gian đối tượng X được định nghĩa:

$$X = \{X_1, X_2, \dots, X_m\}$$

trong đó mỗi X_i biểu diễn một đối tượng. Không gian này có thể là vô hạn. Để tiện xem xét chúng ta chỉ xét tập X là hữu hạn.

**)Không gian diễn dịch*

Không gian diễn dịch là tập các tên gọi của đối tượng. Kết thúc quá trình nhận dạng ta xác định được tên gọi cho các đối tượng trong tập không gian đối tượng hay nói là đã nhận dạng được đối tượng. Một cách hình thức gọi Ω là tập tên đối tượng:

$$\Omega = \{w_1, w_2, \dots, w_k\} \text{ với } w_i, i = 1, 2, \dots, k \text{ là tên các đối tượng}$$

Quá trình nhận dạng đối tượng f là một ánh xạ $f: X \rightarrow \Omega$ với f là tập các quy luật để xác định một phần tử trong X ứng với một phần tử trong Ω . Nếu tập các quy luật và tập tên các đối tượng là biết trước như trong nhận dạng chữ viết (có 26 lớp từ A đến Z), người ta gọi là nhận dạng có thày. Trường hợp thứ hai là nhận dạng không có thày. Đương nhiên trong trường hợp này việc nhận dạng có khó khăn hơn.

3.1.2 Mô hình và bản chất của quá trình nhận dạng

3.1.2.1 Mô hình

Việc chọn lựa một quá trình nhận dạng có liên quan mật thiết đến kiểu mô tả mà người ta sử dụng để đặc tả đối tượng. Trong nhận dạng, người ta phân chia làm 2 họ lớn:

- Họ mô tả theo tham số
- Họ mô tả theo cấu trúc.

Cách mô tả được lựa chọn sẽ xác định *mô hình* của đối tượng. Như vậy, chúng ta sẽ có 2 loại mô hình: *mô hình theo tham số* và *mô hình cấu trúc*.

- **Mô hình tham số:** sử dụng một véctơ để đặc tả đối tượng. Mỗi phần tử của véctơ mô tả một đặc tính của đối tượng. Thí dụ như trong các đặc trưng chức năng, người ta sử dụng các hàm cơ sở trực giao để biểu diễn. Và như vậy ảnh sẽ được biểu diễn bởi một chuỗi các hàm trực giao. Giả sử C là đường bao của ảnh và $C(i,j)$ là điểm thứ i trên đường bao, $i = 1, 2, \dots, N$ (đường bao gồm N điểm).

Giả sử tiếp :

$$x_0 = \frac{1}{N} \sum_{i=1}^N x_i$$

$$y_0 = \frac{1}{N} \sum_{i=1}^N y_i$$

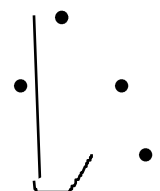
là tọa độ tâm điểm. Như vậy, moment trung tâm bậc p, q của đường bao là:

$$\mu_{pq} = \frac{1}{N} \sum_{i=1}^N (x_i - x_0)^p (y_i - y_0)^q \quad (7.1)$$

Véc tơ tham số trong trường hợp này chính là các moment μ_{ij} với $i=1, 2, \dots, p$ và $j=1, 2, \dots, q$. Còn trong số các đặc trưng hình học, người ta hay sử dụng chu tuyến, đường bao, diện tích và tỉ lệ $T = 4\pi S/p^2$, với S là diện tích, p là chu tuyến.

Việc lựa chọn phương pháp biểu diễn sẽ làm đơn giản cách xây dựng. Tuy nhiên, việc lựa chọn đặc trưng nào là hoàn toàn phụ thuộc vào ứng dụng. Thí dụ, trong nhận dạng chữ (sẽ trình bày sau), các tham số là các dấu hiệu:

- số điểm chạc ba, chạc tư,
- số điểm chu trình,
- số điểm ngoặt,
- số điểm kết thúc,

chẳng hạn với chữ t  có 4 điểm kết thúc, 1 điểm chạc tư,...

• **Mô hình cấu trúc:** Cách tiếp cận của mô hình này dựa vào việc mô tả đối tượng nhờ một số khái niệm biểu thị các đối tượng cơ sở trong ngôn ngữ tự nhiên. Để mô tả đối tượng, người ta dùng một số dạng nguyên thủy như đoạn thẳng, cung, v, ..., v. Chẳng hạn một hình chữ nhật được định nghĩa gồm 4 đoạn thẳng vuông góc với nhau từng đôi một. Trong mô hình này người ta sử dụng một bộ kí hiệu kết thúc V_t , một bộ kí hiệu không kết thúc gọi là V_n . Ngoài ra có dùng một tập các luật sản xuất để mô tả cách xây dựng các đối tượng phù hợp dựa trên các đối tượng đơn giản hơn hoặc đối tượng nguyên thủy (tập V_t). Trong cách tiếp cận này, ta chấp nhận một khẳng định là: cấu trúc một dạng là kết quả của việc áp dụng luật sản xuất theo theo những nguyên tắc xác định bắt đầu từ một dạng gốc ban đầu. Một cách hình thức, ta có thể coi mô hình này tương đương một văn phạm $G = (V_t, V_n, P, S)$ với:

- V_t là bộ ký hiệu kết thúc,
- V_n là bộ ký hiệu không kết thúc,
- P là luật sản xuất,

- S là dạng (ký hiệu bắt đầu).

3.1.2.2 **Bản chất của quá trình nhận dạng**

Quá trình nhận dạng gồm 3 giai đoạn chính:

- Lựa chọn mô hình biểu diễn đối tượng.
- Lựa chọn luật ra quyết định (phương pháp nhận dạng) và suy diễn quá trình học.
- Học nhận dạng.

Khi mô hình biểu diễn đối tượng đã được xác định, có thể là định lượng (mô hình tham số) hay định tính (mô hình cấu trúc), quá trình nhận dạng chuyển sang giai đoạn học. Học là giai đoạn rất quan trọng. Thao tác học nhằm cải thiện, điều chỉnh việc phân hoạch tập đối tượng thành các lớp.

Việc nhận dạng chính là tìm ra quy luật và các thuật toán để có thể gán đối tượng vào một lớp hay nói một cách khác gán cho đối tượng một tên.

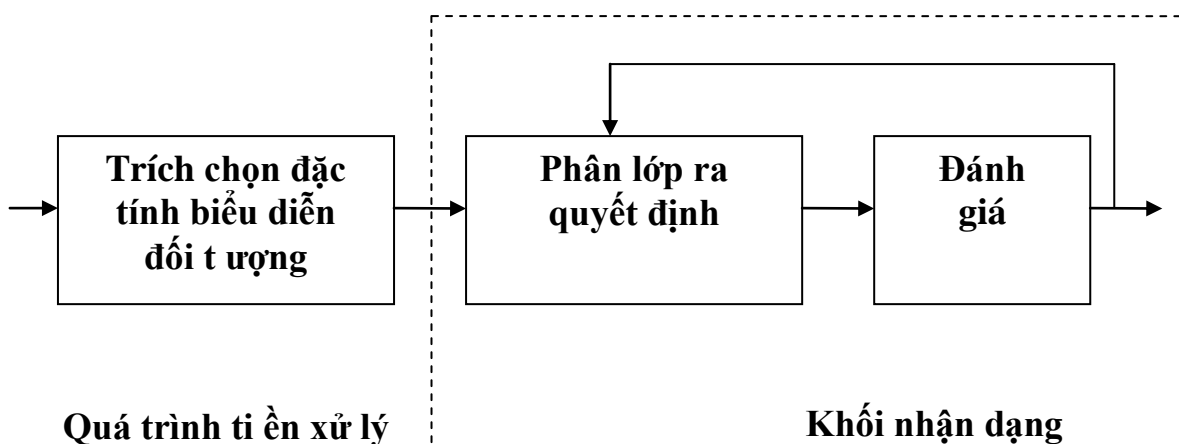
****) Học có thầy (supervised learning)***

Kỹ thuật phân loại nhờ kiến thức biết trước gọi là học có thầy. Đặc điểm cơ bản của kỹ thuật này là người ta có một thư viện các mẫu chuẩn. Mẫu cần nhận dạng sẽ được đem sánh với mẫu chuẩn để xem nó thuộc loại nào. Thí dụ như trong một ảnh viễn thám, người ta muốn phân biệt một cánh đồng lúa, một cánh rừng hay một vùng đất hoang mà đã có các miêu tả về các đối tượng đó. Vấn đề chủ yếu là thiết kế một hệ thống để có thể đối sánh đối tượng trong ảnh với mẫu chuẩn và quyết định gán cho chúng vào một lớp. Việc đối sánh nhờ vào các thủ tục ra quyết định dựa trên một công cụ gọi là *hàm phân lớp* hay *hàm ra quyết định*. Hàm này sẽ được đề cập trong phần sau.

****) Học không có thầy (unsupervised learning)***

Kỹ thuật học này phải tự định ra các lớp khác nhau và xác định các tham số đặc trưng cho từng lớp. Học không có thầy đương nhiên là khó khăn hơn. Một mặt, do số lớp không được biết trước, mặt khác những đặc trưng của các lớp cũng không biết trước. Kỹ thuật này nhằm tiến hành mọi cách gộp nhóm có thể và chọn lựa cách tốt nhất. Bắt đầu từ tập dữ liệu, nhiều thủ tục xử lý khác nhau nhằm phân lớp và nâng cấp dần để đạt được một phương án phân loại.

Nhìn chung, dù là mô hình nào và kỹ thuật nhận dạng ra sao, một hệ thống nhận dạng có thể tóm tắt theo sơ đồ sau:



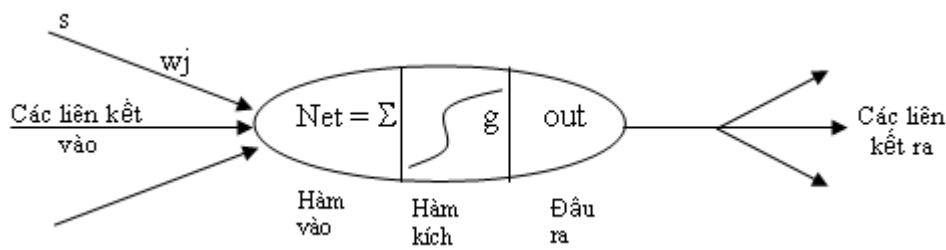
Hình 3.1: Sơ đồ tổng quát một hệ nhận dạng.

3.2 Mô hình mạng nơron nhân tạo

Mạng nơron nhân tạo (Artificial Neural Network) bao gồm các nút (đơn vị xử lý) được nối với nhau bởi các liên kết nơron. Mỗi liên kết kèm theo một trọng số nào đó, đặc trưng cho đặc tính kích hoạt giữa các nơron. Có thể xem trọng số là phương tiện để lưu giữ thông tin dài hạn trong mạng và nhiệm vụ của quá trình huấn luyện (học) mạng là cập nhật các trọng số khi có thêm các thông tin về các mẫu học, hay nói cách khác, các trọng số được điều chỉnh sao cho đúng.

Trong mạng, một số nơron được nối với môi trường bên ngoài như các đầu ra, đầu vào

3.2.1 Mô hình nơron nhân tạo



Hình 3.2: Mô hình nơron nhân tạo

Mỗi nơon được nối với các nơon khác và nhận được các tín hiệu s_j từ chúng với các trọng số w_j . Tổng các thông tin vào có trọng số là:

$$\text{Net} = \sum w_j s_j$$

Người ta gọi đây là thành phần tuyến tính của nơon. Hàm kích hoạt g (còn gọi là hàm chuyển). Đóng vai trò biến đổi từ Net sang tín hiệu đầu ra out.

$$\text{Out} = g(\text{Net})$$

Đây là thành phần phi tuyến của nơon. Có 3 dạng hàm kích hoạt thường được dùng trong thực tế

***) Hàm dạng bước:**

$$\text{step}(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad \text{step}(x) = \begin{cases} 1 & x \geq \theta \\ 0 & x < \theta \end{cases}$$

***) Hàm dấu:**

$$\text{step}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases} \quad \text{step}(x) = \begin{cases} 1 & x \geq \theta \\ -1 & x < \theta \end{cases}$$

***) Hàm sigmoid:** $\text{Sigmoid}(x) = \frac{1}{1 + e^{-\alpha(x+\theta)}}$

Ở đây ngưỡng θ đóng vai trò làm tăng tính thích nghi và khả năng tính toán của mạng nơon. Sử dụng ký pháp vectơ, $S = (s_1, \dots, s_n)$ vectơ tín hiệu vào, $W = (w_1, \dots, w_n)$ vectơ trọng số, ta có

$$\text{out} = g(\text{Net}) \quad \text{Net} = SW$$

Trường hợp xét ngưỡng θ , ta dùng biểu diễn vectơ mới $S = (s_1, \dots, s_n, \theta)$, $W = (w_1, \dots, w_n, -1)$

3.2.2 Mạng Nơon

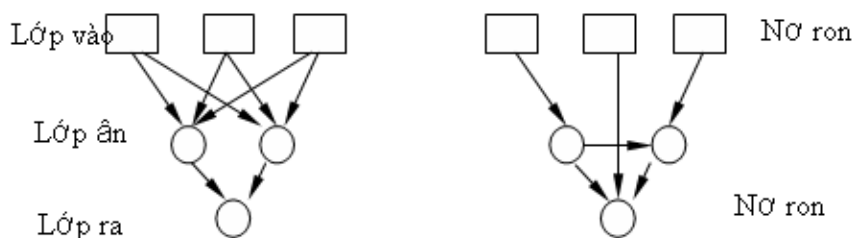
Mạng nơon là hệ thống bao gồm nhiều phần tử xử lý đơn giản (nơon) hoạt động song song. Tính năng của hệ thống này tùy thuộc vào cấu trúc của hệ, các trọng số liên kết nơon và quá trình toán tại các nơon đơn lẻ. Mạng nơon có thể học từ dữ liệu mẫu và tổng quát hóa dựa trên các dữ liệu mẫu

học. Trong mạng nơron, các nơron đón nhận tín hiệu vào gọi là nơron vào và các nơron đưa thông tin ra gọi là nơron ra.

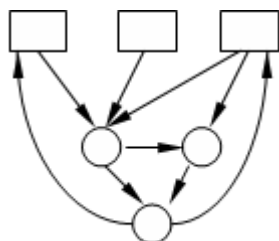
3.2.2.1 Phân loại các mạng nơron

- **Theo kiểu liên kết nơron:** ta có mạng nơron truyền thẳng (feed-forward Neural Network) và mạng nơron qui hồi (recurrent Neural Network). Trong *mạng nơron truyền thẳng*, các liên kết nơron đi theo một hướng nhất định, không tạo thành đồ thị không có chu trình với các đỉnh là các nơron, các cung là các liên kết giữa chúng. Ngược lại, các *mạng qui hồi* cho phép các liên kết nơron tạo thành chu trình. Vì các thông tin ra của các nơron được truyền lại cho các nơron đã góp phần kích hoạt chúng, nên mạng hồi quy còn có khả năng lưu giữ trạng thái trong của nó dưới dạng các ngưỡng kích hoạt ngoài các trọng số liên kết nơron.

- **Theo số lớp:** các nơron có thể tổ chức lại thành các lớp sao cho mỗi nơron của lớp này chỉ được nối với các nơron ở lớp tiếp theo, không cho phép các liên kết giữa các nơron trong cùng một lớp, hoặc từ nơron lớp dưới lên nơron lớp trên. Ở đây cũng không cho phép các liên kết nhảy qua một lớp



Hình 3.3: Mạng nơron truyền thẳng và nhiều lớp



Hình 3.4: Mạng nơron hồi qui

3.2.2.2 Hai chức năng của mạng nơron

- **Mạng nơron như một công cụ tính toán:**

Giả sử mạng nơron Neural network có m nơron vào và n nơron ra, khi đó với mỗi vecto các tín hiệu vào $X=(x_1, \dots, x_m)$, sau quá trình tính toán tại các nơron ẩn, ta nhận được kết quả ra $Y=(y_1, \dots, y_n)$. Theo nghĩa nào đó mạng nơron làm việc với tư cách một bảng tra, mà không cần biết dạng phụ thuộc hàm tường minh giữa Y và X . khi đó ta viết:

$$Y = \text{tinh}(X, NN)$$

Cần lưu các nơron trên cùng một lớp có thể tính toán đồng thời, do vậy độ phức tạp tính toán nói chung sẽ phụ thuộc vào số lớp mạng.

- **Các thông số cấu trúc mạng nơron bao gồm:**

- + Số tín hiệu vào, số tín hiệu ra
- + Số lớp nơron
- + Số nơron trên mỗi lớp ẩn
- + Số lượng liên kết của mỗi nơron (liên kết đầy đủ, liên kết bộ phận và liên kết ngẫu nhiên)
- + Các trọng số liên kết nơron.

- **Mạng nơron như một hệ thống thích nghi có khả năng học:**

Để chỉnh các trọng số liên kết cũng như cấu trúc của mình sao cho phù hợp

với các mẫu học (samples). Người ta phân biệt ba loại kỹ thuật học:

- (a) Học có quan sát (supervised learning)
- (b) Học không quan sát (unsupervised learning)
- (c) Học tăng cường.

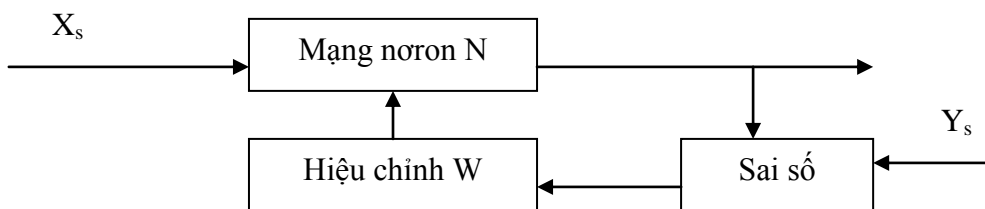
Trong *học giám sát*, mạng được cung cấp một tập mẫu học $\{(X_s, Y_s)\}$ theo nghĩa X_s là các tín hiệu vào, thì kết quả ra đúng của hệ phải là Y_s . Ở mỗi lần học, vecto tín hiệu vào X_s được đưa vào mạng, sau đó so sánh sự sai khác giữa các kết quả ra đúng Y_s với kết quả tính toán out_s . Sai số này sẽ được dùng để hiệu chỉnh lại các trọng số liên kết trong mạng. Quá trình cứ tiếp tục cho đến khi thỏa mãn một tiêu chuẩn nào đó. Có hai cách sử dụng tập mẫu học: hoặc dùng

các mẫu lần lượt, hết mẫu này đến mẫu khác, hoặc sử dụng đồng thời tất cả các mẫu một lúc. Các mạng với cơ chế học không giám sát được gọi là các mạng tự tổ chức. Các kỹ thuật học trong mạng nơron có thể nhằm vào hiệu chỉnh các trọng số liên kết (gọi là học tham số) hoặc điều chỉnh, sửa đổi cấu trúc của mạng bao gồm số lớp, số nơron, kiểu và trọng số các liên kết (gọi là học cấu trúc).

***) Học tham số:**

Giả sử có k nơron trong mạng và mỗi nơron có đúng một liên kết vào với các nơron khác. Khi đó, ma trận trọng số liên kết W sẽ có kích thước kx1. Các thủ tục học tham số nhằm mục đích tìm kiếm ma trận W sao cho

$$Y_s = \text{Tinh}(X_s, W) \text{ đối với mọi mẫu học } s = (X_s, Y_s) \quad (1)$$



Hình 3.5: Học tham số có giám sát

***) Học cấu trúc:**

Với học tham số ta giả định rằng mạng có một cấu trúc cố định. việc học cấu trúc của mạng truyền thẳng gắn với yêu cầu tìm ra số lớp của mạng L và số nơron trên mỗi lớp n_j . Tuy nhiên, với các mạng hồi quy còn phải xác định thêm các tham số ngưỡng θ của các nơron trong mạng. Một cách tổng quát phải xác định bộ tham số $P = (n_1, \dots, n_k, \theta_1, \dots, \theta_k)$ ở đây $k = \sum n_j$ sao cho $Y_s = \text{Tinh}(X_s, P)$ đối với mọi mẫu học $s = (X_s, Y_s)$ (2).

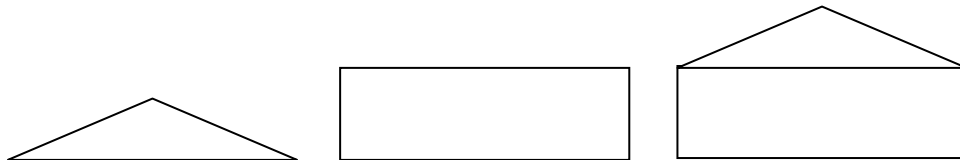
Về thực chất, việc điều chỉnh các vecto tham số W trong (1) hay P trong (2)

đều qui về bài toán tìm kiếm tối ưu trong không gian tham số. Do vậy, có thể áp dụng các cơ chế tìm kiếm kinh điển theo gradient.

3.2.3 Mạng Kohonen

Cách xử lý thông tin trong các mạng ở trên thường chỉ quan tâm tới giá trị và dấu của các thông tin đầu vào, mà chưa quan tâm khai thác các mối liên hệ có tính chất cấu trúc trong lân cận của các vùng dữ liệu mẫu hay toàn thể không gian mẫu.

Chẳng hạn, với 2 thành phần: 1 tam giác, 1 hình chữ nhật,



ta có thể tạo thành hình ngôi nhà khi chúng được phân bố kề giáp với nhau theo một trật tự nhất định.

Teuvo Kohonen (1989) đã đề xuất một ý tưởng rất đáng chú ý về ánh xạ các đặc trưng topo tự tổ chức (theo nghĩa không cần có mẫu học) nhằm bảo toàn trật tự sắp xếp các mẫu trong không gian biểu diễn nhiều chiều sang một không gian mới các mảng nơ ron (một hoặc hai chiều). Trong mạng Kohonen, các vectơ tín hiệu vào gần nhau sẽ được ánh xạ sang các nơ ron trong mạng lân cận nhau.

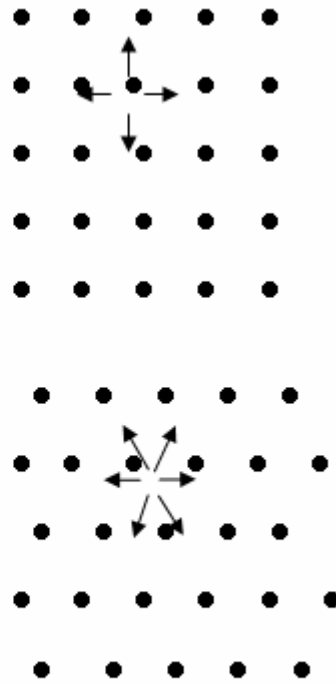
3.2.3.1 Cấu trúc mạng

Mạng Kohonen rất gần gũi với kiểu cấu trúc mạng nơ ron sinh học cả về cấu tạo lẫn cơ chế học. Mạng Kohonen thuộc vào nhóm mạng một lớp các nơ ron được phân bố trong mặt phẳng hai chiều theo kiểu lưới vuông, hay lưới lục giác dưới

Phân bố này phải thoả mãn yêu cầu ; Mỗi nơ ron có cùng số nơ ron trong từng lớp láng giềng. ý tưởng cơ bản của Kohonen là các đầu vào tương tự nhau sẽ kích hoạt các nơ ron gần nhau về khoảng không gian. Mối quan hệ tương tự (theo khoảng cách) có thể tổng quát hoá cho một lớp tương đối rộng các quan hệ tương tự giữa các tín hiệu đầu vào.

Một cách trực quan, có thể xem thuật giải huấn luyện mạng Kohonen nhằm biến đổi không gian tín hiệu vào sang mạng nơ ron giống như các thủ tục kiểu như "làm tròn" hay "tạo hình" dữ liệu.

Để đáp ứng yêu cầu các nơ ron có cùng số nơ ron lân cận trong mỗi lớp láng giềng, người ta thường dùng các phép cuộn chỉ số để đạt được hiệu ứng cái sấm xe. Chẳng hạn, tọa độ (x_i, y_i) của các nơ ron thuộc lớp láng giềng thứ k của nơ ron có tọa độ (x, y) trong mảng nơ ron 2 chiều có kích thước $p \times q$ được cho trong thủ tục sau:



Hình 3.6. Lưới các nơ ron trong mặt phẳng hai chiều

```

for   i:=-k to k do
for   j:=-k to k do
begin xi:=mod(x+i+p-1,p) + 1;
      yi:=mod(y+j+q-1,q) + 1;
      if (i=k) or (j=k) then
          nơ ron  $(x_i, y_i)$  thuộc vào lớp láng giềng thứ  $k$ 
      else
          nơ ron  $(x_i, y_i)$  thuộc vào lớp láng giềng thứ  $r$ 
           $r < k$ ;  $r$  được xác định bởi  $\max(x_i, y_i)$ 
      end;

```

Trường hợp lớp nơ ron Kohonen là một dãy, cách cuộn tròn mảng nơ ron tạo thành một đường tròn.

Tất cả các nơ ron ở lớp kích hoạt có liên kết đầy đủ với lớp vào. Điểm quan trọng nhất trong mạng Kohonen là với một vector tín hiệu vào, nó chỉ cho phép các phản hồi mang tính chất địa phương nghĩa là đầu ra của mỗi nơ ron

không được nối với tất cả các nơ ron khác mà chỉ với một số nơ ron lân cận. Sự phản hồi mang tính địa phương của những điều chỉnh (nếu có) tạo ra hiệu ứng là các nơ ron gần nhau về vị trí sẽ có hành vi tương tự khi có những tín hiệu giống nhau được đưa vào.

3.2.3.2 Huấn luyện mạng

Quá trình học được sử dụng trong mạng Kohonen dựa trên kỹ thuật cạnh tranh, không cần có tập mẫu học. Khác với trường hợp học có giám sát, các tín hiệu đầu ra có thể không biết được một cách chính xác.

Tại mỗi thời điểm chỉ có một nơ ron duy nhất C trong lớp kích hoạt được lựa chọn sau khi đã đưa vào mạng các tín hiệu X_s . Nơ ron này được chọn theo một trong hai nguyên tắc sau:

Nguyên tắc 1 Nơ ron c có tín hiệu ra cực đại

$$\text{out}_c \leftarrow \max_j(\text{out}_j) = \max_{i=1} (\sum x_{si} w_{ji}) \quad (9)$$

Nguyên tắc 2 Vectơ trọng số của nơ ron c gần với tín hiệu vào nhất

$$\text{err}_c \leftarrow \min_j(\text{err}_j) = \min_j (\sum_{i=1} (x_{si} - w_{ji})^2) \quad (10)$$

Sau khi xác định được nơ ron c , các trọng số w_{ci} được hiệu chỉnh nhằm làm cho đầu ra của nó lớn hơn hoặc gần hơn giá trị trọng số mong muốn. Do vậy, nếu tín hiệu vào x_{si} với trọng số w_{ci} tạo kết quả ra quá lớn thì phải giảm trọng số và ngược lại. Các trọng số của các nơ ron láng giềng j cũng phải được hiệu chỉnh giảm, tùy thuộc vào khoảng cách tính từ c . Ta đưa vào hàm tỷ lệ $a(.) = a(dc_j)$, ở đây dc_j là khoảng cách topo giữa nơ ron trung tâm c và nơ ron j đang xét. Trên thực tế hàm $a(.)$ có thể là hằng số, hàm tỷ lệ nghịch hoặc hàm có điểm uốn. Để đảm bảo yêu cầu, do có nhiều mẫu tham gia quá trình huấn luyện, ta đưa vào hệ số $\eta(t)$:

$$f = \eta(t) \cdot a(dc_j),$$

$$\eta(t) = (a_{\max} - a_{\min}) \frac{t_{\max} - t}{t_{\max} - 1} + a_{\min} \quad (11)$$

ở đây t là số đối tượng mẫu đã dùng để luyện mạng

t_{max} là số mẫu tối đa

a_{max} , a_{min} tương ứng là giá trị cực đại, cực tiểu của hàm $a(.)$

Tùy thuộc vào nơ ron trung tâm c được lựa chọn theo nguyên tắc 1 hoặc nguyên tắc 2 ta có cách hiệu chỉnh các trọng số w_{ji} tương ứng:

$$w_{ji} = w_{ji} + \eta(t) a(dc_j)(1 - x_i w_{ji}) \quad (12)$$

hoặc $w_{ji} = w_{ji} + \eta(t) a(dc_j) \frac{w_{ji}}{\sqrt{\sum_{i=1}^n w_{ji}^2}} \quad (13)$

Sau đó, chuẩn hoá các trọng số $\sqrt{\sum_{i=1}^n w_{ji}^2} = 1$

Theo kinh nghiệm, cần phải tạo ra phân bố ngẫu nhiên các trọng số trong khoảng -0.1 đến 0.1 hoặc $-1/m$ đến $1/m$, ở đây m là số trọng số của mạng và chuẩn hoá dữ liệu vào, ra bằng -1 hoặc 1 .

Tuy nhiên cũng phải chú ý một điều là việc lựa chọn tiêu chuẩn chuẩn hoá, định cỡ dữ liệu phụ thuộc rất nhiều vào bản chất bài toán.

3.2.3.3 Sử dụng mạng

Giả sử đã huấn luyện mạng để nhận được ma trận trọng số W . Khi đưa vào mạng một vector X , toàn bộ ma trận W lại được cập nhật theo các công thức (12) hoặc (13) tùy thuộc vào sử dụng nguyên tắc 1 hay nguyên tắc 2.

Như vậy, mạng Kohonen cho chúng ta biết được sự phân bố và quan hệ tương đối về mặt "địa lý" giữa các mẫu trong không gian biểu diễn.

3.2.3.4 Thử nghiệm mạng

Ánh xạ từ không gian 3 chiều sang không gian 2 chiều.

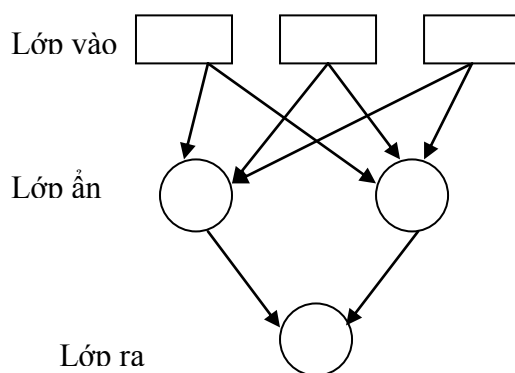
Bài toán đặt ra là tạo ánh xạ từ một mặt cầu đơn vị 3 chiều với 2000 điểm phân bố ngẫu nhiên trong 8 múi cầu sang mặt phẳng các nơ ron được phân bố trong lưới kích thước 15×15 .

Mạng Kohonen được thiết kế có 3 đầu vào, tương ứng với 3 tọa độ và 225 nơ ron, phân bố thành lưới vuông 15×15 . Mỗi nơ ron vào được nối đầy đủ với các nơ ron ra, do vậy tổng cộng có 675 trọng số. Ban đầu nơ ron trung tâm có 7 lớp láng giềng để đảm bảo rằng tất cả các vùng láng giềng kề giáp nhau. Giả sử, hiệu chỉnh cực đại tại nơ ron trung tâm $a(0) = 0.3$ (xem công thức(11)) và tại lớp

- Mạng không chỉ quan tâm đến nội dung tín hiệu vào mà còn xem xét cấu trúc topo của các mẫu.
- Mạng có thể biến đổi từ không gian nhiều chiều sang không gian ít chiều hơn
- Cơ chế học không có giám sát
- Các quan hệ topo được bảo toàn khi ánh xạ.

3.2.4 Mạng nơron nhiều lớp lan truyền ngược sai số

3.2.4.1 Kiến trúc mạng



Hình 3.9: Mạng Nơron 2 lớp

Các nơron lớp thứ t được nối đầy đủ với các nơron lớp thứ $t+1$. Trong nhiều ứng dụng thực tế, để đơn giản, người ta thường sử dụng mạng có một lớp ẩn, số nơron trong lớp ẩn được xác định dựa trên kinh nghiệm, hoặc dựa trên các kỹ thuật tìm kiếm khác.

3.2.4.2 Huấn luyện mạng

Quá trình huấn luyện mạng được trình bày ở đây là quá trình học có giám sát với tập mẫu $\{X_s, Y_s\}$. Quá trình học có thể tóm tắt như dưới:

Mỗi khi một mẫu $X_s = \{x_1, \dots, x_n\}$ vào mạng, ta thực hiện các công việc sau:

- Lan truyền mẫu X_s qua mạng để có $out_s = \text{Tính}(X_s, NN)$,
- Tính sai số Err_s của mạng dựa trên sai lệch $out_s - Y_s$,

- Hiệu chỉnh các trọng số liên kết nơron dẫn tới lớp ra w_{ij} từ nơron j tại lớp ẩn cuối cùng tới nơron i tại lớp ra: $w_{ij} = w_{ij} + \alpha * a_j * \delta_i$ (1)

Với:

α là hệ số học.

a_j là đầu ra của nơron j,

δ_i là sai số mà nơron I ở lớp ra phải chịu trách nhiệm, được xác định theo công thức: $\delta_i = err_i * g'(Net_i)$ (2)

với err_i là sai số thành phần thứ i trong err , Net_i là tổng thông tin vào có trọng số của nơron thứ i $Net_i = \sum w_{ij} * a_j$ và g' là đạo hàm của hàm kích hoạt g được dùng trong các nơron.

Hiệu chỉnh các trọng số liên kết nơron w_{ik} dẫn tới tất cả lớp ẩn từ nơron thứ k sang nơron j (các lớp ẩn được xét từ dưới lên):

Tính tổng sai số tại nơron j phải chịu trách nhiệm

$$\delta_j = g'(Net_j) \sum w_{jk} \delta_k \quad (3)$$

Hiệu chỉnh trọng số $w_{jk} = w_{jk} + \alpha * a_k * \delta_j$ (4)

(Trường hợp xét liên kết từ nơron vào thứ k sang nơron j trên lớp ẩn thứ nhất, ta có $a_k = i_k$) chính là tín hiệu vào).

Chú ý:

a) Trường hợp xét hàm kích hoạt tại các nơron

$$g(x) = \frac{1}{1 + e^{-x}}$$

Ta có hệ thức $g'(x) = g(x) * (1 - g(x))$

b) Từ các công thức (1) và (4) ta có thể viết lại:

$$w_{ij} = w_{ij} + \Delta w_{ij} \quad \text{với} \quad \Delta w_{ij} = \alpha * a_j * \delta_i \quad \text{và}$$

$$w_{jk} = w_{jk} + \Delta w_{jk} \quad \text{với} \quad \Delta w_{jk} = \alpha * a_k * \delta_j$$

Trong thực tế, thường hiệu chỉnh Δw_{ij} theo nguyên tắc có chú ý đến thao tác trước đó. Do vậy:

$$\Delta w_{ij}^{(u)} = \alpha * a_j * \delta_i + \beta \Delta w_{ij}^{(u-1)}$$

ở đây β là hệ số quán tính.

Quá trình huấn luyện mạng cần chú ý tới các yếu tố sau:

- i. Các trọng số ban đầu w_{ij} được gán các giá trị ngẫu nhiên, nhỏ,
- ii. Lựa chọn các hệ số học α và hệ số quán tính β sao cho $\alpha + \beta \approx 1$, với β không lớn hơn α quá nhiều,

Các tín hiệu vào, ra nên được định cỡ chỉ nằm trong khoảng $[0,1]$. Các nghiên cứu thực nghiệm chỉ ra rằng nên ở trong khoảng $[0.2,0.8]$

3.2.4.3 Sử dụng mạng

Giả sử đã huấn luyện mạng như hình ở trên với tập mẫu $\{X_s, Y_s\}$ để được ma trận trọng số W . Quá trình lan truyền trong mạng một vecto tín hiệu vào $X = [x_1, x_2, x_3]$ được cho bởi:

$$\begin{aligned} out &= g(w_{64}a_4 + w_{65}a_5) = g(w_{64}g(w_{41}x_1 + w_{42}x_2 + w_{43}x_3) + w_{65}g(w_{51}x_1 + w_{52}x_2 + w_{53}x_3)) \\ &= F(X, W) \end{aligned}$$

Khả năng tính toán của mạng nhiều lớp

- Với một lớp ẩn, mạng có thể tính toán xấp xỉ một hàm liên tục bất kỳ đối với các biến tương ứng là các tín hiệu vào.
- Với 2 lớp ẩn, mạng có thể tính toán xấp xỉ một hàm bất kỳ. Tuy vậy, số nơron trong các lớp ẩn có thể tăng theo hàm mũ đối với số đầu vào và cho đến nay vẫn chưa có các hàm có thể xấp xỉ nhờ các mạng nhiều lớp

3.3 Sử dụng mạng nơron lan truyền ngược hướng cho nhận dạng ký tự

3.3.1 Nhận dạng bằng mạng nơron lan truyền ngược hướng (kn chung)

Mạng nơron nói chung và mạng lan truyền ngược hướng nói riêng là sự mô phỏng sinh học bằng máy tính bộ não người. Nó có khả năng học từ kinh nghiệm hay từ một tập mẫu. Quá trình học của mạng lan truyền ngược hướng là quá trình học có giám sát với một mẫu $\{X_s, Y_s\}$ cho trước, ở đây X_s là vecto vào (ma trận điểm ảnh của một ký tự) và Y_s là giá trị ASCII của ký tự đó. Thực chất việc học của mạng là biến đổi và ánh xạ topo các ký tự xuống mặt phẳng hai chiều tương ứng với cá nơron. Sau khi huấn luyện, mạng lan truyền ngược hướng hoạt động như một bảng tra với đầu vào là các vecto điểm ảnh của các

ký tự. Một trong những ưu điểm chính của mạng là không đòi hỏi các quá trình tiền xử lý như làm mảnh, làm trơn đường biên hay khử nhiễu.

Quá trình học của mạng lan truyền ngược hướng là quá trình học có giám sát. Do đó nó cần có một tập mẫu chuẩn $\{X_s, Y_s\}$. Trong quá trình học vector vector vào X_s đi vào mạng Kohonen, ở đây diễn ra quá trình học cạnh tranh. Vector lời giải Y_s đi vào lớp ra theo hướng ngược lại làm thay đổi giá trị các trọng số của các nơ ron trên lớp ra. Giả thiết chúng ta có mạng lan truyền ngược hướng gồm N nơ ron trên lớp Kohonen và M nơ ron trên lớp ra. W_{ji} là trọng số thứ i của nơ ron thứ j trên lớp Kohonen. C_{ji} là trọng số của nơ ron thứ i trên lớp ra nối với nơ ron thứ j trên lớp Kohonen. Quá trình học của mạng lan truyền ngược hướng bao gồm các bước sau đây:

- Một đối tượng gồm cặp vector (X_s, Y_s) được lấy ra từ tập mẫu.
- Vector X_s đi vào lớp Kohonen.
- Nơ ron trung tâm được chọn theo phương trình
- Tất cả các trọng số của nơ ron trên lớp Kohonen được điều chỉnh theo phương trình.
- Các trọng số của nơ ron trên lớp ra được điều chỉnh theo phương trình:

$$C_{ji}^{(new)} = C_{ji}^{(old)} + \eta(t).a(dc - dj).(y_i - C_{ji}^{(old)})$$

- Quá trình lặp lại đối với đối tượng tiếp theo.

Mỗi lần tất cả các đối tượng mẫu đã đi qua mạng được gọi là một lượt. Thông thường cần phải thực hiện từ vài trăm đến hàng nghìn lượt để mạng ổn định. Khi chọn được các hằng số đặc trưng của quá trình học a_{max} , a_{min} thích hợp, quá trình học của mạng luôn hội tụ.

3.3.2 Cài đặt mạng lan truyền ngược hướng cho nhận dạng ký tự

Một mạng tổng quát cho việc nhận dạng ký tự được cài đặt trên ngôn ngữ C như một lớp (Class) có tên gọi là Netcount. Các tham số của mạng là các biến thành viên còn các chức năng của mạng được thiết kế cho các hàm thành viên. Mạng chỉ có một nơ ron trên lớp ra và có kiểu là ký tự.


```

Class Netcount
{protected:
int dai, rong, N;
float amax, amin, *W[1600];
char C[1600];
public;
Netcount(int, int);
Void hoc(char*, long T);
Char doan (char*);
};

```

Các trọng số W_{ji} được cấp phát động cho bảng các con trỏ W . Khoảng cách giữa nơ ron có toạ độ k_j, l_j với nơ ron trung tâm k_c, l_c được tính theo công thức:

$$D = \max[\min(|k_j - k_c|, |k_j - k_c + dai|, |k_j - k_c - dai|), \min(|l_j - l_c|, |l_j - l_c + rong|, |l_j - l_c - rong|)]$$

Hàm phụ thuộc topo $a(d_c - d_j)$ được dùng trong chương trình là hàm tam giác:

$$a \left(\frac{d_c - d_j}{D_{\max}} \right) = \begin{cases} 1 & D_{ci} = D_{\max} \\ \frac{D_{\max} - D}{D_{\max}} & D_{ci} < D_{\max} \end{cases}$$

Trong đó: D_{\max} là khoảng cách từ lân cận xa nhất có thể có của mạng:

$$D_{\max} = \max(dai/2, rong/2) + 1;$$

Nhìn chung để cài đặt mạng nơ ron cho nhận dạng ký tự cần:

Tổ chức số liệu

Tập mẫu được tổ chức trong một tệp số liệu. Các cặp (X_s, Y_s) được viết lần lượt theo từng dòng. Một điều đặt ra là phải số thực hoá các vector vào khoảng $[0, 1]$ vì các trọng số của mạng là các số thực. Các nghiên cứu cho thấy việc số thực hoá làm cho mạng có khả năng đoán nhận các ký tự từ các ảnh số sai lệch lớn hơn. Hơn nữa, với việc tổ chức số thực hoá, chúng ta có thể làm giảm kích thước của vector vào và có khả năng làm việc đối với các ký tự có kích

thước ảnh khác nhau. Thực tế chỉ ra các phương pháp số thực hoá khác nhau sẽ ảnh hưởng đến khả năng cực đại mà mạng có thể đoán nhận từ các ảnh sai lệch.

Cấu trúc và các tham số học

Mục đích của việc xây dựng mạng là xác định số lượng nơ ron trên lớp Kohonen. Với số lượng nơ ron trên lớp Kohonen càng lớn, khả năng đoán nhận các ký tự từ các ảnh có tỷ lệ sai lớn hơn. Tuy nhiên, khi tăng số lượng các nơ ron, khả năng nhận biết sẽ tiến sát tới khả năng cực đại mà mạng có thể đoán nhận với các ảnh sai (phụ thuộc vào phương pháp số thực hoá). Chúng ta cũng dễ nhận thấy thời gian học và thời gian đoán nhận, cũng như bộ nhớ của máy tính tăng tỷ lệ, có thể hàm mũ với số lượng nơ ron trên lớp Kohonen. Thực tế, việc xây dựng mạng là công việc thử nghiệm, dần dần tăng kích thước mạng cho đến khi đạt được các chỉ tiêu mong muốn.

Các giá trị trọng số ban đầu thực sự không quan trọng với quá trình học nhưng chúng phải được gán bằng các số ngẫu nhiên từ 0 đến 1.

Các tham số học a_{max} , a_{min} ảnh hưởng không nhiều đến quá trình học nếu chúng thoả mãn các điều kiện sau:

$$a_{max} \in [0.3, 1]; a_{min} \in [0, 0.1].$$

Với giá trị $a_{max} = 0.5$ và $a_{min} = 0.01$ có thể là giá trị tốt cho quá trình học.

3.3.3 Nhận dạng các ký tự sử dụng mạng lan truyền ngược hướng

Một tập mẫu 37 ký tự từ A \rightarrow Z, 0 \rightarrow 9 và ký tự '<' được tách ra từ tệp ảnh quét bởi scanner có kích thước 32 x 32 điểm ảnh.

Ba thử nghiệm được tiến hành là:

- Không số thực hoá
- Lọc các điểm ảnh bằng mặt nạ 3 x 3
- Phân mảnh ảnh thành 64 mảnh. Mỗi vùng có giá trị thực bằng tổng điểm số điểm ảnh đen (giá trị 1) chia cho 16

Bảng 1 thống kê khả năng nhận đúng ký tự từ các ảnh có tỷ lệ sai cực đại của mạng 20 x 20 nơ ron sau 3000 lượt học.

Bảng 2 thống kê sự phụ thuộc của khả năng nhận dạng các ảnh sai vào kích thước với việc số thực hoá là phân 64 mảnh.

Bảng 1

Không số thực hoá	Mặt nạ 3 x 3	Phân 64 mảnh
3%	15%	19%

Bảng 2

10 x 10	20 x 20	30 x 30	40 x 40
3%	19%	24%	25%

Với việc phân bố của các ký hiệu ở hình bên ta dễ nhận thấy mạng đã phát hiện một cách khách quan các đặc trưng topo của các ký tự thường được dùng trong các phương pháp nhận dạng cấu trúc truyền thống. Các ký tự có cấu trúc topo tương đối giống nhau được sắp xếp đặt gần nhau, như các ký tự có điểm kết thúc như nhau {'Z', '2'}, {'5', 'S'}; các ký tự có một chu trình {'O', '0', 'Q', 'R', '9', 'D'}; Các ký tự có hai chu trình {'B', '8'}. Một đặc điểm rất quan trọng là mạng đã phát hiện ra các ký tự có "tiềm năng" giống nhau như các ký tự {'H', 'E', 'W'} rất dễ trở thành có hai chu trình khi ảnh bị sai lớn. Ký tự 'A' khi bị mất góc cuối bên trái có thể trở thành số '4'; Ký tự 'U' rất dễ trở thành có chu trình. Ngoài ra mạng đã phát hiện các ký tự có một hay nhiều phần giống nhau khó có khả năng mô tả trong các chương trình nhận dạng truyền thống như mật độ các điểm đen như {'M', 'X', 'A'}, hay nét cong của đường biên ký tự 'G' và 'O'.

Kết luận

Từ ví dụ nhận dạng 37 ký tự cho thấy việc nhận dạng ký tự bằng mạng lan truyền ngược hướng có hiệu quả, đơn giản và nhanh hơn các phương pháp truyền thống. Nó có khả năng nhận dạng được các ký tự từ các ảnh có chất lượng tồi với số điểm ảnh sai 25%. Lợi thế chính của mạng loại này xuất phát từ khả năng học các đặc trưng topo của các mẫu. Tuy nhiên với một tập mẫu khá lớn, việc sử dụng tài nguyên của máy tính sẽ rất lớn.

PHẦN KẾT LUẬN

Sự phát triển của công nghệ thông tin đã có tác động đến nhiều mặt của đời sống xã hội trong đó phải kể đến lĩnh vực giám sát tự động. Trong giám sát tự động, việc giám sát đối với các phương tiện giao thông là một vấn đề nổi trội. Nhiều chính phủ, thành phố trên thế giới đã xây dựng hệ thống giám sát tự động đối với các phương tiện giao thông của mình. Và các hệ thống giám sát đều lấy biển số xe là mục tiêu giám sát.

Ở nước ta, các hệ thống giám sát tự động nói chung và các hệ thống nhận dạng biển số xe nói riêng chưa được chú ý tới và nó cũng là một lĩnh vực tương đối mới mẻ. Đa phần các công tác quản lý, xử lý đối với các phương tiện giao thông đều cần nhân lực là con người. Báo cáo nhằm mục đích tìm hiểu bài toán giám sát, quản lý các phương tiện giao thông một cách tự động thông qua việc “Phát hiện và nhận dạng chữ, số trong biển số xe”.

Khoá luận đã trình bày một cách hệ thống về bài toán nhận dạng biển số xe và các hướng giải quyết trên cơ sở các bài toán cơ bản: Phát hiện vùng chứa biển số xe và bài toán nhận dạng chữ và số trong vùng được phát hiện.

Với mục đích để tìm hiểu do thời gian có hạn nên em không hoàn thành được sản phẩm ứng dụng của mình. Em hy vọng rằng ở Việt nam không xa, thì các hệ thống này được sử dụng nhiều. Để hỗ trợ một phần công tác giám sát, quản lý các phương tiện giao thông một cách hiệu quả hơn.

TÀI LIỆU THAM KHẢO

- [1] Nhập môn xử lý ảnh số. Ths. Lương Mạnh Bá, Pts. Nguyễn Thanh Thủy. Nxb KHKT 2003.
- [2] Một thuật toán phát hiện vùng và ứng dụng của nó trong quá trình vecto hóa tự động.PGS.TS Đỗ Năng Toàn.Tạp chí Tin học và Điều khiển, Tập 16 số 1 năm 2000
- [3] Machine Vision: Theory, Algorithms and Practicalities. E.Davies. Academic Press 1990
- [4] A robust and fast skew detection algorithm for generic document. B.Yu and A.Jain. Pattern Recognition 1996
- [5] Khoá luận của anh Đào Đình Dũng trường ĐHQGHN khoá 2005
Và 1 số tạp chí tin học khác...