

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

-----o0o-----



15 | 9001:2000

ĐỒ ÁN TỐT NGHIỆP

NGÀNH CÔNG NGHỆ THÔNG TIN

HẢI PHÒNG 2013

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

-----o0o-----

**TÌM HIỂU MỘT SỐ PHƯƠNG PHÁP TRÍCH
CHỌN ĐẶC TRƯNG CHO NHẬN DẠNG
CHỮ VIẾT**

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công nghệ Thông tin

HẢI PHÒNG - 2013

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

-----o0o-----

**TÌM HIỂU MỘT SỐ PHƯƠNG PHÁP TRÍCH
CHỌN ĐẶC TRƯNG CHO NHẬN DẠNG
CHỮ VIẾT**

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công nghệ Thông tin

Giáo viên hướng dẫn: PGS TS Ngô Quốc Tạo

Sinh viên thực hiện: Nguyễn Sơn Hà

Mã số sinh viên: 1351010034

HẢI PHÒNG - 2013

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

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc
-----oOo-----

NHIỆM VỤ THIẾT KẾ TỐT NGHIỆP

Sinh viên: Nguyễn Sơn Hà Mã SV: 1351010034

Lớp: CT1301

Ngành: Công nghệ Thông tin

Tên đề tài: Tìm hiểu một số phương pháp trích chọn đặc trưng cho nhận dạng chữ viết.

NHIỆM VỤ ĐỀ TÀI

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

a. Nội dung

- Giới thiệu về trích chọn đặc trưng: Đặc trưng bất biến, khả năng khôi phục.
- Trích chọn đặc trưng từ ảnh đa cấp xám: Đối sánh mẫu, mẫu biến dạng, biến đổi ảnh Unitar, bất biến hình học, bất biến Zenite.
- Trích chọn đặc trưng từ ảnh nhị phân: Đối sánh mẫu, mẫu biến dạng, biến đổi ảnh Unitar, bất biến hình học.
- Trích chọn đặc trưng từ biên ảnh: Chiều nghiêng, tách vùng, xấp xỉ đường cong, mô tả Fourier.
- Trích chọn đặc trưng từ biểu diễn véc tơ: Đối sánh mẫu, mẫu biến dạng, đặc trưng rời rạc, biểu diễn đồ thị, biểu diễn Fourier.

b. Các yêu cầu cần giải quyết

- Hiểu nội dung trích chọn đặc trưng theo đối sánh
- Viết xong đồ án
- Cài đặt thử nghiệm chương trình trích đặc trưng

CÁN BỘ HƯỚNG DẪN ĐỀ TÀI TỐT NGHIỆP

Người hướng dẫn thứ nhất:

Họ và tên: Ngô Quốc Tạo

Học hàm, học vị: Phó giáo sư Tiến sĩ

Cơ quan công tác: Viện Công nghệ thông tin, Viện Hàn Lâm Khoa học và Công nghệ Việt Nam.

Nội dung hướng dẫn:
.....
.....

Người hướng dẫn thứ hai:

Họ và tên:

Học hàm, học vị:

Cơ quan công tác:

Nội dung hướng dẫn:
.....
.....

Đề tài tốt nghiệp được giao ngày...tháng...năm 2013.

Yêu cầu phải hoàn thành trước ngày...tháng...năm 2013.

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

Nguyễn Sơn Hà

PGS TS Ngô Quốc Tạo

Hải Phòng, ngàytháng.....năm 2013

HIỆU TRƯỞNG

GS.TS.NGŨT Trần Hữu Nghị

PHẦN NHẬN XÉT TÓM TẮT CỦA CÁN BỘ HƯỚNG DẪN

1. Tinh thần thái độ của sinh viên trong quá trình làm đề tài tốt nghiệp:

.....
.....
.....
.....
.....

2. Đánh giá chất lượng của đề tài tốt nghiệp (so với nội dung yêu cầu đã đề ra trong nhiệm vụ đề tài tốt nghiệp)

.....
.....
.....
.....
.....
.....
.....

3. Cho điểm của cán bộ hướng dẫn:

(Điểm ghi bằng số và chữ)

.....
.....

Ngày.....tháng.....năm 2013

Cán bộ hướng dẫn chính

(Ký, ghi rõ họ tên)

**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 lý 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 2013

Cán bộ chấm phản biện

(Ký, ghi rõ họ tên)

LỜI CẢM ƠN

Trước tiên em xin được bày tỏ sự trân trọng và lòng biết ơn đối với thầy giáo PGS.TS. Ngô Quốc Tạo- Trưởng phòng Nhận dạng và Công nghệ tri thức, Viện Công nghệ thông tin, Viện Hàn Lâm Khoa Học và Công nghệ Việt Nam. Trong suốt thời gian làm đồ án tốt nghiệp, thầy đã dành rất nhiều thời gian quý báu để tận tình chỉ bảo, hướng dẫn, định hướng cho em thực hiện đồ án.

Em xin được cảm ơn các thầy cô giáo Trường Đại học Dân lập Hải Phòng đã giảng dạy trong quá trình học tập, thực hành, làm bài tập, giúp em hiểu thấu đáo hơn các nội dung học tập và những hạn chế cần khắc phục trong việc học tập, nghiên cứu và thực hiện bản đồ án này.

Em xin cảm ơn các bạn bè và nhất là các thành viên trong gia đình đã tạo mọi điều kiện tốt nhất, động viên, cổ vũ trong suốt quá trình học tập và đồ án tốt nghiệp.

Do thời gian và kiến thức có hạn nên không tránh khỏi những thiếu sót nhất định. Em rất mong nhận được sự đóng góp quý báu của thầy cô!

Em xin chân thành cảm ơn!

Hải Phòng, ngày tháng năm 2013.

Sinh viên

Nguyễn Sơn Hà

MỤC LỤC

LỜI CẢM ƠN

DANH MỤC HÌNH

DANH MỤC CÁC KÝ HIỆU, CỤM TỪ VIẾT TẮT

PHẦN MỞ ĐẦU.....5

TÓM TẮT ĐỀ TÀI7

CHƯƠNG 1. TỔNG QUAN VỀ CHỮ VIẾT VÀ LÝ THUYẾT NHẬN DẠNG 8

1.1. GIỚI THIỆU8

1.2. MÔ HÌNH TỔNG QUÁT CỦA MỘT HỆ NHẬN DẠNG CHỮ VIẾT TAY 8

1.2.1. Tiền xử lý8

1.2.2. Khôi tách chữ : 13

1.2.3. Trích chọn đặc trưng: 14

1.2.4. Huấn luyện và nhận dạng : 14

1.2.5. Hậu xử lý : 14

1.3. CÁC PHƯƠNG PHÁP NHẬN DẠNG CHỮ VIẾT TAY 15

1.3.1. Đối sánh mẫu: 15

1.3.2. Phương pháp tiếp cận cấu trúc: 16

1.3.3. Mạng nơ ron 17

1.3.4. Mô hình Markov ẩn (HMM - Hidden Markov Model) : 20

1.3.5. Máy véc tơ tựa (SVM) : 21

1.3.6. Kết hợp các kỹ thuật nhận dạng : 23

1.3.7. Kết luận 24

CHƯƠNG 2: TỔNG QUAN VỀ TRÍCH CHỌN ĐẶC TRƯNG VÀ MỘT SỐ PHƯƠNG PHÁP TRÍCH CHỌN ĐẶC TRƯNG TRONG NHẬN DẠNG CHỮ VIẾT25

2.1. Trích chọn đặc trưng25

2.1.1. Biến đổi toàn cục và khai triển chuỗi25

2.1.2. Đặc trưng thống kê26

2.1.3. Đặc trưng hình học và hình thái27

2.2. Đặc trưng bất biến28

2.3. Khả năng khôi phục29

2.4. Trích chọn đặc trưng từ ảnh đa cấp xám29

2.4.1. Giới thiệu29

2.4.2. Đối sánh mẫu30

2.4.3. Mẫu biến dạng30

2.4.4. Biến đổi ảnh Unitar	30
2.4.5. Bất biến Zenite	31
2.5. Trích chọn đặc trưng từ ảnh nhị phân:	31
2.5.1. Giới thiệu	31
2.5.2. Biến đổi ảnh Unitar	32
2.5.3. Bất biến hình học.....	32
2.6. Trích chọn đặc trưng từ biên ảnh:	32
2.6.1. Giới thiệu	32
2.6.2. Tách vùng	32
2.6.3. Xấp xỉ đường cong:	33
2.6.4. Mô tả Fourier:.....	33
2.7. Trích chọn đặc trưng từ biểu diễn véc tơ:	33
2.7.1. Giới thiệu	33
2.7.2. Đối sánh mẫu	34
2.7.3. Mẫu biến dạng	34
2.7.4. Đặc trưng rời rạc.....	34
2.7.5. Biểu diễn Fourier.....	34
CHƯƠNG 3: CHƯƠNG TRÌNH THỬ NGHIỆM	36
3.1. Giới thiệu	36
3.2. Xây dựng giao diện vẽ	36
3.3. Xử lý dữ liệu (phân tích ảnh).....	37
3.4. Kết quả nhận dạng	39
3.5. Đánh giá, nhận xét	39
KẾT LUẬN.....	43
TÀI LIỆU THAM KHẢO	44

DANH MỤC CÁC HÌNH

Hình 1.1. Sơ đồ tổng quát của một hệ nhận dạng chữ viết tay	9
Hình 1.2. . Nhị phân hóa ảnh	10
Hình 1.3. Nhiễu đốm và nhiễu vệt	10
Hình 1.4. Chuẩn hóa kích thước ảnh các ký tự “A” và “P”	11
Hình 1.5. (a) Ảnh gốc, (b) Ảnh sau khi được làm trơn biên	11
Hình 1.6. Làm mảnh chữ.	12
Hình 1.7. Hiệu chỉnh độ nghiêng của văn bản	12
Hình 1.8. Tách dòng chữ dựa trên histogram theo chiều ngang của khối chữ	13
Hình 1.9. Xác định khoảng cách giữa hai ký tự và giữa hai từ dựa trên histogram theo chiều thẳng đứng của dòng chữ.....	14
Hình 2.1. Mô hình nhận dạng chữ viết tay rời rạc	22
Hình 2.2 Quá trình tìm giới hạn ký tự.....	37
Hình 2.3 Quá trình lấy mẫu xuống	38
Hình 2.4. Quá trình ánh xạ từ ma trận điểm sang ma trận giá trị.....	41
Hình 2.5. Giao diện chính của chương trình	41
Hình 2.6. Giao diện của chương trình sau khi nhận dạng chữ viết xong	41
Hình 2.7. Giao diện của ô thêm chữ viết.....	42
Hình 2.8. Giá trị của các ô vùng khi được nhận dạng	42

DANH MỤC CÁC KÝ HIỆU, CỤM TỪ VIẾT TẮT

HMM	Markov Model (Mô hình Markov ẩn)
off-line	ngoại tuyến
on-line	trực tuyến
OCR	Optical Character Recognition (nhận dạng chữ quang học)
SVM	Support Vector Machines (Máy véc tơ tựa)
SOM	Self Organizing Map
USPS	United States Postal service
MNIST	bộ mẫu chữ số viết tay NIST - Viện Công nghệ và Tiêu chuẩn Quốc gia Hoa Kỳ (National Institute of Standard and Technology of the United States)
SV	Support vector (véc tơ tựa)
working set	tập làm việc
k-NN	k – láng giềng gần nhất

PHẦN MỞ ĐẦU

Công nghệ thông tin ngày càng phát triển và có vai trò hết sức quan trọng không thể thiếu trong cuộc sống hiện đại. Con người ngày càng tạo ra những cỗ máy thông minh có khả năng tự nhận biết và xử lý được các công việc một cách tự động, phục vụ cho lợi ích của con người. Trong những năm gần đây, một trong những bài toán nhận được nhiều sự quan tâm và tốn nhiều công sức nhất của lĩnh vực công nghệ thông tin, đó chính là bài toán nhận dạng. Tuy mới xuất hiện chưa lâu nhưng nó đã rất được quan tâm vì tính ứng dụng thực tế của bài toán cũng như sự phức tạp của nó.

Nhận dạng chữ viết là một lĩnh vực đã được quan tâm nghiên cứu và ứng dụng từ nhiều năm nay theo hai hướng chính:

-Nhận dạng chữ in: phục vụ cho công việc tự động hóa đọc tài liệu, tăng tốc độ và hiệu quả nhập thông tin vào máy tính trực tiếp từ các nguồn tài liệu.

-Nhận dạng chữ viết tay: với những mức độ ràng buộc khác nhau về cách viết, kiểu chữ... phục vụ cho các ứng dụng đọc và xử lý chứng từ, hóa đơn, phiếu ghi, bản thảo viết tay... Nhận dạng chữ viết tay được tách thành hai hướng phát triển: nhận dạng chữ viết tay trực tuyến (on-line) và nhận dạng chữ viết tay ngoại tuyến (off-line).

Đến thời điểm này, bài toán nhận dạng chữ in đã được giải quyết gần như trọn vẹn. Tuy nhiên trên thế giới cũng như ở Việt Nam, bài toán nhận dạng chữ viết tay vẫn còn là vấn đề thách thức lớn đối với các nhà nghiên cứu. Bài toán này chưa thể giải quyết trọn vẹn vì nó phụ thuộc quá nhiều vào người viết và sự biến đổi quá đa dạng trong cách viết và trạng thái tinh thần của từng người viết. Đặc biệt đối với việc nghiên cứu nhận dạng chữ viết tiếng Việt lại càng gặp nhiều khó khăn hơn do bộ ký tự tiếng Việt có thêm phần dấu, rất dễ nhầm lẫn với các nhĩều.

Trích chọn đặc trưng trong nhận dạng chữ viết là việc tìm và chọn ra các đặc trưng đặc thù của mỗi chữ viết, qua đó là đầu vào cho quá trình nhận dạng. Bên cạnh việc lựa chọn một thuật toán nhận dạng phù hợp thì việc tìm ra phương pháp trích chọn đặc trưng thích hợp sẽ nâng cao độ chính xác và giảm thời gian nhận dạng cho toàn bộ hệ thống,

Với sự hấp dẫn của bài toán và những thách thức còn đang ở phía trước, với niềm đam mê công nghệ hiện đại và những ứng dụng thực tế tuyệt vời của nó, với khát khao khám phá và chinh phục những tri thức mới mẻ.. em đã chọn đề tài

nghiên cứu: Tìm hiểu một số phương pháp trích chọn đặc trưng cho nhận dạng chữ viết làm đề tài nghiên cứu và bảo vệ luận văn tốt nghiệp của mình.

- Về lý thuyết :
 - Tìm hiểu khái quát về chữ viết và lý thuyết nhận dạng.
 - Tìm hiểu về trích chọn đặc trưng cho nhận dạng chữ viết.
- Về thực tiễn :
 - Cài đặt thử nghiệm chương trình đã tìm hiểu được.
- Bài báo cáo được trình bày trong 3 chương:
 - Chương 1: Trình bày tổng quan về chữ viết và lý thuyết nhận dạng.
 - Chương 2: Tổng quan về trích chọn đặc trưng và một số phương pháp trích chọn đặc trưng cho nhận dạng chữ viết.
 - Chương 3: Chương trình thử nghiệm.
 - Kết luận: Tóm tắt những vấn đề tìm hiểu được trong bài, các vấn đề liên quan và đưa ra hướng phát triển trong tương lai.

TÓM TẮT ĐỀ TÀI

Bài báo cáo đồ án tốt nghiệp của em, nghiên cứu về “Một số phương pháp trích chọn đặc trưng cho nhận dạng chữ viết”. Nội dung nghiên cứu gồm 3 chương như sau:

- CHƯƠNG 1: Tổng quan về chữ viết và lý thuyết nhận dạng

Chương này nghiên cứu tổng quan về chữ viết và lý thuyết nhận dạng. Nhận dạng chữ là lĩnh vực được nhiều nhà nghiên cứu quan tâm và cho đến nay lĩnh vực này cũng đã đạt được nhiều thành tựu lớn lao cả về mặt lý thuyết lẫn ứng dụng thực tế. Đồng thời nêu mô hình tổng quát của một hệ nhận dạng chữ viết..

- CHƯƠNG 2: Tổng quan về trích chọn đặc trưng và một số phương pháp trích chọn đặc trưng cho nhận dạng chữ viết.

Chương này nghiên cứu tổng quan về trích chọn đặc trưng và một số phương pháp trích chọn đặc trưng cho nhận dạng chữ viết. Giới thiệu về trích chọn đặc trưng: Đặc trưng bất biến, khả năng khôi phục ; trích chọn đặc trưng từ ảnh đa cấp xám(Đối sánh mẫu, mẫu biến dạng, biến đổi ảnh Unitar, bất biến hình học, bất biến Zenite); trích chọn đặc trưng từ ảnh nhị phân (Đối sánh mẫu, mẫu biến dạng, biến đổi ảnh Unitar, bất biến hình học); trích chọn đặc trưng từ biên ảnh(Chiếu nghiêng, tách vùng, xấp xỉ đường cong, mô tả Fourier); trích chọn đặc trưng từ biểu diễn vectơ(Đối sánh mẫu, mẫu biến dạng, đặc trưng rời rạc, biểu diễn đồ thị, biểu diễn Fourier).

- CHƯƠNG 3 :Chương trình thử nghiệm.

CHƯƠNG 1. TỔNG QUAN VỀ CHỮ VIẾT VÀ LÝ THUYẾT NHẬN DẠNG

1.1. GIỚI THIỆU

Nhận dạng chữ là lĩnh vực được nhiều nhà nghiên cứu quan tâm và cho đến nay lĩnh vực này cũng đã đạt được nhiều thành tựu lớn lao cả về mặt lý thuyết lẫn ứng dụng thực tế. Lĩnh vực nhận dạng chữ được chia làm hai loại: Nhận dạng chữ in và nhận dạng chữ viết tay.

Đến thời điểm này, nhận dạng chữ in đã được giải quyết gần như trọn vẹn. Tuy nhiên, nhận dạng chữ viết tay vẫn đang là vấn đề thách thức lớn đối với các nhà nghiên cứu. Nhận dạng chữ viết tay được phân ra làm hai loại: *nhận dạng chữ viết tay on-line (trực tuyến)* và *nhận dạng chữ viết tay off-line (ngoại tuyến)*.

Nhận dạng chữ viết tay on-line được thực hiện trên cơ sở lưu lại các thông tin về nét chữ như thứ tự nét viết, hướng và tốc độ của nét viết trong quá trình nó đang viết. Đây chính là cơ sở để máy tính nhận diện được các chữ cái, do đó việc nhận dạng không gặp quá nhiều khó khăn. Ngược lại, đối với nhận dạng chữ viết tay off-line, dữ liệu đầu vào là ảnh văn bản được quét vào nên việc nhận dạng có độ khó cao hơn nhiều so với nhận dạng chữ viết tay on-line. Do dữ liệu đầu vào là ảnh văn bản nên nhận dạng chữ viết tay off-line và nhận dạng chữ in còn được gọi chung là nhận dạng chữ quang học (OCR - Optical Character Recognition).

Khó khăn lớn nhất khi nghiên cứu bài toán nhận dạng chữ viết tay là sự biến thiên quá đa dạng trong cách viết của từng người. Cùng một người viết nhưng đôi khi cũng có nhiều sự khác biệt trong cách viết tùy thuộc vào từng ngữ cảnh, kiểu viết của một người cũng có thể thay đổi theo thời gian hoặc theo thói quen... Điều này gây ra nhiều trở ngại trong việc trích chọn đặc trưng cũng như lựa chọn mô hình nhận dạng.

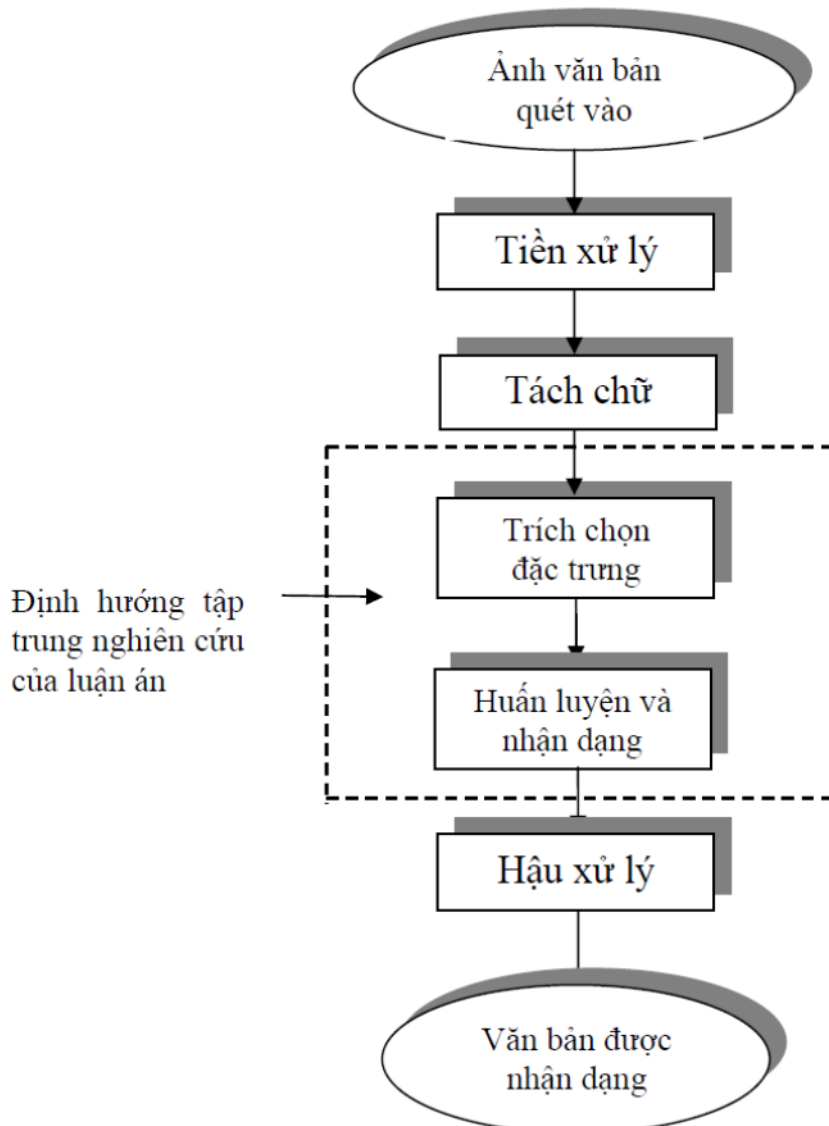
1.2. MÔ HÌNH TỔNG QUÁT CỦA MỘT HỆ NHẬN DẠNG CHỮ VIẾT TAY

1.2.1. Tiền xử lý

Giai đoạn này góp phần làm tăng độ chính xác phân lớp của hệ thống nhận dạng, tuy nhiên nó cũng làm cho tốc độ nhận dạng của hệ thống chậm lại. Vì vậy, tùy thuộc vào chất lượng ảnh quét vào của từng văn bản cụ thể để chọn một hoặc một vài chức năng trong khối này. Nếu cần ưu tiên tốc độ xử lý và chất lượng của

máy quét tốt thì có thể bỏ qua giai đoạn này. Khối tiền xử lý bao gồm một số chức năng:

Nhị phân hóa ảnh, lọc nhiễu, chuẩn hóa kích thước ảnh, làm trơn biên chữ, làm đầy chữ, làm mảnh chữ và xoay văn bản.

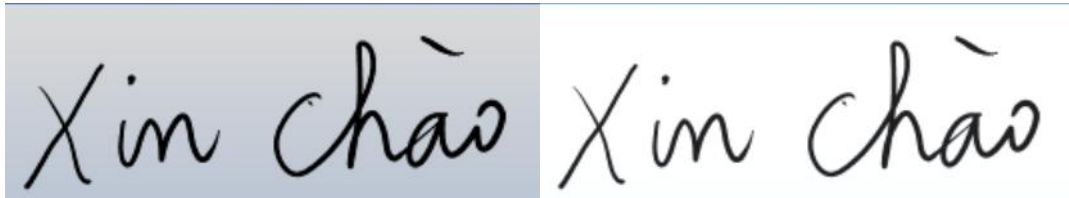


Hình 1.1. Sơ đồ tổng quát của một hệ thống nhận dạng chữ viết tay.

1.2.1.1. Nhị phân hóa ảnh

Nhị phân hóa ảnh là một kỹ thuật chuyển ảnh đa cấp xám sang ảnh nhị phân. Trong bất kỳ bài toán phân tích hoặc nâng cao chất lượng ảnh nào, nó cũng cần thiết để xác định các đối tượng quan trọng. Nhị phân hóa ảnh phân chia ảnh thành 2 phần: phần nền và phần chữ. Hầu hết các phương pháp nhị phân hóa ảnh hiện nay

đều lựa chọn một ngưỡng thích hợp theo cường độ sáng của ảnh và sau đó chuyển tất cả các giá trị độ sáng lớn hơn ngưỡng đó thành một giá trị độ sáng (ví dụ “trắng”) và tất cả các giá trị bé hơn ngưỡng thành một giá trị độ sáng khác (“đen”).



Hình 1.2. Nhị phân hóa ảnh.

1.2.1.2. Lọc nhiễu

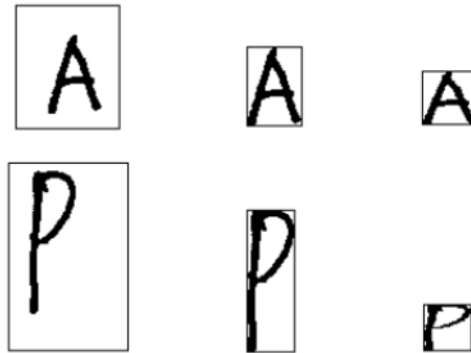
Nhiễu là một tập các điểm sáng thừa trên ảnh. Khử nhiễu là một vấn đề thường gặp trong nhận dạng, nhiễu có nhiều loại (nhiễu đốm, nhiễu vệt, nhiễu đứt nét...).



Hình 1.3. Nhiễu đốm và nhiễu vệt.

Để khử các nhiễu đốm (các nhiễu với kích thước nhỏ), có thể sử dụng các phương pháp lọc (lọc trung bình, lọc trung vị...). Tuy nhiên, với các nhiễu vệt (hoặc các nhiễu có kích thước lớn) thì các phương pháp lọc tỏ ra kém hiệu quả, trong trường hợp này sử dụng phương pháp khử các vùng liên thông nhỏ tỏ ra có hiệu quả hơn.

1.2.1.3. Chuẩn hóa kích thước ảnh



Hình 1.4. Chuẩn hóa kích thước ảnh các ký tự “A” và “P”.

Việc chuẩn hóa kích thước ảnh dựa trên việc xác định trọng tâm ảnh, sau đó xác định khoảng cách lớn nhất từ tâm ảnh đến các cạnh trên, dưới, trái, phải của hình chữ nhật bao quanh ảnh. Thông qua khoảng cách lớn nhất đó, có thể xác định được một tỷ lệ co, giãn của ảnh gốc so với kích thước đã xác định, từ đó hiệu chỉnh kích thước ảnh theo tỷ lệ co, giãn này. Như vậy, thuật toán chuẩn hóa kích thước ảnh luôn luôn đảm bảo được tính cân bằng khi co giãn ảnh, ảnh sẽ không bị biến dạng hoặc bị lệch.

1.2.1.4. Làm trơn biên chữ:

Đôi khi do chất lượng quét ảnh quá xấu, các đường biên của chữ không còn giữ được dáng điệu trơn tru ban đầu mà hình thành các đường răng cưa giả tạo. Trong các trường hợp này, phải dùng các thuật toán làm trơn biên để khắc phục.



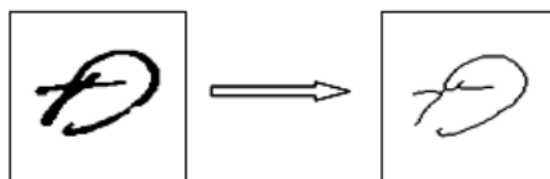
Hình 1.5. (a) Ảnh gốc, (b) Ảnh sau khi được làm trơn biên.

1.2.1.5. Làm đầy chữ

Chức năng này được áp dụng với các ký tự bị đứt nét một cách ngẫu nhiên. Ảnh đứt nét gây khó khăn cho việc tách chữ, dễ bị nhầm hai phần liên thông của ký tự thành hai ký tự riêng biệt, tạo nên sai lầm trong quá trình nhận dạng.

1.2.1.6. Làm mảnh chữ

Đây là một bước quan trọng nhằm phát hiện khung xương của ký tự bằng cách loại bỏ dần các điểm biên ngoài của các nét. Tuy nhiên, quá trình làm mảnh chữ rất nhạy cảm với việc khử nhiễu.



Hình 1.6. Làm mảnh chữ.

1.2.1.7. Điều chỉnh độ nghiêng của văn bản

Do trang tài liệu quét vào không cẩn thận hoặc do sự cố in ấn, các hàng chữ bị lệch so với lề chuẩn một góc α , điều này gây khó khăn cho công đoạn tách chữ, đôi khi không thể tách được. Trong những trường hợp như vậy, phải tính lại tọa độ điểm ảnh của các chữ bị sai lệch.

Có nhiều kỹ thuật để điều chỉnh độ nghiêng, kỹ thuật phổ biến nhất dựa trên cơ sở biểu đồ chiếu (projection profile) của ảnh tài liệu; một số kỹ thuật dựa trên cơ sở các phép biến đổi Hough và Fourier.



Hình 1.7. Hiệu chỉnh độ nghiêng của văn bản

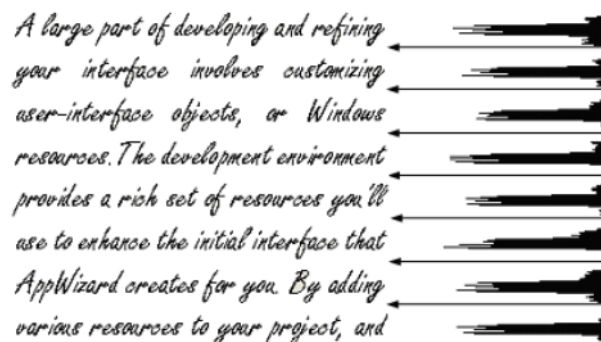
1.2.2. Khối tách chữ :

Khối này có nhiệm vụ tách từng ký tự ra khỏi văn bản. Chỉ khi nào văn bản được tách và cô lập đúng từng ký tự đơn ra khỏi tổng thể văn bản thì hệ thống mới có thể nhận dạng đúng ký tự đó. Sau đây là một số phương pháp tách chữ thông dụng:

1.2.2.1. Tách chữ theo chiều nằm ngang và thẳng đứng

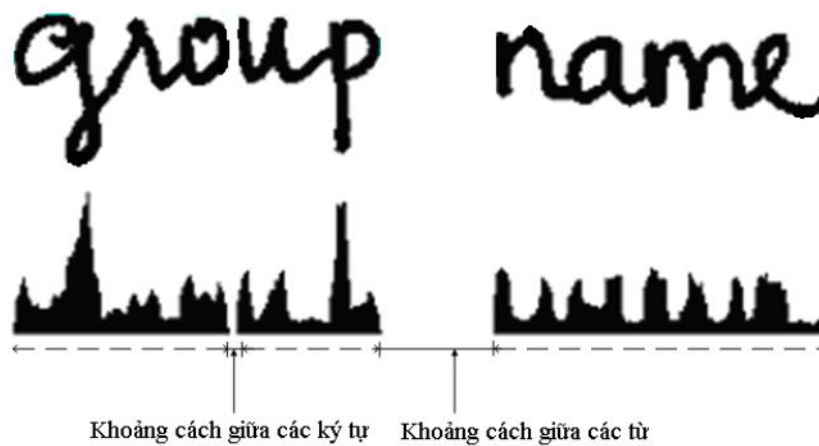
Phương pháp này thường áp dụng cho chữ in. Khác với chữ viết tay, kích thước và kiểu chữ cố định, phải tuân theo một số quy định in ấn, các chữ phải nằm gọn trong một khung nên việc cô lập một ký tự đơn có thể đồng nhất với việc tìm ra khung bao của chữ đó tại vị trí của nó trong văn bản. Tách chữ theo chiều nằm ngang và thẳng đứng là tìm một hình chữ nhật có cạnh thẳng đứng và nằm ngang chứa trọn một ký tự ở bên trong.

1.2.2.2. Tách chữ dùng lược đồ sáng



Hình 1.8. Tách dòng chữ dựa trên histogram theo chiều ngang của khối chữ

Đối với chữ viết tay thì việc tìm đường phân cách giữa các dòng và các ký tự trong văn bản thường rất khó khăn. Trong trường hợp này, không thể tìm đường phân cách theo nghĩa thông thường mà phải hiểu là đường phân cách với số điểm cắt hai dòng là ít nhất. Khi đó phải xây dựng lược đồ sáng của các dòng chữ, từ đó các đoạn thấp nhất trên lược đồ chính là đường phân cách cần tìm (hình 1.8 và 1.9).



Hình 1.9. Xác định khoảng cách giữa hai ký tự và giữa hai từ dựa trên histogram theo chiều thẳng đứng của dòng chữ.

1.2.3. Trích chọn đặc trưng:

Trích chọn đặc trưng đóng vai trò cực kỳ quan trọng trong một hệ thống nhận dạng. Trong trường hợp đơn giản nhất, ảnh đa cấp xám hoặc ảnh nhị phân được sử dụng cho việc nhận dạng. Tuy nhiên, trong hầu hết các hệ nhận dạng, để giảm độ phức tạp và tăng độ chính xác của các thuật toán phân lớp thì đòi hỏi các đặc trưng được trích chọn phải rút gọn lại càng nhỏ càng tốt nhưng vẫn phải đảm bảo được thông tin của ký tự. Với mục tiêu này, một tập các đặc trưng được trích chọn cho mỗi lớp sao cho có thể phân biệt được với các lớp khác.

1.2.4. Huấn luyện và nhận dạng :

Đây là giai đoạn quan trọng nhất, giai đoạn này quyết định độ chính xác của hệ thống nhận dạng. Có nhiều phương pháp phân lớp khác nhau được áp dụng cho các hệ thống nhận dạng chữ viết tay.

1.2.5. Hậu xử lý :

Đây là công đoạn cuối cùng của quá trình nhận dạng. Có thể hiểu hậu xử lý là bước ghép nối các ký tự đã nhận dạng thành các từ, các câu, các đoạn văn nhằm tái hiện lại văn bản đồng thời phát hiện ra các lỗi nhận dạng sai bằng cách kiểm tra chính tả dựa trên cấu trúc và ngữ nghĩa của các từ, các câu hoặc các đoạn văn. Việc phát hiện ra các lỗi, các sai sót trong nhận dạng ở bước này góp phần đáng kể vào việc nâng cao chất lượng nhận dạng.

Cách đơn giản nhất để kết nối các thông tin ngữ cảnh là tận dụng một từ điển để điều chỉnh các lỗi của hệ thống nhận dạng.

1.3. CÁC PHƯƠNG PHÁP NHẬN DẠNG CHỮ VIẾT TAY

Có nhiều phương pháp nhận dạng mẫu khác nhau được áp dụng rộng rãi trong các hệ thống nhận dạng chữ viết tay. Các phương pháp này có thể được tích hợp trong các hướng tiếp cận sau: Đối sánh mẫu, thống kê, cấu trúc, mạng nơ ron và SVM.

1.3.1. Đối sánh mẫu:

Kỹ thuật nhận dạng chữ đơn giản nhất dựa trên cơ sở đối sánh các nguyên mẫu (prototype) với nhau để nhận dạng ký tự hoặc từ. Nói chung, toán tử đối sánh xác định mức độ giống nhau giữa hai vé tơ (nhóm các điểm, hình dạng, độ cong...) trong một không gian đặc trưng. Các kỹ thuật đối sánh có thể nghiên cứu theo ba hướng sau:

Đối sánh trực tiếp: Một ký tự đầu vào là ảnh đa cấp xám hoặc ảnh nhị phân được so sánh trực tiếp với một tập mẫu chuẩn đã được lưu trữ. Việc so sánh dựa theo một độ đo về sự tương đồng nào đó (chẳng hạn như độ đo Euclide) để nhận dạng. Các kỹ thuật đối sánh này có thể đơn giản như việc so sánh một – một hoặc phức tạp hơn như phân tích cây quyết định. Mặc dù phương pháp đối sánh trực tiếp đơn giản và có một cơ sở toán học vững chắc nhưng kết quả nhận dạng của nó cũng rất nhạy cảm với nhiễu.

Các mẫu biến dạng và Đối sánh mềm: Một phương pháp đối sánh khác là sử dụng các mẫu biến dạng, trong đó một phép biến dạng ảnh được dùng để đối sánh một ảnh chưa biết với một cơ sở dữ liệu ảnh đã biết.

Ý tưởng cơ bản của đối sánh mềm là đối sánh một cách tối ưu mẫu chưa biết với tất cả các mẫu có thể mà các mẫu này có thể kéo giãn ra hoặc co lại. Chỉ một không gian đặc trưng được thành lập, các véc tơ chưa biết được đối sánh bằng cách sử dụng quy hoạch động và một hàm biến dạng.

Đối sánh giảm nhẹ: Đây là một kỹ thuật đối sánh ảnh ở mức độ tượng trưng, kỹ thuật này sử dụng hình dáng đặc trưng cơ bản của ảnh ký tự. Thứ nhất, các vùng đối sánh đã được nhận biết. Sau đó, trên cơ sở một số vùng đối sánh được đánh giá tốt, các phần tử của ảnh được so sánh với các vùng đối sánh này. Công việc này đòi hỏi một kỹ thuật tìm kiếm trong một không gian đa chiều để tìm cực đại toàn cục của một số hàm.

Các kỹ thuật đối sánh mẫu chỉ áp dụng tốt đối với nhận dạng chữ in, còn đối với chữ viết tay thì các kỹ thuật này tỏ ra kém hiệu quả.

1.3.2. Phương pháp tiếp cận cấu trúc:

Cách tiếp cận của phương pháp này dựa vào việc mô tả đối tượng nhờ một số khái niệm biểu diễn 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,... Mỗi đối tượng được mô tả như một sự kết hợp của các dạng nguyên thủy.

Các quy tắc kết hợp các dạng nguyên thủy được xây dựng giống như việc nghiên cứu văn phạm trong một ngôn ngữ, do đó quá trình quyết định nhận dạng là quá trình phân tích cú pháp. Phương pháp này đặt vấn đề để giải quyết bài toán nhận dạng chữ tổng quát. Tuy vậy, cho đến nay còn nhiều vấn đề liên quan đến hệ nhận dạng cú pháp chưa được giải quyết độc lập và chưa xây dựng được các thuật toán phổ dụng. Hiện nay, nhận dạng theo cấu trúc phổ biến là trích chọn các đặc trưng của mẫu học, phân hoạch bằng ký tự dựa trên các đặc trưng này, sau đó ảnh cần nhận dạng sẽ được trích chọn đặc trưng, sau đó so sánh trên bảng phân hoạch để tìm ra ký tự có các đặc trưng phù hợp.

Đối với nhận dạng chữ viết tay rời rạc dựa theo cấu trúc xương và đường biên, công việc này đòi hỏi phải xây dựng các đặc trưng của chữ, đặc biệt là đặc trưng về các điểm uốn, điểm gấp khúc và đặc trưng của các nét. Sau khi tiến hành công đoạn tiền xử lý, công việc tách các nét được tiến hành thông qua các điểm chạc. Sau đó trích chọn đặc trưng cấu trúc xương của chữ, mỗi nét đặc trưng bởi cặp chỉ số đầu và cuối tương ứng với thứ tự của điểm chạc đầu và điểm chạc cuối. Cuối cùng là xây dựng cây tìm kiếm, dựa vào đặc trưng về cấu trúc xương và cấu trúc biên để phân tập mẫu học thành các lớp. Quá trình tìm kiếm để phân lớp được tiến hành qua hai bước: Xác định lớp tương ứng với mẫu vào và tìm kiếm trong lớp đó mẫu nào gần giống với mẫu vào nhất.

Các phương pháp cấu trúc áp dụng cho các bài toán nhận dạng chữ được phát triển theo hai hướng sau:

1.3.2.1. Phương pháp ngữ pháp (Grammatical Methods)

Giữa thập niên 1960, các nhà nghiên cứu bắt đầu xét các luật của ngôn ngữ học để phân tích tiếng nói và chữ viết. Sau đó, các luật đa dạng của chính tả, từ vựng và ngôn ngữ học đã được áp dụng cho các chiến lược nhận dạng. Các phương pháp ngữ pháp khởi tạo một số luật sinh để hình thành các ký tự từ một tập các công

thức ngữ pháp nguyên thủy. Các luật sinh này có thể kết nối bất kỳ kiểu đặc trưng thống kê và đặc trưng hình thái nào dưới một số cú pháp hoặc các luật ngữ nghĩa. Giống như lý thuyết ngôn ngữ, các luật sinh cho phép mô tả các cấu trúc câu có thể chấp nhận được và trích chọn thông tin theo ngữ cảnh về chữ viết bằng cách sử dụng các kiểu ngữ pháp khác nhau.

Trong các phương pháp này, việc huấn luyện được thực hiện bằng cách mô tả mỗi ký tự bằng một văn phạm Gi. Còn trong pha nhận dạng thì chuỗi, cây hoặc đồ thị của một đơn vị viết bất kỳ (ký tự, từ hoặc câu) được phân tích để quyết định văn phạm của mẫu đó thuộc lớp nào. Các phương pháp ngữ pháp hầu hết được sử dụng trong giai đoạn hậu xử lý để sửa các lỗi mà khối nhận dạng đã thực hiện sai.

1.3.2.2. Phương pháp đồ thị (Graphical Methods)

Các đơn vị chữ viết được mô tả bởi các cây hoặc các đồ thị. Các dạng nguyên thủy của ký tự (các nét) được lựa chọn bởi một hướng tiếp cận cấu trúc. Đối với mỗi lớp, một đồ thị hoặc cây được thành lập trong giai đoạn huấn luyện để mô tả các nét, các ký tự hoặc các từ. Giai đoạn nhận dạng gán một đồ thị chưa biết vào một trong các lớp bằng cách sử dụng một độ đo để so sánh các đặc điểm giống nhau giữa các đồ thị.

Có rất nhiều hướng tiếp cận khác nhau sử dụng phương pháp đồ thị, tiêu biểu là hướng tiếp cận đồ thị phân cấp được dùng trong việc nhận dạng chữ viết tay Trung Quốc và Hàn Quốc.

1.3.3. Mạng nơ ron

Một mạng nơ ron được định nghĩa như một cấu trúc tính toán bao gồm nhiều bộ xử lý “nơ ron” được kết nối song song chằng chịt với nhau. Do bản chất song song của các nơ ron nên nó có thể thực hiện các tính toán với tốc độ cao hơn so với các kỹ thuật phân lớp khác. Một mạng nơ ron chứa nhiều nút, đầu ra của một nút được sử dụng cho một nút khác ở trong mạng và hàm quyết định cuối cùng phụ thuộc vào sự tương tác phức tạp giữa các nút. Mặc dù nguyên lý khác nhau, nhưng hầu hết các kiến trúc mạng nơ ron đều tương đương với các phương pháp nhận dạng mẫu thống kê.

Các kiến trúc mạng nơ ron có thể được phân thành hai nhóm chính: mạng truyền thẳng và mạng lan truyền ngược. Trong các hệ thống nhận dạng chữ, các mạng nơ ron sử dụng phổ biến nhất là mạng perceptron đa lớp thuộc nhóm mạng

truyền thẳng và mạng SOM (Self Organizing Map) của Kohonen thuộc nhóm mạng lan truyền ngược.

Mạng perceptron đa lớp được đề xuất bởi Rosenblatt được nhiều tác giả sử dụng trong các hệ nhận dạng chữ viết tay. Hầu hết các nghiên cứu phát triển nhận dạng chữ viết tay hiện nay đều tập trung vào mạng SOM. SOM kết hợp trích chọn đặc trưng và nhận dạng trên một tập lớn các ký tự huấn luyện. Mạng này chứng tỏ rằng nó tương đương với thuật toán phân cụm k-means.

Với thuật toán đơn giản nhưng rất hiệu quả, cùng với thành công của mô hình này trong các ứng dụng thực tiễn, mạng nơ ron hiện đang là một trong các hướng nghiên cứu của lĩnh vực học máy. Mạng nơ ron tỏ ra phù hợp với các bài toán đối sánh, phân loại mẫu, xấp xỉ hàm, tối ưu hoá, lượng tử hoá véc tơ và phân hoạch không gian dữ liệu, trong khi các phương pháp truyền thống không đủ khả năng giải quyết các vấn đề nêu trên một cách hiệu quả. Đặc biệt trong các hệ thống nhận dạng sử dụng mạng nơ ron đã đạt được tỉ lệ nhận dạng khá chính xác, có thể so sánh với các phương pháp nhận dạng cấu trúc, thống kê, ...

1.3.3.1. Đặc trưng của mạng nơron

Tính phi tuyến

Một nơron có thể tính toán một cách tuyến tính hay phi tuyến. Một mạng nơron cấu thành bởi các nơron phi tuyến thì cũng sẽ có tính phi tuyến. Hơn nữa, điều đặc biệt là tính phi tuyến này được phân tán trên toàn mạng. Tính phi tuyến là một thuộc tính rất quan trọng, nhất là khi các cơ chế vật lý sinh ra các tín hiệu đầu vào (ví dụ tín hiệu tiếng nói) vốn là phi tuyến.

1.3.3.1.1. Tính chất tương ứng đầu vào, đầu ra

Mặc dù khái niệm “học” (hay “huấn luyện”) chưa được bàn đến nhưng để hiểu được mối quan hệ đầu vào-đầu ra của mạng nơron, chúng ta sẽ đề cập sơ qua về khái niệm này. Một mô hình học phổ biến được gọi là “học có thầy” (hay “học có giám sát”), liên quan đến việc thay đổi các trọng số liên kết của mạng nơron bằng việc áp dụng một tập hợp các mẫu tích lũy. Mỗi một mẫu bao gồm một tín hiệu đầu vào và một đầu ra mong muốn tương ứng. Các trọng số liên kết của mạng được biến đổi sao cho có thể cực tiểu hoá sự sai khác giữa đầu ra mong muốn và đầu ra thực sự của mạng theo một tiêu chuẩn thống kê thích hợp. Sự tích lũy của mạng được lặp lại với nhiều ví dụ trong tập hợp cho tới khi mạng đạt tới một trạng thái ổn định mà ở đó không có một sự thay đổi đáng kể nào của các trọng số liên kết. Các ví dụ tích lũy được áp dụng trước có thể được áp dụng lại trong thời gian

của phiên tích lũy nhưng theo một thứ tự khác. Như vậy, mạng nơron học từ các ví dụ bằng cách xây dựng nên một tương ứng đầu vào - đầu ra cho vấn đề cần giải quyết.

1.3.3.1.2. Tính chất thích nghi

Các mạng nơron có một khả năng mặc định là có thể biến đổi các trọng số liên kết tùy theo sự thay đổi của môi trường xung quanh. Đặc biệt, một mạng nơron đã được tích lũy để hoạt động trong một môi trường xác định có thể được tích lũy lại một cách dễ dàng khi có những thay đổi nhỏ của các điều kiện môi trường.

1.3.3.1.3. Tính chất đưa ra lời giải có bằng chứng

Trong ngữ cảnh phân loại mẫu, một mạng nơron có thể được thiết kế để đưa ra thông tin không chỉ về mẫu được phân loại, mà còn về sự tin cậy của quyết định đã được thực hiện. Thông tin này có thể được sử dụng để loại bỏ các mẫu mơ hồ hay nhập nhằng.

1.3.3.1.4. Tính chất chấp nhận sai sót

Một mạng nơron, được cài đặt dưới dạng phần cứng, vốn có khả năng chấp nhận lỗi, hay khả năng tính toán thô (chứ không nhạy cảm lỗi). Tức là, tính năng của mạng nơron chỉ thoái hoá (chứ không đổ vỡ) khi có những điều kiện hoạt động bất lợi. Ví dụ, nếu một nơron hay các liên kết kết nối của nó bị hỏng, việc nhận dạng lại một mẫu được lưu trữ sẽ suy giảm về chất lượng. Tuy nhiên, do bản chất phân tán của thông tin lưu trữ trong mạng nơron, sự hỏng hóc cũng được trải ra trên toàn mạng. Như vậy, về cơ bản, trong trường hợp này một mạng nơron sẽ thể hiện một sự thoái hoá về tính năng hơn là sự đổ vỡ trầm trọng.

1.3.3.1.5. Khả năng cài đặt VLSI (Very-Large-Scale-Intergrated)

Bản chất song song đồ sộ của một mạng nơron làm cho mạng nơron đó rất nhanh trong tính toán đối với một số công việc. Đặc tính này cũng làm cho một mạng nơron có khả năng phù hợp trong việc cài đặt sử dụng kỹ thuật Very -large-scale-intergrated (VLSI). Kỹ thuật này cho phép xây dựng những mạch cứng tính toán song song quy mô lớn. Chính vì vậy, ưu điểm nổi bật của VLSI là mang lại những phương tiện hữu hiệu để có thể xử lý được những hành vi có độ phức tạp cao.

1.3.3.1.6. Tính chất đồng dạng trong phân tích và thiết kế

Về cơ bản, các mạng nơron có tính chất chung như là các bộ xử lý thông tin. Điều này áp dụng cho tất cả các lĩnh vực có liên quan tới việc ứng dụng mạng nơron. Đặc tính này thể hiện ở một số điểm như sau:

- Các nơron, dưới dạng này hoặc dạng khác, biểu diễn một thành phần chung cho tất cả các mạng nơron.
- Tính thống nhất này đem lại khả năng chia sẻ các lý thuyết và các thuật toán học trong nhiều ứng dụng khác nhau của mạng nơron.
- Các mạng tổ hợp (modular) có thể được xây dựng thông qua một sự tích hợp các mô hình khác nhau.

1.3.3.1.7. Một số ứng dụng chính của mạng nơron

- Aerospace: Phi công tự động, giả lập đường bay, các hệ thống điều khiển lái máy bay, bộ phát hiện lỗi.
- Automotive: Các hệ thống dẫn đường tự động cho ô tô, các bộ phân tích hoạt động của xe.
- Banking: Bộ đọc séc và các tài liệu, tính tiền của thẻ tín dụng.
- Defense: Định vị - phát hiện vũ khí, dò mục tiêu, phát hiện đối tượng, nhận dạng nét mặt, các bộ cảm biến thế hệ mới, xử lý ảnh radar,...
- Electronics: Dự đoán mã tuần tự, sơ đồ chip IC, điều khiển tiến trình, phân tích nguyên nhân hỏng chip, nhận dạng tiếng nói, mô hình phi tuyến.
- Entertainment: Hoạt hình, các hiệu ứng đặc biệt, dự báo thị trường.
- Financial: Định giá bất động sản, cho vay, kiểm tra tài sản cầm cố, đánh giá mức độ hợp tác, phân tích đường tín dụng, chương trình thương mại qua giấy tờ, phân tích tài chính liên doanh, dự báo tỷ giá tiền tệ.
- Insurance: Đánh giá việc áp dụng chính sách, tối ưu hóa sản phẩm.

1.3.4. Mô hình Markov ẩn (HMM - Hidden Markov Model) :

HMM là một mô hình xác suất hữu hạn trạng thái theo kiểu phát sinh tiến trình bằng cách định nghĩa xác suất liên kết trên các chuỗi quan sát. Mỗi chuỗi quan sát được sinh ra bởi một chuỗi các phép chuyển trạng thái, bắt đầu từ trạng thái khởi đầu cho đến khi thu được trạng thái kết thúc. Tại mỗi trạng thái thì một phần tử của chuỗi quan sát được phát sinh ngẫu nhiên trước khi chuyển sang trạng thái tiếp theo.

1.3.5. Máy véc tơ tựa (SVM) :

1.3.5.1. Giới thiệu:

Cho đến nay, việc nhận dạng chữ viết tay vẫn chưa có được một giải pháp tổng thể, các ứng dụng của nó cũng chỉ giới hạn trong phạm vi hẹp. Các kết quả chủ yếu về lĩnh vực này chỉ tập trung trên các tập dữ liệu chữ số viết tay chuẩn như USPS và MNIST, bên cạnh đó cũng có một số công trình nghiên cứu trên các hệ chữ cái tiếng La tinh, Hy Lạp, Trung Quốc, Việt Nam... tuy nhiên các kết quả đạt được cũng còn nhiều hạn chế.

Các giải pháp tiếp cận để giải bài toán nhận dạng chữ viết tay khá phong phú, một số phương pháp học máy thường được áp dụng như: mô hình Markov ẩn, mạng nơ ron hay phương pháp máy véc tơ tựa (SVM - Support Vector Machines). Trong đó SVM được đánh giá là phương pháp học máy tiên tiến đang được áp dụng rộng rãi trong các lĩnh vực khai phá dữ liệu và thị giác máy tính... SVM gốc được thiết kế để giải bài toán phân lớp nhị phân, ý tưởng chính của phương pháp này là tìm một siêu phẳng phân cách sao cho khoảng cách lề giữa hai lớp đạt cực đại. Khoảng cách này được xác định bởi các véc tơ tựa (SV - Support Vector), các SV này được lọc ra từ tập mẫu huấn luyện bằng cách giải một bài toán tối ưu lồi.

1.3.5.2. Mô hình nhận dạng chữ viết tay rời rạc:

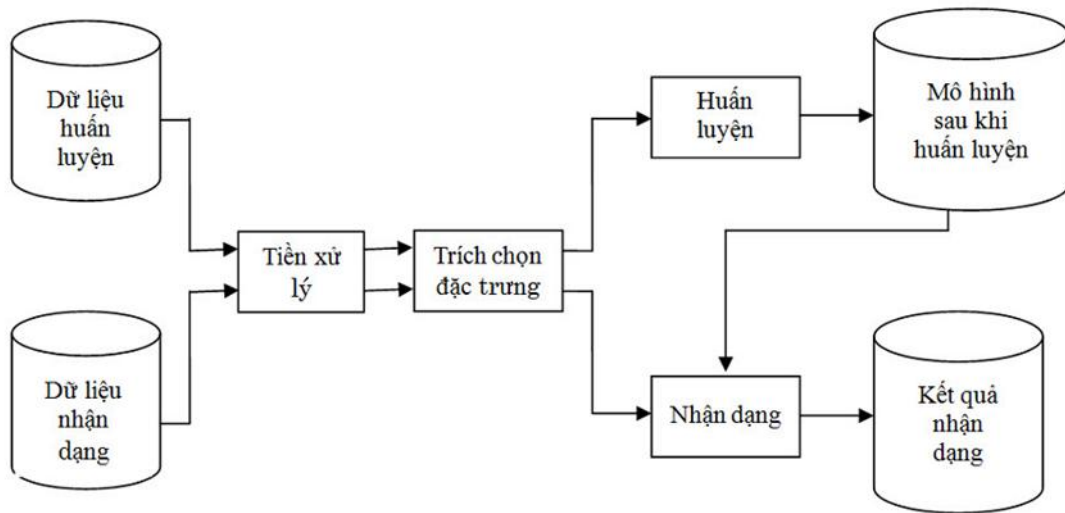
Trong phần này, sẽ tập trung xây dựng mô hình nhận dạng chữ viết tay rời rạc theo phương pháp phân lớp SVM. Công việc được thực hiện theo hai bước chính sau đây:

Bước 1: Xây dựng mô hình huấn luyện.

Tập dữ liệu huấn luyện sau khi qua các khâu tiền xử lý và trích chọn đặc trưng sẽ được đưa vào máy huấn luyện phân lớp SVM. Sau khi kết thúc quá trình huấn luyện, hệ thống sẽ lưu lại giá trị các tham số của hàm quyết định phân lớp để phục vụ cho việc nhận dạng sau này. Quá trình huấn luyện tiêu tốn khá nhiều thời gian, tốc độ huấn luyện nhanh hay chậm tùy thuộc vào từng thuật toán huấn luyện, chiến lược phân lớp SVM cũng như số lượng mẫu tham gia huấn luyện.

Bước 2: Phân lớp nhận dạng.

Dựa vào giá trị các tham số của hàm quyết định thu được ở Bước 1, một mẫu mới x sau khi đã qua các khâu tiền xử lý và trích chọn đặc trưng sẽ được đưa vào tính toán thông qua hàm quyết định để xác định lớp của mẫu x (Hình 2.1).



Hình 2.1. Mô hình nhận dạng chữ viết tay rời rạc.

1.3.5.3. Đánh giá hiệu quả phân lớp SVM

SVM là một phương pháp học máy tiên tiến có cơ sở toán học chặt chẽ và đạt độ chính xác phân lớp cao. Tuy nhiên, hạn chế lớn nhất của SVM là tốc độ phân lớp chậm, tùy thuộc vào số lượng vectơ tựa thu được sau khi huấn luyện. Một hạn chế khác của SVM là pha huấn luyện đòi hỏi không gian nhớ lớn, vì vậy việc huấn luyện đối với các bài toán có số lượng mẫu lớn sẽ gặp trở ngại trong vấn đề lưu trữ.

Bản chất nhị phân cũng là một hạn chế của SVM, việc mở rộng khả năng của SVM để giải quyết các bài toán phân loại nhiều lớp là vấn đề không đơn giản. Có nhiều chiến lược được đề xuất để mở rộng SVM cho bài toán phân loại nhiều lớp với những điểm mạnh, yếu khác nhau tùy thuộc vào từng loại dữ liệu cụ thể. Cho đến nay, việc lựa chọn các chiến lược phân lớp vẫn thường được tiến hành trên cơ sở thực nghiệm.

Bài toán huấn luyện SVM thực chất là bài toán qui hoạch toàn phương (QP) trên một tập lồi, do đó luôn luôn tồn tại nghiệm toàn cục và duy nhất, đây là điểm khác biệt rõ nhất giữa SVM so với mạng nơ ron, vì mạng nơ ron vốn tồn tại nhiều

cực trị địa phương. Bản chất của SVM là việc phân lớp được thực hiện gián tiếp trong không gian đặc trưng với số chiều cao hơn số chiều của không gian đầu vào thông qua hàm nhân. Do đó, hiệu quả phân lớp của SVM phụ thuộc vào hai yếu tố: giải bài toán QP và lựa chọn hàm nhân. Việc giải bài toán QP luôn luôn đạt được giải pháp tối ưu nên mọi cố gắng trong nghiên cứu lý thuyết SVM tập trung vào việc lựa chọn hàm nhân. Lựa chọn hàm nhân và các tham số của nó như thế nào để SVM phân lớp tốt nhất vẫn là một bài toán mở.

Tốc độ phân lớp của SVM bị đánh giá là chậm so với các phương pháp phân lớp khác, tùy thuộc vào số lượng vectơ tựa thu được sau khi huấn luyện. Vì vậy, có nhiều công trình tập trung nghiên cứu để giảm tối đa số lượng vectơ tựa nhằm tăng tốc độ phân lớp của SVM, một số kết quả nghiên cứu có giá trị về SVM đã được công bố trong các công trình .

Muốn áp dụng kỹ thuật phân lớp SVM vào bài toán nhận dạng chữ viết tay tiếng Việt, cần phải có những giải pháp để tránh bùng nổ số phân lớp cũng như giảm tối đa số vectơ tựa để tăng tốc độ nhận dạng.

1.3.6. Kết hợp các kỹ thuật nhận dạng :

Các phần đã trình bày ở trên cho thấy rằng có nhiều phương pháp phân lớp có thể áp dụng đối với các hệ nhận dạng chữ viết tay. Tất cả các phương pháp trên đều có những ưu điểm và nhược điểm riêng. Vấn đề đặt ra là các phương pháp trên có thể kết hợp với nhau theo một cách nào đó để nâng cao chất lượng nhận dạng hay không? Nhiều công trình nghiên cứu kiến trúc phân lớp theo ý tưởng kết hợp các phương pháp phân lớp đã nêu trên. Các hướng tiếp cận kiến trúc kết hợp để phân lớp có thể chia thành ba nhóm sau: Kiến trúc tuần tự, kiến trúc song song và kiến trúc lai ghép.

1.3.6.1. Kiến trúc tuần tự

Kiến trúc này chuyển kết quả đầu ra của một máy phân lớp thành đầu vào của máy phân lớp tiếp theo. Có bốn chiến lược cơ bản được sử dụng trong kiến trúc tuần tự, đó là dãy, chọn lựa, boosting và thác nước.

Trong chiến lược về dãy, mục tiêu của mỗi giai đoạn là thu gọn số lớp mà mẫu đầu vào có thể thuộc về các lớp đó. Số lớp có thể thu gọn tại mỗi giai đoạn sinh ra nhãn của mẫu ở giai đoạn cuối cùng .

Trong chiến lược chọn lựa, đầu tiên máy phân lớp gán mẫu chưa biết vào một nhóm ký tự gần giống nhau. Các nhóm này tiếp tục được phân lớp ở các giai

đoạn sau đó theo một cây phân cấp. Tại mỗi mức của cây, nhánh con cùng mẹ là giống nhau theo một độ đo nào đó. Vì vậy, các máy phân lớp thực hiện phân lớp từ thô đến tinh dần trong các nhóm nhỏ .

Đối với chiến lược boosting, mỗi máy phân lớp điều khiển một số lớp, các máy phân lớp ở phía trước không thể điều khiển được các lớp của các máy phân lớp ở phía sau .

Cuối cùng, trong chiến lược thác nước, các máy phân lớp được kết nối từ đơn giản đến phức tạp. Các mẫu không thỏa mãn ở một mức độ tin cậy nào đó thì phải thông qua một máy phân lớp mạnh hơn trong một giới hạn nào đó của các đặc trưng hoặc các chiến lược nhận dạng khác .

1.3.6.2. Kiến trúc song song

Kiến trúc này kết nối kết quả của các thuật toán phân lớp độc lập bằng cách sử dụng nhiều phương pháp khác nhau. Trong số các kiến trúc này, tiêu biểu nhất là phương pháp bỏ phiếu và luật quyết định Bayes.

1.3.6.3. Kiến trúc lai ghép

Kiến trúc này là một sự lai ghép giữa hai kiến trúc tuần tự và song song. Ý tưởng chính là kết hợp các điểm mạnh của cả hai kiến trúc trên và chặn bớt những khó khăn trong việc nhận dạng chữ viết.

1.3.7. Kết luận

Chương này đã giới thiệu một cách tổng quan về lĩnh vực nhận dạng chữ viết. Cho đến nay các kết quả nghiên cứu nhận dạng chữ viết tay vẫn còn hạn chế, các ứng dụng chủ yếu chỉ tập trung ở một số lĩnh vực hẹp. Đặc biệt có rất ít kết quả liên quan đến nhận dạng chữ viết tay tiếng Việt, các kết quả nghiên cứu cũng chỉ tập trung vào chữ Việt viết tay on-line .

Có nhiều kỹ thuật tiên tiến đang được áp dụng cho bài toán nhận dạng chữ viết tay như HMM, mạng nơ ron, k-láng giềng gần nhất, luật quyết định Bayes, SVM... Trong số các kỹ thuật này thì SVM được đánh giá là phương pháp có độ chính xác phân lớp cao và phương pháp luận của nó được xây dựng dựa trên một nền tảng toán học rất chặt chẽ.

CHƯƠNG 2: TỔNG QUAN VỀ TRÍCH CHỌN ĐẶC TRƯNG VÀ MỘT SỐ PHƯƠNG PHÁP TRÍCH CHỌN ĐẶC TRƯNG TRONG NHẬN DẠNG CHỮ VIẾT

2.1. Trích chọn đặc trưng

Trích chọn đặc trưng đóng vai trò cực kỳ quan trọng trong một hệ thống nhận dạng. Trong trường hợp đơn giản nhất, ảnh đa cấp xám hoặc ảnh nhị phân được sử dụng cho việc nhận dạng. Tuy nhiên, trong hầu hết các hệ nhận dạng, để giảm độ phức tạp và tăng độ chính xác của các thuật toán phân lớp thì đòi hỏi các đặc trưng được trích chọn phải rút gọn lại càng nhỏ càng tốt nhưng vẫn phải đảm bảo được thông tin của ký tự. Với mục tiêu này, một tập các đặc trưng được trích chọn cho mỗi lớp sao cho có thể phân biệt được với các lớp khác. Có hàng trăm phương pháp trích chọn đặc trưng cho ảnh văn bản, nhưng chung quy lại, các phương pháp này được gom lại thành ba nhóm chính sau:

2.1.1. Biến đổi toàn cục và khai triển chuỗi

Một tín hiệu liên tục thường chứa nhiều thông tin và chúng có thể sử dụng làm các đặc trưng cho mục đích phân lớp. Các đặc trưng được trích chọn cũng có thể đúng đối với việc xấp xỉ các tín hiệu liên tục thành các tín hiệu rời rạc. Một cách dễ

biểu diễn một tín hiệu là sử dụng một tổ hợp tuyến tính của một dãy các hàm đơn giản hơn. Các hệ số của tổ hợp tuyến tính cung cấp một tri thức giải mã vừa đủ, chẳng hạn như các phép biến đổi hoặc khai triển chuỗi. Một số biến dạng khác như các phép dịch chuyển và phép quay là bất biến dưới các phép biến đổi toàn cục và khai triển chuỗi. Sau đây là một số phương pháp biến đổi và khai triển chuỗi thường được áp dụng trong lĩnh vực nhận dạng chữ:

Biến đổi Fourier: Một trong những tính chất nổi bật nhất của phép biến đổi Fourier là khả năng nhận dạng các ký tự có sự thay đổi về các tư thế khác nhau, các phép biến đổi này đã được áp dụng để nhận dạng ký tự theo nhiều cách khác nhau .

Biến đổi Wavelet: Phép biến đổi này là một dãy các kỹ thuật khai triển cho phép mô tả đặc trưng của ảnh ở các mức độ khác nhau. Các công đoạn tách chữ

thành các ký tự hoặc từ được mô tả bằng các hệ số wavelet theo các mức độ khác nhau đối với từng giải pháp. Sau đó các hệ số wavelet được chuyển qua một máy phân lớp để phục vụ cho việc nhận dạng .

Phương pháp mô men: Theo phương pháp này, ảnh gốc sẽ được thay thế bằng một tập các đặc trưng vừa đủ của để nhận dạng các đối tượng bất biến đối với các phép thay đổi tỷ lệ, tịnh tiến hoặc quay. Các mô men được xét như các dãy khai triển đặc trưng vì ảnh gốc có thể xây dựng lại một cách đầy đủ từ các hệ số mô men.

Khai triển Karhunen-Loeve: Việc khai triển này nhằm phân tích các véc tơ riêng để rút gọn số chiều của tập đặc trưng bằng cách tạo ra các đặc trưng mới là tổ hợp tuyến tính của các đặc trưng gốc. Đây chỉ là một phép biến đổi tối ưu trong một số giới hạn nào đó của việc nén thông tin. Khai triển Karhunen-Loeve được dùng trong một số bài toán nhận dạng mẫu như nhận dạng mặt người, nó cũng được sử dụng trong hệ thống OCR của Viện Công nghệ và Tiêu chuẩn Quốc gia Hoa Kỳ (NIST – National Institute of Standards and Technology of the United States). Vì việc khai triển này đòi hỏi phải sử dụng các thuật toán có khối lượng tính toán rất lớn nên việc sử dụng các đặc trưng Karhunen-Loeve trong các bài toán nhận dạng chữ không được phổ biến rộng rãi. Tuy nhiên, để tăng tốc độ tính toán cho các máy phân lớp, các đặc trưng này trở nên thiết thực hơn cho các hệ nhận dạng chữ trong những năm gần đây.

2.1.2. Đặc trưng thống kê

Các đặc trưng thống kê của ảnh văn bản bảo toàn các kiểu biến đổi đa dạng về hình dáng của chữ. Mặc dù các kiểu đặc trưng này không thể xây dựng lại ảnh gốc, nhưng nó được sử dụng để thu nhỏ số chiều của tập đặc trưng nhằm tăng tốc độ và giảm thiểu độ phức tạp tính toán. Sau đây là một số đặc trưng thống kê thường dùng để biểu diễn ảnh ký tự:

Phân vùng (zoning): Khung chứa ký tự được chia thành một vài vùng chồng nhau hoặc không chồng nhau. Mật độ của các điểm ảnh trong các vùng khác nhau được phân tích và tạo thành các đặc trưng .

Các giao điểm và khoảng cách: Một đặc trưng thống kê phổ biến là số giao điểm giữa chu tuyến của chữ với một đường thẳng theo một hướng đặc biệt nào đó.

Các phép chiếu: Các ký tự có thể được biểu diễn bằng cách chiếu các giá trị mức xám của từng điểm lên trên các dòng theo các hướng khác nhau. Các đặc trưng này tạo ra dãy tín hiệu một chiều từ ảnh hai chiều.

Đặc trưng hướng: Các ký tự bao gồm các nét chữ, các nét này là các đoạn thẳng có hướng, các cung hoặc các đường cong. Hướng của các nét đóng vai trò quan trọng trong việc so sánh sự khác nhau giữa các ký tự. Các ký tự được mô tả như các véc tơ mà các phần tử của nó là các giá trị thống kê về hướng. Để trích chọn các đặc trưng này, góc định hướng của nét chữ phải được phân chia thành một số vùng cố định và số các đoạn của nét chữ trong mỗi vùng góc được chọn như một giá trị đặc trưng. Vì vậy, tập các số lượng của các đoạn định hướng sẽ tạo thành một biểu đồ được gọi là biểu đồ hướng và các đặc trưng về biểu đồ hướng có thể gọi chung là đặc trưng hướng. Các ảnh ký tự được phân rã thành các mặt phẳng định hướng và một độ đo khoảng cách được tính giữa các mặt phẳng đó với mẫu của mỗi lớp. Hướng nét chữ cục bộ của một ký tự có thể được xác định bằng nhiều cách khác nhau: hướng của xương, phân đoạn nét chữ, mã hóa chu tuyến, hướng đạo hàm. Hiện nay, các đặc trưng mã hóa chu tuyến và hướng đạo hàm được áp dụng rộng rãi vì chúng dễ cài đặt và xấp xỉ bất biến với sự biến đổi đa dạng của các nét chữ.

2.1.3. Đặc trưng hình học và hình thái

Các tính chất cục bộ và toàn cục khác nhau của các ký tự có thể được biểu diễn bằng các đặc trưng hình học và hình thái. Các kiểu đặc trưng này cũng có thể giải mã một số tri thức về cấu trúc của đối tượng ảnh hoặc có thể cung cấp một số tri thức như sắp xếp các thành phần để tạo ra đối tượng. Các loại đặc trưng này có thể phân thành các nhóm sau:

Trích chọn và đếm các cấu trúc hình thái: trong nhóm đặc trưng này, một cấu trúc đã xác định được tìm kiếm trong một ký tự hoặc một từ. Số lượng vị trí hoặc quan hệ vị trí của các cấu trúc trong ký tự này tạo thành các đặc trưng biểu diễn ký tự. Thông thường, các cấu trúc nguyên thủy (các đoạn thẳng, các cung) là các nét tạo ra ký tự. Các ký tự và các từ có thể được mô tả bằng cách trích chọn và đếm nhiều loại đặc trưng về hình thái như các điểm cực đại và cực tiểu, các điểm chớp trên và chớp dưới của một ngưỡng nào đó, mở rộng cho các điểm trái, phải, trên,

dưới và các giao điểm, các điểm nhánh, điểm cuối đoạn thẳng, hướng của một nét từ một điểm đặc biệt, các điểm cô lập... đã tạo nên các ký tự .

Do và xấp xỉ các tính chất hình học: Trong nhiều công trình nghiên cứu , các ký tự được biểu diễn bằng độ đo của các đại lượng hình học như tỷ số giữa chiều rộng và chiều cao của hộp chứa ký tự, quan hệ khoảng cách giữa hai điểm, so sánh độ dài giữa hai nét, độ rộng của một nét, khối lượng chữ hoa và chữ thường của các từ, độ dài từ. Một độ đo tiêu biểu rất quan trọng nữa là độ cong hoặc thay đổi độ cong . Các đại lượng hình học đo được có thể xấp xỉ bởi một tập các đặc trưng hình học vừa đủ và thuận tiện hơn .

Đồ thị và cây: Đầu tiên, các từ hoặc các ký tự được phân chia thành một tập các đối tượng nguyên thủy như các nét, các điểm chạc... Sau đó, các thành phần nguyên thủy được thay thế bằng các thuộc tính hoặc các đồ thị liên quan . Có hai loại đặc trưng ảnh được mô tả bằng đồ thị. Loại thứ nhất sử dụng các tọa độ của hình dáng ký tự . Loại thứ hai là một đặc trưng trừu tượng, các nút của đồ thị tương ứng với các nét chữ và các cạnh của đồ thị tương ứng với các mối quan hệ giữa các nét chữ . Cây cũng có thể dùng để biểu diễn các từ và các ký tự với một tập các đặc trưng theo một quan hệ phân cấp .

Trích chọn đặc trưng hầu hết được thực hiện trên ảnh nhị phân. Tuy nhiên, việc nhị phân hóa ảnh đa cấp xám có thể xóa đi một số thông tin quan trọng của các ký tự. Trong trường hợp này, cũng có một số công trình nghiên cứu để trích chọn các đặc trưng trực tiếp từ các ảnh đa cấp xám .

Cuối cùng, mục đích chính của việc trích chọn đặc trưng là lựa chọn một tập đặc trưng phục vụ cho việc phân lớp sao cho hệ thống nhận dạng đạt độ chính xác cao nhất với số lượng phần tử được trích chọn ít nhất.

L luận án chỉ tập trung nghiên cứu một số đặc trưng thống kê và đặc trưng wavelet cho bài toán nhận dạng chữ Việt viết tay rời rạc.

2.2. Đặc trưng bất biến

Để nhận ra nhiều biến thể của cùng một đối tượng, tính năng là bất biến cho một số biến đổi trên các đối tượng cần phải được sử dụng. Bất biến là các tính năng có giá trị xấp xỉ như nhau, Ví dụ:dịch, thu nhỏ, xoay, kéo dài, sai lệch hoặc nhân đôi .Tuy nhiên, không phải tất cả các biến thể giữa các đối tượng trong cùng một lớp(ví dụ: tiếng ồn, suy thoái và sự vắng mặt hay hiện diện của serifs) có thể được mô hình sử dụng bất biến. Kích thước và bất biến dịch có thể dễ dàng đạt được. Việc phân

định các đặc điểm cá nhân có thể tự cung cấp các ước tính về kích thước và vị trí, nhưng các tính năng phương pháp khai thác thường có thể cung cấp chính xác hơn ước tính.

Quay bất biến là quan trọng nếu các đối tượng được công nhận có thể xảy ra trong định hướng bất kỳ. Tuy nhiên, nếu tất cả các đối tượng được dự kiến sẽ có cùng luân chuyển, sau đó tính năng quay biến thể nên được sử dụng để phân biệt giữa các ký tự như "6" và "9", và "n" và "u". Một lựa chọn khác là sử dụng tính năng bất biến, thêm vào các phát hiện rota-góc hóa. Nếu góc quay bị hạn chế giữa -45 và 45 °, các đối tượng đó, nói 180 ° quay của nhau có thể được phân biệt. cùng nguyên tắc có thể được sử dụng cho các tính năng kích thước bất biến, nếu một muốn nhận dấu chấm câu ngoài nhân vật và muốn phân biệt giữa, nói rằng, "." "o" và "O", và "," và "9".

2.3. Khả năng khôi phục

Đối với một số phương pháp, khôi phục lại chính xác có thể yêu cầu một số lượng lớn tính năng, xấp xỉ hợp lý của hình dạng đối tượng ban đầu thường có thể thu được bằng cách chỉ sử dụng một số lượng nhỏ các tính năng với mức cao nhất nội dung thông tin. Hy vọng là các tính năng này cũng có quyền phân biệt đối xử cao. Bởi khôi phục lại hình ảnh trong tính năng trích xuất, một trực quan có thể kiểm tra xem một số tính năng được sử dụng để nắm bắt được cơ bản cấu trúc của các đối tượng. Khôi phục cũng có thể được sử dụng để kiểm soát chính thức rằng việc thực hiện là chính xác hay không.

2.4. Trích chọn đặc trưng từ ảnh đa cấp xám

2.4.1. Giới thiệu

Một thách thức lớn trong phương pháp này là xác định vị trí, địa điểm của đối tượng. Người ta có thể sử dụng một phương pháp thích ứng để có được một hình ảnh nhị phân tốt và sử dụng kết nối thành phần của kích thước đối tượng dự kiến để xác định vị trí đối tượng cần xác định. Tuy nhiên, phương pháp này thường được sử dụng khi công nhận dựa trên ảnh nhị phân đại diện, do đó vấn đề vẫn chưa được giải quyết cho những hình ảnh khó. Người ta có thể phải dùng đến phương pháp tiếp cận tất cả các địa điểm có thể trong hình ảnh. Tuy nhiên, sau đó người ta phải giả định một kích thước tiêu chuẩn cho một hình ảnh, là sự kết hợp của tất cả các kích thước các đối tượng cần xác định và địa điểm dựa trên sự tính toán. Cách tiếp cận không thể được sử dụng nếu kích thước đối tượng dự kiến thay đổi.

2.4.2. Đối sánh mẫu

Chúng ta không nhận thức được hệ thống OCR sử dụng mẫu phù hợp với hình ảnh xám. Tuy nhiên, kể từ khi mẫu phù hợp là một hình ảnh khá chuẩn kỹ thuật xử lý, chúng ta đã bao gồm hoàn chỉnh. Trong mẫu phù hợp với các bước tính năng khai thác được bỏ hoàn toàn và hình ảnh chính được sử dụng như là một "vector". Trong giai đoạn công nhận, một biện pháp giữa các mẫu T_J và hình ảnh Z được tính toán. Các thông số T_k, trong đó có các biện pháp tương tự cao nhất, là xác định và nếu sự tương đồng này là trên một quy định ngưỡng, sau đó các hình ảnh được gán nhãn lớp k. Ngược lại, các nhân vật còn lại không được phân loại. Trong trường hợp một biện pháp đồng dạng, mẫu T_k có biện pháp đồng dạng thấp nhất được xác định và nếu dưới mức quy định, các đối tượng được dán lớp nhãn k.

2.4.3. Mẫu biến dạng

Mẫu biến dạng đã được sử dụng rộng rãi trong nhiều công nhận đối tượng. Các mẫu được sử dụng là bộ xương nhân vật. Nó không phải là rõ ràng cách vị trí ban đầu của các mẫu được chọn. Nếu tất cả các vị trí có thể có trong các hình ảnh đã được xét xử, thì thời gian tính toán sẽ được ngăn cấm.

2.4.4. Biến đổi ảnh Unitar

Trong mẫu phù hợp, tất cả các điểm ảnh trong hình ảnh xám được sử dụng như một biến đổi Unitar để hình ảnh giảm về số lượng các tính năng trong khi bảo quản hầu hết các thông tin về hình dạng. Trong khoảng thời gian chuyển đổi, các điểm ảnh được sắp xếp và các điểm ảnh với mức cao nhất được sử dụng như tính năng. Biến đổi Unitar phải được áp dụng để có được ước tính của phương sai của các điểm ảnh trong không gian chuyển đổi.

Các tính năng được lấy từ biến đổi Unitar là không bất biến, vì vậy những hình ảnh đầu vào phải được luân chuyển đến một định hướng tiêu chuẩn nếu xoay ký tự có thể xảy ra. Hơn nữa, những hình ảnh đầu vào có được chính xác kích thước tương tự, do đó, một rộng hoặc tái lấy mẫu là cần thiết nếu kích thước có thể khác nhau. Các đơn nhất biến đổi là không chiếu sáng bất biến, nhưng đối với các Fourier chuyển đổi hình ảnh giá trị tại gốc là tỷ lệ thuận với giá trị điểm ảnh trung bình của đầu vào hình ảnh, do đó tính năng này có thể bị xóa để có được sang Ness bất biến. Cho tất cả các biến đổi đơn nhất, một nghịch đảo chuyển đổi tồn tại, vì vậy hình ảnh nhân vật ban đầu có thể được xây dựng lại.

2.4.5. Bất biến Zenite

Zernike khoảnh khắc đã được sử dụng bởi một số tác giả cho nhận dạng ký tự nhị phân rắn ký tự.

2.5. Trích chọn đặc trưng từ ảnh nhị phân:

2.5.1. Giới thiệu

Một ảnh raster nhị phân thu được bằng một hoặc toàn cầu binarization địa phương thích nghi của đầu vào hình ảnh xám. Trong nhiều trường hợp, các phân đoạn của đối tượng là thực hiện chỉ đơn giản bằng cách cách ly các thành phần được kết nối. Tuy nhiên, cho hình ảnh khó khăn, một số đối tượng có thể chạm vào hoặc chồng chéo nhau đối tượng in khác hoặc khác. Một vấn đề khác xảy ra khi đối tượng được phân mảnh thành hai hoặc nhiều thành phần được kết nối, vấn đề có thể được giảm nhẹ phần nào bằng cách chọn một phương pháp thích nghi tốt hơn, nhưng Trier và Taxt đã chỉ ra rằng ngay cả tốt nhất phương pháp binarization địa phương thích ứng có thể vẫn còn không dẫn đến đối tượng bị cô lập hoàn toàn. Phương pháp phân đoạn đối tượng cảm động là do Westall và Narasimha, Fujisawa và trong các cuộc điều tra. Tuy nhiên, các phương pháp giả định rằng các đối tượng xuất hiện trong chuỗi cùng nội dung và có định hướng được biết đến. Trong bản đồ thủy văn), ví dụ, một số đối tượng chạm vào hoặc chồng chéo lên nhau đường, hoặc các ký tự liên lạc từ một dòng văn bản. Một phương pháp dựa trên xám được tích hợp và phân khúc. Phương pháp này cho một tốt hơn hiệu suất, kể từ khi thông tin thu được trong topo - bước phân tích đồ họa được sử dụng trong phân chia các ảnh nhị phân. Bước phân khúc cũng xử lý xoay đối tượng và đối tượng chạm vào từ văn bản khác nhau đây. Đại diện raster nhị phân của một đối tượng là đơn giản hóa của các đại diện xám. các chức năng chụp ảnh $Z(x, y)$ bây giờ có trên hai giá trị (nói, 0 và 1) thay vì, nói 256 giá trị màu xám cấp. này có nghĩa là tất cả các phương pháp phát triển cho màu xám - đại diện quy mô được áp dụng cho các nhị phân rắn raster đại diện là tốt. Vì vậy, chúng ta sẽ không lặp lại mô tả đầy đủ của mỗi phương pháp, nhưng chỉ chỉ ra sự đơn giản hóa trong các tính toán trong volved cho mỗi phương pháp khai thác tính năng này. Nói chung, bất biến để chiếu sáng không còn phù hợp, nhưng các invariances khác. Một ảnh nhị phân rắn có thể được chuyển đổi sang khác cơ quan đại diện, chẳng hạn như các đường viền ngoài của char bộ phận công tác, các cấu hình đường viền, hoặc các bộ xương và tính năng có thể được chiết xuất từ một trong các đại diện là tốt. Với mục đích thiết kế

OCR hệ thống, mục tiêu của việc chuyển đổi này là để bảo vệ thông tin liên quan về các hình và loại bỏ một số thông tin không cần thiết.

2.5.2. Biến đổi ảnh Unitar

NIST dựa trên hình thức hệ thống nhận dạng tay-in sử dụng Karhunen-Loeve chuyển đổi để trích xuất các tính năng từ các đại diện raster nhị phân. Hiệu quả của nó là tuyên bố được tốt, và hệ thống OCR này có sẵn trong phạm vi công cộng.

2.5.3. Bất biến hình học

Một hình ảnh nhị phân có thể được coi là một trường hợp đặc biệt một hình ảnh màu xám cấp với $Z(x, y) = 1$ cho các điểm ảnh và in ấn $Z(z_i, Y_i) = 0$ cho các điểm ảnh nền.

2.6. Trích chọn đặc trưng từ biên ảnh:

2.6.1. Giới thiệu

Đường cong đường viền bên ngoài khép kín của một ảnh là một đường cong tuyến tính piecewise khép kín đi qua trung tâm của tất cả các điểm ảnh mà bốn kết nối đến bên ngoài nền và không có điểm ảnh khác. Sau đường cong, các điểm ảnh được truy cập trong, nói, ngược kim đồng hồ trật tự và đường cong có thể truy cập vào một điểm ảnh cạnh hai lần địa điểm mà đối tượng là rộng một điểm ảnh. mỗi dòng bộ phận là một đường thẳng từ tâm điểm ảnh tới 8 láng giềng.

2.6.2. Tách vùng

Kimura và Shridhar được sử dụng phân vùng trên đường viền đường cong. Trong mỗi khu vực, các phân đoạn đường đồng mức được- tween pixel lân cận đã được nhóm lại theo định hướng: ngang (0°), thẳng đứng (90°) và hai đường chéo định hướng ($45, 135^\circ$). Số đoạn đường mỗi định hướng được tính.

Takahashi cũng được sử dụng biểu đồ định hướng từ khu vực, nhưng sử dụng theo chiều dọc, ngang và đường chéo lát như khu vực thêm vào bên ngoài đường viền khi thực hiện các biểu đồ. Hơn nữa, Takahashi xác định điểm cong cao dọc theo hai đường viền bên ngoài và bên trong. Đối với mỗi điểm, giá trị độ cong, tiếp tuyến đường viền và vị trí của khu, điểm được trích xuất. thời gian này một mạng lưới thường xuyên được sử dụng như khu vực. Khi các đường cong đường viền là gần với biên giới khu vực, sự thay đổi nhỏ trong các đường viền cong có thể dẫn đến sự thay đổi lớn trong các khai thác các tính năng. Họ đã cố gắng để bù đắp cho điều này bằng cách sử dụng biên giới mờ. Điểm gần biên giới khu vực là cho

các giá trị thành viên mờ để hai hoặc bốn khu vực và các thành viên mờ giá trị số tiền cho một .

2.6.3. Xấp xỉ đường cong:

Trên các đường viền nhân vật bên ngoài và áp gần đúng đường cong giữa hai điểm dừng với một chức năng . Sau đó , cả hai Đường cong và các thông số đường cong được sử dụng như các tính năng .

Xấp xỉ đường cong đường viền bên ngoài với một đường cong spline, mà sau đó được làm nhẵn. các tron spline đường cong được chia thành các phần bằng nhau M -chiều dài đường cong . Đối với mỗi phần , độ cong trung bình được tính bằng công thức xác định . Ngoài ra, khoảng cách từ arithme - tic nghĩa của điểm uốn cong để N như nhau điểm khoảng cách đều nhau trên các đường viền được đo . bằng cách nhân rộng spline đường cong xấp xỉ của nhân vật để một tiêu - kích thước Sở NN & PTNT trước khi tính năng này được đánh giá , các tính năng sẽ trở thành kích thước bất biến. Các tính năng đã dịch bất biến của tự nhiên, nhưng phụ thuộc vào quay.

2.6.4. Mô tả Fourier:

Trước khi Kuhl ,Giardina và Lin và Hwang, mô tả Fourier khác đã được phát triển bởi Zahn ,Roskies và Granlund. .Trong Zahn và Roskies ,các góc cạnh khác nhau-là 1 nghiệm .

2.7. Trích chọn đặc trưng từ biểu diễn véc tơ:

2.7.1. Giới thiệu

Một số lượng áp đảo của các thuật toán mỏng tồn tại và một số nghiên cứu đánh giá gần đây cho thấy mỗi để họ giá trị và bất lợi. Nhiệm vụ của một trong những quyền thường liên quan đến một sự thỏa hiệp , một muốn một điểm ảnh rộng bộ xương tám kết nối với chi nhánh giả hoặc các nút di dời , một số loại mạnh mẽ để luân chuyển và tiếng ồn và tại đồng thời một nhanh chóng và dễ dàng thực hiện các thuật toán mỏng dường như là một ứng cử viên tốt, mặc dù thực hiện của nó là một việc phức tạp .

Một biểu đồ đối tượng có thể được bắt nguồn từ các bộ xương bởi nó xấp xỉ với một số đường thẳng phân đoạn và các điểm giao nhau . Vòng cung có thể được sử dụng cho phần cong của bộ xương .

Phương pháp này rất hữu ích khi thậm chí tốt nhất có sẵn phương pháp không thể bảo toàn hình dạng bộ phận làm việc trong các hình ảnh.

2.7.2. Đối sánh mẫu

Mẫu phù hợp ở dạng tinh khiết của nó rất không phù hợp cho bộ xương đối tượng vật, vì rất có thể là nhỏ

2.7.3. Mẫu biến dạng

Mẫu biến dạng đã được sử dụng bởi Burr và Wakahara công nhận nhân vật skel - etons . Trong cách tiếp cận Wakahara , mỗi mẫu là de - hình thành trong một số bước nhỏ , được gọi là affine địa phương biến đổi (LAT) để phù hợp với mô hình đầu vào ứng cử viên

Số lượng và kiểu chuyển đổi trước khi trận đấu thu được có thể được sử dụng như một không giống nhau - biện pháp ity giữa các bản mẫu và các đầu vào mẫu .

2.7.4. Đặc trưng rời rạc

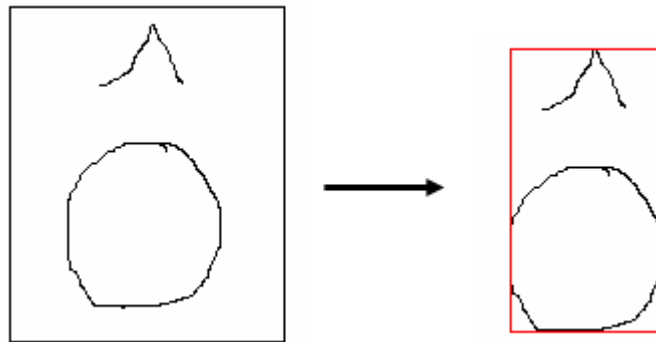
Từ các nhân vật mỏng , các tính năng sau đây có thể được trích xuất: < 1,82 s số lượng của các vòng , số lượng T- khớp , số lượng X - khớp , số lượng uốn cong điểm, chiều rộng và chiều cao của tỷ lệ kèm theo hình chữ nhật ;sự hiện diện của một dấu chấm bị cô lập , tổng số lượng thiết bị đầu cuối và số lượng thiết bị đầu cuối trong mỗi bốn hướng N , S , W và E , số lượng bán vòng tròn trong mỗi bốn hướng , và số lượng các giao cắt theo chiều dọc và trục ngang , tương ứng, các trục đặt trên các trung tâm của lực hấp dẫn .Người ta có thể sử dụng giao cắt với nhiều chồng dòng như tính năng , và trong thực tế , điều này đã được thực hiện trong đầu OCR hệ thống , tl) Tuy nhiên , các tính năng này một mình làm không dẫn đến hệ thống nhận dạng mạnh mẽ , như số lượng dòng chồng được tăng lên, các tính năng kết quả ít mạnh mẽ với sự thay đổi trong các phong chữ (dùng cho máy ký tự in) và biến đổi trong hình dạng nhân vật và văn bản phong cách (ký tự viết tay)

2.7.5. Biểu diễn Fourier

Các phương pháp mô tả Fourier cho đường nét nhân vật cũng có thể sử dụng cho bộ xương nhân vật hoặc đồ thị nhân vật , kể từ bộ xương hoặc đồ thị có thể được đi qua để tạo thành một đường cong đóng cửa ,nhấn mạnh các biểu đồ nhân vật với hai dòng kết thúc , không có các nút và không có vòng, , trong khi đối với đồ thị với các nút giao hoặc vòng, tất cả các mô tả sẽ có giá trị khác không .

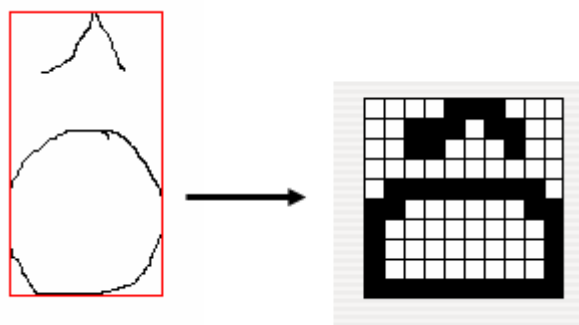
Các đặc tính cho phép quay và kích thước bất biến tính năng này cũng đã được tìm thấy , trường hợp của cùng một nhân vật mà xảy ra được khác nhau đối với các loại trên với sẽ có được vector tính năng rất khác nhau . Giải pháp là trước khi phân loại các đồ thị ký tự vào một trong ba loại và sau đó sử dụng một phân loại riêng cho từng loại .

- Tìm giới hạn trên, giới hạn dưới, trái, phải của ký tự sau khi được vẽ.
- Cắt bỏ tất cả các khoảng trắng xung quanh ký tự được vẽ thu được ảnh chỉ có hoàn toàn ký tự. Như thế quy về xử lý ảnh chỉ có mỗi ký tự, không quan tâm ký tự được vẽ ở vị trí nào: bên trên hay dưới, bên trái hay bên phải



Hình 2.2 Quá trình tìm giới hạn ký tự

- Thực hiện lấy mẫu xuống ảnh của ký tự bằng cách ánh xạ ký tự vào một lưới có kích thước 10*10. Bằng cách này ta quy ký tự về một kích thước cố định, như vậy không cần quan tâm đến kích thước ký tự được vẽ vào là to hay nhỏ.

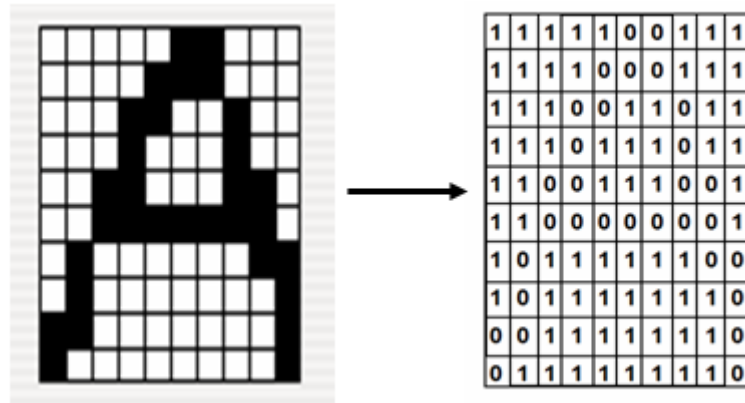


Hình 2.3 Quá trình lấy mẫu xuống

3.3. Xử lý dữ liệu (phân tích ảnh)

Quá trình phân tích ảnh là quá trình tìm ra đặc tính của ký tự. Các điểm ảnh này chính là giá trị đầu vào cho cả quá trình huấn luyện và quá trình kiểm tra. Quá trình phân tích này dựa trên việc nhận dạng từ ảnh người sử dụng vẽ vào bằng

phương pháp tìm các điểm đen RGB(0,0,0) hoặc điểm trắng RGB(255,255,255). Kết quả thu được là ma trận điểm ảnh. Sau đó ma trận này sẽ được ánh xạ thành vector đầu vào có kích thước 100 phần tử. Trước khi đưa vào mạng vector này còn được xử lý làm nổi điểm ảnh thành vector gồm các phần tử 0.5 và -0.5.



Hình 2.4. Quá trình ánh xạ từ ma trận điểm sang ma trận giá trị

- . **Ký tự:** Chứa các chữ mẫu dùng để trích chọn
- . **Giá trị:** Lưu trữ giá trị các chữ mẫu dùng để trích chọn
- . **Nhận dạng:** Vùng để vẽ chữ lên.
- . **Tập mẫu:** Chứa các chữ mẫu dùng để trích chọn.
- . **Tải mẫu:** Tải tập chữ mẫu đã tạo sẵn trong file: TapMau.xml
- . **Thêm chữ:** Ghi lại tập chữ mẫu vào file: TapMau.xml
- . **Trích chọn đặc trưng:** Trích chọn đặc trưng trọng số vùng chữ viết
- . **Xóa chữ:** Xóa một chữ mẫu được chọn ra khỏi file TapMau.xml
- . **Nhận dạng:** Nhận dạng chữ vẽ vào.
- . **Thông tin:** Giới thiệu sơ bộ về chương trình

Chúng ta có thể tải tập chữ ra từ file tạo sẵn TapMau.xml bằng cách nhấn nút “Tải mẫu” hoặc chúng ta có thể tự nhập mẫu để trích chọn bằng cách vẽ chữ vào vùng vẽ ký tự rồi nhấn “Thêm chữ”, khi đó có một hộp thoại nhắc bạn đánh vào chữ tương ứng để khi nhận dạng nó sẽ chỉ ra chữ tương ứng với chữ vừa vẽ.

Sau khi nhập xong tập mẫu thì nhấn nút “Trích chọn đặc trưng”. Quá trình trích chọn hoàn thành, có thể thực hiện nhận dạng chữ bằng cách vẽ chữ rồi nhấn “Nhận dạng”. Khi đó chữ được hiển thị trên lưới kích thước 10*10.

Có thể xóa chữ hoặc xóa mẫu trong tập mẫu bằng cách chọn chữ cần xóa rồi nhấn nút “Xóa chữ”.

3.4. Kết quả nhận dạng

- Để nhận dạng chữ ta vẽ chữ cần nhận dạng vào vùng vẽ chữ (vùng chữ nhật màu trắng, ngay bên dưới phần “Nhận dạng”).

- Ký tự đồng thời được hiển thị trên lưới kích 10*10 (vùng ô vuông màu xanh, bên phải giao diện chương trình).

- Nhấn nút nhận dạng.

- Chữ nhận dạng tương ứng với chữ cần nhận dạng đã được vẽ sẽ được hiển thị ra màn hình.

3.5. Đánh giá, nhận xét

Chương trình có thể nhận dạng với độ chính xác rất cao. Nếu chữ nhận dạng chưa được chính xác ta có thể thêm mẫu và trích chọn lại cho chữ đó đến khi nhận dạng được chữ đó chính xác. Nếu các chữ gần giống nhau khi nhận dạng dễ lẫn sang nhau thì khi trích chọn ta thêm mẫu tương ứng với các chữ đến khi nào đạt được tỷ lệ nhận dạng cao.

Chương trình nhận dạng chính xác hơn với các chữ không dấu đến 96%, đối với chữ có dấu thì độ chính xác lớn hơn 85%. Đối với chữ có dấu độ chính xác giảm hơn so với trường hợp không có dấu vì trường hợp này các chữ khác nhau rất ít, nên số lượng mẫu trích chọn cho các chữ có dấu cần nhiều hơn.

Để nhận dạng được nhiều kiểu viết khác nhau thì tập mẫu cho mỗi chữ phải có nhiều dạng phong phú, tương ứng với các cách viết mà người sử dụng có thể viết ra. Quá trình huấn luyện này giải quyết được khía cạnh phức tạp của bài toán nhận dạng chữ viết tay đó là sự đa dạng của chữ viết tay, đặc biệt là chữ viết tay tiếng Việt.

Như vậy chương trình minh họa cho nhận dạng chữ viết tay tiếng Việt trực tuyến đã đạt được mục tiêu của bài toán phức đặt ra. Nghiên cứu về trích chọn đặc trưng đã là một hướng nghiên cứu hiện nay còn rất mới mẻ và có nhiều hứa hẹn.

Với hướng tiếp cận mới này áp dụng cho giải quyết bài toán nhận dạng chữ viết tay tiếng Việt trực tuyến vốn rất phức tạp ở những đặc điểm:

- Chữ viết tay trực tuyến phải thêm phần tạo giao diện để vẽ chữ, so với việc nhập đầu vào là file ảnh của chữ viết tay.

- Chữ viết tay thì hình dạng của chữ rất phong phú đa dạng phụ thuộc nhiều vào người viết, bởi mỗi người có một cách viết khác nhau.

- Chữ viết tay tiếng Việt có số lượng lớn, lớn hơn rất nhiều so với các chữ theo các chuẩn khác. Ví dụ: chuẩn tiếng Việt có 89 ký tự gồm ký tự không dấu và có dấu, chuẩn tiếng Anh chỉ gồm 26 ký tự không dấu. Mặt khác các ký tự lại rất giống nhau nên dễ nhầm lẫn trong quá trình nhận dạng.

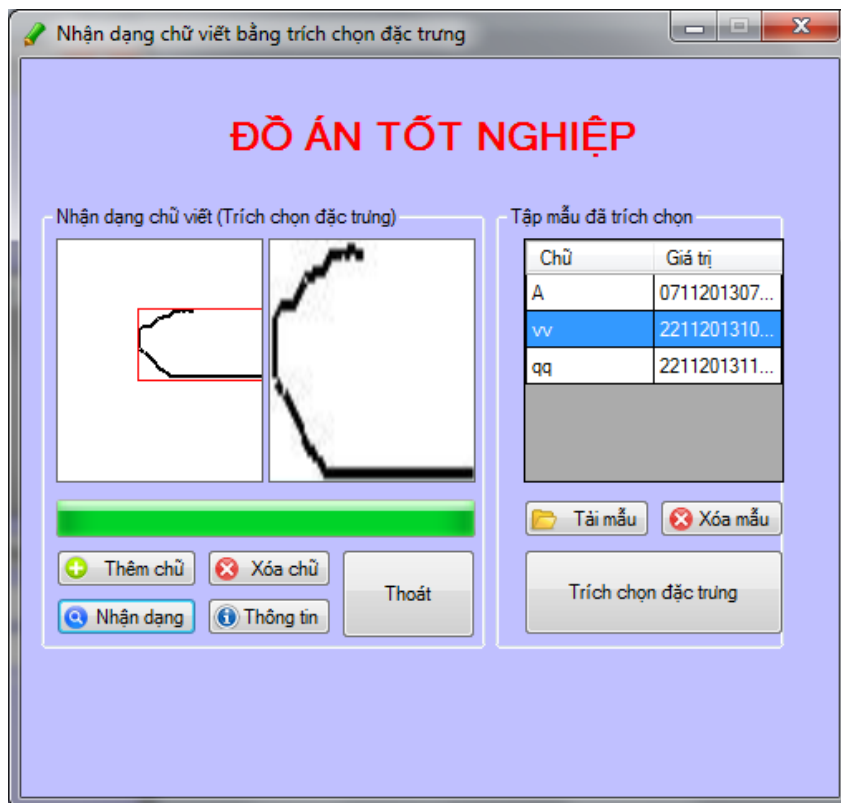
Như vậy khi giải quyết bài toán ở đây phải đáp ứng được các yêu cầu:

- Giải quyết được sự phức tạp của việc xử lý dữ liệu đầu vào.
 - Giải quyết được việc xử lý khối lượng dữ liệu lớn.
 - Giải quyết được độ chuẩn xác trong quá trình nhận dạng.
 - Giải quyết được mức độ tổng quát, đa dạng, phong phú trong quá trình xây dựng và trích chọn để đạt được độ chính xác cao khi nhận dạng.
- Chương trình thử nghiệm được xây dựng bằng ngôn ngữ C# sử dụng bộ Visual Studio 2012.

Một số giao diện chương trình



Hình 2.5. Giao diện chính của chương trình



Hình 2.6. Giao diện của chương trình sau khi nhận dạng chữ viết xong

KẾT LUẬN

A. Kết quả đạt được:

Trong quá trình nghiên cứu và hoàn thành bài báo cáo đồ án tốt nghiệp “Tìm hiểu một số phương pháp trích chọn đặc trưng cho nhận dạng chữ viết”, em đã thu nhận được thêm rất nhiều kiến thức và cũng nhận thấy trích chọn đặc trưng cho nhận dạng chữ viết là một lĩnh vực rộng lớn và được ứng dụng rất mạnh mẽ. Hơn thế nữa nó còn rất nhiều vấn đề mà chúng ta cần khám phá. Trong đề tài của mình em đã cố gắng tìm hiểu và nghiên cứu tổng quan về lý thuyết nhận dạng chữ viết và phương pháp trích chọn đặc trưng cho việc nhận dạng chữ viết. Cài đặt thử nghiệm chương trình trích đặc trưng cho việc nhận dạng các chữ cái không dấu.

Do thời gian thực hiện hạn chế và kiến thức còn hạn chế nên em mới chỉ nghiên cứu được một số lý thuyết cơ bản trong trích chọn đặc trưng. Còn nhiều lý thuyết cũng như kỹ thuật em vẫn chưa tìm hiểu, khai thác và ứng dụng vào các bài toán thực tế.

Mặc dù đã rất cố gắng, song do năng lực và trình độ có hạn nên trong quá trình thực hiện bài thực tập em đã không tránh khỏi những thiếu sót. Kính mong các thầy cô và các bạn quan tâm giúp đỡ chỉ bảo để chương trình của em một hoàn thiện hơn.

B. Hướng phát triển trong tương lai

Trong thời gian tới em sẽ cố gắng tiếp tục nghiên cứu, tìm hiểu thêm một số kiến thức về trích chọn đặc trưng cho việc nhận dạng chữ viết nói riêng và lý thuyết nhận dạng nói chung.

TÀI LIỆU THAM KHẢO

Tiếng Việt

- [1] Hoàng Kiếm, Nguyễn Hồng Sơn, Đào Minh Sơn, “Ứng dụng mạng nơ ron nhân tạo trong hệ thống xử lý biểu mẫu tự động”, Kỷ yếu hội nghị kỷ niệm 25 năm thành lập Viện Công nghệ Thông tin, 2001, tr. 560-567.
- [2] Bùi Minh Trí, “Quy hoạch toán học”, Nhà xuất bản Khoa học và kỹ thuật, Hà nội, 2006.
- [3] Lê Hoài Bắc, Lê Hoàng Thái, “Neural Network & Genetic Algorithm in Application to Handwritten Character Recognition”, Tạp chí Tin học và Điều khiển học, Tập 17, số 4, 2001, tr. 57-65.
- [4] Nguyễn Thị Thanh Tân, Ngô Quốc Tạo, “Một cấu trúc mạng nơ ron thích hợp cho việc nhận dạng chữ số viết tay”, Kỷ yếu hội thảo FAIR03, NXB KH&KT Hà Nội, 2004, tr. 200-210.
- [5] Nguyễn Thị Thanh Tân, Lương Chi Mai, “Phương pháp nhận dạng từ viết tay dựa trên mô hình mạng nơ ron kết hợp với thống kê từ vựng”, Tạp chí Tin học và Điều khiển học, Tập 22, số 2, 2006, tr. 141-154.
- [5.1]. Phạm Anh Phương, Ngô Quốc Tạo, Lương Chi Mai, “Ứng dụng SVM cho bài toán phân lớp nhận dạng”, *Kỷ yếu Hội thảo khoa học Quốc gia lần thứ ba về nghiên cứu, phát triển và ứng dụng Công nghệ thông tin và Truyền thông (ICT.rda'06)*, nhà xuất bản Khoa học và Kỹ thuật, Hà nội, trang 393-400, 20-21/05/2006.
- [5.2]. Phạm Anh Phương, Ngô Quốc Tạo, Lương Chi Mai, “Trích chọn đặc trưng wavelet Haar kết hợp với SVM cho việc nhận dạng chữ viết tay tiếng Việt”, *Tạp chí Công nghệ Thông tin và Truyền thông*, ISSN 0866-7039, kỳ 3, số 20, 10-2008, tr 36-42.
- [5.3]. Phạm Anh Phương, “Áp dụng một số chiến lược SVM đa lớp cho bài toán nhận dạng chữ viết tay hạn chế”, *Tạp chí khoa học Đại học Huế*, ISSN 1859-1388, số 45, 2008, tr. 109-118.
- [5.4]. Nguyễn Thị Thanh Tân, Lương Chi Mai, “Phương pháp nhận dạng từ viết tay dựa trên mô hình mạng nơ ron kết hợp với thống kê từ vựng”, *Tạp chí Tin học và Điều khiển học*, Tập 22, số 2, 2006, tr. 141-154.

Tiếng Anh

- [6] ivind Due Trier, Anil K.Jain, Feature Extraction methods for character recognition:Survey,Pattern Recognition 29(4):641-662(1996).
- [7] Mark S, Nixon Alberto S. Aguado, Feature Extraction & Image Processing, 2nd Edition from Mark Nixon, Alberto S Aguato. ISBN-9780080556727, 2002.