

MỘT SỐ PHƯƠNG PHÁP TRÍCH CHỌN ĐẶC TRƯNG HIỆU QUẢ CHO BÀI TOÁN NHẬN DẠNG CHỮ VIẾT TAY RỜI RẠC

Phạm Anh Phương

Trường Đại học Khoa học, Đại học Huế

TÓM TẮT

Trích chọn đặc trưng là một mắt xích rất quan trọng ảnh hưởng trực tiếp đến chất lượng của một hệ thống nhận dạng. Bài báo này so sánh một số phương pháp trích chọn đặc trưng đơn giản nhưng hiệu quả, có thể áp dụng cho bài toán nhận dạng chữ viết tay rời rạc. Các kết quả thực nghiệm trên tập dữ liệu chữ số viết tay MNIST cho thấy một số đặc trưng mà chúng tôi lựa chọn cho độ chính xác cao hơn so với các kết quả đã được công bố ở [9].

I. Giới thiệu

Nhận dạng chữ viết tay đang là vấn đề thách thức lớn đối với các nhà nghiên cứu. Cho đến nay, bài toán nhận dạng chữ viết tay vẫn chưa có được một giải pháp tổng thể. Các kết quả chủ yếu chỉ tập trung trên các tập dữ liệu chữ số viết tay chuẩn như USPS và MNIST [4,5,8,9], 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 cũng chỉ giới hạn trong một phạm vi hẹp [2,7,10].

Trong lĩnh vực nhận dạng, vấn đề trích chọn đặc trưng là một bước rất quan trọng, nó có ảnh hưởng lớn đến chất lượng cũng như tốc độ của một hệ thống nhận dạng. Trích chọn đặc trưng như thế nào để vẫn đảm bảo không mất mát thông tin và/hoặc thu gọn kích thước của dữ liệu đầu vào là điều mà nhiều nhà nghiên cứu đang quan tâm. Trong bài báo này, chúng tôi cài đặt thử nghiệm và so sánh một số phương pháp trích chọn đặc trưng cho bài toán nhận dạng chữ viết tay rời rạc. Mô hình nhận dạng của chúng tôi được xây dựng trên cơ sở phương pháp véc tơ tựa (SVM – Support Vector Machines), đây là một phương pháp máy học tiên tiến đã có nhiều thành công trong các lĩnh vực khai phá dữ liệu cũng như nhận dạng [3,6].

Phần còn lại của bài báo này có cấu trúc như sau: Phần 2 trình bày kiến trúc của mô hình nhận dạng chữ viết tay rời rạc. Phần 3 trình bày một số phương pháp trích chọn đặc trưng có thể áp dụng cho chữ viết tay. Phần 4 trình bày một số kết quả thực nghiệm trên tập dữ liệu chữ số viết tay MNIST. Cuối cùng là phần kết luận.

II. Kiến trúc của mô hình nhận dạng

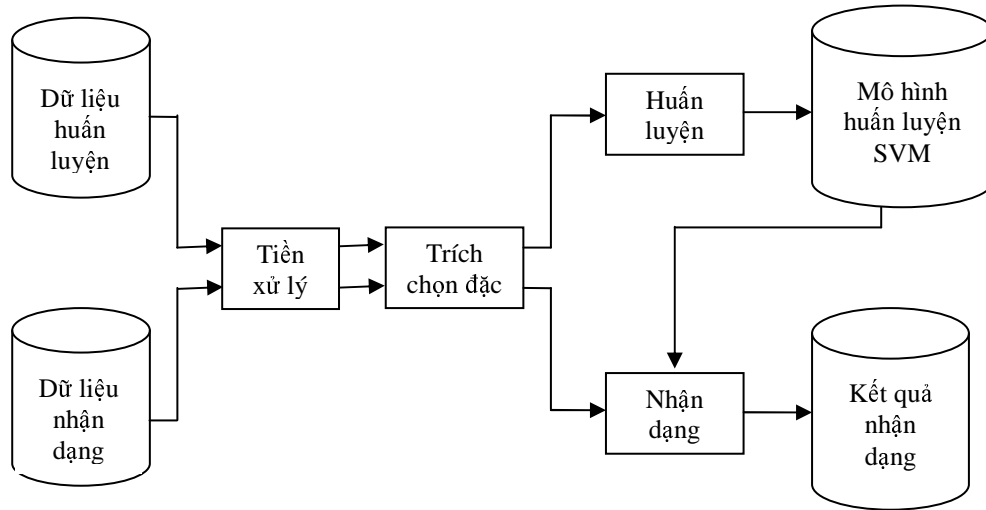
Trong mô hình nhận dạng (hình 1), công đoạn tiền xử lý bao gồm các chức năng lọc nhiễu và chuẩn hóa ảnh đầu vào. Việc chuẩn hóa ảnh bao gồm các bước sau:

Bước 1: Chuyển đổi ảnh đầu vào sang ảnh trắng đen.

Bước 2: Tìm hình chữ nhật R bé nhất chứa các điểm đen trên ảnh.

Bước 3: Lấy vùng ảnh I nằm trong vùng hình chữ nhật R.

Bước 4: Chuẩn hóa kích thước ảnh I thành 16×16 .



Hình 1: Mô hình nhận dạng chữ viết tay rời rạc

Mô hình phân lớp được xây dựng từ các SVM nhị phân theo chiến lược một chống một (OVO – One versus One) với các tham số $C=100$ và sử dụng hàm nhân là hàm Gaussian với $\sigma=0,5$ [9].

Trong mô hình nhận dạng ở hình 1, bài báo này chỉ tập trung vào giai đoạn trích chọn đặc trưng để đánh giá một số loại đặc trưng được sử dụng cho bài toán nhận chữ viết tay. Phần tiếp theo, chúng tôi sẽ trình bày một số phương pháp trích chọn đặc trưng đơn giản, có thể áp dụng cho việc nhận dạng ký tự viết tay rời rạc.

III. Một số phương pháp trích chọn đặc trưng

Trong phần này, chúng tôi sẽ giới thiệu một số phương pháp trích chọn đặc trưng đơn giản nhưng hiệu quả, có thể áp dụng cho các tập dữ liệu chữ viết tay rời rạc.

3.1. Trọng số vùng (Zoning)



Hình 2. Trích chọn đặc trưng trọng số vùng

Ảnh ký tự được chia thành $N \times N$ vùng (zones). Tổng số điểm đen của mỗi vùng sẽ được chọn để tạo thành véc tơ đặc trưng.

Trong thực nghiệm, với ảnh kích thước 16×16 , chúng tôi chọn $N=8$, như vậy có $8 \times 8 = 64$ đặc trưng.

3.2. Biểu đồ chiếu (Projection histograms)

Ý tưởng cơ bản của phương pháp trích chọn đặc trưng này là chiếu các điểm đen trên ảnh 2 chiều theo các hướng ngang, dọc và hai đường chéo thành một dãy các tín hiệu 1 chiều. Ưu điểm của các đặc trưng này là không phụ thuộc vào nhiễu, tuy nhiên nó vẫn phụ thuộc vào độ nghiêng của chữ.



Hình 3. Trích chọn các biểu đồ chiếu ngang, dọc và 2 đường chéo

Trong thực nghiệm, với ảnh kích thước 16×16 , chúng tôi chọn 16 ngang + 16 dọc + 2×31 chéo = 94 đặc trưng.

3.3. Trích chọn chu tuyến (Contour profiles)



Hình 4. Trích chọn các khối bên ngoài của chữ

Phần được trích chọn là khoảng cách từ biên của khung chứa ảnh tới điểm đen đầu tiên của chữ trên cùng 1 dòng quét. Phương pháp trích chọn này mô tả tốt các khối bên ngoài của chữ và cho phép phân biệt một số lượng lớn các ký tự.

Trong thực nghiệm, với ảnh kích thước 16×16 , có 16 trái + 16 phải + 16 trên + 16 dưới = 64 đặc trưng.

3.4. Trích chọn đặc trưng wavelet Haar

Chúng tôi sử dụng ý tưởng của phương pháp trích chọn đặc trưng wavelet Haar [1] để chọn tập đặc trưng cho mỗi ảnh ký tự đầu vào.

Từ ảnh nhị phân kích thước $2^n \times 2^n$ (Hình 5), quá trình trích chọn đặc trưng được mô tả theo thuật toán sau:

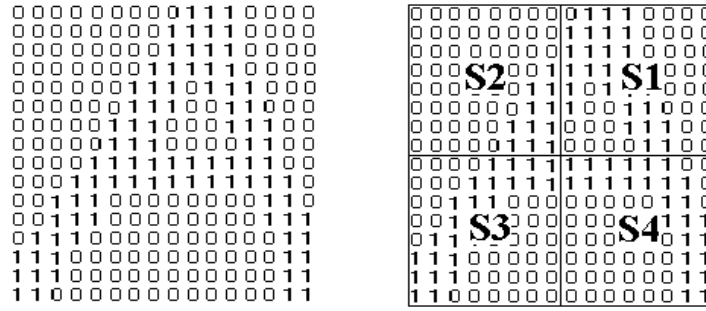
Procedure HaarFeature

Input Ma trận vuông (A,n) cấp 2^n .

Output Tập các đặc trưng $\{F_1, F_2, \dots, F_{2^n \times 2^n}\}$.

Method

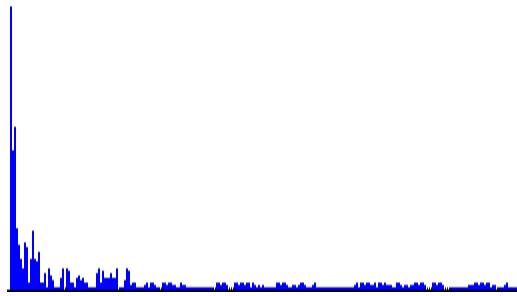
1. Khởi tạo: Queue = \emptyset ;
 $i = 1$;
2. - Tính F_i = Tổng các điểm đen trong toàn bộ ma trận (A,n);
- PUSH((A,n), Queue);
3. While Queue $\neq \emptyset$ Do
{
 - POP(Queue, (A,n));
 - if (n>1)
 {
 Chia ảnh thành 4 phần: A_1, A_2, A_3, A_4 ;
 for (j=1; $i \leq 4$; j++)
 PUSH((A_j, n div 2), Queue);
 }
 - Gọi S1, S2, S3, S4 là tổng các điểm đen tương ứng với A_1, A_2, A_3, A_4 ;
 - Tính $F_{i+1} = S1 + S2$;
 $F_{i+2} = S2 + S3$;
 $F_{i+3} = S4$;
 - $i = i + 3$;
}



Hình 5. Trích chọn đặc trưng wavelet Haar

Tính bất biến của đặc trưng này đã được chứng minh trong [8].

Trong thực nghiệm, với phần chữ chúng tôi chọn $n=4$, như vậy ta có: $1 + 3 + 4 \times 3 + 4 \times 4 \times 3 + 4 \times 4 \times 4 \times 3 = 256$ đặc trưng, còn với phần dấu chúng tôi chọn $n=3$, như vậy có tất cả 64 đặc trưng.

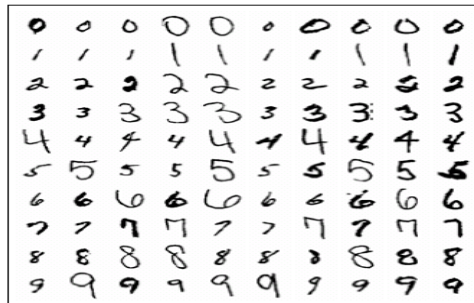


Hình 6. Dãy đặc trưng wavelet Haar

Phương pháp trích chọn đặc trưng này sẽ tạo ra một dãy số các đặc trưng giảm dần. Với cùng một chữ thì các giá trị lớn ở đầu dãy tương đối ổn định, có thể đại diện cho hình dạng khái quát của chữ; còn các giá trị ở cuối dãy nhỏ dần và không ổn định, thể hiện sự đa dạng trong từng chi tiết của chữ (Hình 6).

IV. Kết quả thực nghiệm

Chúng tôi đã cài đặt và thực nghiệm trên tập dữ liệu chữ số viết tay MNIST. Tập dữ liệu MNIST bao gồm 60.000 mẫu huấn luyện và 10.000 mẫu khác để nhận dạng, mỗi mẫu là một ảnh xám kích thước 28×28 (Hình 7).



Hình 7. Các mẫu chữ số viết tay trích từ tập dữ liệu MNIST

Chúng tôi sử dụng mô hình SVM đa lớp theo chiến lược OVO, chọn hàm nhân RBF($\sigma = 0,05$) với các phương pháp trích chọn đặc trưng khác nhau trên tập dữ liệu MNIST (xem *Bảng 1*).

Bảng 1. Kết quả nhận dạng trên tập dữ liệu MNIST với các đặc trưng khác nhau

Đặc trưng	Số đặc trưng	Độ chính xác
<i>Ma trận nhị phân [9]</i>	256	97,2%
Zone 2×2	64	97,9%
Projection	94	97,0%
Contour Profile	64	95,5%
Haar wavelet	256	97,8%

Kết quả ở bảng 1 cho thấy sử dụng các đặc trưng Zone 2×2 và Haar wavelet đạt độ chính xác cao hơn khi sử dụng đặc trưng ma trận nhị phân [9]. Việc sử dụng các đặc trưng Projection và Contour Profile đạt độ chính xác không cao bằng [9] nhưng sử dụng số lượng đặc trưng ít hơn nhiều, vì vậy tốc độ nhận dạng cũng tăng lên đáng kể.

V. Kết luận

Bài báo đã đề xuất một mô hình nhận dạng chữ viết tay rời rạc dựa trên cơ sở phương pháp véc tơ tựa kết hợp với một số phương pháp trích chọn đặc trưng đơn giản nhưng hiệu quả cho việc nhận dạng chữ viết tay rời rạc. Các kết quả thực nghiệm cho thấy các đặc trưng được lựa chọn áp dụng vào bài toán nhận dạng chữ viết tay rời rạc đạt độ chính xác tương đối cao. Việc sử dụng các đặc trưng Zone 2×2, Projection và Contour Profile làm giảm bớt khá nhiều kích thước của dữ liệu đầu vào, làm tăng tốc độ của hệ thống nhận dạng.

Khi sử dụng phương pháp véc tơ tựa thì kết quả của việc phân lớp luôn xấp xỉ ngưỡng tối ưu, vì vậy chất lượng của hệ thống nhận dạng phụ thuộc rất nhiều vào việc xử lý dữ liệu đầu vào. Bài báo này chỉ mới quan tâm đến các phương pháp trích chọn đặc trưng cho dữ liệu chữ viết tay. Trong tương lai, chúng tôi sẽ tiếp tục nghiên cứu tiếp các công đoạn tiền xử lý để chuẩn hóa dữ liệu đầu vào tốt hơn nhằm nâng cao độ chính xác của hệ thống nhận dạng.

TÀI LIỆU THAM KHẢO

1. Viola, P., Jones, M., *Rapid object detection using a boosted cascade of simple features*, Proc. Intl. Conf. on Computer Vision and Pattern Recognition (CVPR), Volume 1, (2001), 511–518
2. 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), 57- 65

3. Chih-Chung Chang and Chil-Jen Lin, *LIBSVM: a Library for Support Vector Machines*, National Taiwan University, 2004.
4. Gorgevik D., Cakmakov D., *An Efficient Three-Stage Classifier for Handwritten Digit Recognition*, Proceedings of 17th Int. Conference on Pattern Recognition, ICPR2004, Vol. 4, pp. 507-510, IEEE Computer Society, Cambridge, UK, 2004.
5. Cakmakov D., Gorgevik D., *Handwritten Digit Recognition Using Classifier Cooperation Schemes*, Proceedings of the 2nd Balkan Conference in Informatics, BCI, Ohrid, (2005), 23-30
6. 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, (2006), 393- 400.
7. G. Vamvakas, B. Gatos, I. Pratikakis, N. Stamatopoulos, A. Roniotis and S.J. Perantonis, *Hybrid Off-Line OCR for Isolated Handwritten Greek Characters*, The Fourth IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA 2007), ISBN: 978-0-88986-646-1, Innsbruck, Austria, (2007), 197-202.
8. 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, (2008), 36-42.
9. 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), 109-118.
10. Pham Anh Phuong, Ngo Quoc Tao, Luong Chi Mai, *An Efficient Model for Isolated Vietnamese Handwritten Recognition*, The Fourth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2008), Harbin, China, August 15 – 17, (2008), 358-361.

SOME EFFICIENT FEATURE EXTRACTION METHODS FOR ISOLATED HANDWRITTEN RECOGNITION

Pham Anh Phuong
College of Sciences, Hue University

SUMMARY

The feature extraction is an important phase in a handwritten recognition system. This paper presents some efficient feature extraction methods for isolated handwritten recognition. Our experiments on the benchmark databases MNIST show that selected feature extraction methods can be successfully used to handwritten recognition with high exact rates.