

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

-----*-----**

NGÔ THANH TÂN

**NHẬN DẠNG KHUÔN MẶT
SỬ DỤNG PCA – SVM**

LUẬN VĂN THẠC SĨ KỸ THUẬT

HÀ NỘI - 2016

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

-----***-----

NGÔ THANH TÂN

**NHẬN DẠNG KHUÔN MẶT
SỬ DỤNG PCA – SVM**

**Chuyên ngành: CÔNG NGHỆ THÔNG TIN
LUẬN VĂN THẠC SĨ KỸ THUẬT**

**GIÁO VIÊN HƯỚNG DẪN
TS. NGUYỄN THỊ OANH**

HÀ NỘI - 2016

LỜI CAM ĐOAN

Luận văn thạc sỹ “Nhận dạng khuôn mặt sử dụng PCA – SVM”, chuyên ngành Công nghệ thông tin là công trình của cá nhân tôi. Các nội dung nghiên cứu và kết quả trình bày trong luận văn là trung thực, rõ ràng. Các tài liệu tham khảo đã được trích dẫn đầy đủ và ghi rõ nguồn gốc.

Hà Nội, ngày 18/01/2016

TÁC GIẢ

Ngô Thanh Tân

MỤC LỤC

LỜI CAM ĐOAN	1
MỤC LỤC.....	2
DANH MỤC CÁC TỪ VIẾT TẮT	4
DANH MỤC HÌNH VẼ, ĐỒ THỊ	5
1. Lý do chọn đề tài	6
2. Lịch sử nghiên cứu.....	6
3. Mục đích nghiên cứu của luận văn, đối tượng, phạm vi nghiên cứu.....	8
4. Tóm tắt cô đọng các luận điểm cơ bản và đóng góp mới của tác giả.....	9
5. Phương pháp nghiên cứu	9
6. Nội dung luận văn.....	9
CHƯƠNG 1. PHÁT BIỂU BÀI TOÁN.....	11
1.1. Hệ thống sinh trắc học	11
1.2 Nhận dạng người qua khuôn mặt.....	12
1.3 Ứng dụng của hệ thống nhận dạng mặt người.....	12
CHƯƠNG 2. CÁC PHƯƠNG PHÁP NHẬN DẠNG KHUÔN MẶT	17
2.1 Phương pháp trích chọn đặc trưng Principal Components Analysis – PCA ..	17
2.1.1 Giới thiệu phương pháp trích chọn đặc trưng PCA.....	17
2.1.2 Các bước thực hiện trích chọn đặc trưng bằng PCA	22
2.2 Máy vector hỗ trợ SVM.....	26
2.2.1 SVM tuyến tính [4],[6],[9]	26
2.2.2 SVM phi tuyến	34

2.2.3 SVM cho vấn đề nhận dạng đa lớp	35
CHƯƠNG 3. ĐỀ XUẤT PHƯƠNG ÁN NHẬN DẠNG MẶT NGƯỜI	38
3.1 Phát hiện khuôn mặt (face detection)	38
3.2 Tiền xử lý ảnh (preprocessing)	39
3.3 Trích rút đặc trưng (feature extraction)	40
3.4 Nhận dạng/Phân lớp (recognition/classification)	40
3.5 Dữ liệu hệ thống nhận dạng	40
CHƯƠNG 4. ĐÁNH GIÁ THỬ NGHIỆM	42
4.1 Độ đo đánh giá hiệu năng	42
4.2 Bộ dữ liệu thử nghiệm	42
4.2.1 Bộ dữ liệu AT&T	42
4.2.2 Bộ dữ liệu ảnh YaleB	43
4.3 Kết quả thử nghiệm.....	43
4.3.1 Thử nghiệm với bộ ảnh AT&T.....	43
4.3.2 Thử nghiệm với bộ ảnh YaleB	44
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	47
TÀI LIỆU THAM KHẢO.....	48

DANH MỤC CÁC TỪ VIẾT TẮT

PCA	Principal Component Analysis
SVM	Support Vector Machine
US visits	United States Visitors and Immigrant Status Indicator Technology
FBI	Federal Bureau of Investigation
NGI	Next Generation Identification
KKT	Karush Kuhn Tucker

DANH MỤC HÌNH VẼ, ĐỒ THỊ

Hình 1.1 Các đặc tính sinh trắc học của người	11
Hình 1.2 Hệ thống xác minh nhân thân của công dân nhập cảnh [19]	14
Hình 2.1 Lựa chọn các trục tọa độ mới để biểu diễn dữ liệu [11]	18
Hình 2.2 Phân cách theo siêu phẳng (w,b) trong không gian hai chiều của tập mẫu [4]	27
Hình 2.3 Siêu mặt phân cách tuyến tính cho trường hợp phân cách được [4]	28
Hình 2.4 Không thể phân hoạch tập mẫu bằng một siêu phẳng [6]	31
Bảng 2.5 Một số hàm nhân thường dùng	35
Hình 2.6 Mẫu cần nhận dạng là SVM 2-vs-rest vì có giá trị bé nhất [10]	36
Hình 2.7 Cấu trúc 1 cây nhị phân cho nhận dạng khuôn mặt 8 lớp [10]	37
Hình 3.1 Mô hình hệ thống nhận dạng khuôn mặt	38
Hình 3.2 Phát hiện khuôn mặt trong ảnh	39
Hình 4.1 Một số ảnh trong bộ dữ liệu AT&T	42
Hình 4.2 Một số ảnh trong bộ dữ liệu YaleB	43
Biểu đồ 4.3 Kết quả thử nghiệm với bộ dữ liệu AT&T	44
Biểu đồ 4.4 Kết quả thử nghiệm trên bộ dữ liệu YaleB	45

PHẦN MỞ ĐẦU

1. Lý do chọn đề tài

Ngày nay, khoa học kỹ thuật ngày càng phát triển, lượng thông tin ngày càng lớn thì nhu cầu bảo vệ thông tin lại càng được chú trọng. Nó đòi hỏi mọi lúc, mọi nơi thông tin phải được đảm bảo an ninh, an toàn. Một trong những phương pháp bảo vệ an ninh là xác thực người sử dụng trong hệ thống thông tin đó. Có nhiều phương pháp để xác thực người sử dụng. Nhận dạng mặt người là một phương pháp để xác nhận quyền truy nhập của một người có hợp pháp hay không. Vì nhận dạng mặt người là một trong những cách mà con người sử dụng để nhận biết nhau. Ngày nay, việc thu thập và xử lý thông tin qua ảnh đang được quan tâm và ứng dụng rộng rãi. Với phương pháp này chúng ta có thể thu thập được nhiều thông tin mà ít tác động đến đối tượng, vì đối tượng sẽ được nhận dạng một cách thụ động.

2. Lịch sử nghiên cứu

Nhận dạng mặt người là một vấn đề đầy thách thức và hứa hẹn trong lĩnh vực phân tích hình ảnh và thị giác máy tính, do đó nó đã nhận được rất nhiều sự chú ý trong thời gian gần đây vì nó có nhiều ứng dụng trong các lĩnh vực khác nhau. Kỹ thuật nhận dạng mặt người có thể được phân chia thành ba loại dựa trên phương pháp thu thập dữ liệu mặt: phương pháp hoạt động dựa trên cường độ hình ảnh; xử lý với chuỗi video; và những yêu cầu dữ liệu giác quan khác như thông tin 3D hoặc hình ảnh hồng ngoại.

Có lẽ ví dụ đầu tiên nổi tiếng nhất của một hệ thống nhận dạng khuôn mặt là do T.Kohonen (1989) [21], người đã chứng minh rằng một mạng neural đơn giản có thể thực hiện nhận dạng khuôn mặt cho ra hình ảnh khuôn mặt đã được sắp xếp và chuẩn hóa. Hệ thống của T.Kohonen mô tả khuôn mặt bằng các vector riêng của ma trận tương quan hình ảnh khuôn mặt, là các eigenfaces. Hệ thống của T.Kohonen không phải là một thành công trong thực tế, tuy nhiên, nó thực sự cần thiết cho mối liên hệ giữa độ chính xác và chuẩn hóa. Trong những năm sau nhiều nhà nghiên cứu đã thử chương trình nhận dạng khuôn mặt dựa trên các cạnh, khoảng cách giữa các

thuộc tính, và cách tiếp cận mạng neural khác. Trong khi một nhà nghiên cứu khác đã thành công với cơ sở dữ liệu ảnh có quy mô nhỏ, chưa ai giải quyết thành công các vấn đề thực tế với về cơ sở dữ liệu ảnh quy mô lớn, hay vị trí của khuôn mặt là không rõ ràng.

M.Turk và A.Pentland (1991) [17] sau đó chứng minh rằng các lỗi còn sót lại khi mã hóa bằng cách sử dụng eigenfaces có thể được sử dụng để phát hiện khuôn mặt trong hình ảnh tự nhiên lộn xộn, cũng như để xác định vị trí chính xác của các khuôn mặt trong ảnh. Sau đó, họ đã chứng minh rằng bằng cách kết hợp phương pháp này để phát hiện và khoanh vùng khuôn mặt với các phương pháp nhận dạng eigenface, người ta có thể đạt được độ tin cậy cao, thời gian xác thực khuôn mặt trong một môi trường bị hạn chế khá nhanh.

Wenyi Zhao, Arvinth Krishnaswamy, Rama Chellappa, Danie L.Swets, John Weng (1998) [22] sử dụng phương pháp PCA (phân tích thành phần chính) kết hợp LDA (phân tích độc lập tuyến tính). Bước 1, ánh xạ khuôn mặt từ không gian ảnh thô sang không gian các ảnh khuôn mặt (mỗi lớp khuôn mặt được nhận dạng sẽ được mô hình hóa bằng một không gian khuôn mặt) dùng PCA. Bước 2, sử dụng phương pháp LDA để tạo bộ phân loại tuyến tính có khả năng phân lớp các lớp khuôn mặt.

Emmanuel Viennet và Francoise Fogelman Soulie (1998) [8], sử dụng phương pháp mạng neural nhân tạo để xử lý và nhận dạng khuôn mặt.

Antonio J.Colmenarez và Thomas S.Huang (1998) [1], sử dụng kỹ thuật học thị giác và phù hợp mẫu 2-D. Ông quan niệm bài toán dò tìm khuôn mặt là thao tác phân loại khuôn mặt trong đó khuôn mặt thuộc về một lớp và các đối tượng khác thuộc về lớp còn lại bằng cách ước lượng mô hình xác suất cho mỗi lớp, và việc dò tìm sử dụng luật quyết định Maximum-likelihood.

Kazunori Okada, Johannes Steffens, Thomas Maurer, Hai Hong, Egor Elagin, Hartmut Neven, and Christoph (1998) [13], nhận dạng khuôn mặt dựa vào sóng Gabor và phương pháp phù hợp đồ thị bó. Với ý tưởng dùng đồ thị để biểu

diễn khuôn mặt, ảnh khuôn mặt được đánh dấu tại các vị trí đã được xác định trước trên khuôn mặt, gọi các vị trí này chính là các vị trí chuẩn. Khi thực hiện thao tác so khớp đồ thị với một ảnh, các điểm chuẩn (Jets) sẽ trích ra từ ảnh và so sánh các điểm chuẩn này với tất cả các điểm chuẩn tương ứng trong các đồ thị khác nhau, và đồ thị nào phù hợp nhất với ảnh sẽ được chọn.

Jeffrey Huang, Chengjun Liu, và Harry Wechsler (1998) [12], đề xuất thuật toán căn cứ trên tính tiến hóa (Evolutionary computation) và di truyền (Genetic) cho các tác vụ nhận dạng khuôn mặt. Đối với cách tiếp cận này, hai mắt sẽ được dò tìm trước tiên và thông tin này được xem là vết để quan sát khuôn mặt, trình xử lý dò tiếp mắt bằng cách sử dụng một thuật toán lai để kết hợp thao tác học và tiến hóa trong quá trình học.

Daniel Bgraham và Nigel M Allinson (1998) [7], sử dụng phương pháp được gọi là tạo bản sao không gian đặc trưng để biểu diễn và nhận dạng hướng di chuyển của khuôn mặt.

Guodong Guo, Stan Z.Li, Kap Luk Chan (17/1/2001) [9], dùng phương pháp SVM để nhận dạng khuôn mặt. Sử dụng chiến lược kết hợp nhiều bộ phân loại nhị phân để xây dựng bộ phân loại SVM đa lớp.

3. Mục đích nghiên cứu của luận văn, đối tượng, phạm vi nghiên cứu.

Mục đích nghiên cứu của luận văn

Nghiên cứu đề tài này nhằm mục đích tìm hiểu bài toán nhận dạng khuôn mặt, từ đó xây dựng các hệ thống ứng dụng trong thực tiễn như: điểm danh, giám sát người ra vào,...

Đối tượng, phạm vi áp dụng

Đề tài tập trung tìm hiểu một số phương pháp nhận dạng mặt người phổ biến hiện nay và đưa ra phương án nhận dạng cho bài toán nhận dạng mặt người.

Để đạt được mục tiêu trên, đề tài tập trung tìm hiểu các nội dung sau:

✓ Tìm hiểu phương pháp trích chọn đặc trưng PCA, cơ sở toán học của PCA.

✓ Tìm hiểu phương pháp phân lớp dữ liệu SVM, cơ sở toán học của SVM.

4. Tóm tắt cô đọng các luận điểm cơ bản và đóng góp mới của tác giả

Luận văn đã nghiên cứu một số phương pháp nhận dạng khuôn mặt, đề xuất mô hình kết hợp hai phương pháp PCA – SVM. Đây cũng là cơ sở cho việc thực hiện kết hợp một số phương pháp khác nhau cho bài toán nhận dạng khuôn mặt, là nền tảng cho việc hiện thực hóa mô hình bài toán vào ứng dụng thực tiễn.

5. Phương pháp nghiên cứu

Phương pháp nghiên cứu lý thuyết: Nghiên cứu tài liệu, ngôn ngữ và công nghệ liên quan. Tổng hợp các tài liệu lý thuyết về các phương pháp nhận dạng ảnh.

Phương pháp nghiên cứu thực nghiệm: Phân tích bài toán nhận dạng khuôn mặt người. Cài đặt mô phỏng, thử nghiệm trên một bộ dữ liệu tham khảo từ thực tế.

6. Nội dung luận văn

Luận văn gồm 04 chương, cụ thể như sau:

CHƯƠNG 1. PHÁT BIỂU BÀI TOÁN

Giới thiệu các cách thức nhận dạng người. Vì sao nên nhận dạng người bằng khuôn mặt. Tầm quan trọng của bài toán trong thực tiễn. Một số ứng dụng thực tiễn của bài toán nhận dạng khuôn mặt.

CHƯƠNG 2. CÁC PHƯƠNG PHÁP NHẬN DẠNG KHUÔN MẶT

Giới thiệu về hai phương pháp nhận dạng khuôn mặt được sử dụng trong luận văn là phương pháp trích chọn đặc trưng và máy vector hỗ trợ.

CHƯƠNG 3. ĐỀ XUẤT PHƯƠNG ÁN NHẬN DẠNG MẶT NGƯỜI

Đưa ra phương án xây dựng bài toán, mô hình bài toán, các bước thực hiện.

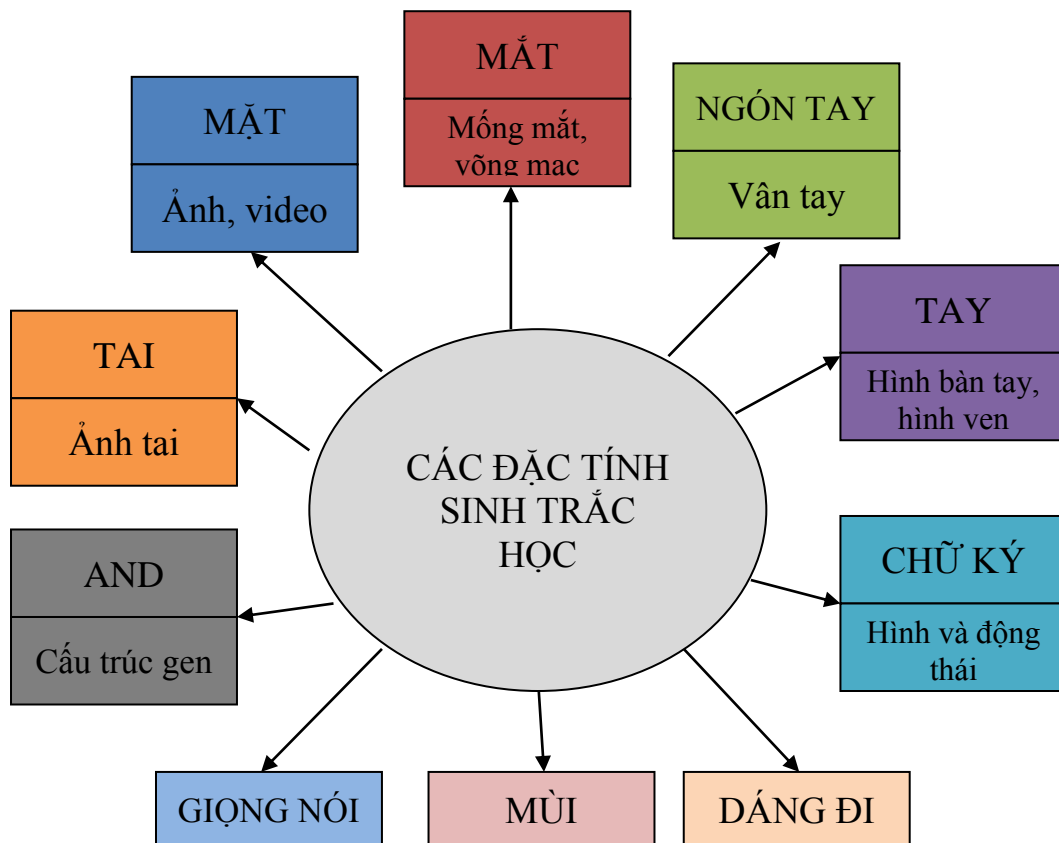
CHƯƠNG 4. ĐÁNH GIÁ THỬ NGHIỆM

Đưa ra các độ đo đánh giá tính đúng đắn của chương trình dựa trên các bộ dữ liệu thử nghiệm là AT&T, YaleB.

CHƯƠNG 1. PHÁT BIỂU BÀI TOÁN

1.1. Hệ thống sinh trắc học

Có rất nhiều phương pháp được sử dụng để nhận dạng người, sinh trắc học là một phương pháp hiệu quả trong việc nhận dạng người. Sinh trắc học hay Công nghệ sinh trắc học (Biometric) là công nghệ sử dụng những thuộc tính vật lý, đặc điểm sinh học riêng của mỗi cá nhân như vân tay, móng mắt, khuôn mặt... để nhận diện. Đây được coi là công cụ xác thực nhân thân hữu hiệu và phổ biến nhất hiện nay do mỗi người có những đặc điểm sinh trắc học duy nhất [18].



Hình 1.1 Các đặc tính sinh trắc học của người

Những năm gần đây, kỹ thuật sinh trắc học trong nhận dạng cá nhân nổi lên một cách đầy hứa hẹn, thay vì chứng thực người và cho phép họ truy cập vào các hệ thống dựa vào các phương thức như thông qua mật khẩu, thẻ thông minh,..., thì phương pháp kiểm tra một đặc tính sinh lý cá nhân và/hoặc hành vi để xác định

đanh tính của người dùng. Mật khẩu thì khó nhớ và có thể bị đánh cắp hoặc đoán ra. Thẻ thông minh, chìa khóa và những cái tương tự có thể bị làm giả, đánh cắp hay thất lạc. Tuy nhiên, đặc điểm sinh học của cá nhân không thể bị mất, bị đoán ra hay giả mạo. Hệ thống sinh trắc học nhận dạng dựa trên đặc điểm sinh lý như khuôn mặt, dấu vân tay, hình ngón tay, hình học mắt, móng mắt, võng mạc và giọng nói; hay đặc điểm hành vi như dáng đi, chữ ký ...

Tuy nhiên, hầu hết các phương pháp sinh trắc học yêu cầu phải có một hoặc nhiều hành động chủ quan của người sử dụng, ví dụ như người sử dụng cần đặt tay lên thiết bị quét vân tay để vân tay, hình tay được phát hiện hoặc phải đứng ở một vị trí cố định ở phía trước của một máy ảnh cho móng mắt hoặc võng mạc được xác định. Bên cạnh đó, việc thu dữ liệu cho hệ thống nhận dạng bằng vân tay có thể trở nên vô dụng khi mô biểu bì bị hư hỏng bởi một lý do nào đó (như bị thương, khuyết tay, ...); việc thu dữ liệu cho hệ thống nhận dạng bằng móng mắt, võng mạc có thể không thể thực hiện khi đối tượng cần nhận diện chuyển động.

1.2 Nhận dạng người qua khuôn mặt

Đây cũng là một phương pháp nhận dạng người dựa trên những đặc trưng sinh học của mỗi cá nhân. Khác với những phương pháp khác trong hệ thống nhận dạng người bằng sinh trắc học, nhận dạng mặt người qua khuôn mặt có thể thực hiện một cách thụ động mà không cần bất kỳ hành động rõ ràng, hoặc sự chủ động tham gia của người sử dụng kể từ khi camera thu được hình ảnh khuôn mặt từ xa. Điều này đặc biệt có lợi cho mục đích an ninh, giám sát. Việc sử dụng hệ thống nhận dạng bằng mặt người còn hạn chế được việc lây nhiễm virus, các vấn đề ảnh hưởng tới sức khỏe có thể xảy ra trong các hệ thống nhận diện sinh trắc học khác [18].

1.3 Ứng dụng của hệ thống nhận dạng mặt người

- Xác nhận cá nhân (one to one matching): khi đưa ra một khuôn mặt của cá nhân và yêu cầu nêu ra danh tính của họ, đòi hỏi hệ thống trả về kết quả họ là ai.

- Xác định (one to many matching): Khi đưa ra hình ảnh của một cá nhân không rõ danh tính, xác định danh tính người đó bằng cách so sánh (có thể sau khi mã hóa) với một cơ sở dữ liệu hình ảnh của các cá nhân đã biết [18].

Có rất nhiều lĩnh vực ứng dụng của hệ thống nhận dạng mặt người, trong đó nhận dạng khuôn mặt có thể được khai thác cho hai mục đích trên, một vài ứng dụng tiêu biểu đã được triển khai ở một số quốc gia như:

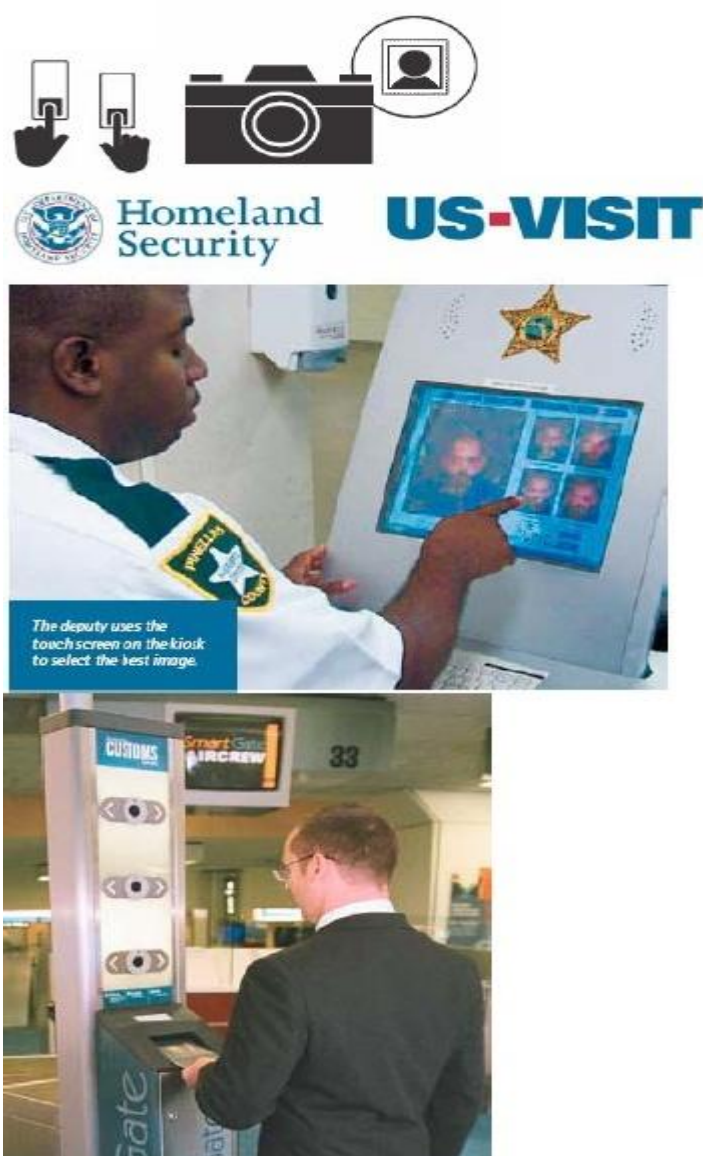
Các ứng dụng chuyên biệt cho ngành hàng không

Thời gian gần đây, vấn đề an toàn hàng không đang được cả thế giới quan tâm, là một đề tài nóng bỏng của xã hội. Có nhiều vụ tấn công khủng bố được thực hiện đối với ngành hàng không do việc đảm bảo an ninh, an toàn hàng không có những thiếu sót. Một số nước trên thế giới đã ứng dụng công nghệ nhận dạng khuôn mặt trong lĩnh vực hàng không như [19]:

Một chương trình tên US visits (United States Visitors and Immigrant Status Indicator Technology) áp dụng cho công dân du lịch tới Mỹ hoặc xin nhập cư vào Mỹ. Chương trình này được bắt đầu thực hiện từ tháng 5 năm 2004 tại các lãnh sự quán của Mỹ ở nước ngoài. Khi một công dân đến lãnh sự quán xin cấp visa thì họ sẽ được thu thập thông tin sinh trắc học, ảnh khuôn mặt để kiểm tra, đối chiếu với một cơ sở dữ liệu của bọn tội phạm nổi tiếng và những nghi can khủng bố. Nếu quá trình kiểm tra, đối chiếu được thông qua thì công dân này sẽ được cấp visa để tới Mỹ. Khi khách tới cửa khẩu, sân bay thì các thông tin sinh trắc học của họ sẽ được sử dụng để xác minh xem họ có phải là chủ sở hữu của visa không, có phải tội phạm nguy hiểm, nghi can khủng bố không. Chương trình US visits tăng cường sự an toàn của công dân Mỹ và du khách bằng cách xác minh nhận dạng của khách có visa, đồng thời nó tạo điều kiện đi lại và thương mại hợp pháp bằng cách tận dụng công nghệ và phát triển sử dụng sinh trắc học để tiến hành giám sát ngay tại biên giới, trước khi nhập cảnh vào Mỹ.

Ứng dụng tương tự được sử dụng tại Úc là SmartGate. Nó đã được thử nghiệm bởi Hải quan sân bay quốc tế Sydney từ năm 2002. Mục tiêu chính của thử

nghiệm này là để phát triển và giới thiệu một hệ thống tự xử lý sử dụng công nghệ sinh trắc học nhận dạng khuôn mặt để xác nhận danh tính và tinh giản thủ tục hải quan cho công dân nhập cảnh. Mất khoảng 17 giây để SmartGate thực hiện so sánh một hình ảnh thực của công dân được hệ thống tự động chụp tại điểm giao dịch SmartGate với một hoặc nhiều hình ảnh của họ đã được lưu trữ. Nếu những bức ảnh này cho kết quả là trùng khớp thì họ được chấp nhận nhập cảnh vào Úc.



Hình 1.2 Hệ thống xác minh nhân thân của công dân nhập cảnh [19]
Nhận dạng khuôn mặt được sử dụng kèm với thẻ truy cập

Tại các nước phát triển, hầu như mọi người dân đều dùng thẻ tín dụng để mua bán, rút tiền [2], trao đổi hàng hóa; trong tương lai gần, Việt Nam chúng ta có lẽ cũng áp dụng hình thức thanh toán này. Điều này rất nguy hiểm khi thẻ truy cập này bị người khác nhặt được hay biết được mật khẩu của chủ sở hữu thẻ này. Do đó vấn đề tiềm ẩn đó là không phải chủ thẻ cũng có thể sử dụng được thẻ. Để đảm bảo an toàn hơn cho việc phổ cập sử dụng thẻ trong toàn xã hội là chúng ta có thể dùng thêm một phương pháp nữa để xác minh tính xác thực của người sử dụng thẻ song song với mật khẩu, đó là sử dụng khuôn mặt như là một mật khẩu thứ hai để truy cập vào hệ thống cùng với thông tin từ thẻ truy cập. Để rút được tiền, người dùng cần thực hiện các bước sau:

- Đưa thẻ vào hệ thống
- Đứng trước camera để nhận dạng
- Xác minh người này có phải là chủ sở hữu của thẻ hay không?

Nếu thông tin về người sử dụng thẻ khớp với thông tin của chủ thẻ thì cho thực hiện việc rút tiền, ngược lại thì không cho rút tiền và tự động thông báo tới nhà chức trách.

Lần dấu vết đi tìm kẻ phạm tội

Từ những bức ảnh hay những đoạn video đã được ghi lại tự động tại hiện trường trước khi vụ khủng bố xảy ra, lực lượng an ninh có thể tìm được những kẻ khả nghi xuất hiện trong những bức ảnh hay đoạn video này. Nếu trong cơ sở dữ liệu của hệ thống có thông tin về những người xuất hiện trong đó, chúng ta có thể dễ dàng hơn trong việc tìm ra họ vì chúng ta biết họ là ai [24].

Tháng 2 năm 2012 FBI triển khai chương trình NGI Facial Recognition Pilot là để tiến hành tìm kiếm nhận dạng khuôn mặt dựa trên hình ảnh của kho lưu trữ quốc gia của FBI (có khoảng 12,8 triệu bức ảnh được lưu trữ) và cung cấp danh sách những đối tượng khả nghi cho cơ quan điều tra. Ảnh truy vấn là các hình ảnh thu được từ hiện trường, từ các trang mạng... và các yêu cầu truy vấn được hệ thống

xử lý kín (không có sự can thiệp của con người), và kết quả được trả về cho cơ quan điều tra như là một gợi ý về đối tượng tình nghi.

Hệ thống giám sát công nhân và chấm công tự động

Hiện nay trong các khu công nghiệp hay những công ty sản xuất lớn có hàng ngàn công nhân vào ra mỗi ngày nên việc giám sát kẻ gian vào công ty cũng như công việc chấm công rất phức tạp. Vậy làm thế nào để nhận ra từng nhân viên của công ty? Người ta đã xây dựng hệ thống có thể nhận dạng và chấm công tự động cho công nhân khi họ đi qua khu vực làm việc của hệ thống chấm công tự động.

CHƯƠNG 2. CÁC PHƯƠNG PHÁP NHẬN DẠNG KHUÔN MẶT

2.1 Phương pháp trích chọn đặc trưng Principal Components Analysis – PCA

2.1.1 Giới thiệu phương pháp trích chọn đặc trưng PCA

Phương pháp trích chọn đặc trưng được phát minh năm 1901 bởi Karl Pearson, và được phát triển một cách độc lập bởi Hotelling (1933). Hiện nay nó được sử dụng như một công cụ để phân tích dữ liệu nghiên cứu và thực hiện các mô hình dự đoán. PCA còn bao gồm cả việc tính toán phân tích các giá trị đặc trưng của một ma trận tương quan dữ liệu hay phân tích các giá trị đơn của ma trận dữ liệu thường sau khi tính trung bình dữ liệu của mỗi thuộc tính [11],[15].

PCA là phương pháp đơn giản nhất phân tích đa biến dựa trên các vector đặc trưng. Thông thường hoạt động của nó có thể được hiểu nhằm khám phá ra cấu trúc bên trong của dữ liệu. Nếu một tập dữ liệu đa biến được xem xét như tập các tọa độ trong một không gian dữ liệu nhiều chiều (mỗi trục biểu diễn một biến) thì phương pháp PCA cung cấp cho chúng ta một bức ảnh ít chiều, một cái bóng của vật thể khi quan sát từ chính những đặc trưng cơ bản nhất của vật thể đó.

Mục tiêu của phương pháp PCA là thực hiện giảm số chiều nhưng vẫn đảm bảo tối đa sự phân tán dữ liệu. Có thể nói phương pháp trích chọn đặc trưng tìm cách giữ lại những thành phần thống kê quan trọng nhất của tập dữ liệu ban đầu.

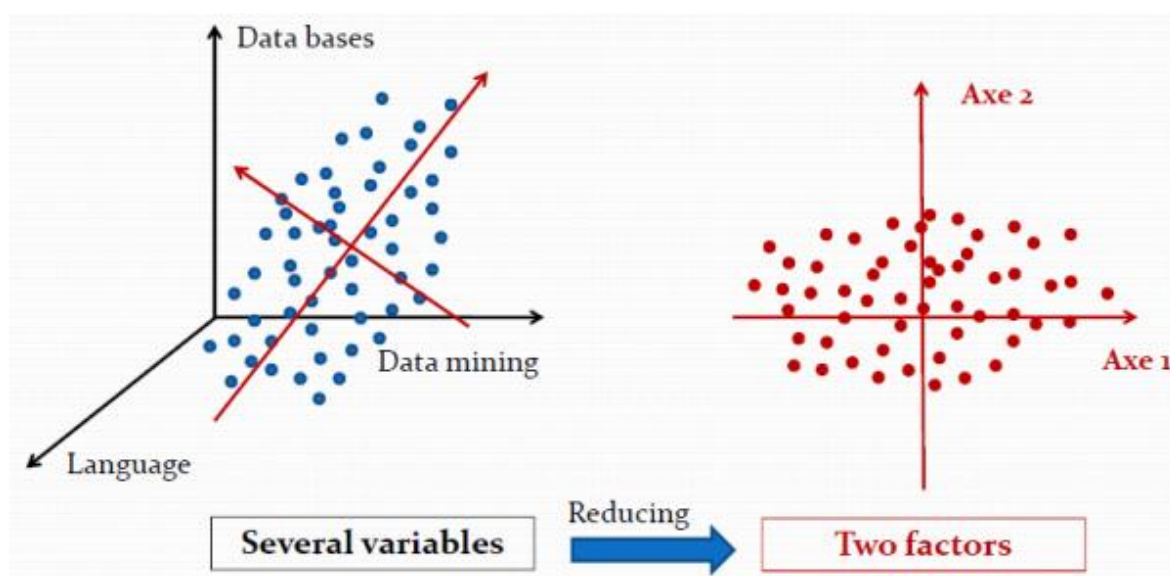
Giả sử ta cần giảm số chiều của dữ liệu từ N chiều xuống còn K ($K < N$) chiều nghĩa là ta cần tìm một ánh xạ từ không gian N chiều sang không gian K chiều.

Trong vấn đề đang xét của luận văn này, PCA là một thuật toán được sử dụng để tạo ra một ảnh mới từ ảnh ban đầu. Ảnh mới này có kích thước nhỏ hơn rất nhiều so với ảnh đầu vào và vẫn mang những đặc trưng cơ bản nhất của ảnh cần nhận dạng. Phương pháp PCA không cần quan tâm đến việc tìm ra các đặc điểm cụ thể của thực thể cần nhận dạng và mối quan hệ giữa các đặc điểm đó. Tất cả các chi tiết của thực thể đều được thể hiện ở ảnh mới được tạo ra từ PCA.

Bản chất của PCA là tìm ra một không gian mới theo hướng biến thiên mạnh nhất của một tập hợp các vector trong không gian cho trước. Ở không gian mới, ta hi vọng rằng việc phân loại sẽ cho ta bộ dữ liệu tốt hơn so với bộ dữ liệu trong không gian ban đầu.

Ví dụ: giả sử tập dữ liệu ban đầu được quan sát trong không gian ba chiều như hình bên trái. Rõ ràng ba trục này (các trục có tên Databases, Data minning, Language trong hình 2.4) không biểu diễn được tốt nhất mức độ biến thiên của dữ liệu. Phương pháp PCA sẽ tìm hệ trục tọa độ mới (là hệ trục không có tên trong hình bên trái) để biểu diễn tốt nhất mức độ biến thiên của dữ liệu. Sau khi tìm được không gian mới, dữ liệu sẽ được chuyển sang không gian này để được biểu diễn như trong hình bên phải. Rõ ràng hình bên phải chỉ cần hai trục tọa độ nhưng biểu diễn tốt hơn độ biến thiên của dữ liệu so với hệ trục ba chiều ban đầu.

Một ưu điểm của PCA là các trục tọa độ trong không gian mới luôn đảm bảo trực giao từng đôi một, mặc dù trong không gian ban đầu các trục có thể không trực giao với nhau.



Hình 2.1 Lựa chọn các trục tọa độ mới để biểu diễn dữ liệu [11]

Xem tập mẫu gồm K vectors trong không gian M chiều [5],[17]:

$$O = \{f_1, f_2, \dots, f_k, \dots, f_K\} \quad (2.1)$$

Giả sử cơ sở của không gian tập mẫu là:

$$B_M = \{\psi_1, \psi_2, \dots, \psi_M\} \quad (2.2)$$

Ta có:

$$f_1 = \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ \dots \\ f_{1M} \end{pmatrix}, f_2 = \begin{pmatrix} f_{21} \\ f_{22} \\ f_{23} \\ \dots \\ f_{2M} \end{pmatrix}, \dots, f_k = \begin{pmatrix} f_{k1} \\ f_{k2} \\ f_{k3} \\ \dots \\ f_{kM} \end{pmatrix} \quad (2.3)$$

Trong đó f_{km} là thành phần thứ m của vector f_k . Mỗi vector f_k có thể viết dưới dạng:

$$f_k = f_{k1}\psi_1 + f_{k2}\psi_2 + \dots + f_{kM}\psi_M \quad (2.4)$$

Như vậy mỗi phần tử của tập mẫu bây giờ như là một vector trong hệ cơ sở B_M .

Bằng cách sử dụng M vectors cơ sở ta có thể biểu diễn mỗi quan sát. Điều này cũng đúng khi ta chọn cơ sở gồm M-vectors trực giao được thể hiện trong 2.2.

Trong thực tế chúng ta không thể đưa tất cả các cơ sở M-chiều do nhiều lý do khác nhau như M quá lớn hay có chứa một số thông tin không quan trọng. Vì vậy chúng ta chuyển sang không gian có số chiều nhỏ hơn là N chiều. Khi đó tập mẫu là:

$$O = \{\tilde{f}_1, \tilde{f}_2, \dots, \tilde{f}_k, \dots, \tilde{f}_K\} \quad (2.5)$$

Cơ sở tập mẫu khi đó là B_N , lúc này mỗi phần tử của mẫu được viết dưới dạng:

$$\tilde{f}_k = \tilde{f}_{k1} \psi_1 + \tilde{f}_{k2} \psi_2 + \dots + \tilde{f}_{kN} \psi_N \quad (2.6)$$

Như vậy, sai số của mỗi phần tử của mẫu là:

$$\left\| f_k - \tilde{f}_k \right\| \quad (2.7)$$

Ở đây chúng ta quan tâm đến việc tìm một cơ sở trực giao (OrthNormal-ON)

B_N :

$$B_N = \{\psi_1, \psi_2, \dots, \psi_N\} \text{ với } \langle \psi_i, \psi_j \rangle = \delta_{ij} \quad (2.8)$$

$$\delta_{ij} = 0 \text{ nếu } i \neq j, \delta_{ij} = 1 \text{ nếu } i = j.$$

B_N được gọi là cắt (truncated) của B_M khi nó nhận ít vector cơ sở hơn. Vậy ta thể hiện O trong cơ sở nhỏ hơn với sai số:

$$\frac{1}{K} \sum_{k=1}^K \|f_k - \tilde{f}_k\|^2 \quad (2.9)$$

Ta có $\langle \psi_m, f_k \rangle = \psi_m^T f_k$ vậy

$$K \sum_{m=1}^N \psi_m^T S \psi_m \leq K \sum_{m=1}^N \lambda_{(m)} \quad (2.10)$$

Ma trận phân bố các phần tử của tập mẫu O là:

$$S = \frac{1}{K} \sum_{k=1}^K f_k f_k^T \quad (2.11)$$

Để sai số xấp xỉ bé nhất thì ta phải chọn B_N là N vectors riêng có ý nghĩa nhất của S thỏa mãn công thức sau:

$$\begin{aligned}\frac{1}{K} \sum_{k=1}^K \|f_k - f_k^{\sim}\|^2 &= \frac{T}{K} - \frac{1}{K} \sum_{k=1}^K \|f_k^{\sim}\|^2 \\ &\geq \frac{T}{K} - \sum_{m=1}^N \lambda_{(m)} = \sum_{m=N+1}^M \lambda_{(m)}\end{aligned}\quad (2.12)$$

$$\text{Với } \frac{T}{K} = \frac{1}{K} \sum_{k=1}^K \|f_k\|^2 = \lambda_1 + \lambda_2 + \dots + \lambda_M \quad (2.13)$$

Khi ta chọn N-vectors riêng và sai số xấp xỉ là nhỏ nhất và bằng tổng M-N trị riêng bé nhất của S thì ta được:

$$S = \frac{1}{K} \sum_{k=1}^K f_k f_k^T = \frac{1}{K} [f_1, f_2, \dots, f_K] \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_K \end{bmatrix} = \frac{1}{K} \mathbf{O} \mathbf{O}^T \quad (2.14)$$

Do S là đối xứng nửa xác định riêng nên những vector riêng là trực giao điều này đảm bảo cơ sở tối ưu là trực giao.

Như vậy cơ sở mới được xây dựng từ cơ sở quan sát ban đầu theo phương trình:

$$(\mathbf{O}^{\sim})^T = \mathbf{O}^T \mathbf{B}_N \quad (2.15)$$

Với

$$\mathbf{O}^{\sim} = [f_1^{\sim}, f_2^{\sim}, \dots, f_K^{\sim}], \mathbf{O} = [f_1, f_2, \dots, f_N], \mathbf{B}_N = [\psi_1, \psi_2, \dots, \psi_N] \quad (2.16)$$

Với $[\psi_1, \psi_2, \dots, \psi_K]$ là K-vectors riêng tương ứng với K-trị riêng lớn nhất của S, là ma trận hiệp phương sai của tập quan sát ban đầu.

Như ta biết phép biến đổi trực giao không làm thay đổi Trace-Vết của ma trận mà phép biến đổi cơ sở này giữ lại K-vectors riêng ứng với K-trị riêng lớn nhất. Nghĩa là sự phân bố các mẫu trong tập dữ liệu mới thu được luôn là lớn nhất.

Theo các kết quả nghiên cứu thông thường ta chọn K sao cho.

$$\frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^N \lambda_i} \geq \text{nguong}(e.g., 0.90 \text{ or } 0.95) \quad (2.17)$$

Với λ_i là các trị riêng của ma trận hiệp phương sai và $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$

Tóm lại, phương pháp PCA ánh xạ một vector từ không gian M chiều xuống không gian N chiều sẽ đi tìm các giá trị riêng và vector riêng của ma trận hiệp phương sai C của tập mẫu và giữ lại N vector riêng tương ứng với N giá trị riêng lớn nhất làm cơ sở cho không gian N chiều này.

2.1.2 Các bước thực hiện trích chọn đặc trưng bằng PCA

Giả sử ta có N ảnh khuôn mặt, là tập ảnh huấn luyện X_1, X_2, \dots, X_N

Biểu diễn mỗi ảnh thành ma trận $M \times 1$ có dạng:

$$X_i = (x_{i1}, x_{i2}, \dots, x_{iM})^T \text{ với } i = 1, \dots, N \quad (2.18)$$

Bước 1: Tính vector khuôn mặt trung bình của tập ảnh huấn luyện

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (2.19)$$

Bước 2: Tính vector độ lệch của mỗi khuôn mặt so với vector khuôn mặt trung bình

$$\Phi_i = X_i - \bar{X} \text{ với } i = 1, \dots, N \quad (2.20)$$

Bước 3: Tạo thành ma trận $M \times N$

$$A = [\Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_N] \quad (2.21)$$

Sau đó tính ma trận hiệp phương sai $M \times M$

$$C = \frac{1}{N} A A^T \quad (2.22)$$

Bước 4: Tính các giá trị riêng của ma trận hiệp phương sai C ta được

$$\lambda_1, \lambda_2, \dots, \lambda_M$$

Các giá trị riêng được sắp xếp theo chiều giảm dần. Ta chỉ giữ lại K giá trị riêng lớn nhất, thể hiện K đặc trưng quan trọng nhất là:

$$\lambda_1, \lambda_2, \dots, \lambda_K, \quad K \ll M$$

Có hai cách để xác định K :

Cách 1: Sắp xếp theo thứ tự giảm dần của các giá trị riêng tìm được. Theo dõi sự biến thiên của dãy trên, khi dãy không còn biến thiên hoặc biến thiên xấp xỉ bằng 0 thì lúc đó ta đã chọn đủ K .

Cách 2: Ta có thể chọn K theo công thức sau:

$$\frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^M \lambda_i} > \text{nguong}, \quad (0.9 \leq \text{nguong} \leq 0.95)$$

Phương châm ở đây là chọn làm sao cho số lượng thành phần chính là thấp nhất, đủ để giải thích khả năng phân tán của tập mẫu thành các lớp mẫu riêng cần thiết.

Bước 5: Tính các vector riêng của ma trận hiệp phương sai C

$$\psi_1, \psi_2, \dots, \psi_K$$

Chú ý 1: Ta nên chuẩn hóa vector riêng nghĩa là:

$$\Psi_i := \frac{\Psi_i}{\|\Psi_i\|}, i = 1, \dots, K \quad (2.23)$$

Chú ý 2: Nếu $M \gg N$ thì việc tính các giá trị riêng của ma trận hiệp phương sai C sẽ gặp khó khăn do C có kích thước lớn. Khi đó ta tính các giá trị riêng của ma trận $C' = \frac{1}{N} A^T A$ có kích thước $N \times N$. Sau đó ta chọn K giá trị riêng lớn nhất theo quy tắc của bước 4. Lúc này ta được K vector riêng tương ứng với K giá trị riêng này là:

$$v_1, v_2, \dots, v_K$$

Cuối cùng ta tính được K vector riêng của ma trận C theo công thức:

$$\Psi_i = A v_i \quad (2.24)$$

Bước 6: Giảm số chiều, chỉ giữ lại những thuộc tính tương ứng với các giá trị riêng lớn nhất (biểu diễn ảnh khuôn mặt trong không gian mới với K đặc trưng quan trọng nhất)

Trong không gian mới, với các vector cơ sở là $\psi_1, \psi_2, \dots, \psi_K$ mỗi ảnh khuôn mặt trong tập huấn luyện được biểu diễn thành tổ hợp tuyến tính của các vector cơ sở trên như sau:

$$X_i = x'_{i1}\psi_1 + x'_{i2}\psi_2 + \dots + x'_{iK}\psi_K \text{ trong đó } K \ll M \quad (2.25)$$

Khi đó tọa độ của X_i được tính theo công thức:

$$\begin{bmatrix} x'_{i1} \\ x'_{i2} \\ \dots \\ x'_{iK} \end{bmatrix} = \begin{bmatrix} \Psi_1^T \\ \Psi_2^T \\ \dots \\ \Psi_K^T \end{bmatrix} \cdot \Phi_i \quad (2.26)$$

* Đối với bài toán nhận dạng ảnh sử dụng PCA, cho ảnh khuôn mặt chưa biết là ai (ảnh có dạng giống các ảnh trong tập huấn luyện). Ta tiến hành các bước sau:

Giả sử ảnh đưa vào là X' :

Bước 1: Tính độ lệch của ảnh khuôn mặt X' so với vector khuôn mặt trung bình:

$$\Phi' = X' - \bar{X} \quad (2.27)$$

Bước 2: Biểu diễn ảnh X' trong không gian mới như sau:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_K \end{bmatrix} = \begin{bmatrix} \Psi_1^T \\ \Psi_2^T \\ \dots \\ \Psi_K^T \end{bmatrix} \cdot \Phi' \quad (2.28)$$

Bước 3: Ta tìm ảnh thứ p trong tập huấn luyện có khoảng cách gần nhất với khuôn mặt cần nhận dạng. Nếu khoảng cách này nhỏ hơn hoặc bằng 1 ngưỡng chấp nhận được thì ta kết luận ảnh đưa vào nhận dạng là người có ảnh thứ p trong tập huấn luyện, ngược lại kết quả là không tìm thấy người đó trong hệ thống.

$$e_r = \min_p \| X' - X_p \| \quad (2.29)$$

Cách để nhận được các thành phần chính

Các thành phần chính có thể nhận được bằng cách chiếu các vector dữ liệu có nhiều biến động vào không gian mở rộng từ các vector đặc trưng.

Khi lấy số đặc trưng càng về sau thì khả năng biến động càng thấp, có nghĩa là mối quan hệ giữa các phần tử càng cao, khi đó sự giao nhau giữa các lớp mẫu trong tập mẫu càng lớn.

Ngược lại, khi ta lấy số lượng thành phần chính không đủ, thì khả năng phân tán của tập mẫu càng cao.

2.2 Máy vector hỗ trợ SVM

Thuật toán SVM ban đầu được tìm ra bởi Vladimir N. Vapnik và dạng chuẩn hiện nay sử dụng lẽ mềm được tìm ra bởi Vapnik và Corinna Cortes năm 1995. Phương pháp này ngày càng được sử dụng rộng rãi trong nhiều lĩnh vực, đặc biệt là lĩnh vực phân loại và nhận dạng mẫu [6].

Phương pháp SVM ngày càng được sử dụng phổ biến do:

- SVM rất hiệu quả trong việc giải quyết bài toán với dữ liệu có số chiều lớn.
- SVM giải quyết tốt vấn đề overfitting, là dữ liệu có nhiều và tách rời nhóm, hoặc dữ liệu huấn luyện ít.
- SVM là phương pháp phân lớp nhanh.
- SVM có hiệu suất tổng hợp tốt và khả năng tính toán cao.

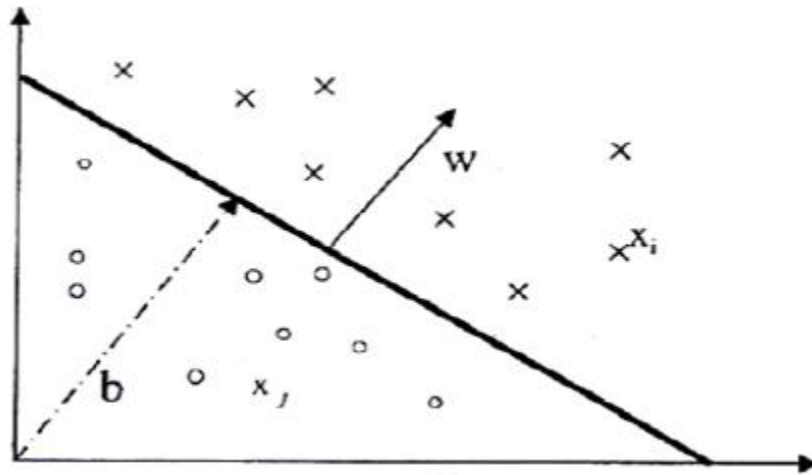
Ta khảo sát l mẫu quan sát. Mỗi mẫu quan sát gồm một cặp: một vector $x_i \in R_n$, $i = 1, \dots, l$ với một giá trị xác định y_i mà giá trị của nó xuất phát từ việc gán chủ quan từ người tổ chức dữ liệu. Gọi $P(x,y)$ là hàm phân phối xác suất giữa x và y và chưa được xác định tường minh. Cách tổ chức dữ liệu như trên có tính tổng quát cao hơn so với việc chúng ta kết hợp cứng giữa mỗi y với một x , điều này cho phép tính được phân phối của y dựa vào dữ liệu x cho trước. Tuy nhiên, sau phần này, ta thừa nhận cố định y với x cho trước.

Nhiệm vụ của hệ máy học là học ánh xạ $x_i \rightarrow y_i$, được định nghĩa từ một tập hợp các ánh xạ $x \rightarrow f(x,a)$, trong đó hàm $f(x,a)$ được gán nhãn bởi các tham số a (a có thể hiệu chỉnh được trong quá trình xử lý trên tập học). Hệ máy học có thể xem như là một hệ quyết định. Với dữ liệu đầu vào là x cho trước, chọn ra một a thích hợp, và kết xuất sẽ là $f(x,a)$. Ta có thể chọn a theo nhiều cách khác nhau, ở đây chúng ta sẽ tiếp cận theo phương pháp máy học.

2.2.1 SVM tuyến tính [4],[6],[9]

2.2.1.1 SVM trong trường hợp tập mẫu có thể phân cách tuyến tính

Máy học được huấn luyện trên dữ liệu có thể phân loại tuyến tính, đây là trường hợp đơn giản nhất của hệ máy học. Gán nhãn dữ liệu huấn luyện $\{x_i, y_i\}$, với $i = 1, \dots, l$, $x_i \in R^d$, $y_i \in \{-1, 1\}$. Giả sử có các siêu mặt phẳng phân loại mẫu dương với mẫu âm (gọi là “siêu mặt phân cách”). Điểm x nằm trên siêu mặt thỏa phương trình $w x + b = 0$, trong đó w là pháp tuyến của siêu mặt phân cách, $|b|$ là khoảng cách từ siêu mặt đến gốc toạ độ, và $\|w\|$ là độ lớn (Euclide) của w . Đặt $d_+(d_-)$ là khoảng cách ngắn nhất từ siêu mặt phân cách đến mẫu dương (âm) gần nhất. Định nghĩa “bờ” (lề - margin) của siêu mặt phân cách (kí hiệu là r), là $(d_+) + (d_-)$.



Hình 2.2 Phân cách theo siêu phẳng (w, b) trong không gian hai chiều của tập mẫu [4]

Với trường hợp tập mẫu có thể phân loại tuyến tính, thuật toán SVM chỉ đơn giản là tìm siêu mặt có khoảng cách bờ là cực đại. Các mô tả trên đây được công thức hoá như sau: giả sử mọi điểm trong tập học thỏa các ràng buộc:

$$w x_i + b \geq 1 \text{ với } y_i = +1 \quad (2.30)$$

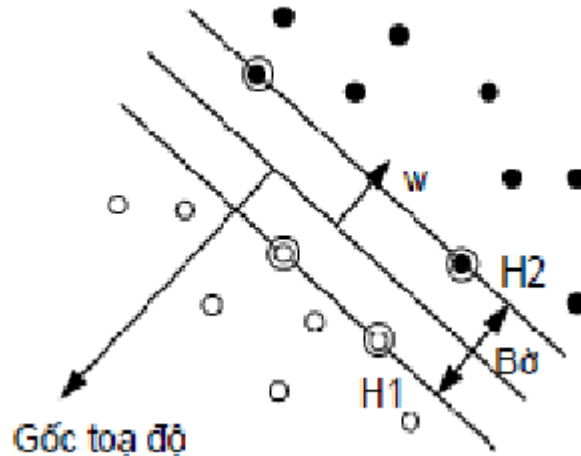
$$w x_i + b \leq -1 \text{ với } y_i = -1 \quad (2.31)$$

Kết hợp hai bất phương trình trên, ta được bất đẳng thức ràng buộc:

$$y_i (w x_i + b) - 1 \geq 0 \forall i \quad (2.32)$$

Các dữ liệu mẫu thỏa (2.30) nằm trên siêu mặt $H_1: wx_1 + b = 1$ có pháp tuyến là vector w và khoảng cách đến gốc tọa độ là $|1-b|/\|w\|$. Tương tự, các dữ liệu mẫu thỏa công thức (2.31) nằm trên siêu mặt $H_2: wx_2 + b = -1$ có pháp tuyến là w và khoảng cách đến gốc tọa độ là $|-1-b|/\|w\|$. Lúc này $(d_+) = (d_-) = (1/\|w\|)$ và bờ là $r = 2/\|w\|$.

Do H_1 và H_2 song song với nhau và giữa chúng không có điểm dữ liệu nào nên ta có thể tìm cặp siêu mặt phân cách có bờ r cực đại, bằng việc cực tiểu $\|w\|$ với ràng buộc (2.32).



Hình 2.3 Siêu mặt phân cách tuyến tính cho trường hợp phân cách được [4]

Như hình trên, những điểm huấn luyện thỏa (2.32) là những điểm nằm trên một trong hai siêu mặt H_1, H_2 , việc loại bỏ những điểm này làm thay đổi lời giải, được gọi là các vector hỗ trợ. Chúng là những điểm được bao bằng hình tròn (có 2 điểm trên H_1 và 2 điểm trên H_2).

Để giải quyết bài toán này, chúng ta có thể dùng hàm Lagrange bởi vì các lý do sau. Thứ nhất, ràng buộc bất đẳng thức (2.32) sẽ được thay thế bằng ràng buộc trên hệ số nhân Lagrange để dễ làm việc hơn. Thứ hai, dữ liệu huấn luyện sẽ chỉ xuất hiện dưới dạng phép nhân vô hướng giữa các vector, điều này cho phép tổng quát hóa trong trường hợp phi tuyến.

Với ràng buộc bất đẳng thức (2.32) có vế trái ≥ 0 thì ta nhân với hệ số Lagrange $\alpha_i > 0$ và phương trình ràng buộc sau khi nhân với hệ số Lagrange thì sẽ bị trừ khỏi hàm mục tiêu. Với các ràng buộc là đẳng thức thì hệ số nhân Lagrange không bị ràng buộc. Khi đó hàm Lagrange có dạng sau:

$$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i y_i (wx_i + b) + \sum_{i=1}^l \alpha_i \quad (2.33)$$

Mục tiêu của chúng ta là phải cực tiểu L_p theo pháp tuyến w và b , và đạo hàm của L_p triệt tiêu $\forall \alpha_i \geq 0$ (gọi tập ràng buộc này là C_1). Hay nói cách khác là giải bài toán đối ngẫu đó tìm cực đại L_p với điều kiện đạo hàm của L_p triệt tiêu với w , b và cũng với ràng buộc $\alpha_i \geq 0$ (gọi tập ràng buộc này là C_2).

$$\text{Đạo hàm } L_p \text{ triệt tiêu với } w \text{ và } b \text{ ta có các điều kiện: } w = \sum_{i=1}^l \alpha_i x_i y_i \quad (2.34)$$

$$\text{Vì đây là các ràng buộc tuyến tính nên } \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.35)$$

$$\text{Thay vào (2.33) ta được: } L_D = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j x_i x_j y_i y_j \quad (2.36)$$

Trong trường hợp tuyến tính và có thể phân loại, việc huấn luyện SVM là làm cực đại L_D theo $\alpha_i > 0$ với ràng buộc (2.35), với lời giải được tính theo (2.34). Khi đó, các mẫu thỏa điều kiện $\alpha_i > 0$ được gọi là “vector hỗ trợ - SVM” và nằm trên một trong hai siêu mặt H_1 , H_2 . Các điểm dữ liệu còn lại có $\alpha_i = 0$ và nằm trên H_1 hoặc H_2 thỏa mãn trường hợp dấu bằng xảy ra của (2.32); hoặc chúng nằm về một phía của H_1 hoặc H_2 trong trường hợp thỏa mãn bất đẳng thức (2.32). Khi đó, các vector hỗ trợ là các thành phần tới hạn của tập huấn luyện. Nếu mọi điểm huấn luyện khác bị loại bỏ và việc huấn luyện được lặp lại thì các siêu mặt phân cách đã tìm thấy không bị thay đổi.

Nếu số thuộc tính của các mẫu dữ liệu là lớn, chúng ta có thể đơn giản hóa bằng cách chuyển bài toán với điều kiện Karush-Kuhn-Tucker, tương đương với

phương trình Lagrange dạng kép. Điều kiện Karush-Kuhn-Tucker có vai trò quan trọng trong bài toán tối ưu ràng buộc. Với bài toán trên, điều kiện Karush-Kuhn-Tucker có thể phát biểu như sau:

$$\frac{\partial}{\partial w_v} L_p = w_v - \sum_i \alpha_i y_i x_{iv} = 0 \text{ với } v = 1, \dots, d \quad (2.37)$$

$$\frac{\partial}{\partial b} L_p = -\sum_i \alpha_i y_i = 0 \quad (2.38)$$

$$y_i(w x_i + b) - 1 \geq 0 \text{ với } i = 1, \dots, l \quad (2.39)$$

$$\alpha_i \geq 0 \quad \forall i \quad (2.40)$$

Giải bài toán này ta tìm được α và từ đó tính được w theo công thức (2.34).
 Tính b , vận dụng điều kiện KKT cho bài toán gốc, ta có được:

$$\alpha_i [y_i(w x_i + b) - 1] = 0 \quad \forall i = 1, 2, \dots, l \quad (2.42)$$

Do đó, với mỗi i ta có 2 trường hợp:

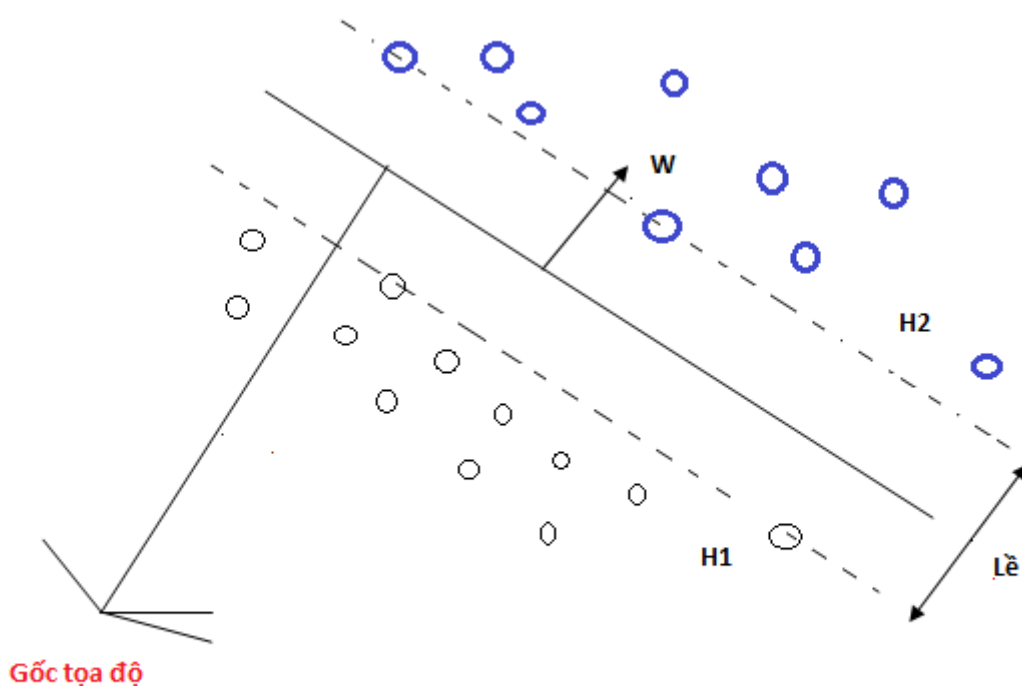
✓ $\alpha_i = 0$: trong trường hợp này $y_i(w x_i + b) - 1 > 0$ suy ra x_i không nằm trên siêu phẳng H1 hay H2. Vì $\alpha_i = 0$ nên x_i không tham gia vào việc cấu trúc w theo công thức (2.49). Những x_i này là không cần thiết và có thể được bỏ đi mà không ảnh hưởng đến w .

✓ $\alpha_i > 0$: trong trường hợp này $y_i(w x_i + b) - 1 = 0$ suy ra x_i nằm trên siêu phẳng biên H1 hay H2, x_i được gọi là support vector và tham gia vào việc tính toán w . Thông thường số lượng support vector nhỏ hơn nhiều số lượng mẫu.

Do vậy, để tính b ta chỉ cần chọn một giá trị i mà khi đó $\alpha_i > 0$, lúc đó ta lần lượt tính các giá trị của ngưỡng b và lấy giá trị trung bình của những ngưỡng b thu được đó.

2.2.1.2 SVM trong trường hợp tập mẫu không thể phân cách tuyến tính

Trường hợp tập mẫu có thể phân cách tuyến tính chỉ là một trường hợp đặc biệt. Trong trường hợp tổng quát, tập mẫu là tập không thể phân cách tuyến tính bằng một siêu phẳng.



Hình 2.4 Không thể phân hoạch tập mẫu bằng một siêu phẳng [6]

Tuy nhiên, ta vẫn muốn dùng siêu phẳng để phân cách tập mẫu này. Để có thể áp dụng được phương pháp ở trên, ta phải gán cho mỗi mẫu x_i một sai số ξ_i không âm với $i = 1, \dots, l$ để xem như có thể phân cách tuyến tính. Ta được các biểu thức như sau:

$$wx_i + b \geq +1 - \xi_i \text{ với } y_i = +1 \quad (2.43)$$

$$wx_i + b \leq -1 + \xi_i \text{ với } y_i = -1 \quad (2.44)$$

$$\xi_i \geq 0 \quad \forall i = 1, 2, \dots, l \quad (2.45)$$

Với cách đặt như vậy thì sai số thực nghiệm Empirical Risk là: $\sum_{i=1}^l \xi_i^2$

Giống như trên, nếu cực tiểu $\|w\|$ thì sẽ làm cho chiều VC nhỏ, và từ đó VC confidence nhỏ. Do đó, khi cực tiểu cùng lúc Empirical Risk và VC confidence, ta có thể cực tiểu biểu thức sau:

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad (2.46)$$

Trong đó C là một hằng số dương tùy chọn. Nếu chọn C lớn sẽ làm cho Empirical Risk nhỏ do tập trung vào việc cực tiểu ξ . Còn nếu chọn C nhỏ thì sẽ làm cho giá trị VC confidence nhỏ do cực tiểu w (cách này được gọi là C-SVM do dựa trên giá trị của khoảng cách lề, ngoài ra còn một phương án khác là tối thiểu số điểm lề, được gọi là v-SVM).

Ta có thể phát biểu lại bài toán như sau:

$$\text{Cực tiểu: } f(w, b, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad (2.47)$$

$$\text{Với } y_i(wx_i + b) - 1 + \xi_i \geq 0 \quad \xi_i \geq 0, i=1, 2, \dots, l$$

Bài toán đối ngẫu của bài toán trên sẽ là:

Cực đại $\Phi(\alpha, v)$ dựa trên điều kiện $\alpha \in R^l, v \in R^l, \alpha \geq 0, v \geq 0$ dựa trên

$$\Phi(\alpha, v) = \inf \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i [y_i(wx_i + b) - 1 + \xi_i] - \sum_{i=1}^l v_i \xi_i \right\} \quad (2.48)$$

Cũng theo điều kiện Fermat cực tiểu của

$$L(w, b, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i [y_i(wx_i + b) - 1 + \xi_i] - \sum_{i=1}^l v_i \xi_i \quad (2.49)$$

Xảy ra tại w, b, ξ sao cho:

$$\frac{\partial L(w, b, \xi)}{\partial w} = w - \sum_{i=1}^l \alpha_i x_i y_i = 0 \Rightarrow w = \sum_{i=1}^l \alpha_i x_i y_i \quad (2.50)$$

$$\frac{\partial L(w, b, \xi)}{\partial b} = - \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.51)$$

$$\frac{\partial L(w, b, \xi)}{\partial \xi_i} = C - \alpha_i - v_i = 0 \Rightarrow 0 \leq \alpha_i \leq C \quad (2.52)$$

Lúc đó giá trị cực tiểu của L là:

$$L_o(w, b, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i [y_i (wx_i + b) - 1 + \xi_i] - \sum_{i=1}^l v_i \xi_i \quad (2.53)$$

$$L_o(w, b, \xi) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j x_i x_j y_i y_j \quad (2.54)$$

Như vậy, bài toán đối ngẫu được viết lại thành tìm cực đại của:

$$F(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j x_i x_j y_i y_j \quad (2.55)$$

Dựa trên điều kiện $0 \leq \alpha_i \leq C, i = 0, 1, 2, \dots, l; \sum_{i=1}^l \alpha_i y_i = 0$

Bài toán này hoàn toàn giống bài toán mẫu có thể phân cách tuyến tính với điều kiện mở rộng $\alpha_i \leq C$. Giải bài toán này ta tìm được các thành phần của công thức. Từ đó ta tính được w, trong đó x_i là các support vector và các vector lỗi ứng với $0 \leq \alpha_i \leq C$.

Để tìm b, sử dụng điều kiện Karush-Kuhn-Tucker cho bài toán gốc, ta có:

$$\alpha_i [y_i (wx_i + b) - 1 + \xi_i] = 0 \quad (2.56)$$

$$v_i \xi_i = 0 \quad (2.57)$$

Chọn một giá trị α_i sao cho $0 < \alpha_i < C$ lúc đó $y_i (wx_i + b) - 1 + \xi_i = 0$ để điều kiện xảy ra.

Do $C - \alpha_i - v_i = 0$ suy ra $v_i > 0$ nên có thêm điều kiện $\xi_i = 0$ để bất đẳng thức xảy ra. Lúc đó ta có thể tính được

$$b = y_i - wx_i$$

Như vậy, ta đã tìm được siêu phẳng phân chia một tập dữ liệu không thể phân cách tuyến tính. Siêu phẳng này được gọi là siêu phẳng khoảng cách phân hoạch mềm (soft margin hyperplane).

2.2.2 SVM phi tuyến

Trong nhiều ứng dụng, một bộ phân lớp phi tuyến có độ chính xác cao hơn bộ phân lớp tuyến tính. Tuy nhiên, phân lớp tuyến tính có lợi thế đó là thuật toán đơn giản. Điều này đặt ra câu hỏi có cách nào phân lớp tuyến tính có thể mở rộng cho phân lớp phi tuyến không. Ta thấy rằng dữ liệu trong bài toán huấn luyện xuất hiện dưới dạng tích vô hướng $x_i x_j$. Giả sử ta ánh xạ dữ liệu mẫu sang không gian Euclide H khác [6]:

$$\Phi: R^d \rightarrow H$$

Lúc này thuật toán huấn luyện chỉ phụ thuộc vào dữ liệu qua tích vô hướng trong không gian H, tức là hàm có dạng $\Phi(x_i) \cdot \Phi(x_j)$. Nếu có một hàm Kernel (K) mà $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, khi đó ta sẽ chỉ cần dùng hàm K trong thuật toán huấn luyện mà không cần quan tâm tới Φ . Ví dụ:

$$K(x_i, x_j) = e^{\frac{-\|x_i - x_j\|^2}{2\sigma^2}} \quad (2.58)$$

Như vậy là chỉ cần biết dạng của hàm hạt nhân K mà không cần biết cụ thể ánh xạ Φ . Lúc đó hàm nhận dạng trở thành:

$$f(x) = \sum_{i=1}^{N_s} \alpha_i y_i \Phi(s_i) \cdot \Phi(x) + b = \sum_{i=1}^{N_s} \alpha_i y_i K(s_i, x) + b \quad (2.59)$$

Trong đó s_i là vector hỗ trợ, từ (2.61) ta có $\Phi(s_i) \cdot \Phi(x) = K(s_i, x)$

Tuy nhiên không thể chọn tùy ý hàm K mà phải chọn hàm K sao cho tồn tại Φ mà $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$. Điều kiện để đảm bảo vấn đề này là điều kiện Mercer.

Như vậy, phương pháp SVM phi tuyến là tìm một hàm nhân K , sau đó giải bài toán soft-margin hyperplane với việc thay $x_1 x_2 = K(x_1, x_2)$ để tìm ra w và b .

Tóm lại, xây dựng hàm nhân với một mẫu thử x^* nào đó là:

$$f(x^*) = \text{sign}[\sum_{i=1}^l \alpha_i y_i K(x_i, x^*) + b] \quad (2.60)$$

Một số hàm nhân thường được dùng trong bảng 2.3

Kiểu hàm nhân	Công thức
Linear kernel	$K(x_i, x_j) = x_i^T x_j + C$
Polynomial kernel	$K(x_i, x_j) = (\alpha x_i^T x_j + 1)^d$
Radial basic tangent kernel	$K(x_i, x_j) = e^{\frac{-\ x_i - x_j\ ^2}{2\sigma^2}}$
Hyperbolic tangent kernel (sigmoid kernel)	$K(x_i, x_j) = \tanh(\alpha x_i^T x_j + C)$

Bảng 2.5 Một số hàm nhân thường dùng

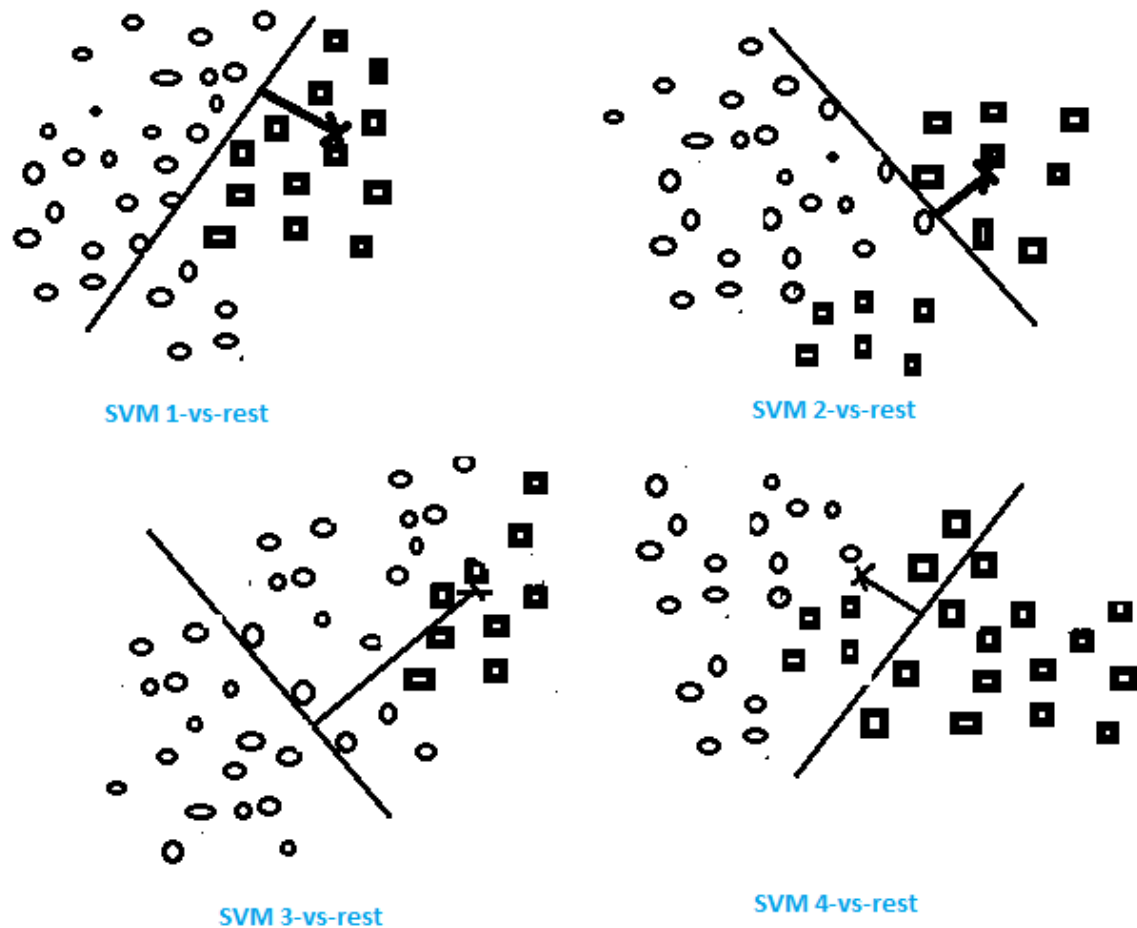
2.2.3 SVM cho vấn đề nhận dạng đa lớp

Trong các phần trên đã đề cập đến vấn đề SVM nhận dạng cho hai lớp, $y_i = \{-1, 1\}$. Phương pháp SVM chỉ được xây dựng trên nền tảng nhận dạng hai lớp, bắt đầu từ lý thuyết học thống kê. Đối với nhận dạng đa lớp, ta có thể kết hợp nhiều SVM hai lớp lại với nhau từng đôi một theo một số cách sau [10],[14]:

2.2.3.1 Chiến lược One-vs-all (One-vs-rest)

Giả sử cần phải nhận dạng s lớp thì sử dụng s SVM. SVM thứ i sẽ phân biệt lớp i với lớp còn lại (nếu i là lớp 1 thì lớp còn lại là -1). Một mẫu thử x sẽ được cho

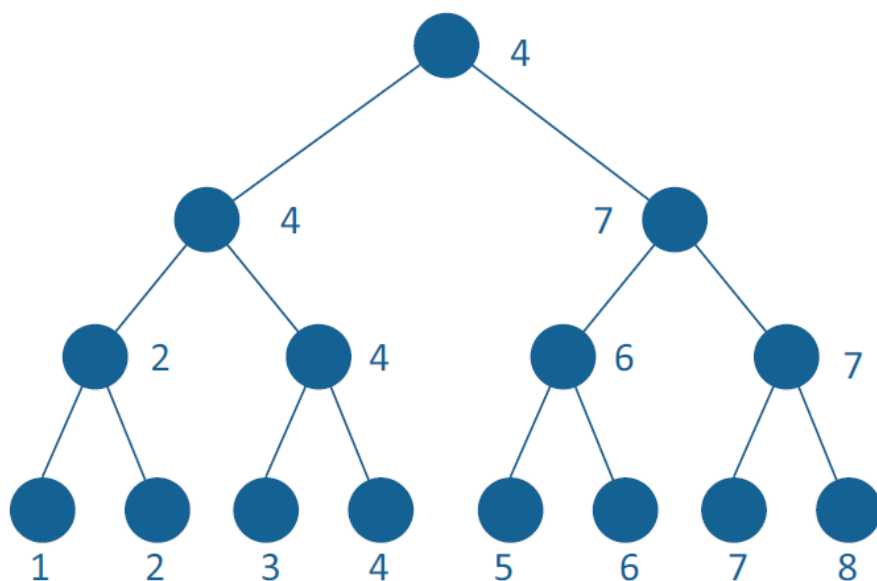
kiểm tra trên s SVM này. Hàm nhận dạng của SVM thứ i nào cho ra kết quả bé nhất thì mẫu thử thứ x thuộc lớp thứ i đó.



Hình 2.6 Mẫu cần nhận dạng là SVM 2-vs-rest vì có giá trị bé nhất [10]

2.2.3.2 Chiến lược One-vs-one

Để giải quyết trường hợp này, ta có thể dùng phương pháp DAG-SVM (A Directed Acyclic Graph – đồ thị có hướng không chu trình) hoặc sử dụng cây nhị phân.



Hình 2.7 Cấu trúc 1 cây nhị phân cho nhận dạng khuôn mặt 8 lớp [10]

Giả sử có 8 lớp trong tập dữ liệu, cây quyết định được biểu diễn như trong hình 2., trong đó các số 1-8 mã hóa các lớp. Bằng cách so sánh từng cặp, 1 lớp đại diện cho người chiến thắng giữa 2 lớp sẽ được chọn. Các lớp được chọn (từ cấp thấp nhất của cây quyết định) sẽ lên cấp trên, rồi lại được so sánh tiếp. Cuối cùng một lớp duy nhất sẽ xuất hiện ở đỉnh của cây.

Nếu s (số lượng lớp) không là lũy thừa của 2 thì ta phân tích:

$$s = 2^{n_1} + 2^{n_2} + \dots + 2^{n_l} \text{ với } n_1 \geq n_2 \geq \dots \geq n_l$$

Bởi vì bất kì một số tự nhiên có thể được chia thành các số nguyên dương hữu hạn là bội số của 2. Nếu s chẵn thì $n_1 = 0$ và ngược lại. Lưu ý rằng ta có thể phân tích s bằng nhiều cách khác nhau. Sau khi phân tích, việc nhận dạng được thực hiện trong từng cây nhị phân, các lớp đầu ra của các cây nhị phân này được dùng để tạo ra cây nhị phân khác. Nó được lặp lại cho tới khi chỉ còn một đầu ra.

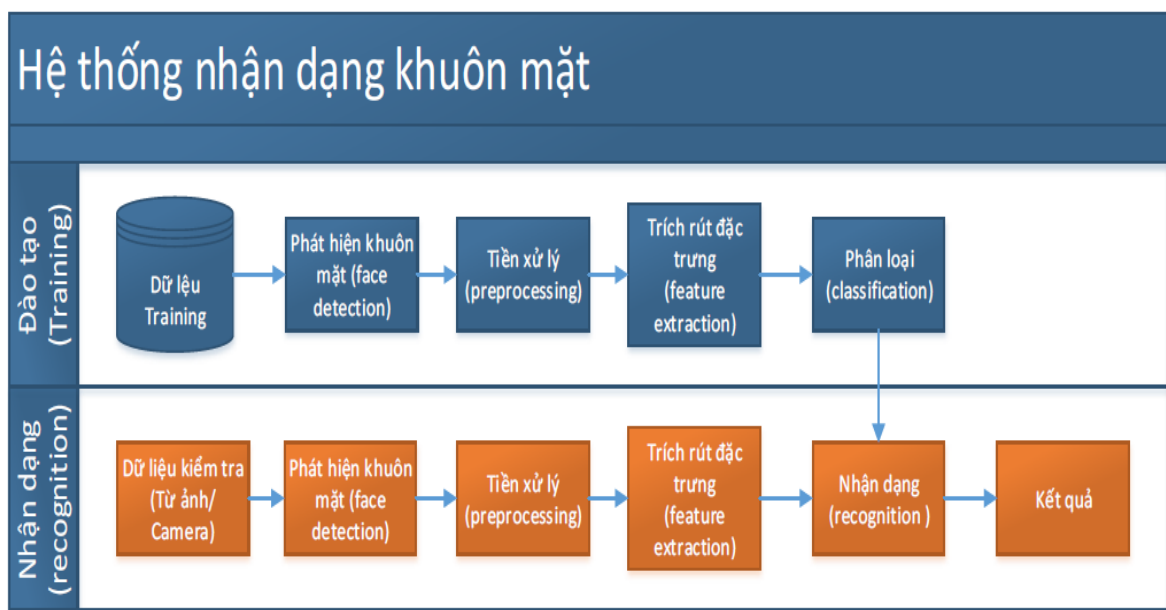
SVM học $\frac{s(s-1)}{2}$ hàm phân biệt trong giai đoạn huấn luyện và thực hiện ($s-$

1) phép so sánh theo cấu trúc cây nhị phân đã tạo ra.

CHƯƠNG 3. ĐỀ XUẤT PHƯƠNG ÁN NHẬN DẠNG MẶT NGƯỜI

Có nhiều cách để kết hợp các phương pháp nhận dạng khuôn mặt nhằm đạt hiệu quả cao, trong luận văn này tác giả xin được trình bày phương pháp kết hợp PCA – SVM.

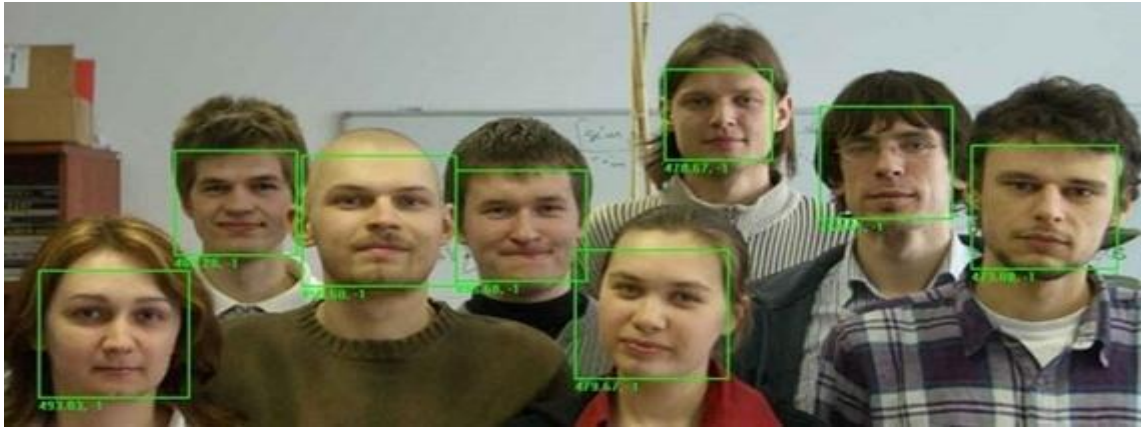
Ảnh đưa vào nhận dạng, sau khi sử dụng Haar cascade để phát hiện và lấy ra khuôn mặt trong ảnh, sẽ tiến hành trích chọn đặc trưng sử dụng PCA và sau đó sử dụng SVM để phân lớp/nhận dạng.



Hình 3.1 Mô hình hệ thống nhận dạng khuôn mặt

3.1 Phát hiện khuôn mặt (face detection)

Khi đưa vào một bức ảnh, trong đó có thể có một hoặc nhiều khuôn mặt, cũng có thể không có khuôn mặt nào trong ảnh. Do đó ta cần phải kiểm tra xem trong bức ảnh ta đưa vào có khuôn mặt hay không. Phát hiện khuôn mặt sẽ tách tất cả khuôn mặt trong ảnh. Có rất nhiều kỹ thuật khác nhau được sử dụng để phát hiện khuôn mặt như AdaBoost hay phân loại Haar cascade. Trong luận văn này sẽ sử dụng thuật toán Haar cascade có sẵn trong thư viện mở OpenCV.



Hình 3.2 Phát hiện khuôn mặt trong ảnh

3.2 Tiền xử lý ảnh (preprocessing)

Bước này nhằm mục đích lọc nhiễu, nâng cao chất lượng ảnh, chuẩn hóa kích cỡ ảnh. Nâng cao chất lượng ảnh là làm cho ảnh tốt hơn theo ý đồ sử dụng. Thường là ảnh thu được có nhiễu, cần phải loại bỏ nhiễu hay ảnh không sắc nét, bị mờ cần làm rõ các chi tiết. Có nhiều cách để nâng cao chất lượng ảnh như [24]:

Ảnh với độ tương phản thấp có thể do điều kiện sáng không đủ hay không đều, hoặc do tính không tuyến tính hay biến động nhỏ của bộ cảm nhận ảnh. Để điều chỉnh lại độ tương phản của ảnh, cần điều chỉnh lại biên độ trên toàn dải hay trên dải có giới hạn bằng cách biến đổi tuyến tính biên độ đầu vào.

Tách nhiễu là trường hợp đặc biệt của dẫn độ tương phản. Tách nhiễu được ứng dụng có hiệu quả để giảm nhiễu khi biết tín hiệu vào trên khoảng xác định. Để làm trơn nhiễu hay tách nhiễu, người ta sử dụng các bộ lọc tuyến tính (lọc trung bình, thông thấp) hay lọc phi tuyến (trung vị, giả trung vị).

Trích chọn bit: Mỗi điểm ảnh thường được mã hóa trên N bit, nếu $N = 8$ thì ta có ảnh 256 mức xám. Trong các bit mã hóa này, người ta chia làm bit bậc thấp và bit bậc cao. Với bit bậc cao, độ bảo toàn thông tin cao hơn so với bit bậc thấp. Các bit bậc thấp thường biểu diễn nhiễu hay nền, do đó ta chọn n bit có nghĩa nhất.

Trong nội dung luận văn, với bộ ảnh thử nghiệm là các bộ ảnh chuẩn của AT&T, YaleB do đó không thực hiện bước này.

3.3 Trích rút đặc trưng (feature extraction)

Ảnh đưa vào sau khi đã được thực hiện tiền xử lý vẫn có số lượng thành phần ảnh lớn, do đó ta cần thực hiện trích rút đặc trưng của ảnh, chỉ giữ lại những thành phần quan trọng nhất nhằm mục đích giảm độ lớn ảnh, tăng tốc độ xử lý của hệ thống. Sử dụng phương pháp PCA để thực hiện trích rút đặc trưng, ta tiến hành như mục 2.1.2.

Đầu vào của PCA là các vector cột có M thành phần biểu diễn ảnh trong tập huấn luyện, đầu ra là các vector cột có K thành phần biểu diễn ảnh đã được trích rút đặc trưng.

3.4 Nhận dạng/Phân lớp (recognition/classification)

Bước nhận dạng hay phân lớp tức là xác định danh tính (identity) hay nhãn của ảnh (label) – đó là ảnh của ai. Ở bước nhận dạng/phân lớp, ta sử dụng phương pháp SVM (Support Vector Machine). SVM sẽ tiến hành phân lớp ảnh trong tập huấn luyện, khi đưa ảnh vào nhận dạng sẽ được so sánh, tìm ra ảnh đó thuộc vào lớp nào.

3.5 Dữ liệu hệ thống nhận dạng

Dữ liệu của hệ thống nhận dạng được chia làm hai tập:

- ✓ Tập huấn luyện (training set)
- ✓ Tập để nhận dạng (probe set)

Tập training gồm các ảnh được dùng để huấn luyện, thông thường tập này được dùng để sinh ra một không gian con (sub projection space) là một ma trận. Bước training nhằm hai mục đích: giảm số chiều (dimension reduction) của các ảnh trong tập huấn luyện vì các ảnh này thường có số lượng thuộc tính lớn nên nếu để nguyên thì việc tính toán sẽ rất lâu, thứ hai là làm tăng tính phân biệt giữa các ảnh khác lớp (định danh khác nhau, mỗi định danh sẽ có số lượng ảnh nhất định), ngoài ra còn có thể làm giảm tính phân biệt giữa các ảnh thuộc cùng một lớp. Sau khi thực hiện chiếu các ảnh vào không gian con, hệ thống lưu lại kết quả là một ma trận với mỗi cột của ma trận là một vector tương ứng với ảnh định danh đã biết để thực hiện

nhận dạng hay phân lớp. Nhận dạng hay phân lớp được thực hiện với tập các ảnh probe, sau khi tiền xử lý xong, mỗi ảnh sẽ được áp dụng phương pháp trích chọn đặc trưng (như các ảnh thuộc tập training) và chiếu vào không gian con.

Việc phân lớp sẽ dựa trên phương pháp SVM, định danh của ảnh cần nhận dạng sẽ được gán định danh của ảnh có khoảng cách gần với nó nhất.

CHƯƠNG 4. ĐÁNH GIÁ THỬ NGHIỆM

4.1 Độ đo đánh giá hiệu năng

Ta có thể đánh giá hiệu năng của chương trình theo nhiều cách, trong luận văn này, tác giả đánh giá hiệu năng của chương trình theo độ chính xác của ảnh tìm được. Tức là với những ảnh đưa vào, hệ thống sẽ cho ta bao nhiêu kết quả chính xác, theo công thức sau:

$$\text{Độ chính xác} = (\text{số ảnh nhận đúng} / \text{tổng số ảnh đưa vào nhận dạng}) \times 100\%$$

4.2 Bộ dữ liệu thử nghiệm

Trong luận văn này, chương trình sẽ được tiến hành thực nghiệm trên hai bộ dữ liệu là AT&T và YaleB.

4.2.1 Bộ dữ liệu AT&T



Hình 4.1 Một số ảnh trong bộ dữ liệu AT&T

Đây là bộ dữ liệu thử nghiệm được công bố bởi Phòng thí nghiệm Nghiên cứu Olivetti tại Cambridge, Vương quốc Anh, theo link bên dưới.

http://www.cl.cam.ac.uk/Research/DTG/attarchive/pub/data/att_faces.zip

Bộ dữ liệu ảnh này có 10 hình ảnh khác nhau của 40 đối tượng riêng biệt. Đối với một số đối tượng, các hình ảnh được chụp vào những thời điểm khác nhau, ánh sáng thay đổi nhẹ. Biểu hiện khuôn mặt bao gồm các trạng thái: mở mắt, nhắm mắt, mỉm cười, không mỉm cười; các chi tiết trên khuôn mặt như: đeo kính, không đeo kính.

Các tập tin có định dạng PGM. Kích thước của mỗi hình ảnh là 92x112, mức độ xám 8-bit. Những hình ảnh được tổ chức trong 40 thư mục, mỗi một đối tượng được đặt trong một thư mục riêng biệt.

4.2.2 Bộ dữ liệu ảnh YaleB



Hình 4.2 Một số ảnh trong bộ dữ liệu YaleB

Đây là bộ dữ liệu được phát triển bởi đại học Yale, Hoa Kỳ. Bộ ảnh này chứa ảnh của 39 người với 9 tư thế chụp khác nhau, mỗi tư thế có 65 ảnh.

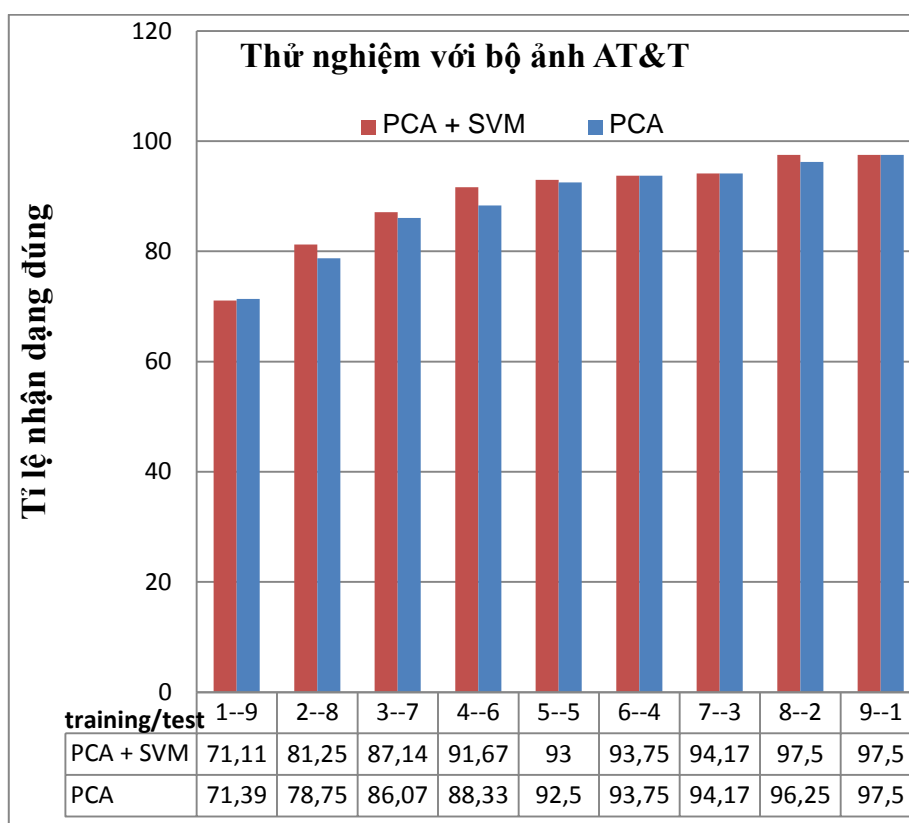
Dữ liệu thử nghiệm của chương trình lấy ảnh của 39 người, mỗi người 65 ảnh chụp trực diện, tổng cộng 2535 ảnh. Ảnh của mỗi người có sự thay đổi về hướng nguồn sáng, độ cao đặt camera. Đường dẫn của bộ ảnh thử nghiệm:

<http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>

4.3 Kết quả thử nghiệm

4.3.1 Thử nghiệm với bộ ảnh AT&T

Tiến hành thử nghiệm với bộ ảnh AT&T, với 10 ảnh cho mỗi người ta chia thành các trường hợp thử nghiệm khác nhau với số lượng ảnh huấn luyện/ảnh để nhận dạng thay đổi. Ảnh trong tập huấn luyện thay đổi từ 1 tới 9, số ảnh trong tập để nhận dạng thay đổi từ 9 tới 1, áp dụng cho tập ảnh của mỗi người đồng thời. Ta thu được kết quả như bảng sau:



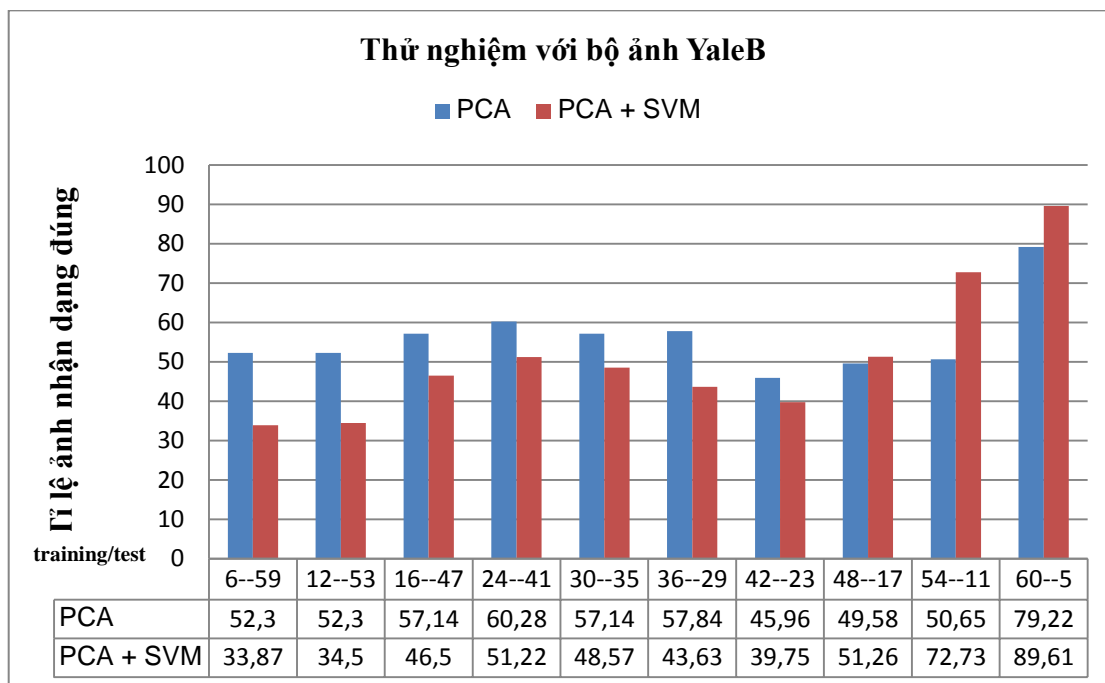
Biểu đồ 4.3 Kết quả thử nghiệm với bộ dữ liệu AT&T

Với bộ ảnh AT&T, khi tập huấn luyện chỉ có một ảnh duy nhất thì kết quả của phương pháp PCA tốt hơn phương pháp PCA – SVM. Phương pháp PCA – SVM cho hiệu quả nhận dạng đúng tốt hơn khi chỉ thực hiện nhận dạng đơn thuần dùng PCA với số lượng ảnh tăng lên của bộ dữ liệu huấn luyện. Thử nghiệm với bộ dữ liệu AT&T đạt kết quả tốt ngay cả khi số ảnh huấn luyện nhỏ.

4.3.2 Thử nghiệm với bộ ảnh YaleB

Với bộ ảnh YaleB (trường hợp ảnh được chụp trực diện), ta tiến hành thử nghiệm với các trường hợp ảnh huấn luyện/ảnh để nhận dạng thay đổi, mỗi lần thay đổi là đồng thời thực hiện tăng 6 ảnh trong tập huấn luyện và giảm 6 ảnh trong tập để nhận dạng. Số ảnh trong tập học thay đổi từ 6 tới 60, số ảnh trong tập kiểm tra thay đổi từ 59 tới 5, áp dụng cho tập ảnh của mỗi người đồng thời. Với bộ dữ liệu ảnh YaleB, với cách tách bộ ảnh huấn luyện theo phương pháp lấy những bức ảnh ở

đầu danh sách đưa vào bộ huấn luyện, những ảnh còn lại đưa vào bộ ảnh nhận dạng. Ta có kết quả như bảng sau:



Biểu đồ 4.4 Kết quả thử nghiệm trên bộ dữ liệu YaleB

Ta nhận thấy, với trường hợp số lượng ảnh huấn luyện ít thì phương pháp PCA và phương pháp PCA – SVM cho ta kết quả không cao. Khi tỉ lệ ảnh huấn luyện/ảnh nhận dạng là tương đương nhau thì kết quả của chương trình đạt mức trung bình. Lúc này, độ chính xác của hai phương pháp là tương đương nhau. Cho tới khi tỉ lệ ảnh huấn luyện/ảnh nhận dạng là 54/11 thì độ chính xác của phương pháp kết hợp PCA – SVM cho thấy sự thay đổi rõ rệt, tốt hơn hẳn phương pháp PCA thuần túy. Phương pháp SVM là việc học máy, do đó khi số lượng ảnh trong tập huấn luyện lớn thì việc học máy cho hiệu quả cao khi thực hiện phân lớp/nhận dạng. Ngược lại, khi tập ảnh huấn luyện có số lượng ảnh bé thì sẽ tạo ra những bộ phân loại không chính xác, do đó kết quả nhận dạng không cao.

Vậy, không phải lúc nào ta cũng kết sử dụng phương pháp kết hợp PCA – SVM. Với bài toán có số lượng ảnh huấn luyện ít thì ta dùng phương pháp PCA cũng cho kết quả tốt, thực hiện nhận dạng tốc độ khá nhanh. Khi số lượng ảnh trong

tập huấn luyện lớn thì ta nên áp dụng phương pháp kết hợp PCA –SVM, lúc này phương pháp sẽ phát huy hiệu quả tốt.

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Qua thời gian nghiên cứu về các phương pháp nhận dạng khuôn mặt, tác giả đã tìm hiểu được một số phương pháp và áp dụng các phương pháp này cho bài toán nhận dạng khuôn mặt. Những kết quả chính được tổng kết như sau:

- ✓ Giới thiệu chi tiết về phương pháp trích chọn đặc trưng PCA và máy vector hỗ trợ SVM.
- ✓ Áp dụng kết hợp hai phương pháp này cho bài toán nhận dạng khuôn mặt.
- ✓ Nhận xét và đánh giá những kết quả đạt được của các giải thuật trong bài toán nhận dạng khuôn mặt.

Bên cạnh những kết quả đã đạt được, còn có những vấn đề mà thời điểm hiện tại khóa luận chưa giải quyết được, định hướng phát triển:

- ✓ Tăng hiệu năng của chương trình khi tập huấn luyện có số ảnh ít, độ khác nhau giữa các ảnh trong tập huấn luyện với tập dùng để nhận dạng lớn.
- ✓ Nghiên cứu về một số các giải thuật trích chọn đặc trưng và phân lớp dữ liệu khác.
- ✓ Xây dựng một chương trình hoàn chỉnh có giao diện tương tác với người sử dụng ứng dụng trong thực tiễn.

TÀI LIỆU THAM KHẢO

[1] Antonio J.Colmenarez and Thomas S.Huang (1997), Face Detection and Recognition, *Face Recognition from Theory and Applications*, Department of Electrical and Computer Engineering, Coordinated Science Laboratory, and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 N. Mathews Ave, USA, pp. 174-185.

[2] Aru, Okereke Eze, Ihekweaba Gozie, Facial Verification Technology for Use In Atm Transactions, *American Journal of Engineering Research (AJER) e-ISSN: 2320-0847 p-ISSN: 2320-0936 Volume-02, Issue-05*, pp-188-193.

[3] Baback Moghaddam and Alex Pentland (1998), Beyond Linear Eigenspaces: Bayesian Matching for Face Recognition, *Face Recognition from Theory and Applications*, MIT Media Laboratory, 20 Ames St. Cambridge, MA 02139, USA, pp. 230-243.

[4] B. Heisele, P. Ho, T. Poggio, Face Recognition with Support Vector Machines: Global versus Component-based Approach, *Proc. of the Eighth IEEE International Conference on Computer Vision, ICCV 2001*, Vol. 2, 09-12 July 2001, Vancouver, Canada, pp. 688-694

[5] Alaa Eleyan and Hasan Demirel, *PCA and LDA based Neural Networks for Human Face Recognition*, Eastern Mediterranean University, Northern Cyprus.

[6] Christopher J.C. Burges (1998), A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery* 2, pp. 121-167.

[7] Daniel Bgraham và Nigel M Allinson (1998), Characterising Virtual Eigensignatures for General Purpose Face Recognition, *Face Recognition from Theory and Applications, Image Engineering and Neural Computing Group*, Department of Electrical Engineering and Electronics, University of Manchester Institute of Science and Technology. Manchester M60 1QD, UK, pp. 446-456.

- [8] Emmanuel Viennet and Francoise Fogelman Soulie (1998), Connectionists Methods for Human face Rrocessing, *Face Recognition from Theory and Applications*, University Paris 13, 93430 Villetaneuse, France, pp. 124-156.
- [9] Guodong Guo, Stan Z.Li, Kap Luk Chan (2001), Learning Similarity for Texture Image Retrieval, *School of EEE*, Nanyang, Technological University Nanyang Avenue, Singapore.
- [10] Bernd Heisele, Purdy Ho and Tomaso Poggio, *Face Recognition with Support Vector Machines: Global versus Component-based Approach*, Massachusetts Institute of Technology Center for Biological and Computational Learning Cambridge, MA 02142.
- [11] H. Moon, P.J. Phillips, Computational and Performance aspects of PCA-based Face Recognition Algorithms, *Perception*, Vol. 30, 2001, pp. 303-321
- [12] Jeffrey Huang, Chengjun Liu, and Harry Wechsler, Eye Detection and Face Recognition Using Evolutionary Computation, *Face Recognition from Theory and Applications*, *Department of Computer Science*, George Mason University, pp. 348-377.
- [13] Kazunori Okada, Johannes Steffens, Thomas Maurer, Hai Hong, Egor Elagin, Hartmut Neven, and Christoph (1998), Computer Science Department and Center for Neural Engineering, *Face Recognition from Theory and Applications*, University of Southrn California Los Angeles, USA, pp. 186-205.
- [14] K. Jonsson, J. Matas, J. Kittler, Y.P. Li, Learning Support Vectors for Face Verification and Recognition, *Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition*, 26-30 March 2000, Grenoble, France, pp. 208-213
- [15] M.A. Turk, A.P. Pentland, Face Recognition Using Eigenfaces, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, USA, 3-6 June 1991, pp. 586-591

[16] M.Kirby and L.Sirovich, Application of the karhunen-loeve procedure for the characterization of human faces, *IEEE Pattern Analysis and Machine Intelligence*, vol.12, no. 1, 1990, pp.103-108.

[17] M. Turk, A. Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991, pp. 71-86

[18] Rabia Jafri and Hamid R. Arabnia, A Survey of Face Recognition Techniques, *Journal of Information Processing Systems*, Vol.5, No.2, June 2009, pp. 44-68.

[19] Setiawan Hadi, Iping Supriana Suwardi, and Farid Wazdi, *Technology of Face Recognition for Security System*, Mathematics Department. UNPAD, Jl. Dipati Ukur 35 Bandung.

[20] Steven W.Smith (1999), *Chapter 27: Data compression*, The Scientist and Engineer's Guide to Digital Signal Processing, California Technical Publishing San Diego, California, pp. 481-502.

[21] T.Kohonen, Self-organization and Associative Memory, *Springer-Verlag*, Berlin, 1989.

[22] Wenyi Zhao, Arvinth Krishnaswamy, Rama Chellappa, Danie L.Swets, John Weng (1998), Discriminant Analysis of Principal Components for Face Recognition, *Face Recognition from Theory and Applications*, Centre for Automation Research, University of Maryland, pp. 73-85.

[23] Nguyễn Quang Hoan, Xử lý ảnh, Học viện Công nghệ Bưu chính viễn thông, 2006

[24] <http://www.fbi.gov>