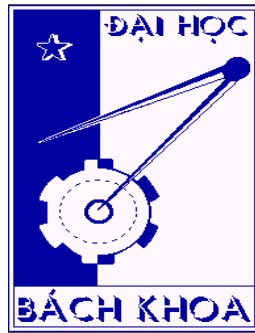


TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

-----o0o-----



BÀI TẬP LỚN
XỬ LÝ ẢNH SỐ

Đề tài 44: Tìm hiểu các đặc trưng sinh trắc ảnh khuôn mặt, nghiên cứu ứng dụng của phép biến đổi KL và phân tích thành các thành phần chính (PCA) trong trích chọn đặc trưng khuôn mặt (Face Feature Extraction).

Nhóm sinh viên thực hiện: **Đinh Thị Am**
 Lê Thanh Bình
 Lê Thế Văn
Lớp **:** **HTTT-k52**

Hà Nội – tháng 11/2010

MỤC LỤC

I. Đặc trưng sinh trắc ảnh khuôn mặt.	3
1. Phương pháp sinh trắc học.....	3
2. Ứng dụng	4
3. Đặc trưng sinh trắc khuôn mặt người và một số phương pháp xác định.	7
a. Các đặc trưng khuôn mặt	7
b. Kết cấu.....	11
c. Đa đặc trưng.....	12
II. Ứng dụng của phép biến đổi KL và phân tích thành các thành phần chính (PCA) trong trích chọn đặc trưng ảnh khuôn mặt.....	15
1. Sơ lược đại số tuyến tính trong thống kê.	15
Vecto riêng, trị riêng và sự chéo hóa ma trận.....	15
a.Kì vọng.....	16
b.Ma trận hiệp phương sai.....	Error! Bookmark not defined.
2. Phép biến đổi KL	17
a.Cơ sở lý thuyết của KL.	17
b.Biến đổi KL	19
3.Phương pháp phân tích thành phần chính PCA trong trích chọn đặc trưng khuôn mặt.....	20
a.Yêu cầu	20
b.Trích đặc trưng bằng phương pháp PCA	20
c.Kỹ thuật trích đặc trưng bằng PCA	23
d.Các đánh giá quan trọng về rút trích đặc trưng bằng phương pháp PCA.....	25
II. Phân công công việc và các khó khăn trong quá trình thực hiện.	25
1. Những khó khăn trong quá trình thực hiện.....	25
2. Phân công công việc trong nhóm.	26
III. Các tài liệu tham khảo.	27

I. Đặc trưng sinh trắc ảnh khuôn mặt.

1. Phương pháp sinh trắc học.

Nhận dạng sinh trắc học (biometrics) là công nghệ được sử dụng như một công cụ giúp xác thực người dùng dựa trên bạn là ai, chứ không phải dựa trên thứ mà bạn mang theo (các loại thẻ) hoặc cái mà bạn biết hoặc nhớ (mật khẩu). Dựa trên các đặc điểm riêng biệt của mỗi cá nhân về sinh lý và hành vi như đặc trưng về vân tay, khuôn mặt, giọng nói, chữ ký... các hệ thống nhận dạng sinh trắc học phân tích, đánh giá nhân dạng của đặc trưng đó. Với nhu cầu bảo mật ngày càng cao của các ứng dụng như kiểm soát truy nhập, kiểm soát vào ra, kiểm soát xuất nhập cảnh... nhận dạng sinh trắc học đã chứng minh tiềm năng ứng dụng trong các hệ thống quản lý nhân dạng số lượng lớn.

Việc nhận dạng sinh trắc học dựa trên các đặc trưng sinh trắc của con người. Các đặc trưng sinh trắc được chia làm **hai loại**:

Đặc trưng sinh lý: là các đặc trưng liên quan đến hình dạng, cấu tạo của cơ thể, ví dụ vân tay, khuôn mặt, vân lòng bàn tay, tĩnh mạch ngón tay, tĩnh mạch lòng bàn tay, tròng mắt, hình dạng tay, tai, cấu tạo răng, mùi cơ thể, ADN...

Đặc trưng hành vi: là các đặc trưng liên quan đến hành động, ví dụ dáng đi, giọng nói, chữ ký, hình thức gõ phím...

Vân tay, khuôn mặt và giọng nói, chữ ký là những đặc trưng sinh trắc đã được sử dụng từ lâu và rất quen thuộc với mọi người. Khuôn mặt và giọng nói được sử dụng bởi con người cho mục đích xác định nhân dạng của bạn bè, người quen... Chữ ký vốn được sử dụng phổ biến như một phương pháp xác thực của cá nhân trên các giấy tờ, công văn, sổ sách... Trong khi đó, vân tay từ lâu vốn đã được sử dụng trong chứng minh thư nhân dân, bằng lái xe... Các ứng dụng quản lý ra vào, quản lý nhân công bằng vân tay cũng đã trở nên rất phổ biến trong các công ty, tổ chức. Bên cạnh đó, những đặc trưng sinh trắc như tròng mắt, tĩnh mạch lòng bàn tay, tĩnh mạch ngón tay gần đây nhận được nhiều sự quan tâm và đang trên đường trở thành những giải pháp nhận dạng nhanh chóng hơn, chính xác hơn.

2. Ứng dụng

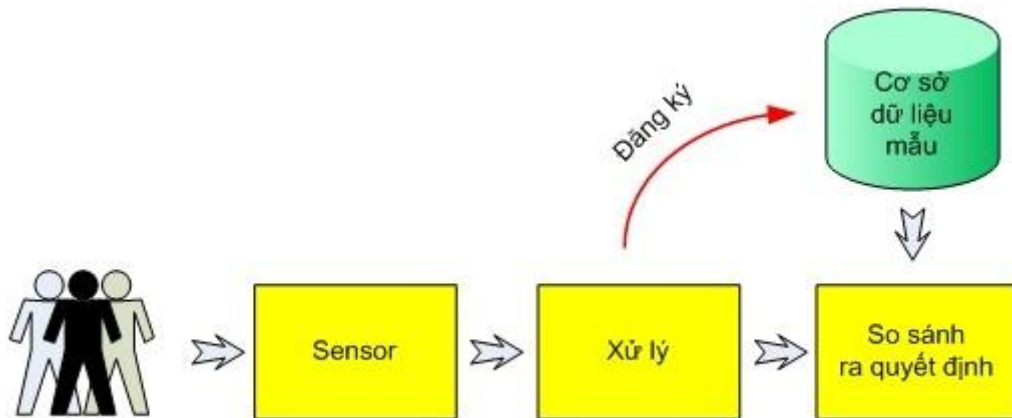
Với các ưu điểm về tính an toàn, tiện lợi so với các phương pháp xác thực truyền thống, các hệ thống sinh trắc đang ngày càng trở nên phổ biến đối với các ứng dụng cần xác thực danh tính của người sử dụng. Các ứng dụng của nhận dạng sinh trắc học rất đa dạng, được áp dụng rộng rãi trong cả các hoạt động của chính phủ cũng như các công ty, tổ chức thương mại, bao gồm từ việc quản lý nhân công, quản lý khách hàng, quản lý vào ra, tới quản lý xuất nhập cảnh, quản lý tội phạm... Theo International Biometric Group, các ứng dụng của nhận dạng sinh trắc học có thể được liệt kê như sau:

- Thi hành pháp luật: công nghệ nhận dạng sinh trắc học đã được sử dụng từ lâu như một phương tiện an toàn để xác thực danh tính của tội phạm. Một trong các ứng dụng này là thu thập vân tay tại hiện trường các vụ án, so sánh với các mẫu vân tay có sẵn trong cơ sở dữ liệu để xác định danh tính của người cần điều tra. Hiện nay, cơ sở dữ liệu vân tay lớn nhất thuộc về Cục điều tra liên bang Mỹ FBI với khoảng 70 triệu mẫu vân tay.
- Giám sát: các hệ thống nhận dạng sinh trắc học được sử dụng để tự động định vị, theo dõi và định danh người trong một khu vực nhất định. Hiện nay, các hệ thống này bao gồm một số camera giám sát kết hợp với các đặc trưng sinh trắc để giám sát. Khuôn mặt là đặc trưng sinh trắc được sử dụng nhiều nhất trong loại này. Những hệ thống giám sát gần đây nhất đã có thể xác định được danh tính của người từ khoảng cách 200m sử dụng khuôn mặt. Tròng mắt cũng đang được ứng dụng để xác định danh tính từ khoảng cách xa. So với khuôn mặt, tròng mắt cho độ chính xác cao hơn nhưng vì kích thước nhỏ nên việc thu nhận tròng mắt yêu cầu khoảng cách gần hơn. Những hệ thống gần đây đã cho phép nhận dạng người sử dụng tròng mắt từ khoảng cách 15m.
- Xuất nhập cảnh: việc tự động hóa và tăng cường an ninh trong việc xác thực danh tính của người xuất nhập cảnh đang ngày càng được quan tâm khi số lượng người xuất nhập cảnh đang tăng lên nhanh chóng. Hiện nay, hộ chiếu điện tử đã trở thành một tiêu chuẩn quốc tế ICAO và được áp dụng rộng rãi tại hơn 70 nước bao gồm Mỹ, Liên minh Châu Âu (Anh, Pháp, Đức, Italia, Hà

Lan...), Úc, Hàn Quốc, Singapore... Hộ chiếu điện tử là một loại thẻ thông minh có bộ nhớ lưu trữ các thông tin về đặc trưng sinh trắc của cá nhân có thể bao gồm vân tay, khuôn mặt, tròng mắt.

- Chống gian lận: công nghệ nhận dạng sinh trắc học có thể được sử dụng trong các ứng dụng công cộng nhằm kiểm soát việc một cá nhân hưởng lợi từ việc đăng ký nhiều danh tính khác nhau. Hiện nay, liên hiệp quốc đã và đang sử dụng vân tay để kiểm soát việc trợ cấp lương thực tránh trường hợp một người có thể gian lận trong việc nhận trợ cấp nhiều lần khi khai báo nhiều danh tính khác nhau.
- Khách du lịch tin cậy: các ứng dụng này cho phép khách du lịch đăng ký các đặc trưng sinh trắc như vân tay, tròng mắt với chương trình giúp cho những lần du lịch tiếp theo đơn giản, nhanh chóng hơn nhiều khi chỉ phải kiểm tra nhân dạng tại các kios, điển hình như chi nhánh Disneyland ở Florida và Hồng Kông đã thực hiện đưa nhận dạng vân tay vào việc bán vé.
- Quản lý vào ra: công nghệ nhận dạng sinh trắc học có thể được sử dụng nhằm xác định hoặc xác thực nhân dạng của người được quyền vào ra ở những khu vực cụ thể.
- Quản lý nhân công: vân tay đã được sử dụng rất phổ biến trong việc quản lý thời gian đi, thời gian đến, giám sát sự có mặt của nhân công.
- Quản lý khách hàng: các ứng dụng loại này cần xác thực danh tính của khách hàng trước khi thực hiện các giao dịch. Thay cho các phương thức truyền thống như mật khẩu, số PIN, thẻ, chữ ký, khách hàng có thể sử dụng các đặc trưng sinh trắc để xác thực danh tính của mình một cách nhanh chóng, thuận tiện và an toàn. Rất nhiều các ngân hàng trên thế giới đã đưa nhận dạng vân tay, khuôn mặt, tròng mắt và tĩnh mạch lòng bàn tay vào việc thực hiện các giao dịch với khách hàng.
- Bảo vệ tài sản: các ứng dụng này cho phép người dùng bảo vệ các thông tin, tài sản trước những người sử dụng khác. Ví dụ bao gồm dùng vân tay để truy cập vào máy tính xách tay, dùng vân tay thay cho khóa tủ, hoặc dùng giọng nói để khởi động xe ô tô.

Các ứng dụng trên đã cho thấy công nghệ nhận dạng sinh trắc học thường được sử dụng yêu cầu tính bảo mật, hiệu năng và tiện lợi. Với đa dạng các ứng dụng, các hệ thống nhận dạng sinh trắc học đang phát triển nhanh chóng. Theo số liệu từ tổ chức International Biometric Group, doanh thu của các hệ thống nhận dạng sinh trắc năm 2009 là hơn 3 tỷ USD và sẽ tăng gấp 3 vào năm 2014 lên hơn 9 tỷ USD.



Hình 1. Quá trình xử lý nhận dạng bằng sinh trắc học.

Cấu tạo cơ bản của một hệ thống nhận dạng sinh trắc học bao gồm các thành phần sau:

- Thiết bị thu nhận đặc trưng (sensor): đây là thiết bị tương tác với người dùng nhằm thu nhận các đặc điểm sinh trắc của người đó. Một số loại thiết bị thu nhận điển hình gồm camera nhằm chụp ảnh khuôn mặt, tròng mắt, hình dáng tai; micro dùng thu âm giọng nói; thiết bị thu nhận vân tay; thiết bị thu nhận tĩnh mạch; thiết bị thu nhận ADN...
- Xử lý: đây là khối nhằm trích và chọn ra các đặc trưng riêng biệt của người và lưu lại thành các mẫu. Mỗi người có một mẫu riêng, chính sự duy nhất của mỗi đặc trưng sinh trắc của mỗi người được thể hiện ở sự duy nhất của mẫu tạo ra này. Nếu là lần đầu tiên người sử dụng đăng ký với hệ thống, mẫu tạo ra sẽ được cập nhật vào cơ sở dữ liệu mẫu. Nếu là những lần đăng nhập sau, mẫu này sẽ được so sánh với các mẫu có sẵn để xác định danh tính của người có mẫu đó.

- Cơ sở dữ liệu mẫu: lưu trữ dữ liệu về các mẫu sinh trắc của các cá nhân nhằm phục vụ cho việc đối sánh.
- So sánh và ra quyết định: từ mẫu của người vừa thu thập được, mẫu này sẽ được so sánh với các mẫu có sẵn trong cơ sở dữ liệu để xác định xem mẫu này trùng với mẫu lưu sẵn nào. Nếu việc so sánh cho thấy có một mẫu trùng hợp, hệ thống sẽ ra quyết định dựa trên việc xác thực được danh tính của mẫu mới thu nhận.

3. Đặc trưng sinh trắc khuôn mặt người và một số phương pháp xác định.

Dựa trên nhận xét thực tế, con người dễ dàng nhận biết các khuôn mặt và các đối tượng trong các tư thế khác nhau và điều kiện ánh sáng khác nhau, thì phải tồn tại các thuộc tính hay đặc trưng không thay đổi. Có nhiều nghiên cứu đầu tiên xác định các đặc trưng khuôn mặt rồi chỉ ra có khuôn mặt trong ảnh hay không. Các đặc trưng như: lông mày, mắt, mũi, miệng, và đường viền của tóc được trích bằng phương pháp xác định cạnh. Trên cơ sở các đặc trưng này, xây dựng một mô hình thống kê để mô tả quan hệ của các đặc trưng này và xác định sự tồn tại của khuôn mặt trong ảnh. Một vấn đề của các thuật toán theo hướng tiếp cận đặc trưng cần phải điều chỉnh cho phù hợp điều kiện ánh sáng, nhiễu, và bị che khuất. Đôi khi bóng của khuôn mặt sẽ tạo thêm cạnh mới, mà cạnh này lại rõ hơn cạnh thật sự của khuôn mặt, vì thế nếu dùng cạnh để xác định sẽ gặp khó khăn.

a. Các đặc trưng khuôn mặt

Sirohey đưa một phương pháp xác định khuôn mặt từ một ảnh có hình nền phức tạp. Phương pháp dựa trên cạnh (dùng phương pháp Candy và heuristics để loại bỏ các cạnh để còn lại duy nhất một đường bao xung quanh khuôn mặt. Một hình ellipse dùng để bao khuôn mặt, tách biệt vùng đầu và hình nền. Tỷ lệ chính xác của thuật toán là 80%. Cũng dùng phương pháp cạnh như Sirohey, Chetverikov và Lerch dùng một phương pháp dựa trên blob và streak (hình dạng giọt nước và sọc xen kẽ), để xác định theo hướng các cạnh. Hai ông dùng hai blob tối và ba blob sáng để mô tả hai mắt, hai bên gò má, và mũi. Mô hình này dùng các streak để mô tả hình dáng ngoài của khuôn mặt, lông mày, và môi. Dùng ảnh có độ phân giải thấp theo biến đổi Laplace để xác định khuôn mặt thông qua blob. Graf đưa ra một phương pháp xác

định đặc trưng rồi xác định khuôn mặt trong ảnh xám. Dùng bộ lọc để làm nổi các biên, các phép toán hình thái học (morphology) được dùng để làm nổi các vùng có cường độ cao và hình dáng chắc chắn (như mắt). Thông qua histogram để tìm các đỉnh nổi bật để xác định các ngưỡng chuyển ảnh xám thành hai ảnh nhị phân. **Các thành phần dính nhau đều xuất hiện trong hai ảnh nhị phân thì được xem là vùng của ứng viên khuôn mặt rồi phân loại xem có phải là khuôn mặt không.**

Phương pháp được kiểm tra trên các ảnh chỉ có đầu và vai của người. Tuy nhiên còn vấn đề, làm sao sử dụng các phép toán morphology và làm sao xác định khuôn mặt trên các vùng ứng viên. Leung trình bày một mô hình xác suất để xác định khuôn mặt ở trong ảnh có hình nền phức tạp trên cơ sở một bộ xác định đặc trưng cục bộ và so khớp đồ thị ngẫu nhiên. Ý chính là xem bài toán xác định khuôn mặt như là bài toán tìm kiếm với mục tiêu là tìm thứ tự các đặc trưng chắc chắn của khuôn mặt để tạo thành giống nhất một mẫu khuôn mặt. Dùng năm đặc trưng (hai mắt, hai lỗ mũi, phần nối giữa mũi và miệng) để mô tả một khuôn mặt. Luôn tính quan hệ khoảng cách với các đặc trưng cặp (như mắt trái, mắt phải), dùng phân bố Gauss để mô hình hóa. Một mẫu khuôn mặt được đưa ra thông qua trung bình tương ứng cho một tập đa hướng, đa tỷ lệ của bộ lọc đạo hàm Gauss. Từ một ảnh, các đặc trưng ứng viên được xác định bằng cách so khớp từng điểm ảnh khi lọc tương ứng với vector mẫu (tương tự mỗi tương quan), chọn hai ứng viên đặc trưng đứng đầu để tìm kiếm cho các đặc trưng khác của khuôn mặt. Giống như xây dựng một đồ thị quan hệ mỗi node của đồ thị tương ứng như các đặc trưng của một khuôn mặt, đưa xác suất vào để xác định. Tỷ lệ xác định chính xác là 86%.

Bên cạnh tính khoảng cách liên quan để mô tả quan hệ giữa các đặc trưng như Leung, Kendall và dùng lý thuyết xác suất thống kê về hình dáng. Dùng hàm mật độ xác suất (Probability Density Function - PDF) qua N điểm đặc trưng, tương ứng (x_i, y_i) là đặc trưng thứ i với giả sử dựa vào phân bố Gauss có $2N$ -chiều. Các tác giả áp dụng phương thức cực đại khả năng (Maximum-Likelihood - ML) để xác định vị trí khuôn mặt. Một thuận lợi của phương pháp này là các khuôn mặt bị che khuất vẫn có thể xác định được. Nhưng phương pháp không xác định được đa khuôn mặt trong ảnh. Yow và Cipolla trình bày một phương thức dựa vào đặc trưng, dùng số lượng lớn các dấu

hiệu từ ảnh và cả dấu hiệu về ngữ cảnh. Đầu tiên dùng bộ lọc đạo hàm Gauss thứ hai, xác định các điểm mấu chốt ở tại cực đại địa phương trong bộ lọc, rồi chỉ ra nơi có thể là đặc trưng. Giai đoạn hai, kiểm tra các cạnh xung quanh điểm mấu chốt và nhóm chúng lại thành các vùng. Tiêu chuẩn để nhóm các cạnh là gần và tương tự hướng và cường độ. Đo lường các đặc tính vùng như: chiều dài cạnh, cường độ cạnh, và biến thiên cường độ được lưu trong một vector đặc trưng. Từ dữ liệu đặc trưng khuôn mặt đã được huấn luyện, sẽ tính được giá trị trung bình và ma trận hiệp phương sai của mỗi đặc trưng khuôn mặt. Một vùng là ứng viên khuôn mặt khi khoảng cách Mahalanobis giữa các vector đặc trưng đều dưới một ngưỡng. Rồi thông qua mạng Bayes để xác định ứng viên có phải là khuôn mặt không. Tỷ lệ chính xác là 85% , tuy nhiên mức độ sai là 28%, và chỉ hiệu quả với hình khuôn mặt có kích thước 60x60 điểm ảnh.

Phương pháp này được dùng thêm với mô hình đường viền linh hoạt. Takacs và Wechsler trình bày một phương pháp dựa trên tích đặc trưng võng mạc và cử động theo dao động nhỏ của mắt. Thuật toán hoạt động trên bản đồ hay vùng của các mấu chốt, mô hình hóa lưới võng mạc. Đầu tiên tính toán ước lượng thô vùng khuôn mặt trên cơ sở bộ lọc. Giai đoạn thứ hai tinh chế trên độ phân giải mịn hơn. Tỷ lệ sai là 4.69%. Han phát triển một kỹ thuật trên cơ sở morphology để trích các đoạn giống mắt (eyeanalogue) để xác định khuôn mặt người. Ông nói rằng mắt và lông mày là đặc trưng nổi bật nhất và ổn định nhất của khuôn mặt con người, và nó rất hữu dụng để xác định khuôn mặt người. Ông định nghĩa các đoạn giống mắt như là các cạnh trên đường viền của mắt. Đầu tiên, các phép toán morphology như đóng, cắt bỏ sai khác, và phân ngưỡng để trích các điểm ảnh có giá trị cường độ thay đổi đáng kể. Các điểm ảnh này sẽ trở thành các điểm ảnh giống mắt. Sau đó một tiến trình gán nhãn để sinh các đoạn giống mắt. Các đoạn này được dùng để chỉ dẫn tìm kiếm các vùng tiềm năng có thể là khuôn mặt qua kết hợp các đặc tính hình học của mắt, mũi, lông mày, và miệng. Các vùng này sẽ được một mạng neural xem xét có phải là khuôn mặt không, giống. Theo tác giả tỷ lệ chính xác là 94%.

Amit đưa ra phương thức xác định khuôn mặt dựa trên hình dáng và áp dụng cho các khuôn mặt chụp thẳng. Có hai giai đoạn để xác định khuôn mặt người: tập

trung và phân loại chi tiết. Làm có thứ tự các mảnh cạnh, các mảnh này được trích từ bộ xác định cạnh đơn giản thông qua sự khác biệt cường độ là quá trình tập trung. Khi có các ứng viên từ quá trình trên, dùng thuật toán CART để xây dựng một cây phân loại từ các ảnh để huấn luyện, để xem xét ứng viên nào là khuôn mặt người. Jin dùng cấu trúc hình học của khuôn mặt người để tìm ứng viên khuôn mặt trong ảnh xám và hình nền không phức tạp. Mỗi ảnh chỉ có một khuôn mặt người, nhưng tư thế điều kiện ánh sáng, không cố định. Tỷ lệ chính xác khoảng 94.25% và thời gian khá nhanh. Chan và Lewis dùng kỹ thuật lọc để loại bớt tác động của ánh sáng, sau đó phân đoạn để tìm vị trí các ứng viên là con mắt. Từ các ứng viên này xây dựng mạng neural như Rowley để xác định khuôn mặt người. Phương pháp này có thể xác định nhiều khuôn mặt trong một ảnh, các khuôn mặt này có thể có các tư thế, vị trí, tỷ lệ khác nhau. Tỷ lệ chính xác là 53%. Kruppa dùng sắc màu của da người để tìm ứng viên, nhưng ông không xử lý cho từng điểm ảnh theo cách thông thường, mà ông dùng mô hình màu da người trên từng phần nhỏ rồi xử lý phân đoạn trên đó. Sau khi có ứng viên khuôn mặt, ông dùng một số đặc tính về hình dáng để xác định khuôn mặt người. Tỷ lệ chính xác là 85%.

Park dùng Gaze để tìm ứng viên góc mắt, miệng và tâm mắt. Ông xây dựng SVM đã được học trước đó để xác định các vị trí ứng viên có phải là góc mắt, miệng, và tâm mắt hay không để theo vết con mắt người. Sato dùng quan hệ đường viền ở cằm của khuôn mặt. Tác giả chia làm hai trường hợp: thon dài và tròn để xem xét. Tác giả dùng GA để xem xét mối tương quan của đường cong, hình dáng khuôn mặt để xác định khuôn mặt. Chai và Ngan xây dựng phương pháp xác định khuôn mặt người dựa trên đặc trưng về: quan hệ hình học, mật độ, độ chói trong ảnh màu chỉ có đầu và vai của ứng viên để xác định. Kim cũng phân đoạn để tìm ứng viên khuôn mặt, nhưng xác thực khuôn mặt thông qua các cấu trúc các đặc trưng mắt, mũi, miệng, và đường viền của ứng viên. Jang dùng phân bố màu da để phân đoạn tìm ứng viên rồi dùng các đặc trưng hình học để xác định khuôn mặt. Christian và Jonh xây dựng một loại đặc trưng mới, đó là đặc trưng về độ cong của các đường trên khuôn mặt để giải quyết vấn đề điều kiện ánh sáng. Từ đặc trưng cong này, hai ông quay lại phương pháp PCA để xác định khuôn mặt. Juan và Narciso xây dựng một không gian màu mới $YCg'Cr'$ để lọc các vùng là ứng viên khuôn mặt dựa trên sắc thái của màu da người. Sau khi có

ứng viên, hai ông dùng các quan hệ về hình dáng khuôn mặt, mức độ cân đối của các thành phần khuôn mặt để xác định khuôn mặt người. Tương tự, Chang và Hwang cũng dùng một phương thức như, tỷ lệ chính xác hơn 80% trong ảnh xám.

Dae và Nam xem xét các đặc trưng không thay đổi khi thay đổi tư thế của khuôn mặt bằng cách xem xét các quan hệ hình học. Sau đó ước lượng các tư thế của khuôn mặt rồi xây dựng dữ liệu để xác định thông qua PCA. Tỷ lệ chính xác là 76%. Jin xây dựng một bộ lọc để xác định ứng viên khuôn mặt người theo màu da người. Từ ứng viên này tác giả xác định khuôn mặt người theo hình dáng khuôn mặt và các quan hệ đặc trưng về thành phần khuôn mặt, với mắt phải được chọn làm gốc tọa độ để xét quan hệ. Tỷ lệ chính xác cho khuôn mặt chụp thẳng trên 80%.

b. Kết cấu

Khuôn mặt con người có những kết cấu riêng biệt mà có thể dùng để phân loại so với các đối tượng khác. Augusteijn và Skufca cho rằng hình dạng của khuôn mặt dùng làm kết cấu phân loại, gọi là kết cấu giống khuôn mặt (face-like texture). Tính kết cấu qua các đặc trưng thống kê thứ tự thứ hai (SGLD) trên vùng có kích thước 16x16 điểm ảnh. Có ba loại đặc trưng được xem xét: màu da, tóc, và những thứ khác. Hai ông dùng mạng neural về mối tương quan cascade cho phân loại có giám sát các kết cấu và một ánh xạ đặc trưng tự tổ chức Kohonen để gom nhóm các lớp kết cấu khác nhau. Hai tác giả đề xuất dùng phương pháp bầu cử khi không quyết định được kết cấu đưa vào là kết cấu của da hay kết cấu của tóc. Dai và Nakano dùng mô hình SGLD để xác định khuôn mặt người. Thông tin màu sắc được kết hợp với mô hình kết cấu khuôn mặt. Hai tác giả xây dựng thuật giải xác định khuôn mặt trong không gian màu, với các phần tựa màu cam để xác định các vùng có thể là khuôn mặt người. Một thuận lợi của phương pháp này là có thể xác định khuôn mặt không chỉ chụp thẳng và có thể có râu và có kính. Mark và Andrew dùng phân bố màu da và thuật toán DoG (a Difference of Gauss) để tìm các ứng viên, rồi xác thực bằng một hệ thống học kết cấu của khuôn mặt. Manian và Ross dùng biến đổi wavelet để xây dựng tập dữ liệu kết cấu của khuôn mặt trong ảnh xám thông qua nhiều độ phân giải khác nhau kết hợp xác

suất thông kê để xác định khuôn mặt người. Mỗi mẫu sẽ có chín đặc trưng. Tỷ lệ chính xác là 87%, tỷ lệ xác định sai là 18%.

c. Đa đặc trưng

Gần đây có nhiều nghiên cứu sử dụng các đặc trưng toàn cục như: màu da người, kích thước, và hình dáng để tìm các ứng viên khuôn mặt, rồi sau đó sẽ xác định ứng viên nào là khuôn mặt thông qua dùng các đặc trưng cục bộ (chi tiết) như: mắt, lông mày, mũi, miệng, và tóc. Tùy mỗi tác giả sẽ sử dụng tập đặc trưng khác nhau. Yachida đưa ra một phương pháp xác định khuôn mặt người trong ảnh màu bằng lý thuyết logic mờ. Ông dùng hai mô hình mờ để mô tả phân bố màu da người và màu tóc trong không gian màu CIE XYZ. Năm mô hình hình dạng của đầu (một thẳng và bốn xoay xung quanh) để mô tả hình dáng của mặt trong ảnh. Mỗi mô hình hình dạng là một mẫu 2-chiều bao gồm các ô vuông có kích thước $m \times n$, mỗi ô có thể chứa nhiều hơn một điểm ảnh. Hai thuộc tính được gán cho mỗi ô là: tỷ lệ màu da và tỷ lệ tóc, chỉ ra tỷ lệ diện tích vùng da (tóc) trong ô so với diện tích của ô. Mỗi điểm ảnh sẽ được phân loại thành tóc, khuôn mặt, tóc/khuôn mặt, và tóc/nền trên cơ sở phân bố của mô hình, theo cách đó sẽ có được các vùng giống khuôn mặt và giống tóc. Mô hình hình dáng của đầu sẽ được so sánh với vùng giống khuôn mặt và giống tóc. Nếu tương tự, vùng đang xét sẽ trở thành ứng viên khuôn mặt, sau đó dùng các đặc trưng mắt-lông mày và mũi-miệng để xác định ứng viên nào sẽ là khuôn mặt thật sự. Sobottka và Pitas dùng các đặc trưng về hình dáng và màu sắc để xác định khuôn mặt người. Dùng một ngưỡng để phân đoạn trong không gian màu HSV để xác định các vùng có thể là màu da người (vùng giống màu da người), các tiền ứng viên. Các thành phần dính nhau sẽ được xác định bằng thuật toán tăng vùng ở độ phân giải thô. Xem xét tiền ứng viên nào vừa khớp hình dạng ellipse sẽ được chọn làm ứng viên của khuôn mặt. Sau đó dùng các đặc trưng bên trong như: mắt và miệng, được trích ra trên cơ sở các vùng mắt và miệng sẽ tối hơn các vùng khác của khuôn mặt, sau cùng phân loại dựa trên mạng neural để biết vùng ứng viên nào là khuôn mặt người và vùng nào không phải khuôn mặt người. Tỷ lệ chính xác là 85%.

Dựa vào mức độ cân xứng của các mẫu khuôn mặt người để xác định khuôn mặt người. Một bộ phân loại màu da/không phải màu da dùng trong không gian màu

YES cho phép làm mịn các vùng kề có đường cong không mịn, sau khi lọc các vùng có thể là màu da người. Một mẫu khuôn mặt dạng ellipse được dùng để xem xét mức độ tương tự của các vùng có cùng màu da người với mẫu này thông qua khoảng cách Hausdorff. Sau cùng, xác định tâm mắt thông qua các hàm tính giá trị dựa trên quan hệ cân đối của khuôn mặt và vị trí hai mắt. Đỉnh của mũi và tâm của miệng được ước lượng qua khoảng cách tâm mắt. Mặt hạn chế của phương pháp này là chỉ xác định trên ảnh chụp thẳng khuôn mặt, chỉ có duy nhất một khuôn mặt trong ảnh, và xác định được vị trí của cả hai mắt. Cũng có tác giả dùng phương pháp tương tự để giải quyết. Trái ngược với phương pháp xử lý trên điểm ảnh, một phương pháp được xây dựng trên cấu trúc, màu sắc, và liên quan hình học đã được đề nghị. Đầu tiên dùng phân đoạn đa tỷ lệ để trích các vùng đồng đều trong ảnh dựa vào mô hình màu da người theo Gauss để có được các vùng có màu cùng với màu da người, gom các vùng này vào trong các vùng có hình dạng ellipse. Một vùng có hình dạng ellipse được xác định là một khuôn mặt người nếu tồn tại mắt miệng trong vùng đó.

Tác giả cho biết có thể xác định các khuôn mặt ở các hướng khác nhau khi có thêm các đặc trưng phụ như: râu, mắt kính. Kauth trình bày một biểu diễn dạng blob để trích đặc trưng, mà đặc trưng này dùng tô tả có ý nghĩa cấu trúc của đa phổ của ảnh chụp từ vệ tinh. Mỗi vector đặc trưng tại một điểm ảnh bao gồm các tọa độ của điểm ảnh và liên quan theo các thành phần phổ (hay các thành phần kết cấu). Các điểm ảnh này được gom nhóm bằng cách dùng vector đặc trưng để có các vùng dính liền nhau, hoặc có dạng blob. Mỗi vector đặc trưng bao gồm tọa độ ảnh và sắc màu được chuẩn hóa. Dùng một thuật toán tạo các vùng liên kết lại với nhau để tăng kích thước của blob và xem xét nếu ứng viên dạng blob nào thỏa mãn hình dáng kích thước khuôn mặt thì xem đó là khuôn mặt.

Phạm vi và màu sắc được Kim dùng để xác định khuôn mặt người. Tính biểu đồ chênh lệch rồi phân đoạn dựa trên biểu đồ histogram với giả thuyết các điểm ảnh là nền sẽ có cùng độ sâu và số lượng sẽ nhiều hơn các điểm ảnh trong đối tượng. Dùng phân bố Gauss trong không gian màu RGB đã được chuẩn hóa, được các ứng viên rồi dùng phân loại để xác định cuối cùng ứng viên nào là khuôn mặt người. Cùng các tiếp cận này có Darrell. Hsu được xem là người khá thành công khi xác định khuôn mặt

người trong ảnh màu . Ông xây dựng một bộ phân loại để xác định các vị trí của ứng viên mắt và miệng dựa trên sắc màu đặc trưng của mắt và miệng. Trên quan hệ về khoảng cách của hai mắt và miệng để xác định ứng viên nào sẽ là khuôn mặt thông qua biến đổi Hough để có ứng viên nào gần giống dạng ellipse nhất. Jesorsky xác định cạnh của các đối tượng trong ảnh rồi so sánh hình dáng kết hợp dùng khoảng cách Hausdorff để đo mức độ tương tự của khuôn mặt người với các mẫu. Sau đó Kirchberg cải tiến dùng mô hình Gen (Genetic Model) để phát sinh mô hình khuôn mặt người từ dữ liệu lộn xộn sau khi phân đoạn trong ảnh xám kết hợp khoảng cách Hausdorff. Mức độ chính xác khoảng 85%. Yen và Nithianandan dùng GA để trích các đặc trưng khuôn mặt, như mắt (lông mày), mũi, và miệng. Áp dụng hình thái khuôn mặt giống hình ellipse để xác định khuôn mặt bằng GA trong ảnh màu. Phương pháp này cho phép giải quyết trong điều kiện ánh sáng khác nhau, tư thế khuôn mặt khác nhau. Chang xem xét tính đa dạng về mặt của khuôn mặt người. Từ đây ông xây dựng mạng wavelet tích cực (Active Wavelet Network) để trích các đặc trưng của khuôn mặt rồi dùng hai phương pháp làm giảm số chiều của không gian đặc trưng là LLE (Locally Linear Embedding) và LE (Lipschitz Embedding) và học cấu trúc đa dạng này để xác định khuôn mặt. Daidi và Irek trích các đặc trưng của khuôn mặt bằng sơ đồ phân bố tham số để xác định khuôn mặt người. Tỷ lệ chính xác cho ảnh xám và khuôn mặt được chụp thẳng là 91.4%. Ehsan và Jonh dùng tập hệ số Gabor wavelet ở các hướng khác nhau để trích các đặc trưng của khuôn mặt. Sau đó dùng entropy cục bộ để xác định khuôn mặt trong ảnh xám và khuôn mặt được chụp thẳng hay tựa thẳng nhưng có các vị trí khác nhau. Tỷ lệ chính xác là 94%. Bao dùng sắc thái màu da người để xác định ứng viên trong ảnh màu. Tác giả đã xây dựng các luật mờ dựa vào hai loại đặc trưng: (1) bên ngoài và (2) bên trong. Đặc trưng bên ngoài gồm: tỷ lệ chiều cao, diện tích, chu vi, mức độ tròn, ... Đặc trưng bên trong gồm: quan hệ mức độ cân đối của hai mắt và miệng cũng như tỷ lệ khoảng cách với khuôn mặt. Phương pháp này cho phép xác định khuôn mặt ở nhiều tư thế, vị trí, mức độ nghiêng khác nhau trong môi trường phức tạp. Đặc biệt, tác giả đã xây dựng bộ điều khiển mờ để tách các khuôn mặt dính lẫn nhau. Tỷ lệ chính xác khoảng 87% - 89%.

II. Ứng dụng của phép biến đổi KL và phân tích thành các thành phần chính (PCA) trong trích chọn đặc trưng ảnh khuôn mặt.

1. Sơ lược đại số tuyến tính trong thống kê.

Vecto riêng, trị riêng và sự chéo hóa ma trận.

Xét một toán tử tuyến tính \mathbf{f} trong không gian \mathbf{R}^n với các vectơ cơ sở:

$$\mathbf{e}_i = [0 \dots 1 \dots 0]^T \text{ (với giá trị 1 nằm tại vị trí thứ } i \text{)} \quad (1)$$

Toán tử tuyến tính này sẽ được biểu diễn bởi một ma trận vuông \mathbf{T} kích thước $\mathbf{n} \times \mathbf{n}$.

Nếu tồn tại một đại lượng vô hướng λ và một vectơ $\mathbf{x} \neq \mathbf{0}$ sao cho thỏa mãn điều kiện:

$$\mathbf{f}(\mathbf{x}) = \lambda \mathbf{x} \quad (2)$$

$$\text{hay } \mathbf{T} \cdot \mathbf{x} = \lambda \mathbf{x} \quad (3)$$

Khi đó λ được gọi là trị riêng của \mathbf{f} , và vectơ \mathbf{x} được gọi là vectơ riêng của \mathbf{f} hay của \mathbf{T} , ứng với trị riêng λ . Ma trận \mathbf{T} với kích thước $\mathbf{n} \times \mathbf{n}$ sẽ có tối đa \mathbf{n} trị riêng và \mathbf{n} vectơ riêng tương ứng. Một ma trận \mathbf{T} khả nghịch sẽ có đủ \mathbf{n} trị riêng (kể cả trị riêng bội) và \mathbf{n} vectơ riêng tương ứng.

Nếu tồn tại một cơ sở trong không gian \mathbf{R}^n sao cho ma trận \mathbf{T} biểu diễn trong cơ sở đó có dạng chéo (các phần tử ngoài đường chéo bằng 0) thì ma trận \mathbf{T} (biểu diễn trong không gian \mathbf{R}^n với các vectơ cơ sở \mathbf{e}_i nêu trên) sẽ chéo hóa được.

Ví dụ: Khảo sát trên không gian \mathbf{R}^5 với ma trận chéo 5×5 .

$$\mathbf{T} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 \end{bmatrix}$$

Giả sử \mathbf{C} là ma trận các vectơ cơ sở mới $\{\mathbf{u}_i\}$ được biểu diễn trong cơ sở $\{\mathbf{e}_i\}$.

Ở đây, ma trận \mathbf{T} được chuyển từ cơ sở $\{\mathbf{e}_i\}$ sang cơ sở mới $\{\mathbf{u}_i\}$ nên ma trận chuyển đổi cơ sở từ $\{\mathbf{e}_i\}$ sang $\{\mathbf{u}_i\}$ cũng là \mathbf{C} . Nếu \mathbf{T} chéo hóa được tức là tồn tại ma trận \mathbf{C} khả nghịch (tức là \mathbf{C} tạo được một cơ sở trong \mathbf{R}_n) sao cho :

$$\mathbf{T}_c = \mathbf{C}^{-1} \mathbf{T} \mathbf{C} \quad (4)$$

có dạng chéo.

Nếu ta có \mathbf{C} là một ma trận có các cột là các vectơ cơ sở đã được chuẩn hóa của không gian \mathbf{R}_n thì $\mathbf{C}^T = \mathbf{C}^{-1}$, khi đó từ (4.1.4) ta có thể suy ra :

$$\mathbf{T}_c = \mathbf{C}^T \mathbf{T} \mathbf{C}. \quad (5)$$

Ta có thể tìm được ma trận \mathbf{C} để chéo hóa một ma trận \mathbf{T} bằng cách tìm các vector riêng của ma trận \mathbf{T} . Lúc đó ma trận \mathbf{C} chính là ma trận có các cột là các vector riêng của \mathbf{T} .

a. Kỳ vọng

Kỳ vọng

Đối với thống kê đa chiều, mỗi một mẫu thống kê là một vector đa chiều. Giả sử ta có một biến ngẫu nhiên \mathbf{X} trong không gian tuyến tính n chiều.

$$\mathbf{X} = [x_1 \ x_2 \ \dots \ x_n]^T \quad (6)$$

Khi đó kỳ vọng của biến ngẫu nhiên \mathbf{X} cũng là một vector n chiều, trong thống kê, kỳ vọng $E[\mathbf{X}]$ của một biến ngẫu nhiên \mathbf{X} có thể được ước lượng bằng trung bình mẫu \bar{X} , và được tính bằng công thức:

$$\bar{X} = \frac{1}{M} \sum_{i=1}^M X_i \quad (7)$$

Trong đó M là tổng số mẫu có trong thống kê.

b. Ma trận hiệp phương sai

\mathbf{X} là một biến véc tơ ngẫu nhiên gồm n thành phần x_i , $i = 1, 2, \dots, n$. Mỗi thành phần x_i là giá trị ngẫu nhiên. Ng-ời ta định nghĩa:

- Kỳ vọng toán học (*Trung bình số học*) $E[x] = \int_{-\infty}^{\infty} xP(x)dx$
với $P(x)$ là hàm mật độ xác suất và x là biến ngẫu nhiên liên tục.
- Mô men toán học

$$m_k = \int_{-\infty}^{\infty} x^k P(x)dx = E[x^k]$$

m_k gọi là mô men bậc k của x .

- Tính t-ơng quan: một tín hiệu phụ thuộc vào thời gian
- Hàm tự t-ơng quan của 1 tín hiệu $x[t]$ đ-ợc định nghĩa:

$$\Psi_{xx} = E[x(t).x(t+\tau)]$$

- Hàm t-ơng quan của 2 tín hiệu:

$$\Psi_{xy} = E[x(t).y(t+\tau)]$$

- Cho tập các đối tượng X , *ma trận t-ong quan* của tập các đối tượng ký hiệu là R và đ-ợc định nghĩa $R = E[XX^T] = \langle XX^T \rangle$. Viết d-ới dạng ma trận ta có:

$$R = \begin{bmatrix} E[x_{11}] & E[x_{12}] & \dots & E[x_{1n}] \\ E[x_{21}] & E[x_{22}] & \dots & E[x_{2n}] \\ \dots & \dots & \dots & \dots \\ E[x_{n1}] & E[x_{n2}] & \dots & E[x_{nn}] \end{bmatrix}$$

Ma trận hiệp biến, ký hiệu $A = E[(X-M)(X-M)^T]$

$$= \langle (X-M)(X-M)^T \rangle \quad (3.44)$$

Trong một số tr-ờng hợp $A = \langle XX^T \rangle - \langle MM^T \rangle = R - \langle MM^T \rangle$ (*). Nếu đối tượng không t-ong quan (độc lập) lúc đó ma trận A là ma trận đ-ờng chéo. Có nghĩa là:

$$a_{i,i} = x_i^2 - m_i^2 \neq 0 \text{ còn } a_{i,j} = 0 \text{ với } i \neq j$$

2. Phép biến đổi KL

Biến đổi KL có nguồn gốc từ khai triển chuỗi của các quá trình ngẫu nhiên liên tục. Biến đổi KL cũng còn đ-ợc gọi là biến đổi Hotelling hay ph-ơng pháp thành phần chính. Để tiện theo dõi ta cũng cần nhắc lại một số khái niệm và định nghĩa trong xử lý thống kê.

a. Cơ sở lý thuyết của KL.

Đây là phép biến đổi không gian n chiều thành không gian m chiều, với $m < n$. Mỗi thành phần của véc tơ miêu tả một đặc tính của đối tượng. Nếu ta biến đổi được từ không gian n chiều về không gian m chiều, như vậy ta sẽ làm giảm được thông tin dư thừa (theo thuật ngữ trong xử lý ảnh hay nhận dạng gọi là giảm thứ nguyên). Mục đích của biến đổi KL là chuyển từ không gian n chiều sang không gian trực giao m chiều sao cho sai số bình phương là nhỏ nhất.

Vecto $X = \{x_1, \dots, x_n\} \in R^n$ là vecto dữ liệu, và X sẽ được chuyển thành \overline{X} trong không gian m chiều.

$$\overline{X} = \{x_1, \dots, x_m\} \in R^m$$

Giả sử $U = \{u_1, u_2, \dots, u_n\}$ là tập vecto cơ sở trong không gian trực giao R^n với:

$$\underline{u}_i = \begin{cases} u_{1j} & j=1,2,3\dots n \\ u_{1j} \\ \dots \\ u_{1j} \end{cases}$$

$$\underline{u}_i \cdot \underline{u}_k = \begin{cases} 0 \text{ nếu } i \neq k \\ 1 \text{ nếu } i = k \end{cases}$$

Ta có:

$$X = \varphi_1 u_1 + \varphi_2 u_2 + \dots + \varphi_n u_n = \Phi U$$

$$\rightarrow \Phi = U^T X$$

Sai số trong phép biến đổi $X \rightarrow \bar{X}$:

$$\varepsilon = X - \bar{X} = \sum_{i=1}^n \varphi_i \cdot u_i - \sum_{i=1}^m \varphi_i \cdot u_i = \sum_{i=m+1}^n \varphi_i \cdot u_i$$

Sai số trung bình bình phương:

$$\zeta = E |\varepsilon|^2 = E |(\bar{X} - X)^T (\bar{X} - X)|$$

$$= \sum_{i=m+1}^n \langle \varphi_i^2 \rangle$$

$$\text{Mà } \Phi = U^T X \rightarrow \zeta = \sum_{i=m+1}^n (\underline{u}_i^T X) (\underline{u}_i X)^T$$

$$= \sum_{i=m+1}^n \underline{u}_i^T \langle XX^T \rangle \underline{u}_i$$

$$= \sum_{i=m+1}^n \underline{u}_i^T R \underline{u}_i \text{ với } R = \langle XX^T \rangle$$

Ta có $u_i^T u_i = 1$

$$\begin{aligned}\zeta &= \zeta + \sum_{i=m+1}^n \lambda_i (1 - u_i^T u_i) \\ &= \sum_{i=m+1}^n [u_i^T R u_i + \lambda_i (1 - u_i^T u_i)] \quad (*)\end{aligned}$$

→ ζ đạt min khi (*) min

Ta sử dụng phương pháp đạo hàm để tìm ζ .

Đặt: $L(u_i, \lambda_i) = u_i^T R u_i + \lambda_i (1 - u_i^T u_i)$

Dùng phương pháp đạo hàm dẫn tới các phương trình:

$$\frac{\partial L(u_i, \lambda_i)}{\partial u_i} = 0$$

$$\rightarrow R u_i - \lambda_i u_i = 0$$

$$\rightarrow (R - \lambda_i I) u_i = 0 \quad (**)$$

(**) là phương trình đặc trưng của R với λ_i là trị riêng và u_i là vectơ riêng tương ứng.

Bằng cách giải phương trình trên ta sẽ tìm được các trị riêng λ_i ứng với n vectơ u_i của tập U .

Gọi $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_n]$

$$\bar{X} = \lambda^T X \text{ với } X = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n$$

b. Biến đổi KL

Định nghĩa: Cho u là một véc tơ các số thực ngẫu nhiên; vectơ cơ sở của biến đổi KL là các véc tơ riêng trực giao của ma trận hiệp biến R (định nghĩa trong phần 3.3.2.1) cho bởi phương trình: $R \phi_k = \lambda_k \phi_k$; $0 \leq k \leq N-1$

Biến đổi KL của u là $v = \Phi^* u$ và biến đổi ngược $u = \Phi v = \sum_{k=0}^{N-1} v(k) \Phi_k$

u là véc tơ cột, v là vectơ hàng và Φ_k là cột thứ k của ma trận Φ .

Biến đổi Φ đưa R về dạng đường chéo:

$$\Phi^* R \Phi = \lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_N \end{bmatrix}$$

3. Phương pháp phân tích thành phần chính PCA trong trích chọn đặc trưng khuôn mặt.

a. Yêu cầu

Vấn đề được đặt ra khi thực hiện việc nhận dạng trong không gian đa chiều (với số lượng chiều rất lớn).

Sẽ đạt được một sự cải thiện đáng kể bằng cách sắp xếp dữ liệu vào một không gian đa chiều, nhưng với số chiều nhỏ hơn.

$$x = \begin{bmatrix} a1 \\ a2 \\ \dots \\ an \end{bmatrix} \text{ giảm chiều chuyển thành } y = \begin{bmatrix} b1 \\ b2 \\ \dots \\ bk \end{bmatrix} \quad (\text{với } k < n)$$

b. Trích đặc trưng bằng phương pháp PCA

Mục tiêu của phương pháp PCA là giảm số chiều của một tập các vector sao cho vẫn đảm bảo được tối đa thông tin quan trọng nhất của tập dữ liệu huấn luyện. Có thể nói phương pháp PCA tìm cách giữ lại những thành phần thống kê quan trọng nhất của tập mẫu.

Giả sử ta cần giảm số chiều của tập mẫu huấn luyện từ n chiều:

$$x = a_1 v_1 + a_2 v_2 + \dots + a_n v_n$$

v_1, v_2, \dots, v_n là cơ sở trong không gian N chiều xuống còn k chiều ($k < n$)

$$y = b_1 u_1 + b_2 u_2 + \dots + b_k u_k$$

u_1, u_2, \dots, u_k là cơ sở trong không gian K chiều

Nghĩa là ta cần tìm một ánh xạ từ không gian n chiều xuống không gian nhỏ hơn chỉ có k chiều ($k < n$). Gọi x là một vector trong không gian n chiều, y là một vector trong không gian k chiều. Ta có trung bình bình phương lỗi **MSE** (*mean square error*) $\|x - y\|$ khi loại bỏ một số thành phần trong x để thu được y sẽ bằng tổng phương sai của

những thành phần bị loại bỏ. Phương pháp PCA sẽ tìm một phép biến đổi tuyến tính \mathbf{T} , thỏa :

$$\mathbf{y} = \mathbf{T}^* \mathbf{x} \text{ là ma trận } \mathbf{k} \times \mathbf{n} \text{ (10)}$$

sao cho trung bình bình phương lỗi là bé nhất.

$$\mathbf{b}_1 = t_{11}\mathbf{a}_1 + t_{12}\mathbf{a}_2 + \dots + t_{1n}\mathbf{a}_n$$

$$\mathbf{b}_2 = t_{21}\mathbf{a}_1 + t_{22}\mathbf{a}_2 + \dots + t_{2n}\mathbf{a}_n$$

.....

$$\mathbf{b}_k = t_{k1}\mathbf{a}_1 + t_{k2}\mathbf{a}_2 + \dots + t_{kn}\mathbf{a}_n$$

Hay $\mathbf{y} = \mathbf{T}^* \mathbf{x}$ với:

$$\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1N} \\ t_{21} & t_{22} & \dots & t_{2N} \\ \dots & \dots & \dots & \dots \\ t_{K1} & t_{K2} & \dots & t_{KN} \end{bmatrix}$$

Gọi \mathbf{M} là vector trung bình của các vector \mathbf{x} trong tập huấn luyện \mathbf{X} .

$$\mathbf{M} = \frac{1}{M} \sum_{i=1}^M \mathbf{X}_i \quad \text{Với } M \text{ là số phần tử trong tập huấn luyện (11)}$$

Gọi \mathbf{C} là ma trận hiệp phương sai của các phần tử trong tập \mathbf{X} .

$$\mathbf{C} = \frac{1}{M-1} \sum_{k=1}^M (\mathbf{x}_k - \mathbf{M})(\mathbf{x}_k - \mathbf{M})^T \quad \mathbf{C} \text{ là ma trận đối xứng } \mathbf{n} \times \mathbf{n} \text{ (12)}$$

Người ta chứng minh được rằng nếu \mathbf{T} là một ma trận mà mỗi hàng là một vector riêng của \mathbf{C} và \mathbf{m} vector riêng này (\mathbf{m} hàng của ma trận \mathbf{T}) ứng với \mathbf{m} trị riêng lớn nhất thì \mathbf{T} chính là phép biến đổi tuyến tính thỏa mãn điều kiện để \mathbf{MSE} nhỏ nhất.

Gọi Φ là ma trận vuông $\mathbf{n} \times \mathbf{n}$ mà mỗi cột là một vector riêng của \mathbf{C} đã được chuẩn hóa với phép biến đổi :

$$\mathbf{y} = \Phi^T \mathbf{x} \quad \mathbf{y} = (y^1, y^2, \dots, y^n) \quad (13)$$

được gọi là phép biến đổi **Hotelling**.

Xét theo quan điểm của nhận dạng thì mỗi thành phần y^i của vector \mathbf{y} được xem như là một đặc trưng của vector mẫu \mathbf{x} . Các đặc trưng này là các đặc trưng độc lập với nhau vì ma trận hiệp phương sai của \mathbf{y} là

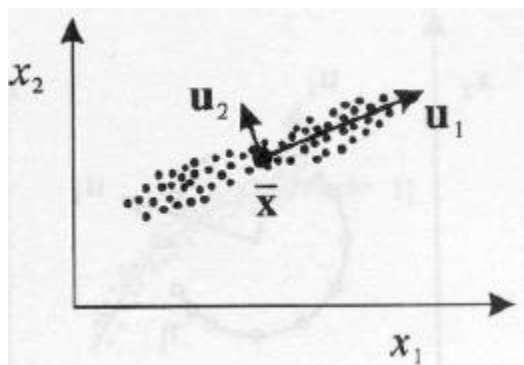
$$\mathbf{C}_y = \Phi^T \mathbf{C} \Phi \quad (14)$$

là một ma trận chéo (đã đề cập tới trong phần : Vector riêng, trị riêng và sự chéo hóa ma trận).

Tóm lại, phương pháp phân tích thành phần chính ánh xạ một vector từ không gian n chiều xuống không gian k chiều sẽ đi tìm các trị riêng và vector riêng của ma trận hiệp phương sai \mathbf{C} của tập \mathbf{X} và giữ lại k vector riêng ứng với k trị riêng lớn nhất làm cơ sở cho không gian k chiều này.

- **Sự diễn giải bằng hình học :**

- PCA chiếu dữ liệu theo hướng mà ở đó dữ liệu khác nhau nhiều nhất.
- Những hướng này được xác định bằng các vector riêng của ma trận hiệp phương sai
- Số lượng các trị riêng tương ứng với sự khác biệt của dữ liệu theo các hướng của vector riêng



Hình vẽ minh họa hướng của vectơ riêng.

- Như chúng ta đã thấy ở trên rằng vector gốc \mathbf{x} có thể được xây dựng lại bằng cách sử dụng những thành phần chính của nó :

$$\hat{\mathbf{x}} - \bar{\mathbf{x}} = \sum_{i=1}^K b_i u_i \quad \text{hay} \quad \hat{\mathbf{x}} = \sum_{i=1}^K b_i u_i + \bar{\mathbf{x}}$$

- Điều đó cho thấy rằng cơ sở trong không gian có số chiều ít được dựa trên những thành phần chính sẽ làm tối thiểu hóa lỗi trong quá trình xây dựng lại :

$$e = \|x - \hat{x}\|$$

- Điều đó cho thấy rằng số lỗi sẽ tương đương :

$$e = 1/2 \sum_{i=K+1}^N \lambda_i$$

c. Kỹ thuật trích đặc trưng bằng PCA

Giả sử x_1, x_2, \dots, x_n là các vector $N \times 1$

Bước 1:

$$\bar{x} = \frac{1}{M} \sum_{i=1}^M x_i$$

Bước 2: trừ hai giá trị trung bình

$$\Phi_i = x_i - \bar{x}$$

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

$$A = [\Phi_1 \ \Phi_2 \ \dots \ \Phi_M]$$

Sau đó tính:

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = AA^T$$

Bước 4: tính toán các trị riêng của C

$$\lambda_1 > \lambda_2 > \dots > \lambda_{N-1} > \lambda_N$$

Bước 5: tính toán các vectơ riêng của C: $u_1, u_2, u_3, \dots, u_{N-1}, u_N$

Vì C đối xứng ($N \times N$) nên $u_1, u_2, u_3, \dots, u_{N-1}, u_N$ hợp thành một cơ sở (bất kể một vector x nào đều có thể viết dưới dạng một tổ hợp tuyến tính của các vector riêng)

$$x - \bar{x} = b_1 u_1 + b_2 u_2 + \dots + b_N u_N = \sum_{i=1}^N b_i u_i$$

Bước 6 (bước giảm số chiều): chỉ giữ lại những thuộc tính tương ứng

với các trị riêng lớn nhất. ($K \ll N$)

$$\hat{x} - \bar{x} = \sum_{i=1}^K b_i u_i$$

Do đó, sự biểu diễn của $\hat{x} - \bar{x}$ trong $u_1, u_2, u_3, \dots, u_{N-1}, u_N$ là:

$$\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_K \end{bmatrix}$$

Phép biến đổi tuyến tính $\mathbf{R}^n \rightarrow \mathbf{R}^K$ nhằm làm giảm số chiều sẽ là:

$$\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_K \end{bmatrix} = \begin{bmatrix} u_1^T \\ u_2^T \\ \dots \\ u_K^T \end{bmatrix} (x - \bar{x}) = U^T (x - \bar{x})$$

- **Lưu ý :**

Khi số lượng mẫu \mathbf{M} trong tập \mathbf{X} nhỏ hơn số chiều \mathbf{n} , thay vì tính trực tiếp các vector riêng từ ma trận hiệp phương sai \mathbf{C} , ta có thể tính các vector riêng theo phương pháp sau :

- **Bước 1 :** Tính ma trận kích thước $\mathbf{M} \times \mathbf{M}$, \mathbf{C}' như sau :

$$\mathbf{C}' = \mathbf{Y}^T \cdot \mathbf{Y}$$

với $\mathbf{Y}_{n \times M} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]$ mỗi cột của ma trận là một phần tử \mathbf{x}_i , $i=1..m$

- **Bước 2:**
Tính \mathbf{M} vector riêng $\mathbf{E}\mathbf{M}_i$ và các trị riêng tương ứng của ma trận \mathbf{C}' . Chọn \mathbf{m} vector riêng ứng với \mathbf{m} trị riêng lớn nhất để tiếp tục
- **Bước 3:**
Chiếu các vector riêng \mathbf{M} chiều này về lại không gian \mathbf{n} chiều của các mẫu \mathbf{x}_i bằng cách như sau :

$$\mathbf{E}\mathbf{n}_i = \sum_{k=1}^M \mathbf{E}\mathbf{M}_k x_k$$

Các vector $\mathbf{E}\mathbf{n}_i$ thu được chính là các vector riêng cần tìm của ma trận \mathbf{C} .

- **Cách xác định số thành phần chính hiệu quả nhất**

– Có hai phương pháp hữu ích giúp cho chúng ta lấy số lượng thành phần chính sao cho hiệu quả . Tất cả hai phương pháp này đều dựa trên mối quan hệ giữa các giá trị đặc trưng.

Sắp xếp lại các giá trị đặc trưng tìm được theo thứ tự giảm dần về mặt giá trị **(1,eigenvalue[1]), (2,eigenvalue[2]), ..., (p,eigenvalue[p])** và Thứ tự này vẫn đảm bảo được thứ tự của các vector đặc trưng tương ứng. Theo dõi sự biến thiên của chuỗi giá trị đặc trưng vừa được xếp lại, khi sự biến thiên này tiến đến một điểm ngưỡng (thông thường xấp xỉ bằng không). Thông thường đó chính là lúc mà ta chọn được đủ số lượng thành phần chính cần thiết .

Để chọn K, sử dụng tiêu chuẩn sau đây :

$$\frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^N \lambda_i} > \text{ngưỡng}$$

Với phương châm làm sao số lượng thành phần thấp nhất đủ để giải thích khả năng phân tán tập mẫu học thành các mẫu riêng cần thiết nhất.

d. Các đánh giá quan trọng về rút trích đặc trưng bằng phương pháp PCA

- Khi lấy số đặc trưng càng về sau thì khả năng biến động càng thấp, có nghĩa mối quan hệ giữa các phần tử càng cao, thì sự giao nhau giữa các lớp mẫu trong tập mẫu càng lớn.
- Nhưng nếu lấy không đủ số lượng thành phần chính , thì khả năng phân tán của tập mẫu càng cao (có thể tăng vượt ngoài số lớp mẫu cần thiết trong tập mẫu).

III. Phân công công việc và các khó khăn trong quá trình thực hiện.

1. Những khó khăn trong quá trình thực hiện.

Đề tài liên quan tới sinh trắc học là một trong những đề tài khó. Đề tài số 44 liên quan tới kỹ thuật trích chọn đặc trưng khuôn mặt cũng là một vấn đề tới hiện tại được rất nhiều chuyên gia quan tâm tới. Đề tài này không chỉ có tính ứng dụng rất cao trong các hệ thống nhận dạng, bảo mật..mà độ khó của nó

cũng là một thách thức lớn trong việc tìm ra các phương pháp để đảm bảo độ chính xác nhất trong thực tế. Với việc nghiên cứu phép biến đổi KL và phép phân tích các thành phần chính PCA trong trích chọn đặc trưng khuôn mặt nhóm thực hiện đề tài đã thực sự rất nỗ lực cố gắng trong việc tìm hiểu thuật toán và những tài liệu liên quan. Chúng em gặp khó khăn rất nhiều vì với đề tài này tài liệu tiếng việt rất ít, chúng em đã tham khảo một số tài liệu tiếng Anh để rút ra những phần cơ bản nhất. Do hiểu biết còn hạn chế, và tính phức tạp của bài toán nên không tránh khỏi sự thiếu sót về nội dung trong báo cáo cũng như trong tầm nhận thức bao quát đề tài. Chúng em rất mong được cô thông cảm và góp ý.

2. Phân công công việc trong nhóm.

Về phần tìm hiểu và dịch các tài liệu liên quan:

Đinh Thị Am: Tìm hiểu phương pháp phân tích thành phần chính trong trích chọn đặc trưng khuôn mặt.

Lê Thanh Bình: Tìm hiểu các tài liệu liên quan tới phép biến đổi KL.

Lê Thế Văn: Tìm hiểu các tài liệu liên quan tới sinh trắc học và đặc trưng sinh trắc ảnh khuôn mặt.

Sau khi dịch được một số phần liên quan tới đề tài, cả nhóm chúng em đã cùng tập hợp để thống nhất các thuật ngữ sử dụng, cách hiểu của từng phần và cũng nhau xây dựng báo cáo.

IV. Các tài liệu tham khảo.

1. Lương Mạnh Bá & Nguyễn Thanh Thủy-Nhập môn xử lý ảnh- Nhà xuất bản khoa học và kỹ thuật-2003.
2. Modular PCA Face Recognition Based on Weighted Average-Vol3, No11, November 2009 (www.ccsenet.org/journal.html)
3. Localized Principal Component Analysis Learning for Face Feature Extraction and Recognition (Irwin King and Lei Xu. Localized principal component analysis learning for face feature extraction and recognition in *Proceedings to the Workshop on 3D Computer Vision '97*, pages 124–128, Shatin, Hong Kong, 1997. The Chinese University of Hong Kong.)
4. SVD, PCA, KLT, CCA, and All That (Rutgers University Electrical & Computer Engineering Department 332:525 Optimum Signal Processing 2002–2007)

Một số các forum về chuyên ngành công nghệ thông tin khác.