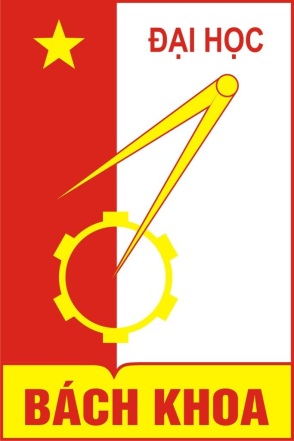
**TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

**VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

\*-------\*-------\*



**BÁO CÁO BÀI TẬP LỚN**

**XỬ LÝ ẢNH**

***Đề số 10:***

**Tìm hiểu phép biến đổi KL và PCA, khảo sát các ứng dụng và phân tích một ứng dụng của của phép biến đổi KL và PCA trong xử lý ảnh màu.**

***Giảng viên hướng dẫn*** : PGS – TS. Nguyễn Thị Hoàng Lan

***Nhóm sinh viên*** : Nhóm 28

*Lê Quang Hiếu* – 2007 1095 – HTTT K52

*Lê Ngọc Minh* – 2007 1946 – KHMT K52

*Lưu Thị Thùy Nhung* – 2007 2167 – HTTT K52

Hà Nội 12/2010

Mục lục

[I. Khái niệm, cơ sở lý thuyết biến đổi KL, PCA 2](#_Toc279705159)

[1. Phép biến đổi KL 2](#_Toc279705160)

[1.1. Cơ sở lý thuyết của phép biến đổi KL 2](#_Toc279705161)

[1.2. Biến đổi KL 3](#_Toc279705162)

[2. Phép biến đổi PCA 4](#_Toc279705163)

[2.1. Lý thuyết biến đổi PCA: 4](#_Toc279705164)

[2.2. Giải bài toán PCA 8](#_Toc279705165)

[II. Khảo sát các ứng dụng của phép biến đổi KL, PCA trong xử lý ảnh màu 9](#_Toc279705166)

[2.1. Khái quát khả năng ứng dụng của biến đổi KL trong xử lý ảnh màu 9](#_Toc279705167)

[2.1.1. Đặc tính ứng dụng của biến đổi KL 9](#_Toc279705168)

[2.1.2. Lưu trữ, nén ảnh trong không gian KL 9](#_Toc279705169)

[2.1.3. Sử dụng xấp xỉ biến đổi KL trong nghiên cứu ảnh màu tự nhiên 11](#_Toc279705170)

[2.2. Ứng dụng thực tế của biến đổi KL, PCA trong xử lý ảnh màu 12](#_Toc279705171)

[2.2.1. Ứng dụng biến đổi KL trong nhận dạng mống mắt (iris recognition) 12](#_Toc279705172)

[2.2.2. Ứng dụng biến đổi KL, PCA trong nhận dạng khuôn mặt (face recognition) 15](#_Toc279705173)

[III. Phân tích ứng dụng của biến đổi KL, PCA trong nhận dạng khuôn mặt 15](#_Toc279705174)

[2.1. Phương pháp nhận dạng khuôn mặt riêng - Eigenface 15](#_Toc279705175)

[2.2. Tính toán khuôn mặt riêng - eigenface 16](#_Toc279705176)

[2.3. Sử dụng khuôn mặt riêng để phân loại một ảnh khuôn mặt 18](#_Toc279705177)

[IV. Cài đặt 21](#_Toc279705178)

[1. Chương trình nhận dạng khuôn mặt KL-based 21](#_Toc279705179)

[2. Chương trình nhận dạng khuôn mặt PCA-based 24](#_Toc279705180)

[V. Tài liệu tham khảo 26](#_Toc279705181)

[VI. Phân công thực hiện 27](#_Toc279705182)

# Khái niệm, cơ sở lý thuyết biến đổi KL, PCA

## Phép biến đổi KL

*Biến đổi KL có nguồn gốc từ khai triển chuỗi của các các quá trình ngẫu nhiên liên tục. Biến đổi KL cũng còn gọi là biến đổi Hoteling hay phương 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.*

### Cơ sở lý thuyết của phép biến đổi KL

Đây là phép biến đổi không gian chiều thành không gian chiều, với . Mỗi thành phần của vectơ 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 ảnh 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. Gọi là tập các vector cơ sở trong không gian trực giao .

Với với và

Mọi véctơ y trong không gian trực giao có thể viết:

với

Gọi là kết quả thu được trong không gian m chiều và

Sai số trong phép biến đổi

Sai số trung bình bình phương:

Mà , do đó

Theo định nghĩa của R, phương trình trở thành:

đạt min khi đặt min.

Đặt .(5)

Như vậy đạt min khi 5 min. Để tìm min của 5 ta dùng phương pháp đạo hàm và dẫn đến việc giải phương trình:

Phương trình 6 gọi là phương trình đặc trưng của R với là các trị riêng và là các véctơ riêng tương ứng. Đây chính là cơ sở lý thuyết của biến đổi KL.

### Biến đổi KL

#### Định nghĩa và khái niệm

Cho là một vectơ các số thực ngẫu nhiên; vectơ cơ sở của biến đổi KL là các véctơ riêng trực giao của ma trận hiệp biến cho bởi phương trình:

Biến đổi KL của là:

và biến đổi ngược:

là véc tơ cột, là véctơ hang và là cột thứ của ma trân .

Biến đổi đưa về dạng đường chéo:

Thường người ta hay làm với ma trận hơn.

#### Biến đổi KL của ảnh

Nếu một ảnh NxN được biểu diễn bởi trường ngẫu nhiên, ma trận A cho bởi:

Thì ảnh cơ sở của biến đổi KL là các hàm riêng, chuẩn và trực giao là lời giải của phương trình:

Theo kí pháp ma trận ta có:

Với là véctơ x1 biểu diễn của và là ma trận X ánh xạ vào véctơ , .

Nếu là tách được thì ma trận X mà các cột là , sẽ tách được:

hay

Biến đổi KL của là

Và biến đổi ngược :

## Phép biến đổi PCA

### Lý thuyết biến đổi PCA:

#### Khái niệm

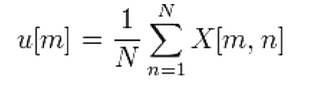
Trong các thực nghiệm khoa học cũng như các thống kê các phép đo đạc thường tạo ra lượng dữ liệu khổng lồ. Phát hiện các quy luật ẩn sau dữ liệu trở thành một công việc rất khó khăn do có quá nhiều yếu tố cần được xét đến. Một cách đơn giản nhưng hiệu quả để giúp các nhà thực nghiệm, các nhà nghiên cứu là PCA.

PCA là viết tắt của Principle Component Analysis (phương pháp phân tích thành phần quan trọng) là phương pháp thay thế các đại lượng của bộ dữ liệu ban đầu bằng các tổ hợp tuyến tính của chúng (gọi là một “thành phần”) và từ đó chọn ra những thành phần quan trọng nhất cho những bước phân tích tiếp theo.

PCA được ứng dụng rộng rãi trong các lĩnh vực nghiên cứu khác nhau: vật lí, sinh học, xã hội... Ngoài ra nhờ khả năng nén dữ liệu (làm giảm khối lượng dữ liệu trong khi vẫn giữ lại phần lớn thông tin), PCA cũng được áp dụng trong một số kĩ thuật nén ảnh, xử lí ảnh, nhận dạng...

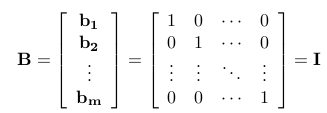
#### Dữ liệu

Giả sử một cuộc khảo sát được thực hiện trên n người, với mỗi người một số m thông số được ghi nhận. Các thông số về một người tạo thành một véc-tơ trong không gian m chiều với m tương đối lớn. Thông số của tất cả các đối tượng khảo sát hợp thành ma trận X có m dòng n cột.

Để đơn giản hoá các phép tính trong các phần tiếp theo, ta giả sử X ở dạng *mean derivation form*. Nghĩa là giá trị của mỗi thông số được trừ đi kì vọng của tất cả các thông số cùng loại (trên tất cả các đối tượng khảo sát khác) sao cho ta có kì vọng của mỗi loại thông số đều bằng 0. Gọi u là ma trận kì vọng kích thước mx1:  
   
Ta thay X bằng X-uh với h là ma trận 1xn chứa toàn số 1.

#### Cơ sở

Mỗi véc-tơ trong không gian m chiều đều là tổ hợp tuyến tính của m véc-tơ cơ sở. Các véc-tơ cơ sở hợp thành cơ sở B kích thước mxm. Một cách chọn đơn giản nhất của cơ sở là ma trận đơn vị I:



Câu hỏi đặt ra là chọn cơ sở như thế nào để “diễn đạt lại” bộ dữ liệu X một cách tốt nhất?

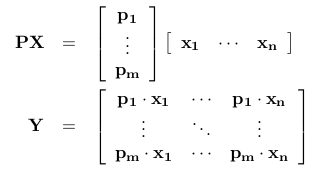
Gọi X và Y là các ma trận mxn liên hệ với nhau bằng toán tử tuyến tính P. X là bộ dữ liệu ban đầu và Y là bộ dữ liệu được “ diễn đạt lại”. Ta có PX=Y (1), kí hiệu:

* pi là các hàng của P
* xi là các cột của X
* yi là các cột của Y

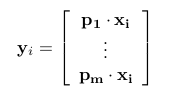
Ta có thể diễn đạt biểu thức (1) bằng nhiều cách:

* P là ma trận biến X thành Y
* P đại diện cho phép quay và phép co giãn hình học chuyển X thành Y
* Các hàng của P, {p1,..., pm} là các véc-tơ cơ sở mới “diễn đạt lại” các cột của X

Xem xét kĩ hơn ta có cách giải thích khác:



Mỗi cột của Y có dạng:



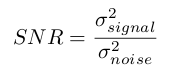
Ta thấy mỗi thành phần của yi đều là tích của xi với một dòng của P. Nói cách khác, phần tử thứ j của yi là kết quả phép chiếu xi lên trục pj. Như vậy mỗi hàng của P là một véc-tơ cơ sớ để diễn đạt lại các cột của X.

#### Mục tiêu

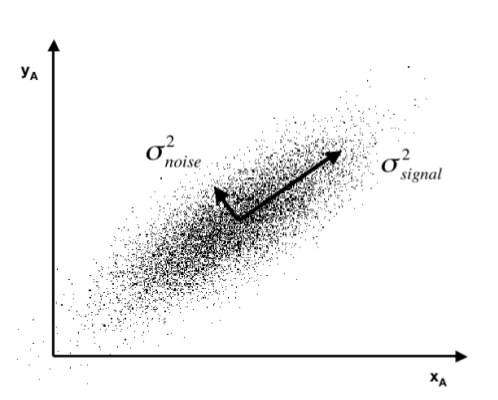
Để trả lời câu hỏi “Chọn cơ sở như thế nào để diễn đạt lại X một cách tốt nhất?” trước hết ta phải tìm hiểu thế nào là một bộ dữ liệu tốt. Trong các hệ tuyến tính, chỉ có hai dạng vấn đề ảnh hưởng xấu đến dữ liệu: nhiễu và dư thừa.

***Nhiễu***

Nhiễu là những tác động ngẫu nhiên làm thay đổi dữ liệu. Trong mọi trường hợp nhiễu cần phải tương đối nhỏ so với tín hiệu để thí nghiệm đó có hiệu quả. Một cách đánh giá nhiễu là tỉ số tín hiệu trên lỗi (signal-to-noise-ratio, SNR) được định nghĩa là:



SNR cao (>> 1) chỉ ra rằng dữ liệu rất chính xác trong khi SNR thấp cho thấy dữ liệu bị ảnh hưởng nặng bởi nhiễu.

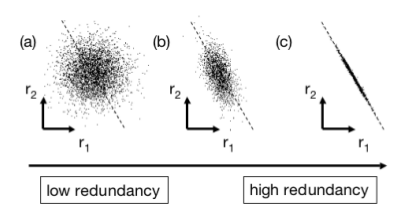


***Hình 1: Dữ liệu thu thập về chuyển động 1 vật trên đường thẳng.***

Hình trên mô tả dữ liệu thu thập được về chuyển động của một vật trên một đường thẳng. Do tác động của nhiễu mà các điểm không thực sự nằm trên một đường thẳng mà phân tán về hai phía tạo thành một hình ô-van. SNR càng lớn thì hình ô-van càng “béo” và ngược lại.

***Dư thừa dữ liệu***

Do trước khi tiến hành khảo sát ta không biết quy luật của hệ thống nên thường đo đạc nhiều dữ liệu hơn cần thiết. Những thông số phụ thuộc lẫn nhau không cho thêm thông tin về hệ thống mà chỉ gây khó khăn cho quá trình nghiên cứu.



***Hình 2.***

Giả sử ta tiến hành hai phép đo r1 và r2, chúng càng phụ thuộc vào nhau thì các điểm càng tập trung quanh một đường thẳng. Hình (a) cho thấy r1 và r2 không phụ thuộc vào nhau, mức độ phụ thuộc tăng dần từ trái sang phải. Trong hình (c) ta thấy rõ ràng thay cho hai phép đo ban đầu ta có thể đo một đại lượng r2-kr1.

#### Ma trận hiệp phương sai

Hiệp phương sai của hai đại lượng A và B được định nghĩa là trung bình của tích của hai đại lượng đo trên cùng một mẫu (giả thiết dữ liệu ở dạng mean derivation form):



Hiệp phương sai cho biết mức độ phụ thuộc của các đại lượng. Giá trị hiệp phương sai bằng 0 cho thấy A và B hoàn toàn không có liên hệ. Giá trị hiệp phương sai đúng bằng phương sai của A và B nếu A=B.

Như vậy, nhiễu được đặc trưng bởi phương sai còn dư thừa được đặc trưng bởi hiệp phương sai. Một cách đơn giản để tổng hợp cả hai tính chất của dữ liệu là tính *ma trận hiệp phương sai* của nó. Giả thiết dữ liệu là mean derivation form, ta có công thức:



SX là ma trận vuông kích thước mxm trong đó:

* Các phần tử trên đường chéo là phương sai của các phép đo
* Các phần tử khác là hiệp phương sai của các phép đo

Dễ thấy SX là ma đối xứng vì hiệp phương sai của A và B cũng giống như hiệp phương sai của B và A.

Ma trận hiệp phương sai cho ta biết tính chất của tất cả các cặp phép đo, từ đó gợi ý ta có thể xác định độ “tốt” của cách biểu diễn lại X dựa trên ma trận hiệp phương sai của Y.

#### Chéo hoá ma trận hiệp phương sai

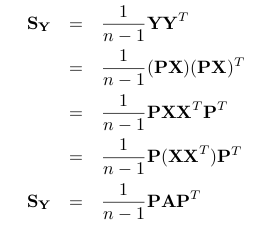
Rõ ràng một bộ dữ liệu tốt là bộ dữ liệu ít nhiễu và không có dư thừa. Ma trận hiệp phương sai của bộ dữ liệu đó có các thành phần nằm ngoài đường chéo đều bằng 0. Các thành phần trên đường chéo cho ta thấy phương sai của các thành phần của dữ liệu, phương sai lớn cho thấy nhiều thông tin chứa đựng trong thành phần đó còn phương sai nhỏ cho thấy thành phần đó có thể là nhiễu. Dựa vào đó ta có thể giữ lại những thành phần quan trọng và loại bỏ những thành phần còn lại.

Tóm lại, mục tiêu của phương pháp PCA là:

*Tìm ma trận trực giao P sao cho với Y=PX thì SY là ma trận chéo.*

### 2.2. Giải bài toán PCA

Đặt A=XXT và biến đổi SY và như sau:



Lời giải của bài toán là ma trận P có các hàng là các véc-tơ riêng của A, khi đó PAPT là ma trận chéo với các thành phần trên đường chéo là các trị riêng tương ứng với các véc-tơ riêng trong P.

# Khảo sát các ứng dụng của phép biến đổi KL, PCA trong xử lý ảnh màu

## Khái quát khả năng ứng dụng của biến đổi KL trong xử lý ảnh màu

### Đặc tính ứng dụng của biến đổi KL

Các hệ thống thu nhận hình ảnh hiện tại dựa trên các định luật màu Red, Green, Blue (R, G, B). Tuy nhiên, mô hình thu nhận ảnh màu này lại không giống như hệ thống thị giác của con người nên để có được sự biểu diễn gần với con người nhất, không gian cường độ và sắc độ (intensity and chrominance space) được sử dụng. Do phương pháp này là không tuyến tính, nó khiến cho màu sắc không được ổn định.

Bên cạnh đó, phân tích ảnh mầu trong không gian RGB và không gian cường độ-sắc độ (intensity – chrominance) cũng cho thấy các thành phần màu có độ tương đồng (correlation) cao. Những thành phần màu này cần được tách biệt (decorrelated) để được phân tích, nghiên cứu độc lập nhau.

Phép biến đổi Karhunen – Loève (KL) cung cấp khả năng phân lập thành phần mầu một cách tối ưu. Biến đổi này làm nổi bật những thành phần màu chìm trong ảnh, từ đó, ta có thể nắm rõ được sự phân phối màu, cấu trúc ảnh, những thành phần dữ liệu quan trọng có giá trị sử dụng trong ảnh.

Có thể nói, không gian lưu trữ ảnh KL cung cấp độ phân lập cao nhất giữa các thành phần màu, đồng thời cho phép độ phân giải tối ưu cùng một tỉ lệ nén cao cho các ảnh thuần nhất (homogeneous images).

### Lưu trữ, nén ảnh trong không gian KL

1. ***Chuyển đổi từ không gian RGB sang không gian KL:***

Quan hệ vector xác định không gian KL:

Với:

* I: “vector ảnh màu gốc”
* : vector trung vị của ảnh màu gốc
* A: ma trận biến đổi
* K: vector biến đổi
* Với k,s =R,G, B

Ma trận A được tạo từ bộ vector cơ sở của ma trận hiệp biến tính trên ba thành phần R,G,B.

N là số pixel của ảnh.

1. ***So sánh đặc tính lưu trữ trong không gian RGB (Red, Green, Blue), HSI (Hue, Saturation, Intensity) và KL***

Vấn đề của phân tích ảnh màu đó là phân biệt độ chói (luminance) và độ màu (sắc độ - chrominance). Do đó, ta cần xác định *hệ số tương đồng chéo trung bình* (mean cross correlation coefficient). Với một không gian xác định thì hệ số tương quan được cho bởi:

* : là hiệp phương sai giữa ảnh và ảnh
* : tương ứng là phương sai của và

Trên thực tế cho thấy:

* Trong không gian RGB các thành phần của ảnh có sự tương quan lớn.
* Trong không gian HSI, hệ số tương quan thấp hơn. Và chúng ta sẽ đạt được kết quả gần như hoàn hảo trong không gian KL.

1. ***Nén ảnh trong không gian KL***

Từ thực tế, ta thấy có thể giảm được kích thước không gian biểu diễn nhờ phép biến đổi KL. Một pixel của ảnh màu được mã hóa ban đầu bởi 24 bits trong không gian RGB. Đối với không gian KL, số lượng bits có thể được giảm xuống mà không làm giảm chất lượng ảnh ban đầu. Trong một số trường hợp, tỷ số nén có thể đạt được 20% đến 30%.

Một số ví dụ về nén ảnh trong không gian KL:

***Ảnh ban đầu Ảnh sau khi nén***





***Hình 3: So sánh các ảnh trước và sau khi nén bằng KL***

### Sử dụng xấp xỉ biến đổi KL trong nghiên cứu ảnh màu tự nhiên

Mặc dù đem lại hiệu quả cao trong phân lập thành phần màu hay nén ảnh, việc sử dụng biến đổi KL cũng có một số nhược điểm : khối lượng tính toán lớn vì phải tính ma trận hiệp biến với mỗi ảnh, tiếp theo là giải phương trình tìm trị riêng và vector riêng để xác định các hệ số.

Để tận dụng được lợi thế của biến đổi KL, đồng thời tăng tốc độ tính toán bằng cách xác định một ma trận biến đổi cho tất cả các ảnh, các xấp xỉ (approximation) của không gian KL được sử dụng.

Đây là những biến đổi tuyến tính được sinh ra từ biến đổi Fourier. Thực chất, những xấp xỉ này sẽ giống như biến đổi KL nếu ta sử dụng lượng vector lớn.

Một số xấp xỉ KL được sử dụng là các biến đổi trực giao: DFT, DOFT, DREFT, DROFT, DEST.

Xét 2 xấp xỉ DEST (Discrete Even Sine Transform) và DCT(Discrete Cosine Transform) :

* DCT

*với u=1 và j=1,..,N*

*với u=2,..,N-1 và j=1,..,N*

* DEST

*với u=2,..,N-1 và j=1,..,N*

*với u=N và j=1,..,N*

Những công thức này giúp khởi tạo ma trận chuyển đổi(ma trận vector riêng của ma trận hiệp biến trong biến đổi KL) một cách đơn giản. Từ đó, ta thu được ba vector trực giao. Nó sẽ xác định ba mặt phẳng ảnh(các thành phần của ảnh màu trong không gian mới). Có thể coi như ba ảnh tạo từ ba trục KL. Mỗi vector có ba thành phần: với u = 1,2,3 và j = 1,2,3.

* ***Kết quả áp dụng:***

Từ kết quả thu được, ta thấy các phép xấp xỉ cũng cho kết quả gần như phép biến đổi KL. So sánh DEST và DCT, DCT có kết quả xấp xỉ KL tốt hơn DEST.

|  |  |  |  |
| --- | --- | --- | --- |
| **Ảnh ban đầu** | **KL** | **DCT** | **DEST** |
|  |  |  |  |

***Hình 4: Kết quả ảnh màu sau khi thực hiện biến đổi hoặc xấp xỉ.***

## Ứng dụng thực tế của biến đổi KL, PCA trong xử lý ảnh màu

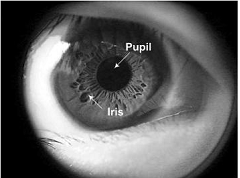
Với khả năng nổi bật là phân lập các thành phần màu phục vụ nghiên cứu và phân tích độc lập, biến đổi KL, PCA được ứng dụng rộng rãi trong các ***công nghệ nhận dạng.***

Ở đây, 2 công nghệ nhận dạng được giới thiệu là ứng dụng trong nhận dạng mống mắt và ứng dụng trong nhận dạng khuôn mặt. Trong đó, ứng dụng nhận dạng khuôn sẽ được khảo sát chi tiết trong phần III.

### Ứng dụng biến đổi KL trong nhận dạng mống mắt (iris recognition)

Nhận dạng mống mắt (iris recognition) là một phương pháp nhận dạng sinh trắc học (biometrics) dựa trên sự duy nhất của mống mắt (iris) con người. Biến đổi KL được đưa vào phương pháp nhằm tách các đặc tính cục bộ của ảnh mống mắt (iris).

Trước hết, hình ảnh mống mắt con người được chụp lại dưới dạng ảnh RGB, nhưng chỉ sử dụng không gian đỏ - red do mống mắt của người là khá nhỏ và sẫm màu với người châu Á.



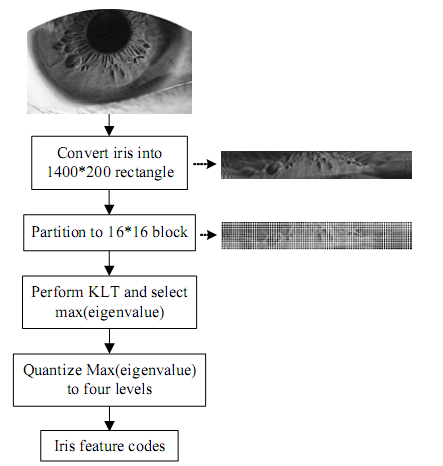
***Hình 5: Mống mắt (iris)***

Hình ảnh sau đó được đưa vào quá trình tiền xử lý đề loại bỏ đi những phần không cần thiết (mi mắt, con ngươi,…).

Sau giai đoạn này, hình ảnh được chuyển tới quá trình phân tách các đặc tính mống mắt. Tại quá trình này, có hai phương pháp có thể được sử dụng. Thứ nhất là phương pháp chiều phân chia(divider dimension) và phương pháp thứ hai là sử dụng biến đổi KL.

Biến đổi KL đem lại kết quả to lớn cho công nghệ nhận dạng mống mắt bằng cách khái quát những đặc tính không tương đồng, trực giao nhằm tránh dư thừa dữ liệu. Bên cạnh đó, biến đổi KL cũng có sai số bình phương trung bình (Mean Square Error – MSE) nhỏ nhất, so sánh với các vector gốc và vector xấp xỉ khác. Nhờ vậy, biến đổi KL được áp dụng vào phân tích hình ảnh mống mắt, từ đó phân tách ra các đặc tính.

Biểu đồ quá trình phân tách đặc tính bằng KLT



Mã đặc tính mống mắt

Lượng tử hóa max(trị riêng) thành 4 mức

Sử dụng KLT và chọn max(trị riêng)

Phân vùng thành 16 \* 16 blocks

Biến đổi mống mắt thành 1400\*200 hình chữ nhật

***Hình 6: Sơ đồ thao tác nhận dạng mống mắt***

1. Để áp dụng được biến đổi KL, các phần mống mắt cục bộ cần được cấu trúc dạng hình chữ nhật.

Do đó, ta phải biến đổi từ hệ tọa độ cực sang hệ tọa độ Cartesian. Vì điểm trung tâm, biên ngoài và biên trong của mống mắt đã được xác nhận, ta có thể dễ dạng chiếu vòng mống mắt sang ảnh hình chữ nhật mống mắt kích cỡ 1400x200 thông qua 1 phép biến đổi nội suy tuyến tính.

1. Nếu ta thực hiện ngay biến đổi KL để phân tích hình ảnh mống mắt dạng chữ nhật, các đặc tính là quá ít để phân định ảnh các tiểu tiết của mống mắt. Do đó, block chữ nhật 1400x200 nên được phân thành 1044 block vuông kích thước 16x16.
2. Sau đó, ta thực hiện biến đổi KL trên mỗi block vuông 16x16 để phân tách các đặc tính cục bộ của mống mắt.
3. Sau khi thực hiện biến đổi KL, ta thu được một tập trị riêng và vector riêng cho mỗi block vuông 16x16. Chọn giá trị maximum *λmax* từ các trị riêng cho mỗi block là đặc tính.
4. Do đó, 1044 đặc tính cho ảnh mống mắt đầu vào được xác định. Mỗi giá trị trong 1044 đặc tính có một vùng dải rộng.

Để giảm không gian nhớ và thời gian tính toán trong thao tác vector đặc tính, ta lượng tử hóa mỗi trị riêng thành 4 mức khác nhau. Từ đó, tổng hợp có 261 bytes 91044x2 bits) cần có để biểu diễn ảnh mống mắt.

### Ứng dụng biến đổi KL, PCA trong nhận dạng khuôn mặt (face recognition)

Trong thế giới mạng công nghệ ngày nay, nhu cầu duy trì bảo mật thông tin và tài sản ngày càng cấp thiết, do đó các công nghệ bảo mật truy nhập đang ngày càng phát triển. Cùng với công nghệ nhận dạng mống mắt, công nghệ nhận dạng khuôn mặt (face recognition) là một bảo mật truy nhập sinh trắc học Biometric đang được phát triển mạnh mẽ. Trong công nghệ này, ứng dụng KL đóng một vai trò hết sức quan trọng.

Về cơ bản, nguyên phương pháp nhận dạng khuôn mặt ở đây là xuất phát từ cơ sở dữ liệu cách khuôn mặt đã có, thiết lập nên các khuôn mặt riêng (eigenfaces) từ các vector riêng, trị riêng, *sử dụng khai triển KL hay PCA.* Các khuôn mặt riêng này đã thu hẹp không gian kiểm thử ban đầu. Khi cần nhận dạng một khuôn mặt nằm ngoài cơ sở dữ liệu, ta chỉ cần so sánh đối chiếu với không gian ảnh mặt (face space) được thiết lập từ các khuôn mặt riêng này, từ đó xác định khuôn mặt tương đồng đã lưu trong cơ sở dữ liệu, hoặc xác định khuôn mặt này là mới.

Phương pháp này được khảo sát, phân tích chi tiết trong phần tiếp theo của báo cáo.

# Phân tích ứng dụng của biến đổi KL, PCA trong nhận dạng khuôn mặt

Công việc nhận dạng khuôn mặt gặp phải nhiều vấn đề khó khăn và phức tạp. Để giải quyết những vấn đề này và tìm ra những điểm bất biến quan trọng phục vụ nhận dạng, các nhà nghiên cứu đã phát triển nhiều công nghệ nhận dạng, và một trong những phương pháp hiệu quả nhất là phương pháp khuôn mặt riêng –Eigenface, áp dụng biến đổi KL hoặc PCA vào quá trình phân tách đặc tính khuôn mặt. Phương pháp này đã giảm thiểu đáng kể chiều đặc tính bề mặt (facial feature dimension) nhưng vẫn giữ được khả năng suy đoán, nhận diện.

## 2.1. Phương pháp nhận dạng khuôn mặt riêng - Eigenface

Xét trên phương diện toán học, ta cần xác định những thành phần chủ yếu tạo nên khuôn mặt, hay những vector riêng (eigenvector) của ma trận phương sai của tập các ảnh khuôn mặt; coi một ảnh là một điểm(hay vector) với trong không gian kích thước lớn. Các vector riêng được sắp xếp, mỗi vector được là một số lượng các đặc tính riêng biệt của các khuôn mặt.

Các vector riêng này có thể coi là một các đặc tính từ đó tạo ra sự khác biệt giữa các khuôn mặt. Các vector riêng có thể được biểu diễn giống với dạng một khuôn mặt, do đó ta có thể gọi đây là các khuôn mặt riêng – eigenfaces. *Như vậy, biến đổi KL hoặc PCA đóng vai trò xác định tổ hợp các vector riêng, trị riêng – khuôn mặt riêng*. Mỗi khuôn mặt có thể được biểu diễn dưới dạng tổ hợp tuyến tính của nhiều khuôn mặt riêng.

Từ đây, ta có thể khái quát các bước của phương pháp như sau:

1. *Yêu cầu 1 tập khởi tạo bao gồm các ảnh khuôn mặt.*
2. *Tính toán các khuôn mặt riêng từ tập đã có, nhưng chỉ giữ lại M ảnh tương ứng với M trị riêng cao nhất, từ đó xác định không gian mặt – face space. Sử dụng biến đổi KL.*
3. *Tính toán trọng số không gian của các nhóm khuôn mặt tương ứng trong cơ sở dữ liệu bằng cách chiếu lên không gian mặt.*
4. *Tính toán tập trong số của khuôn mặt cần nhận dạng bằng cách chiếu lên M khuôn mặt riêng đã có.*
5. *Quyết định hình ảnh đưa vào có phải là khuôn mặt hay không.*
6. *Nếu ảnh đưa vào là khuôn mặt, dựa vào trọng số, xác định có thuộc các nhóm khuôn mặt đã biết hay không.*
7. *Nếu khuôn mặt này tương ứng với nhóm khuôn mặt đã có trong cơ sở dữ liệu, tính toán cập nhật các khuôn mặt riêng và các trọng số riêng (nếu cần).*
8. *Nếu một khuôn mặt lạ xuất hiện nhiều lần, ta có thể tập hợp lại thành nhóm và tạo mới trong cơ sở dữ liệu.*

## Tính toán khuôn mặt riêng - eigenface

Coi một ảnh I(x,y) là một mảng 2 chiều NxN các giá trị cường độ (8bits). Một ảnh có thể coi như một vector chiều N2, do đó một ảnh điển hình kích cỡ 256x256 là một vector chiều 65 536, hay tương ứng với một điểm trong không gian chiều 65 536. Do đó, một tập các ảnh được tham chiếu sang một tập các điểm trong không gian này.

Tuy nhiên, để có được các chi tiết cần thiết, ảnh khuôn mặt có thể được biểu diễn trong một không gian nhỏ hơn. Ý tưởng cơ bản của khai triển KL là tìm những vectors có thể biểu diễn tốt nhất các đặc tính của khuôn mặt trong không gian ảnh. Các vector này xác định một không gian khuôn mặt – face space. Những vector này là vector riêng của ma trận hiệp phương sai , được gọi là các khuôn mặt riêng.

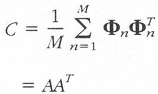
Gọi tập ảnh ban đầu là Γ1, Γ2, Γ3,…, ΓM. Khuôn mặt trung bình của tập được ký hiệu là Ψ = =*n*. Mỗi khuôn mặt chênh lệch với khuôn mặt trung bình một đại lượng là vector Φi = Γ*i* – Ψ. Tập các vector này được phân tích để xác định M vector trực giao un có thể biểu diễn tốt nhất tập dữ liệu. Vector thứ k, uk, được lựa chọn sao cho



đạt giá trị cao nhất, với ràng buộc :



Vector uk và λk tương ứng là vector riêng và trị riêng của ma trận hiệp phương sai:



Với ma trận A = [Φ1 Φ2 … ΦM]. Ma trận C có kích thước N­2 x N­2 và việc xác định N­2 vector riêng và giá trị riêng là khó thực hiện với kích thước ảnh điển hình. Ta cần một phương pháp dễ dàng hơn để tìm các vector riêng này.

Nếu số điểm dữ liệu trong không gian ảnh nhỏ hơn chiều không gian (M<N2), sẽ chỉ có M – 1 vector riêng có ý nghĩa. Ta có thể giải quyết cho các vector riêng N2 chiều bằng cách tính các vector riêng của ma trận MxM, VD: tính 1 ma trận 16 x 16, hơn là ma trận 16 384 x 16 384, và sau đó, lấy các tổ hợp tuyến tính của ảnh mặt Φi. Xét các vector riêng vi của ATA :

ATAvi = μi vi

Nhân 2 vế với A , ta có:

AAT Avi = μi Avi

Từ đó ta thấy Avi là vector riêng của C = AAT.

Tiếp theo hướng phân tích này, ta thiết lập ma trận MxM : L=ATA, với Lmn=Φn, và tìm M vector riêng, vl, của L. Các vector này xác định tổ hợp tuyến tính M ảnh từ tập ban đầu để tạo ra các mặt riêng ul:



Bằng phương pháp này, các phép tính đã được giảm đi đáng kể.

Trong thực tế, tập ảnh khuôn mặt thử nghiệm là tương đối nhỏ (M«N2), và các phép toán trở nên dễ dàng hơn. Các trị riêng tương ứng cho phép ta xếp loại các vector riêng theo mức độ hữu ích của chúng trong biểu thị các điểm khác biệt giữa các ảnh.



***Hình 7 : Tập khuôn mặt đào tạo (training set).***

***Hình 8 : Khuôn mặt trung bình Ψ Hình 9: Các khuôn mặt riêng được lựa chọn***

## Sử dụng khuôn mặt riêng để phân loại một ảnh khuôn mặt

Do các khuôn mặt riêng là khá đầy đủ để mô tả ảnh các khuôn mặt, nên ta có thể sử dụng chúng như một công cụ nhận dạng khuôn mặt. Trong thực tế, một số lượng M’ nhỏ hơn tỏ ra hiệu quả trong nhận dạng, do ta không cần thiết phải tái tạo lại ảnh ban đầu. Các khuôn mặt riêng tạo ra một không gian chiều con M’ từ không gian ảnh N2 ban đầu. Các vector riêng quan trọng M’ của ma trận L được lựa chọn từ những vector có trị riêng lớn nhất. Trong nhiều thực nghiệm, với cơ sở M=16 ảnh khuôn mặt, M’=7 khuôn mặt riêng được sử dụng.

Một ảnh mặt mới (Γ) được biến đổi thành các thành phần khuôn mặt riêng (chiếu vào không gian ảnh mặt) bằng một biến đổi đơn giản



Với k = 1,…, M’.

Các trọng số của một vector ΩT = [ω1, ω2, …, ωM’ ] mô tả các phần liên quan của mỗi khuôn mặt riêng trong biểu diễn ảnh khuôn mặt nhập vào; coi các khuôn mặt riêng như môt tập cơ sở các ảnh khuôn mặt. Vector này có thể được sử dụng trong giải thuật nhận dạng chi tiết chuẩn để tìm ra một số các lớp khuôn mặt định trước. Phương pháp đơn giản nhất để xác định lớp khuôn mặt nào mô tả tốt nhất ảnh khuôn mặt nhập vào là tìm lớp khuôn mặt k làm cực tiểu hóa khoảng cách Euclide:



Với Ωk là một vector mô tả lớp khuôn mặt thứ k. Các lớp khuôn mặt Ωi được tính bằng cách lấy trung bình các kết quả của biểu diễn khuôn mặt riêng trên một số lượng nhỏ ảnh các khuôn mặt của mỗi cá nhân. Một khuôn mặt được coi là thuộc vào lớp k nếu giá trị εk cực tiểu nhỏ hơn một ngưỡng θk chọn trước. Nếu không, khuôn mặt sẽ được xếp vào loại “chưa biết”, và có thể được sử dụng để tạo ra một lớp khuôn mặt mới.

Khoảng cách ϵ giữa hình ảnh và không gian ảnh mặt bình phương khoảng cách giữa giữa ảnh đầu vào trung bình Φ = Γ – Ψ và Φf = i*ui*, hình chiếu của nó lên không gian mặt:



Từ đây, ta có 3 khả năng cho hình ảnh đầu vào và vector mẫu của nó:

1. Gần không gian mặt và gần một lớp khuôn mặt.

* Một cá nhân được nhận dạng và xác định.

1. Gần không gian mặt như không gần một lớp khuôn mặt.

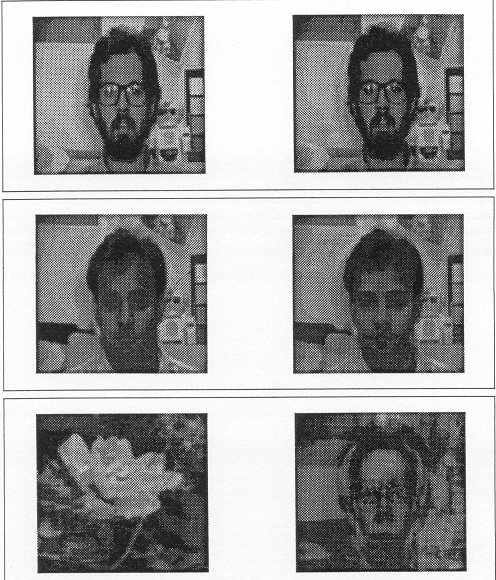
* Phát hiện một cá nhân lạ.

1. Xa không gian mặt và gần một lớp mặt.

* Ảnh đã cho không diễn tả gương mặt

1. Xa không gian mặt và không gần lớp mặt nào.

* Ảnh đã cho không diễn tả khuôn mặt



***Hình 10: 3 ảnh và hình chiếu tương ứng lên không gian ảnh mặt.***

***Khoảng cách tới không gian ảnh mặt tương ứng: (a). 29.8 (b) 58.5 (c)5217.4. Ảnh a và b nằm trong tập ảnh mặt kiểm thử ban đầu.***

# Cài đặt

## Chương trình nhận dạng khuôn mặt KL-based

Dựa theo thuật toán trình bày trong phương pháp Eigenface ở trên, chương trình nhận dạng sau được xây dựng.

Một chương trình được cài đặt sử dụng Matlab với phương pháp tính khuôn mặt riêng và so khớp như đã trình bày ở trên.

* Đọc tập ảnh cho trước (các ảnh .pgm có tên 1 -> 10 lưu trong 3 folder s1,s2,s3, mỗi folder là tập ảnh khuôn mặt của 1 người)

**k = 0;**

**for i=1:1:3**

**for j=1:1:10**

**filename = sprintf('C:\\MATLAB\\att\_faces\\s%d\\%d.pgm',i,j);**

**image\_data = imread(filename);**

**k = k + 1;**

**x(:,k) = image\_data(:);**

**anot\_name(k,:) = sprintf('%2d:%2d',i,j);**

**end;**

**end;**

**nImages = k; %tong so hinh anh**

**imsize = size(image\_data); %dung luong hinh anh**

**nPixels = imsize(1)\*imsize(2); %so pixel tren mot anh**

**x = double(x)/255;**

***Tập các ảnh đầu vào như sau (các biểu cảm khác nhau của 3 người)***









* *Tính khuôn mặt trung bình:*

**avrgx = mean(x')';**

**for i=1:1:nImages**

**x(:,i) = x(:,i) - avrgx;**

**end;**

**subplot(2,2,1); imshow(reshape(avrgx, imsize)); title('mean face')**

* *Tính ma trận hiệp phương sai, đưa ra các khuôn mặt riêng:*

**cov\_mat = x'\*x;**

**[V,D] = eig(cov\_mat); %tri rieng cho ma tran hiep phuong sai**

**V = x\*V\*(abs(D))^-0.5;**

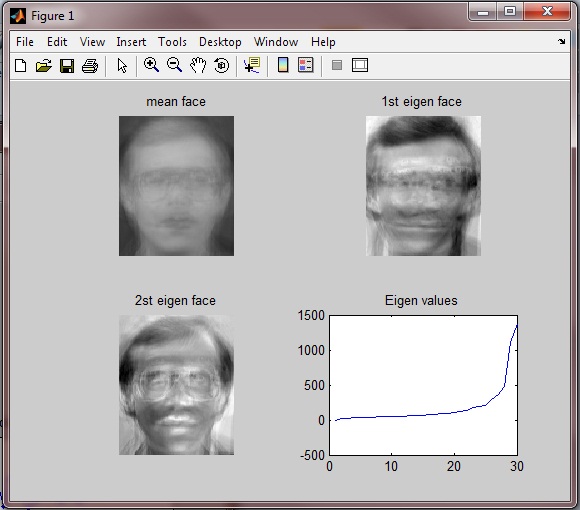
**subplot(2,2,2); imshow(ScaleImage(reshape(V(:,nImages ),imsize))); title('1st eigen face');**

**subplot(2,2,3); imshow(ScaleImage(reshape(V(:,nImages-1),imsize))); title('2st eigen face');**

**subplot(2,2,4); plot(diag(D)); title('Eigen values');**

* *Kết quả thu được:*

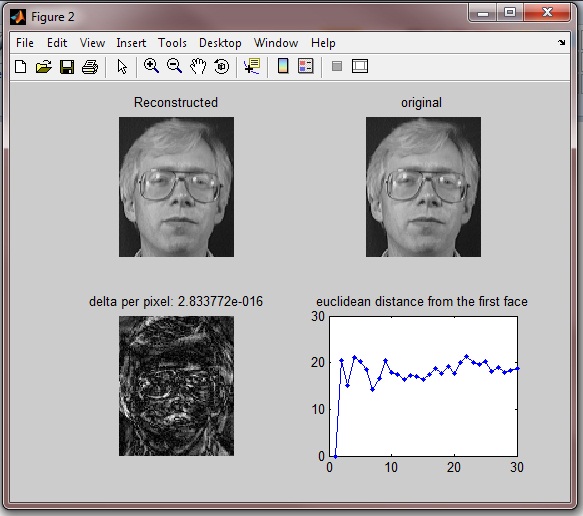
Trong figure 1, ta có 2 eigenface được tổng hợp và khuôn mặt trung bình (mean face hay average face):



* *Tiến hành thử nghiệm:*

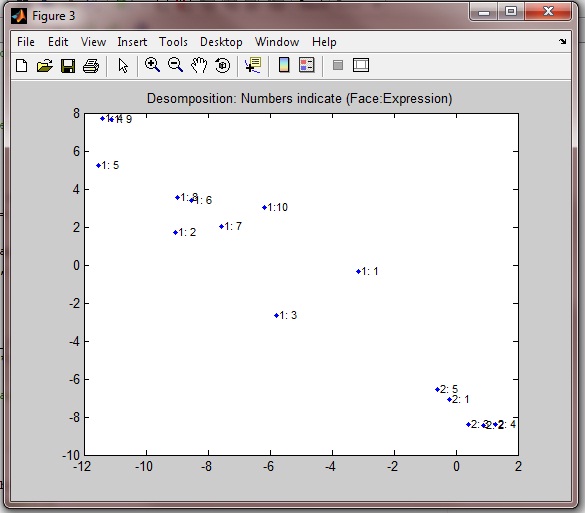
Đưa 1 ảnh test vào kiểm tra. Với chương trình này, ảnh test được chọn với chỉ số lưu vào biến ‘image\_index’ thay đổi trực tiếp trong code. Trong figure 2, ảnh cần nhận dạng là ảnh original.

Sau khi tiến hành so sánh xác định được face class khuôn mặt trong ảnh kiểm thử, nhờ eigenface, ta tái tạo được ảnh reconstructed:



Biểu đồ dưới đây là độ tương đồng khai triển, với dạng : Face:Expression

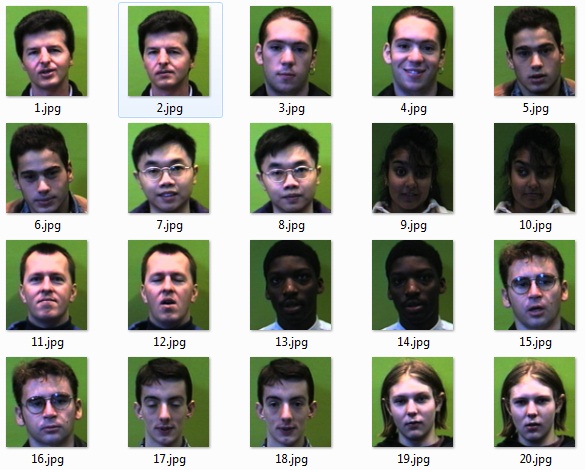
VD: 1:2 có nghĩa là ảnh được lấy từ s1/2.pgm , là biểu cảm 2 của người trong tập 1.

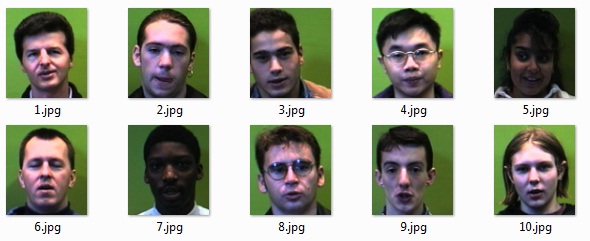


## Chương trình nhận dạng khuôn mặt PCA-based

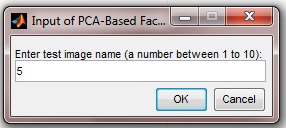
Sử dụng lý thuyết phương pháp Eigenface tương tự, nhưng chương trình sau dựa trên PCA.

Tập đào tạo (training set) là những ảnh trong cơ sở dữ liệu để sinh ra các eigenface như sau:

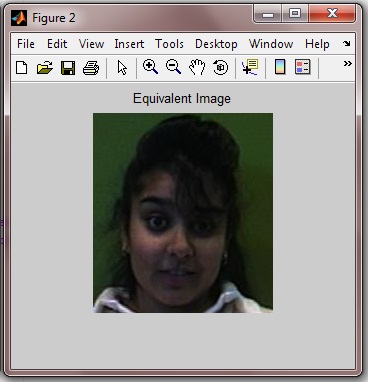
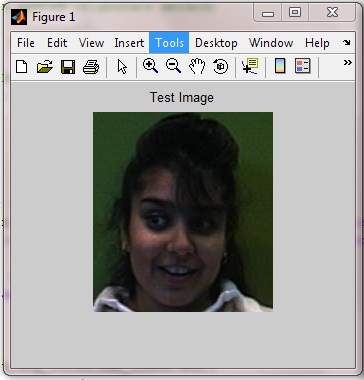


***Tập các giá trị kiểm thử được đưa ra để nhận dạng như sau:***Trong chương trình, chạy file example.m, trước hết, ta cần lựa chọn vị trí đặt tập đào tạo (training set) và sau đó là tập kiểm thử (test set).

Sau đó, ta chọn 1 ảnh trong số các ảnh của tập kiểm thử để nhận dạng:



Chương trình sẽ đưa ra hình ảnh trong tập đào tạo tương đồng với người trong ảnh kiểm thử:



# Tài liệu tham khảo

1. *PGS. TS. Nguyễn Hoàng Lan,* –Viện Công nghệ Thông tin và Truyền thông. Đại học Bách khoa Hà Nội, “***Bài giảng Xử lý ảnh***” .
2. TS. Lương Mạnh Bá, TS. Nguyễn Thanh Thủy – Viện Công nghệ Thông tin và Truyền thông. Đại học Bách Khoa Hà Nội, ***“Nhập môn xử lý ảnh số”.***
3. *R.K.Kouassi, J.C.Devaux, P.Gouton, M.Paindavoine* – Laboratoire d’Electronique, d’Informatique et d’Image. Université de Bourgogne, “***Application of the Karhunen-Loeve Transform for Natural Color Images Analysis”***
4. *Shang-Hung Lin, Ph.D* – IC Media Corporation, “***An Introduction to Face Recognition Technology”*** – Informing Science Special Issue on Multimedia Informing Technologies – Part 2, Volume 3 No 1, 2000.
5. *M.Turk, A. Pentland* –Vision and Modeling Group. The Media laboratory. MIT – “***Eigenfaces for Recognition”*** –Jounal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.
6. *Wen-Shiung Chen, Yung-Lin Lin and Jin-Chan Lio* – VIP-CCLab., Dept. of Electrical Engineering, National Chi Nan University; *Sheng-Wen Shih* – Dept. of Computer Science and Information Engineering, National Chi Nan University, *”****Automatic Iris Recognition Technique based on Divider Dimension and Karhunen-Loeve Transform”*** – Proc. of Computer Vision, Graphics and Image Processing, Taiwan, 2002, 78--85.
7. Anthony Giordano & Michael Uhrig, Regis University, Denver, Colorado, ***“Human Face Recognition Technology Using the Karhunen-Loéve Expansion Technique”***- Rose-Hulman Undergraduate Mathematics Journal ::: Vol. 7, Issue 1, 2006.
8. Kyungnam Kim, Department of Computer Science, University of Maryland, College Park, ***“Face recognition using Principal Components Analysis”***
9. Matlab Central : [www.mathworks.com/mathlabcentral/](http://www.mathworks.com/mathlabcentral/)

# Phân công thực hiện

|  |  |  |
| --- | --- | --- |
| **Sinh viên** | **MSSV** | **Công việc** |
| ***Lưu Thị Thùy Nhung***  [Nhung.luu@gmail.com](mailto:Nhung.luu@gmail.com) | 2007 2167 | Lý thuyết và ứng dụng của biến đổi KL. Khảo sát các ứng dụng của KL, phương pháp Eigenface. |
| ***Lê Quang Hiếu***  [Hieulq89@gmail.com](mailto:Hieulq89@gmail.com) | 2007 1095 | Tìm hiểu lý thuyết phương pháp nhận dạng khuôn mặt của KL và PCA áp dụng vào cài đặt 2 chương trình matlab. |
| ***Lê Ngọc Minh***  [Ngocminh.oss@gmail.com](mailto:Ngocminh.oss@gmail.com) | 2007 1946 | Tìm hiểu lý thuyết và khảo sát ứng dụng của biến đổi PCA vào nhận dạng khuôn mặt Eigenface. |