ĐẠI HỌC QUỐC GIA HÀ NỘI

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

****

**LUẬN VĂN TỐT NGHIỆP**

Đề tài: Tìm hiểu và xây dựng ứng dụng nhận dạng khuôn mặt

Họ và tên: Đặng Khắc Toàn

Mã sinh viên: 14001058

Khoa: Toán cơ tin

Chuyên ngành: Máy tính và khoa học thông tin

Giáo viên hướng dẫn: TS. Nguyễn Bích Thủy

Hà nội 2018

**Lời cảm ơn**

**Mục lục**

[Chương I: Tổng quan về xử lý ảnh và bài toán nhận dạng khuôn mặt 4](#_Toc510644213)

[1.1 Khái quát về xử lý ảnh 4](#_Toc510644214)

[1.1.1 Xử lý ảnh là gì? 4](#_Toc510644215)

[1.1.2 Một số khái niệm cơ bản trong xử lý ảnh 4](#_Toc510644216)

[1.2 Giới thiệu bài toán nhận dạng khuôn mặt 5](#_Toc510644217)

[1.2.1 Tổng quan về nhận dạng khuôn mặt 5](#_Toc510644218)

[1.2.1 Một số phương pháp nhận dạng khuôn mặt hiện nay 6](#_Toc510644219)

[1.3.1 Nhận dạng khuôn mặt dựa vào phương pháp biểu diễn thưa 7](#_Toc510644220)

[Chương II: Thực nghiệm bài toán 12](#_Toc510644221)

[1.1 Bài toán 12](#_Toc510644222)

[1.2 Phân tích, thiết kế chương trình 12](#_Toc510644223)

[1.2.1 Đọc tập training và tạo từ điển training 13](#_Toc510644224)

# Chương I: Tổng quan về xử lý ảnh và bài toán nhận dạng khuôn mặt

## Khái quát về xử lý ảnh

(khoang 50 -60) trang

### Xử lý ảnh là gì?

Xử lý ảnh là một loạt các thao tác và phân tích ảnh bằng máy tính nhằm cải  
thiện chất lượng ảnh cho tốt hơn và xử lý dữ liệu tự động trên máy. Quá trình này  
được xem như là thao tác ảnh đầu vào nhằm cho ra kết quả mong muốn. Kết quả  
đầu ra của quá trình sẽ là một ảnh tốt hơn hoặc một kết luận.

Ảnh có thể xem là tập hợp các điểm ảnh và mỗi điểm ảnh được xem như là  
đặc trưng cường độ sáng hay một dấu hiệu nào đó tại một vị trí xác định của đối tượng  
trong không gian và nó có thẻ xem như một hàm n biến P(c1, c2, c3,..., cn). Do đó,  
ảnh trong xử lý ảnh có thể xem như ảnh n chiều.

Sơ đồ tổng quát của một hệ thống xử lý ảnh:

Thu nhận ảnh

Hậu xử lý

Tiền xử lý

Đối sánh, rút ra kết quả

Tích chọn đặc trưng

### Một số khái niệm cơ bản trong xử lý ảnh

#### Điểm ảnh và ảnh

Gốc của ảnh là ảnh liên tục về không gian và độ sáng. Để xử lý bằng máy  
tính,ảnh cần phải được số hoá. Số hoá ảnh là sự biến đổi gần đúng một ảnh liên tục  
thành một tập điểm phù hợp với ảnh thật về vị trí (không gian) và độ sáng (mức  
xám). Khoảng cách giữa các điểm ảnh đó được thiết lập sao cho mắt người không  
phân biệt được ranh giới giữa chúng. Mỗi một điểm như vậy gọi là điểm ảnh  
(PEL:Picture Element) hay gọi tắt là Pixel. Trong khuôn khổ ảnh hai chiều, mỗi  
pixel ứng với cặp tọa độ (x,y).

Điểm ảnh (Pixel) là một phần tử của ảnh số tại toạ độ (x, y) với độ xám hoặc  
màu nhất định.

Ảnh là tập hợp của các điểm ảnh.

#### Độ phân giải của ảnh

Khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy  
được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ  
phân bố, đó chính là độ phân giải và được phân bố theo trục x và y trong không gian  
hai chiều. Vậy độ phân giải của ảnh là mật độ điểm ảnh được ấn định trên một ảnh  
số được hiển thị.

#### Mức xám của ảnh

Mức xám của điểm ảnh là cường độ sáng của nó được gán bằng giá trị số tại  
điểm đó. Các thang giá trị mức xám thông thường là: 16, 32, 64, 128, 256. Mức xám  
dùng 1 byte biểu diễn: 28=256 mức, tức là từ 0 đến 255.  
Ảnh đen trắng: là ảnh có hai màu đen, trắng (không chứa màu khác) với mức  
xám ở các điểm ảnh có thể khác nhau.

Ảnh nhị phân: ảnh chỉ có 2 mức đen trắng phân biệt tức dùng 1 bit mô tả 21  
mức khác nhau. Nói cách khác mỗi điểm ảnh của ảnh nhị phân chỉ có thể là 0 hoặc  
1

#### Trích chọn đặc trưng

Các đặc trưng của đối tượng được trích chọn tuỳ theo mục đích nhận dạng  
trong quá trình xử lý ảnh. Có thể nêu ra một số đặc trưng của ảnh sau đây:  
Đặc trưng không gian: Phân bố mức xám, phân bố xác suất, biên độ, điểm  
uốn v.v..

Đặc trưng biến đổi: Các đặc điểm loại này được trích chọn bằng việc thực  
hiện lọc vùng (zonal filtering). Các bộ vùng được gọi là “mặt nạ đặc trưng” (feature  
mask) thường là các khe hẹp với hình dạng khác nhau (chữ nhật, tam giác, cung  
tròn v.v..)

Đặc trưng biên và đường biên: Đặc trưng cho đường biên của đối tượng và  
do vậy rất hữu ích trong việc trích trọn các thuộc tính bất biến được dùng khi nhận  
dạng đối tượng. Các đặc trưng này có thể được trích chọn nhờ toán tử gradient, toán  
tử la bàn, toán tử Laplace, toán tử sobel, toán tử canny v.v..

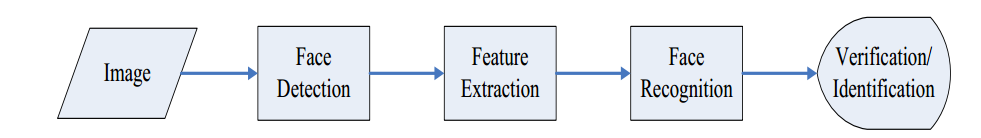
Việc trích chọn hiệu quả các đặc trưng giúp cho việc nhận dạng các đối  
tượng ảnh chính xác, với tốc độ tính toán cao và dung lượng nhớ lưu trữ giảm  
xuống.

## Giới thiệu bài toán nhận dạng khuôn mặt

### Tổng quan về nhận dạng khuôn mặt

Nhận dạng khuôn mặt là một bài toán lâu đời và được nghiên cứu rộng rãi trong khoảng thời gian 30 năm trở lại đây. Bài toán nhận dạng khuôn mặt có thể áp dụng rộng rãi trong nhiều lĩnh vực khác nhau. Các ứng dụng liên quan đến nhận dạng khuôn mặt có thể kể như: hệ thống phát hiện tội phạm, hệ thống theo dõi nhân sự trong tổ chức, hệ thống tìm kiếm thông tin trên hình ảnh, video dựa trên nội dung,… Tuy nhiên bài toán nhận dạng cũng gặp nhiều khó khăn và thách thức ví dụ như hệ thống camera công cộng, chụp hình thì khuôn mặt bị che khuất một phần hay không thể hiện đầy đủ thông tin, ảnh chụp không chính diện hay chất lượng ảnh không thực sự tốt, điều kiện chiếu sáng không đủ, môi trường xung quanh gây nhiễu ảnh. Những yếu tố trên ảnh hưởng không nhỏ đến độ chính xác của các chương trình nhận dạng khuôn mặt.

Bài toán nhận dạng khuôn mặt là một bài toán khá phức tạp và phải trải qua một số bước chính dưới đây: Face detection, feature extraction, face recognition.



#### Face detection

Xác định khuôn mặt người (Face Detection) là một kỹ thuật máy tính để xác định các vị trí và các kích thước của các khuôn mặt người trong các ảnh bất kỳ (ảnh kỹ thuật số). Kỹ thuật này nhận biết các đặc trưng của khuôn mặt và bỏ qua những thứ khác, như: tòa nhà, cây cối, cơ thể, …

Xác định khuôn mặt được nghiên cứu và đặt trong nhiều giả thiết khác nhau, thường những bức ảnh thử nghiệm được chụp trong các môi trường bất kì với điều kiện chiếu sáng và độ nhiễu khác nhau. Một số nhân tố ảnh hưởng đến phát hiện khuôn mặt như góc chụp, biểu cảm khuôn mặt, sự biến dạng của khuôn mặt khi di chuyển camera, những yếu tố che khuất 1 phần khuôn mặt như kính, tóc, râu,… của đối tượng.

Có nhiều nghiên cứu tìm phương pháp xác định khuôn mặt người, từ ảnh xám đến ngày nay là ảnh màu. Dựa vào tính chất của các phương pháp xác định khuôn mặt người trên ảnh. Người ta chia làm bốn hướng tiếp cận chính. Ngoài bốn hướng này, nhiều nghiên cứu có khi liên quan đến không những một hướng tiếp cận mà có liên quan nhiều hơn một hướng chính

* **Hướng tiếp cận dựa trên tri thức**: Mã hóa các hiểu biết của con người về các loại khuôn mặt người thành các luật. Thông thường các luật mô tả quan hệ của các đặc trưng.
* **Hướng tiếp cận dựa trên đặc trưng không thay đổi**: Mục tiêu các thuật toán đi tìm các đặc trưng mô tả cấu trúc khuôn mặt người mà các đặc trưng này sẽ không thay đổi khi tư thế khuôn mặt, vị trí đặt thiết bị thu hình hoặc điều kiện ánh sáng thay đổi.
* **Hướng tiếp cận dựa trên so khớp mẫu**: Dùng các mẫu chuẩn của khuôn mặt người (các mẫu này được chọn lựa và lưu trữ) để mô tả cho khuôn mặt người hay các đặc trưng khuôn mặt (các mẫu này phải chọn làm sao cho tách biệt nhau theo tiêu chuẩn mà các tác giả định ra để so sánh). Các mối tương quan giữa dữ liệu ảnh đưa vào và các mẫu dùng để xác định khuôn mặt người.
* **Hướng tiếp cận dựa trên diện mạo**: Trái ngược hẳn với so khớp mẫu, các mô hình (hay các mẫu) được học từ một tập ảnh huấn luyện trước đó. Sau đó hệ thống (mô hình) sẽ xác định khuôn mặt người. Hay một số tác giả còn gọi hướng tiếp cận này là hướng tiếp cận theo phương pháp học.

### 1.2.1 Một số phương pháp nhận dạng khuôn mặt hiện nay

### 1.3.1 Nhận dạng khuôn mặt dựa vào phương pháp biểu diễn thưa

#### Tổng quát

Bài toán được xây dựng bằng cách xây dựng một tập k các lớp khuôn mặt, mỗi lớp bao gồm các khuôn mặt của một người nào đó. Tập các lớp này được gọi là tập dữ liệu học và vấn đề được đặt ra là khi cần nhận dạng một người nào đó ta phải xác định xem bức ảnh của người đó thuộc vào lớp khuôn mặt nào. Giả sử tập các lớp khuôn mặt là một mà trận mình sẽ tạm gọi là tập traning example A gồm m lớp và mỗi lớp n khuôn mặt, khi đó lớp A­­I = [vi,1,vi,2, ,vi,3, …,vi,n] Rm \* ni.

Tử một tập traning samples của lớp Ai, với mỗi ảnh y Rm của cùng một lớp sẽ xấp xỉ nằm trong mẫu tuyến tính kết hợp với khuôn mặt i:

y = i,1vi,1 + i,2vi,2 + i,3vi,3 + … + i,nvi,n ­­ (1)

Tuy nhiên lớp i là chưa biết, do vậy ta cần định nghĩa mà trận một ma trận A gồm tất cả k tập traning mỗi tập n phần tử.

A = [A1, A2, A3, …, Ak] = [v1,1, v1,2, v1,3, …,vk,nk]

Từ đây sự thể hiện tuyến tính của y có thể viết lại như sau:

y = Ax0 Rm (1)

Trong đó: x0 = [0, …, 0, i,1, i,2i,n, 0, …, 0] Rm là một vector hệ số mà tất cả là 0 ngoại trừ những hệ số của class Ai

Vấn đề đặt ra ở đây là nếu m > n thì phương trình có vô số nghiệm và x0 cần tìm không phải là duy nhất. Thông thường, vấn đề này được giải quyết bằng phương pháp chuẩn 2 l2:

= argmin||x||2

||x||2 =

Tuy nhiên giải pháp chuẩn 2 này thương mang đến số nghiệm khác không khá lớn và không tối ưu khi ứng dụng để giải phương trình (1). Có nghĩa là số phần tử khác không càng lớn sẽ tương đương với nhiều bức ảnh training thuộc các class khác nhau. Để giải quyết khó khăn này chúng ta sử dụng hàm chuẩn 0:

= argmin||x||0 (2)

Giải bài toán y = Ax0 với số phần tử khác 0 là nhỏ nhất

#### Bài toán biểu diễn thưa với chuẩn l1

Những phát triển gần đây trong lý thuyết biểu diễn thưa và lấy mẫu nén chỉ ra rằng nếu x0 được tìm thấy là đủ thưa bài toán l0là tương đương với bài toán l1

(l1): = argmin||x||1

||x||1 = |x1| + |x2| + ... + |xn|

Vấn đề này được giải quyết trong thời gian đa thức bởi phương pháp tuyến tính tiêu chuẩn. Thậm chí còn hiệu quả hơn những phương pháp có sẵn khi biết được x là rất thưa

1. **Phân lớp dựa vào biểu diễn thưa**

Ta đưa vào một ảnh y từ một lớp trong tập traning, trước tiên ta tính sự biết diễn thưa của x1

Với = argmin||x||1

||Ax - y||

Ý tưởng chính ở đây là những phần tử khác 0 của x1 sẽ liên kết với những cột của A từ 1 class i, và từ đó ta có thể dễ dàng chỉ định ảnh test y thuộc class nào. Tuy nhiên sự nhiễu và lỗi mô hình hóa dẫn tới việc một phần nhỏ phần tử khác 0 liên kết với nhiều hơn class. Dựa trên sự đại diện tổng quát, chúng ta có thể thiết kế nhiều phân loại để giải quyết tình trạng này. Để khai thác tốt cấu trúc tuyến tính, chúng ta phân lớp y dựa trên mức độ biểu diễn bằng các hệ số của các ảnh training tạo nên y.

Gọi i: Rn Rn là những hệ số đặc trưng của class i. Với x Rn, i(x) Rn là một vector là một vector mà những phần tử khác 0 là hệ số của x liên kết class i. Với những hệ số của i của class i ta có thể tính gần đúng y với = Ai(). Tiếp theo chúng ta có thể chỉ định y thuộc class nào có sự khác biệt nhỏ nhất giữa y và y1:

mini ri(y)= ||y - A i(). ||2

**Thuật toán 1:**

1: Input: ma trận traning

A = [A1, A2, …Ak] Rm\*n ứng với k class, ảnh test y Rm (Và hệ số )

2: Chuẩn hóa các cột của A với l2-norm

Ax = y

3: Giải bài toán l­1 min

= argmin||x||1

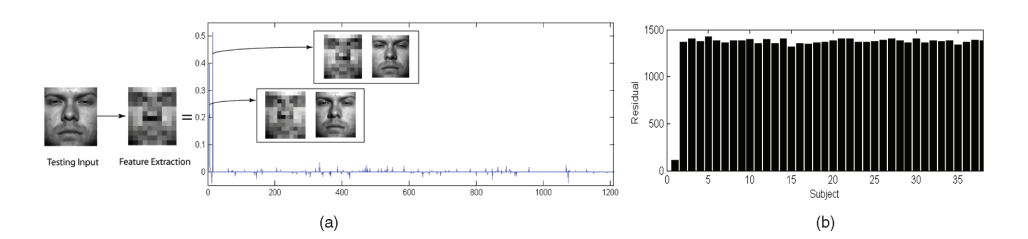
Hoặc:

||Ax - y||

= argmin||x||1

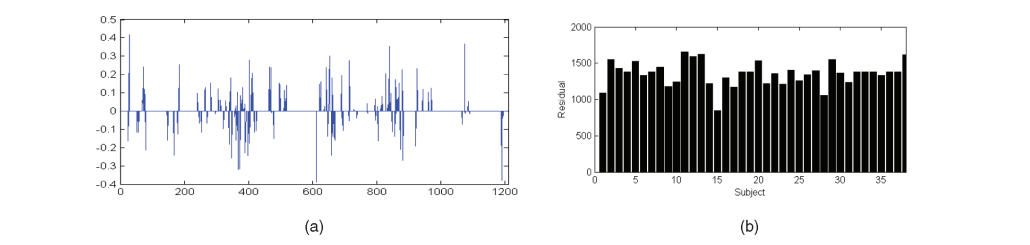
4: Tính ri(y) = ||y - A i(). ||2 với i = 1, …,k.

5: Xác định y = arg mini ri (y).



Hình 3: Nhận dạng ảnh 12\*10 pixels. Ảnh y thuộc về đối tượng 1. Giá trị của những hệ số thưa từ thuật toán 1 được thể hiện ở hình bên phải với 2 ảnh training ứng với 2 ảnh có hệ số thưa lớn nhất. Độ sai lệch ri của ảnh test được thể hiện ở ảnh b. Tỉ lệ giữa 2 hệ số nhỏ nhất là 1:8.6

Ví dụ 1: (So sánh giữa l1minization với l2minization). Để minh họa cho thuật toán 1, ta lựa chọn ngẫu nhiên 1 nửa trong số 2412 ảnh và xem chúng là tập traning cho một nửa còn lại. Ở ví dụ này, chúng ta giảm số chiều từ 192\*168 xuống12 \*10. Giá trị pixel của bức ảnh được sử dụng là 120 như là một cột của ma trận A trong thuật toán. Vì vậy, ma trận A có kích thước 120 \* 1207 và phương trình y = Ax là có vô số nghiệm. Hình 3 cho thấy những đặc trưng và những ảnh gốc tương ứng với hệ số lớn nhất. Hai hệ số lớn nhất là liên kết với 2 ảnh training từ đối tượng 1. Hình 3b chỉ ra độ sai khác i(), i = 1, …, 38. Với thuật toán 1 đạt được độ chính xác 92.1 % với Extended Yale B database



Hình 4: Hình a cho thấy những hệ số từ phương pháp l2minimization cùng sử dụng ảnh test như ở hình 3, kết quả thu được là những hệ số dày đặc, ít thông tin để nhận dạng. Độ sai khác ở ảnh b cũng không chênh lệch nhiều và tỉ lệ giữa 2 sai số nhỏ nhất là 1:1.3

1. **Xác nhận dựa vào biểu diễn thưa**

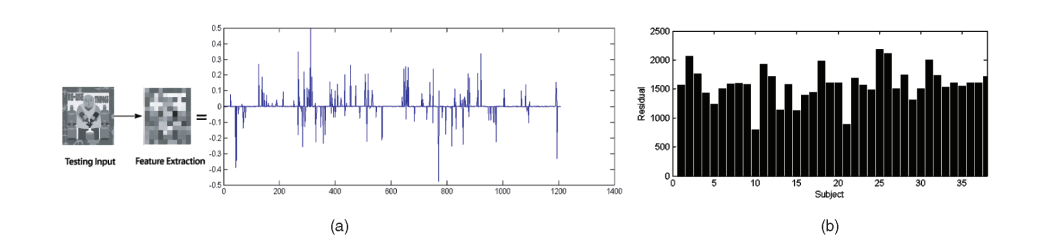
Trước khi phân lớp 1 ảnh test chúng ta đầu tiên cần xem xét nó có phải là 1 ảnh hợp lệ từ 1 trong những lớp data hay không. Khả năng để phát hiện ảnh là không hợp lệ cũng là một phần quan trọng trong hệ thống nhận dạng ở thế giới ngày nay. Với một hệ thống nhận dạng khuôn mặt, ví dụ chúng ta có thể đưa ra một bức ảnh của một đối tượng nào đó nhưng lại không có trong database hay một hình ảnh mà trong đó không chỉ có khuôn mặt mà còn nhiều thông tin gây nhiễu khác.

Hệ thống phân loại cơ bản thông thường như là NN hay NS thường sử dụng hệ số sai khác ri(y) để validation hơn thể nữa là để nhận dạng. Có nghĩa là thuật toán cho phép chấp nhận hay từ chối một ảnh test dựa trên mức độ sai khác. Tuy nhiên, mỗi hệ số sai khác ri(y) được tính mà không có bất cứ kiến thức gì về hình ảnh của đối tượng trong quá trình training mà nó chỉ là sự đo lường giữ các phép thử với các class.

Trong mô hình biểu diễn thưa, hệ số là được tính tổng thể trên toàn bộ các ảnh training. Với một ý nghĩa nào đó, nó có thể khai thác sự phân bố của lớp học để validate. Sự nghiên cứu cho rằng các hệ số của là sự thống kê tốt hơn cho việc validate hơn là sử dụng số dư.

Ví dụ 2: Sự tập trung của các hệ số thưa

Chúng ta sẽ chọn 1 ảnh bất kì không liên quan từ google và sau đó giảm số chiều xuống 12\*10. Sau đó ta tính sự biểu diễn thưa của bức ảnh dựa vào tập training



Hình 5: a) Hệ số biểu diễn thưa của 1 bức ảnh không hợp lệ dựa trên tập training A.

b) Số dư r của bức ảnh qua i() của các hệ số biểu diễn thưa dựa trên thuận toán l1 min. Tỉ lệ giữa 2 hệ số r nhỏ nhất là 1:1.2

So với hệ số của 1 hình ảnh hợp lệ (hình 3), những hệ số này không tập trung vào bất kì 1 đối tượng cụ thể nào mà trải rộng trên toàn tập training. Do vậy sự phân bố các hệ số của mang đến 1 phần thông tin quan trọng để xác định bức ảnh có hợp lệ hay không, 1 bức ảnh hợp lệ sẽ có những hệ số thưa mà những phần tử khác 0 tập trung chủ yếu vào 1 phần tử nào đó trong khi 1 bức ảnh không hợp lệ những hệ số này sẽ trải rộng trên toàn tập training.

Để định mức cho quan sát này ta định nghĩa một mức đo lường sự tập trung của các hệ số trên 1 class trong tập training.

Định nghĩa 1: Sparsity concentration index (SCI)

SCI(x) = [0, 1]

Đối với x trong thuật toán 1, nếu SCI(x) = 1, ảnh test được biểu diễn bởi duy nhất một ảnh của đối tượng, nếu SCI(x) = 0 hệ số của x là trải đều trên toàn bộ classes. Vì thế ta đặt ra 1 ngưỡng t [0, 1] nào đó và chấp nhận ảnh đó là hợp lệ hay không.

SCI(x)

Ở bước thứ 5 của thuật toán 1 ta đối sánh với tiêu chí trên để chấp nhận y có phải là ảnh của đối tượng nào đó hay không.

Thay vì dựa vào 1 sự thống kê của ri(y) để xác định tính hợp lệ của bức ảnh. Bằng phương pháp tiếp cận này ta đã tách thành 2 công việc riêng biệt: số dư r để nhận dạng và hệ số thưa để xác định tính hợp lệ.

Một trong những lợi ích của phương pháp này là làm tăng hiệu năng của chương trình tránh trường hợp 1 đối tượng là tương đương với nhiều class. Ví dụ trong hệ thống nhận dạng khi 1 khuôn mặt khá giống với 1 vài đối tượng và có hệ số r khá nhỏ tương đương với các classes đấy. Khi đó sử dụng số dư để xác thực sẽ dẫn đến sai lầm. Dựa vào đánh giá tính hợp lệ này mà ta có thể xem xét bức ảnh test có phải là một phần của tập training hay không.

#### Tích chọn đặc trưng

Nhiều phương pháp tích chọn đặc trưng được nghiên cứu để tìm ra cách biểu diễn tốt nhất cho mỗi classes với số chiều thấp hơn. Một số phương pháp tích chọn đặc trưng điển hình như: Eigen-faces, Fisherfaces, Laplacianfaces. Một số phương pháp các tích chọn những vừng có ý nghĩa đặc trưng của khuôn mặt như xung quanh vùng mắt và mũi. Thông thường trích chọn đặc trưng được sử dụng kết hợp với sự phân lớp như NN hay NS. Việc lựa chọn đặc trưng là một phần quan trọng để đạt được sự thành công của thuật toán. Điều này dẫn đến sự phát triển rộng lớn và nhanh chóng các phương pháp tích chọn đặc trưng phức tạp hơn bao gồm nonlinear và kernel features.

Lợi ích của tích chọn đặc trưng là giảm thông tin dữ liệu và chi phí tính toán nhưng vẫn mang lại hiệu xuất tốt nhất. Ví dụ như đối với ảnh gốc sự biểu diễn tuyến tính là y = Ax là rất lớn và vượt ra ngoài khả năng tính toán của máy tính thông thường. Sự ánh xạ từng không gian ảnh sang không gian đặc trưng có thể biểu diễn qua ma trận R Rd\*mvới d << m

= Ry = RAx0 ∈ Rd

Bây giờ ta giải quyết thuât toán l1min với

= arg min ||x1||

||RAx –||2 ≤

Với số dư ϵ > 0 và dĩ nhiên ma trận A ở thuật toán 1 bây giơ được thay bằng ma trận RA ∈ Rd\*nvới d là số chiều đặc trưng và ảnh test y được thay bằng ảnh test . Thực nghiệm chỉ ra rằng khi tăng số chiều đặc trưng d thì độ chính xác nhận dạng càng cao tuy nhiên điều đó cũng ảnh hưởng đến thời gian tính toán của chương trình. Việc giảm số chiều không phải là vấn đề của thuật toán l1 min, nó chỉ dòi hỏi ảnh test nằm trong hoặc gần với phạm vi của RAi

# Chương II: Thực nghiệm bài toán

## Bài toán

Với một bức ảnh đầu vào, bài toán đặt ra là chương trình làm thế nào biết được bức ảnh đấy thuộc đối tượng nào trong tập training. Để giải quyết bài toán này chương trình sẽ bao gồm một số công đoạn chính sau đây:

* Chuẩn bị tập dữ liệu để training và test
* Tích chọn đặc trưng, giảm số chiều của ảnh
* Sử dụng thuật toán biểu diễn thưa để nhận dạng đối tượng

## Phân tích, thiết kế chương trình

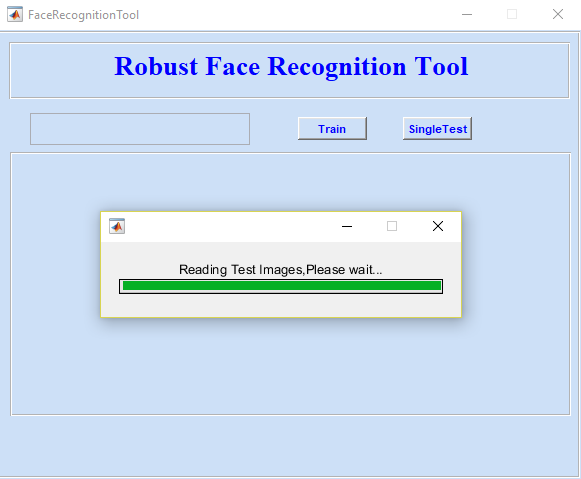
Chương trình được xây dựng bằng ứng dụng MATLAB 2016B với giao diện chính như sau: 

* Nút train: Đọc các ảnh đầu vào và lưu trữ dưỡi dạng ma trận A là bộ training.
* Nút SingleTest: Đọc ảnh test và trả về đối tượng được nhận dạng.

Chúng ta sẽ chuẩn bị tập dữ liệu bao gồm 400 ảnh xám khuôn mặt đã được tiền xử lý, một nửa số đó sẽ được dùng để tạo bộ training cho chương trình và nửa còn lại sẽ là tập test.



### Đọc tập training và tạo từ điển training



Đọc từng ảnh trong tập train và biến đổi thành ảnh xám nếu như chưa phải là ảnh xám:

if length(size(trainImage))==3

Tmp\_Image = rgb2gray(trainImage);

else

Tmp\_Image = trainImage;

end

Sau khi đã biến đổi được ảnh xám ta tiếp tục giảm số chiều của ảnh

imresize(Tmp\_Image,[m n])

Giả sử sau khi đọc ảnh xám đầu vào, giảm số chiều ta thu được ảnh I và biến đổi nó thành vector như sau:

=>

Khi đã đọc hết tất cả các ảnh ta thu được tập training A:

113 86 84 … 110

156 121 104 … 157

148 122 103 … 132

… ….

158 103 97 … 111

Chuẩn hóa từng vector của ma trận A ta được bộ từ điển:

0.18 0.14 0.14 … 0.19

0.25 0.20 0.18 … 0.27

0.23 0.20 0.18 … 0.23

…

0.25 0.17 0.17 … 0.19