

ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN  
**KHOA TOÁN - CƠ - TIN HỌC**

**Đặng Khắc Toàn**

**SỬ DỤNG PHƯƠNG PHÁP BIỂU DIỄN THỪA TRONG BÀI  
TOÁN NHẬN DẠNG KHUÔN MẶT**

Khóa luận tốt nghiệp đại học hệ chính quy  
Ngành máy tính và khoa học thông tin  
(Chương trình đào tạo chuẩn)

**Hà Nội - 2018**

ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN  
KHOA TOÁN - CƠ - TIN HỌC

**Đặng Khắc Toàn**

**SỬ DỤNG PHƯƠNG PHÁP BIỂU DIỄN THỪA TRONG BÀI  
TOÁN NHẬN DẠNG KHUÔN MẶT**

Khóa luận tốt nghiệp đại học hệ chính quy  
Ngành máy tính và khoa học thông tin  
(Chương trình đào tạo chuẩn)

**Cán bộ hướng dẫn: TS Nguyễn Thị Bích Thủy**

**Hà Nội - 2018**

### **Lời cảm ơn**

Trước tiên, tôi xin cảm ơn và gửi lòng biết ơn sâu sắc nhất tới Tiến sĩ Nguyễn Thị Bích Thủy người đã tận tình chỉ bảo và hướng dẫn tôi trong suốt quá trình thực hiện khóa luận tốt nghiệp.

Tôi xin chân thành cảm ơn các thầy, cô đã tận tình hướng dẫn, giảng dạy trong suốt quá trình học tập và rèn luyện tại trường Đại học Khoa Học Tự Nhiên – ĐHQGHN.

Tôi cũng xin gửi lời cảm ơn tới các anh chị, các bạn sinh viên lớp K59 Máy tính và khoa học thông tin đã giúp và hỗ trợ kiến thức những khi gặp khó trong quá trình học tập.

Cuối cùng, tôi muốn gửi lời cảm ơn vô hạn tới người thân, bạn bè luôn bên cạnh động viên tôi trong suốt quá trình thực hiện khóa luận.

Tôi xin chân thành cảm ơn!

Hà Nội, ngày 10 tháng 5 năm 2018

Sinh viên

Đặng Khắc Toàn

## **Tóm tắt**

Với sự phát triển nhanh chóng và mạnh mẽ trong lĩnh vực công nghệ thông tin hiện nay. Các ứng dụng ngày càng được cải tiến và đáp ứng nhu cầu người dùng một cách tốt nhất. Một trong những vấn đề đang được nghiên cứu và phát triển mạnh mẽ trong thời gian gần đây đó là nhận diện khuôn mặt. Tính năng nhận diện khuôn mặt được ứng dụng nhiều trong các hệ thống giám sát nhân viên, hệ thống điều tra tội phạm hay đơn giản chỉ là các tiện ích nhỏ trên những chiếc điện thoại thông minh.

Tuy nhiên để xây dựng được một hệ thống nhận dạng lại không hề đơn giản. Nhiều thuật toán và phương pháp đã được nghiên cứu và áp dụng nhằm đưa ra kết quả tốt nhất và khắc phục những khó khăn khi nhận dạng trong hoàn cảnh thực tế.

Mục tiêu của khóa luận là tìm hiểu các phương pháp, quy trình trong bài toán nhận dạng khuôn mặt hiện nay và ứng dụng phương pháp biểu diễn thưa để giải bài toán nhận dạng khuôn mặt.

Nội dung báo cáo gồm 4 chương:

- Chương I: Tổng quan về xử lý ảnh và bài toán nhận dạng khuôn mặt
- Chương II: Nhận dạng khuôn mặt dựa trên phương pháp biểu diễn thưa
- Chương III: Thực nghiệm bài toán
- Chương IV: Kết luận

## Mục lục

Chương I: Tổng quan về xử lý ảnh và bài toán nhận dạng khuôn mặt .....	8
1.1 Khái quát về xử lý ảnh.....	8
1.1.1 Một số khái niệm cơ bản trong xử lý ảnh.....	8
1.1.2 Xử lý ảnh là gì?.....	9
1.2 Giới thiệu bài toán nhận dạng khuôn mặt.....	11
1.2.1 Tổng quan về nhận dạng khuôn mặt.....	11
1.2.1 Một số phương pháp nhận dạng khuôn mặt hiện nay.....	13
Chương II: Nhận dạng khuôn mặt dựa vào phương pháp biểu diễn thưa .....	17
2.1 Giới thiệu .....	17
2.2 Ý tưởng .....	17
2.3 Bài toán biểu diễn thưa với chuẩn <b><math>l_1</math></b> .....	18
2.4 Trích chọn đặc trưng.....	23
2.5 Giải bài toán chuẩn 1 với phương pháp quy hoạch tuyến tính .....	24
2.5.1 PHÂN LOẠI DẠNG BÀI TOÁN QUY HOẠCH TUYẾN TÍNH.....	24
2.5.2 Thuật toán đơn hình giải bài toán min.....	25
2.5.3 Áp dụng thuật toán đơn hình để giải bài toán phân lớp .....	27
Chương III: Thực nghiệm bài toán.....	29
3.1 Mô tả bài toán .....	29
3.2 Phân tích, thiết kế chương trình .....	29
3.2.1 Xây dựng từ điển .....	30
3.2.2 Nhận dạng 1 ảnh .....	30
3.2.3 Đánh giá kết quả nhận dạng trên toàn bộ tập test.....	31
3.3 Kết quả.....	32
3.3.1 Nhận dạng với ảnh xám.....	32
3.3.2 Nhận dạng với độ chiếu sáng thấp.....	32
3.3.3 Nhận dạng với góc quay khuôn mặt khác nhau.....	33
3.3.4 Nhận dạng với vị trí khuôn mặt khác nhau .....	33
3.3.5 Nhận dạng với khuôn mặt bị che.....	34
3.3.6 Nhận dạng với khuôn mặt có các biểu cảm hay góc nghiêng khác nhau .....	34

3.3.7 Thời gian và độ chính xác trên tập ảnh test .....	35
Chương IV: Kết luận .....	36
4.1 Các kết quả đạt được.....	36
4.2 Hướng phát triển tiếp theo .....	36
Tài liệu tham khảo .....	37

### Danh sách các kí hiệu và từ viết tắt

STT	Từ viết tắt	Từ viết đầy đủ
1	CSDL	Cơ sở dữ liệu
2	PCA	Principal component analysis
3	LDA	Linear discriminant analysis
4	PEL	Picture Element
5	PA	Phương án
6	PATU	Phương án tối ưu
7	QHTT	Quy hoạch tuyến tính

### Danh sách tham chiếu thuật ngữ Anh - Việt

STT	Thuật ngữ tiếng Anh	Thuật ngữ tiếng Việt
1	Principal component analysis	Phân tích thành phần chính
2	Linear discriminant analysis	Phân tích sự khác biệt tuyến tính
3	Test	Kiểm tra
4	Train	Đào tạo
5	Resolution	Độ phân giải
6	Picture Element	Điểm ảnh
7	Class	Lớp
8	Training samples	Mẫu đào tạo
9	Zonal filtering	Lọc vùng
10	Feature mask	Mặt nạ đặc điểm
11	Zero crossing	Chéo không
12	Face detection	Phát hiện khuôn mặt

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

## 1.1 Khái quát về xử lý ảnh

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

#### a. Điểm 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 hay gọi tắt là pixel. Trong khuôn khổ ảnh hai chiều, mỗi pixel ứng với cặp tọa độ  $(x, y)$ .

**Định nghĩa:** Điểm ảnh (Pixel) là một phần tử của ảnh số tại tọa độ  $(x, y)$  với độ xám hoặc màu nhất định. Kích thước và khoảng cách giữa các điểm được chọn thích hợp sao cho mắt người cảm nhận được sự liên tục về không gian và mức xám (hoặc màu) của ảnh số gần như là ảnh thật. Mỗi phần tử trong ma trận được gọi là một phần tử ảnh.

#### b. Độ phân giải của ảnh

**Định nghĩa:** Độ phân giải (Resolution) của ảnh là mật độ điểm ảnh được ấn định trên ảnh số được hiển thị.

Theo định nghĩa, khoảng cách giữa các điểm ảnh phải 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.

Ví dụ: Độ phân giải của màn hình CGA (Color Graphic Adaptor) là một lưới điểm theo chiều ngang màn hình: 320 điểm chiều dọc \* 200 điểm ảnh chiều ngang (320\*200). Rõ ràng, cùng màn hình CGA 12" ta thấy mịn hơn màn hình CGA 17" độ phân giải 320\*200. Lý do: cùng một mật độ (độ phân giải) nhưng diện tích màn hình rộng hơn thì độ mịn (liên tục của các điểm) kém hơn.

#### c. Mức xám của ảnh

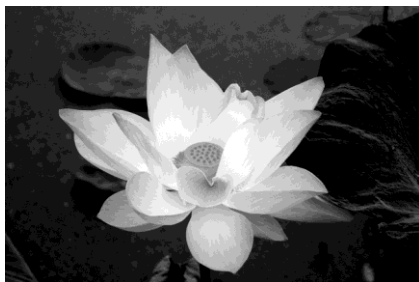
Một điểm ảnh (pixel) có hai đặc trưng cơ bản là vị trí  $(x, y)$  của điểm ảnh và độ xám của nó. Dưới đây chúng ta xem xét một số khái niệm và thuật ngữ thường dùng trong xử lý ảnh.

**Định nghĩa:** 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:** 16, 32, 64, 128, 256 (Mức 256 là mức phổ dụng. Lý do: từ kỹ thuật máy tính dùng 1 byte (8 bit) để biểu diễn mức xám: Mức xám dùng 1 byte biểu diễn:  $2^8 = 256$  mức, tức là từ 0 đến 255).

**Ảnh xám:** 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  $2^1$  mức khác nhau. Nói cách khác: mỗi điểm ảnh của ảnh nhị phân chỉ có thể là 0 hoặc 1.



**Ảnh màu:** Người ta thường dùng 3 byte để mô tả mức màu, khi đó các giá trị màu:  $2^{8*3} = 2^{24} = 16.7$  triệu màu.



#### d. Định nghĩa ảnh số

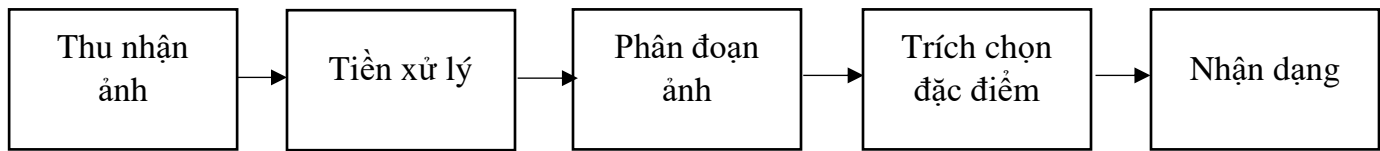
Ảnh số là một biểu diễn của ảnh 2 chiều bao gồm tập hữu hạn các giá trị số (các pixel hay các thành phần của ảnh). Mỗi pixel có 3 giá trị (ảnh màu) hoặc 1 giá trị (ảnh đơn sắc).

##### 1.1.2 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 trên 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 thể xem như một hàm n biến  $P(c_1, c_2, c_3, \dots, c_n)$ . 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:



### a. Thu nhận ảnh

Để thực hiện được quá trình đầu tiên trong hệ thống xử lý ảnh, ta cần sử dụng các thiết bị thu nhận ảnh để chuyển các thông tin dưới dạng hình ảnh thành các cấu trúc lưu trữ được trong máy tính và được hiển thị ra màn hình, máy in, ...

Ảnh có thể thu nhận từ vệ tinh qua các bộ cảm ứng, qua camera, máy chụp ảnh đơn sắc (màu), hay các tranh, ảnh được quét trên máy quét ảnh. Nếu ảnh thu nhận được chưa phải là dạng số hóa ta phải chuyển đổi hay số hóa ảnh trước khi chuyển sang giai đoạn tiếp theo.

### b. Tiền xử lý

Sau khi được thu nhận bởi các thiết bị thu nhận ảnh, ảnh sẽ được cải thiện về độ tương phản, nhiễu, ... bởi các kỹ thuật xử lý ảnh để làm ảnh tốt hơn theo mục đích sử dụng nhằm phục vụ cho quá trình xử lý tiếp theo.

Một số tiến trình trong tiền xử lý là:

- Điều chỉnh độ chiếu sáng: Khắc phục hậu quả của sự chiếu sáng không đồng đều.
- Khử nhiễu: Nhiễu được chia làm hai loại cơ bản là nhiễu hệ thống và nhiễu ngẫu nhiên. Trong đó, nhiễu hệ thống là nhiễu có quy luật có thể khử bằng các phép biến đổi Fourier và loại bỏ các đỉnh điểm. Đối với nhiễu ngẫu nhiên vết bản không rõ nguyên nhân thì được khắc phục bằng các phép lọc (lọc trung bình, lọc trung vị, ...).
- Hiệu chỉnh mức xám: Có thể tăng hay giảm số mức xám nhằm khắc phục tính không đồng bộ gây nên từ hiệu ứng của thiết bị thu nhận hình ảnh hoặc độ tương phản giữa các vùng ảnh.
- Chuẩn hóa độ lớn, hình dạng và màu sắc.
- Nắn chỉnh hình học: Ảnh thu nhận thường bị biến dạng do các thiết bị quang học và điện tử, để khắc phục điều này người ta sử dụng các phép chiếu được xây dựng trên tập các điểm điều kiện.

### c. Phân đoạn ảnh

Phân đoạn ảnh là một quá trình thao tác ở mức thấp trong toàn bộ hệ thống xử lý ảnh. Quá trình này thực hiện việc phân vùng ảnh thành các vùng rời rạc và đồng nhất với nhau hay nói cách khác là xác định các biên của các vùng ảnh đó. Phân đoạn ảnh là chia ảnh thành các vùng không trùng lặp, mỗi vùng gồm một nhóm điểm ảnh liên thông và đồng nhất theo một tiêu chí nào đó.

Ví dụ: Đồng nhất về màu sắc, mức xám, kết cấu, độ sâu của các lớp, ... Sau khi phân đoạn mỗi điểm ảnh chỉ thuộc về một vùng duy nhất.

Để đánh giá chất lượng của quá trình phân đoạn là rất khó, vì vậy cần phải xác định rõ mục tiêu của quá trình phân đoạn là gì? Ví dụ: để nhận dạng chữ (mã vạch) trên phong bì thư cho mục đích phân loại bưu phẩm, cần chia các câu chữ về địa chỉ hoặc tên người gửi thành các từ, các chữ, các số (các vạch) riêng biệt để nhận dạng.

Kết quả của quá trình phân đoạn ảnh thường được cho dưới dạng dữ liệu điểm ảnh thô, trong đó hàm chứa biên của một vùng ảnh hoặc tập hợp tất cả các điểm ảnh thuộc về chính vùng ảnh đó. Trong cả hai trường hợp, sự chuyển đổi dữ liệu thô này thành một dạng thích hợp hơn cho việc xử lý trong máy tính là hết sức cần thiết, nghĩa là nên biểu diễn một vùng ảnh dưới dạng tập hợp các đường biên hay một vùng hoàn chỉnh gồm tất cả những điểm ảnh thuộc về nó.

Biểu diễn dạng biên cho một vùng phù hợp với những ứng dụng chỉ quan tâm đến các đặc trưng hình dạng bên ngoài của đối tượng, ví dụ như các góc cạnh và điểm uốn trên biên.

Biểu diễn dạng vùng lại thích hợp cho những ứng dụng khai thác các tính chất bên trong của đối tượng. Ví dụ như vân ảnh hoặc cấu trúc xương của nó. Và trong một số ứng dụng thì cả hai cách biểu diễn trên đều cần thiết.

#### **d. Trích chọn đặc điểm**

Dựa trên các thông tin thu nhận được qua quá trình phân đoạn, kết hợp với các kỹ thuật xử lý, thủ tục phân tích dữ liệu để đưa ra các đặc điểm đặc trưng, đối tượng ảnh cũng như các thông tin cần thiết trong quá trình xử lý. Nhờ đó việc nhận dạng các đối tượng ảnh chính xác hơn, tốc độ tính toán cao và dung lượng lưu trữ giảm xuống.

Các đặc điểm của đối tượng được trích chọn tùy theo mục đích nhận dạng trong hệ thống xử lý ảnh. Sau đây là một vài đặc điểm của ảnh:

- Đặc điểm không gian: Phân bố mức xám, phân bố xác suất, biên độ, điểm uốn v.v...
- Đặc điểm 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 điểm” (feature mask) thường là các khe hẹp với hình dạng khác nhau (chữ nhật, tam giác, cung tròn v.v..).
- Đặc điểm biên và đường biên: Đặc trưng cho đường biên của đối tượng và do vậy rất hữu ích trong việc trích chọn các thuộc tính bất biến được dùng khi nhận dạng đối tượng. Các đặc điểm này có thể được trích chọn nhờ toán tử gradient, toán tử la bàn, toán tử Laplace, toán tử “chéo không” (zero crossing), ...

#### **e. Nhận dạng**

Nhận dạng ảnh là quá trình cuối cùng của hệ thống xử lý ảnh – quá trình liên quan đến các mô tả đối tượng mà người ta muốn đặc tả nó. Quá trình này thường đi sau quá trình trích chọn đặc điểm trong hệ thống xử lý ảnh.

## **1.2 Giới thiệu bài toán nhận dạng khuôn mặt**

### **1.2.1 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 một số vấn đề dưới đây:

- **Vấn đề về hướng (pose variations):** Các kết quả với các ảnh có hướng thay đổi (lớn hơn 45 độ, không phải chính diện) còn khá khiêm tốn. Để giải quyết vấn đề này người ta đã đưa ra giải pháp nhận dạng trên bộ dữ liệu 3D tuy nhiên để xây dựng hệ thống nhận dạng 3D lại khá phức tạp và tốn kém vì thế giải pháp này chưa được áp dụng nhiều và rộng rãi như hệ thống nhận dạng sử dụng dữ liệu 2D.
- **Vấn đề ảnh có độ phân giải thấp (low resolution):** Ảnh thu được từ các camera giám sát thường có kích thước và chất lượng rất thấp, do đó ảnh hưởng khá lớn đến kết quả trong các hệ thống nhận dạng. Những kết quả nghiên cứu về lĩnh vực này còn chưa nhiều.
- **Làm việc với dữ liệu video (video based face recognition):** Với sự phát triển của các phương tiện truyền thông, thông tin mặt người trong các dữ liệu video là vô cùng nhiều, tuy nhiên hầu hết các phương pháp nhận dạng vẫn làm việc với ảnh tĩnh trích xuất từ dữ liệu video, chưa có phương pháp tận dụng hết ưu thế của dữ liệu video. Với việc nhận dạng dựa trên video đòi hỏi tốc độ xử lý của chương trình khá cao. Việc xây dựng chương trình hay thuật toán để nhận dạng từ dữ liệu video phức tạp hơn nhiều so với dữ liệu ảnh tĩnh.
- **Các hệ thống lớn (very large scale systems):** Các cơ sở dữ liệu ảnh mặt được test bởi các nhà nghiên cứu còn khá nhỏ (vài trăm tới vài chục nghìn ảnh mặt), tuy nhiên trên thực tế các CSDL có thể rất lớn, ví dụ cơ sở dữ liệu ảnh khuôn mặt của cảnh sát của một nước có thể chứa từ hàng triệu tới hơn 1 tỉ ảnh ...

Bài toán nhận dạng khuôn mặt là một bài toán khá phức tạp và phải trải qua một số bước chính dưới đây: Tiền xử lý (Pre-Processing), xác định khuôn mặt (face detection), tích chọn đặc trưng (feature extraction), nhận dạng (face recognition).

#### a. Tiền xử lý

Mục đích chính của bước này là để chuẩn hóa ảnh cần tìm giúp cho việc tìm kiếm được hiệu quả hơn. Các công việc trong bước tiền xử lý có thể là: chuẩn hóa kích cỡ giữa ảnh trong cơ sở dữ liệu và ảnh cần tìm, hiệu chỉnh độ sáng - tối của ảnh, lọc nhiễu, chuẩn hóa về vị trí, tư thế ảnh mặt.

#### b. Phát hiện khuôn mặt

Phát hiện khuôn mặt người (Face detection) là một kỹ thuật máy tính để xác định các vị trí và kích thước của khuôn mặt người trong những 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, ...

Phát hiện 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 kỳ 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 khách quan làm che khuất một phần khuôn mặt như kính, tóc, râu, ... của đối tượng.

Có nhiều nghiên cứu tìm phương pháp phát hiện khuôn mặt người từ ảnh xám đến ngày nay là ảnh màu. Dựa vào tính chất của các phương pháp phát hiện 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.

### c. Trích chọn đặc trưng

Trích chọn đặc trưng là một trong những bước quan trọng trong bài toán nhận dạng. Độ chính xác và tốc độ của chương trình phụ thuộc lớn vào kết quả trích chọn đặc trưng từ bức ảnh khuôn mặt. Ở bước này nhằm tìm ra các đặc trưng chính của ảnh mặt, từ các đặc trưng này hình thành các vector đặc trưng, các vector này sẽ được sử dụng để đối sánh sự giống nhau giữa ảnh mặt cần tìm và ảnh mặt trong cơ sở dữ liệu.

Một số phương pháp trích chọn đặc trưng nổi bật có thể kể đến như: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), ...

### d. Nhận dạng

Thực hiện việc so sánh giữa các vector đặc trưng để chọn ra độ tương tự giữa ảnh cần tìm và ảnh trong cơ sở dữ liệu.

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

Phương pháp nhận dạng hiện nay có 2 loại:

- Nhận dạng dựa trên các đặc trưng của các phần tử trên khuôn mặt (Feature based face recognition).
- Nhận dạng dựa trên xét tổng thể khuôn mặt (Appearance based face recognition).

Ngoài ra còn có một số loại nhận dạng sử dụng mô hình về khuôn mặt, một số phương pháp được dùng cho loại này là nhận dạng 2D (Elastic Bunch Graph, Active Appearance Model) và 3D (3D Morphable Model). Việc so khớp mẫu 2D đang được áp dụng rộng rãi tuy nhiên ngày nay người ta đang dần thay thế và phát triển dựa trên phương pháp so khớp mẫu 3D do độ chính xác dựa trên so khớp mẫu 2D bị ảnh hưởng nhiều bởi sự thay đổi về hướng và độ chiếu sáng. Ưu điểm vượt trội khi so khớp mẫu 3D đó là khắc phục được sự ảnh hưởng của sự thay đổi về hướng và

chiếu sáng. Tuy nhiên, để ứng dụng phương pháp này thì các dữ liệu phải được thu từ camera 3D việc này đòi hỏi quá trình xây dựng dữ liệu và thu nhận ảnh test phức tạp hơn nhiều so với phương pháp 2D. Để giải quyết vấn đề này người ta thường xây dựng mô hình 3D từ các dữ liệu hình ảnh 2D.

### **a. Nhận dạng dựa trên các đặc trưng khuôn mặt:**

Đây là phương pháp nhận dạng khuôn mặt dựa trên việc xác định các đặc trưng hình học của các chi tiết trên một khuôn mặt (như vị trí, diện tích, hình dạng của mắt, mũi, miệng, ...), và mối quan hệ giữa chúng (như khoảng cách của hai mắt, khoảng cách của hai lông mày, ...).

Ưu điểm của phương pháp này là nó gần với cách mà con người sử dụng để nhận biết khuôn mặt. Hơn nữa với việc xác định đặc tính và các mối quan hệ, phương pháp này có thể cho kết quả tốt trong các trường hợp ảnh có nhiễu, bị nghiêng, bị xoay hoặc ánh sáng thay đổi.

Nhược điểm của phương pháp này là cài đặt thuật toán phức tạp do việc xác định mối quan hệ giữa các đặc tính với nhau không dễ dàng. Mặt khác, với các ảnh kích thước bé thì các đặc tính sẽ khó phân biệt do đó khó tính toán được các đặc trưng trên khuôn mặt.

### **b. Nhận dạng dựa trên xét toàn diện khuôn mặt**

Nhận dạng dựa trên toàn bộ khuôn mặt có nghĩa là chúng ta sẽ không xét đến từng thành phần đặc trưng trên khuôn mặt nữa mà sẽ xem khuôn mặt là một không gian cụ thể và tìm những đặc trưng chính trên không gian đó.

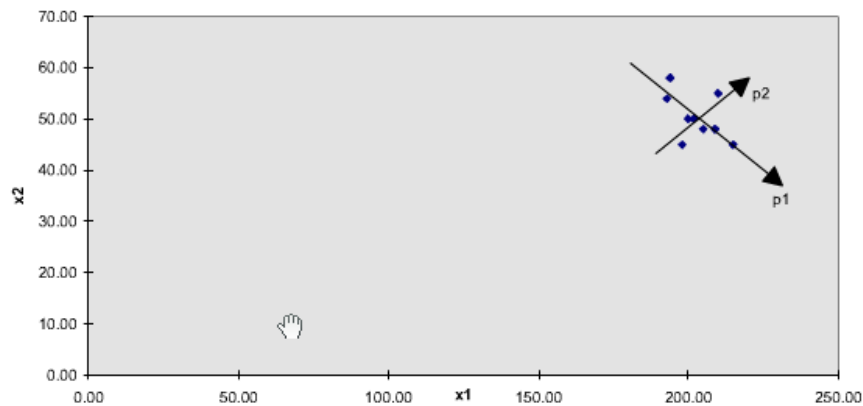
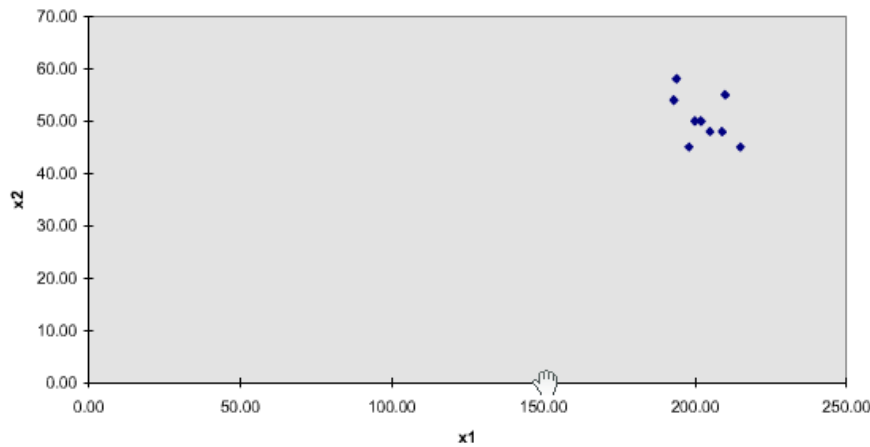
Nội dung chính của hướng tiếp cận này là xem mỗi ảnh có kích thước  $R * C$  là một vector trong không gian  $R * C$  chiều. Ta sẽ xây dựng một không gian mới có chiều nhỏ hơn sao cho khi biểu diễn trong không gian đó các đặc điểm chính của khuôn mặt không bị mất đi.

Với hướng tiếp cận này việc thiết kế chương trình tương đối đơn giản hơn do chúng ta không cần phải phân tích nhiều về đặc điểm khuôn mặt nữa. Do vậy phương pháp tiếp cận này tương đối dễ cài đặt hơn và cũng khá hiệu quả trong việc nhận dạng.

Hai phương pháp chủ yếu được sử dụng trong cách tiếp cận này đó là phân tích thành phần chính (PCA), và phân tích sự khác biệt tuyến tính (LDA)

**Phân tích thành phần chính (PCA):** PCA (Principal Components Analysis) là một thuật toán được sử dụng để tạo ra ảnh mới từ ảnh ban đầu. Ảnh mới này có kích thước nhỏ hơn rất nhiều so với ảnh đầu vào và vẫn mang những đặc trưng cơ bản nhất của ảnh cần nhận dạng. PCA không cần quan tâm đến việc tìm ra các đặc điểm cụ thể của thực thể cần nhận dạng và mối quan hệ giữa các đặc điểm đó. Tất cả các chi tiết đó được thể hiện ở ảnh mới được tạo ra từ PCA.

Về bản chất, PCA tìm ra một không gian mới theo hướng biến thiên mạnh nhất của một tập hợp các vector trong không gian cho trước. Trong không gian mới, người ta hy vọng rằng việc phân loại sẽ mang kết quả tốt hơn so với không gian ban đầu.



*Không gian mới ( $p1, p2$ ) theo hướng phân bố mạnh nhất của các vector trong không gian ( $x1, x2$ ) theo PCA*

Ưu điểm:

- Tìm được các đặc tính tiêu biểu của đối tượng nhận dạng mà không cần phải xác định các thành phần và mối quan hệ giữa các thành phần đó.
- Thuật toán có thể thực hiện tốt với các ảnh có độ phân giải cao, do PCA sẽ thu gọn ảnh thành một ảnh có kích thước nhỏ hơn.
- PCA có thể kết hợp với các phương pháp khác như mạng Noron, Support Vector Machine, ... để mang lại hiệu quả nhận dạng cao hơn.

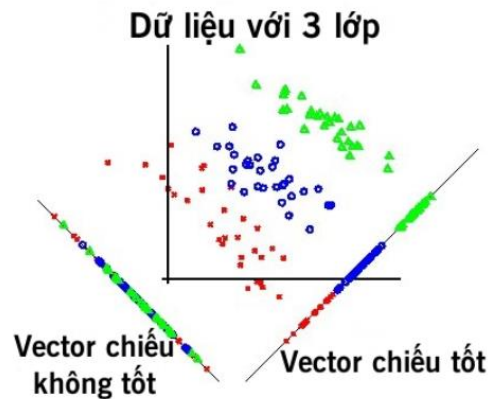
Nhược điểm:

- PCA phân loại theo chiều phân bố lớn nhất của tập vector. Tuy nhiên, chiều phân bố lớn nhất không phải lúc nào cũng mang lại hiệu quả tốt nhất cho bài toán nhận dạng.
- PCA rất nhạy với nhiễu.

**Phân tích sự khác biệt tuyến tính (LDA):** LDA là phương pháp dùng để nhận diện khuôn mặt dựa trên một phép chiếu tuyến tính từ không gian hình ảnh vào một chiều không gian thấp hơn bằng cách tối đa giữa các lớp tán xạ và giảm thiểu phân tán trong lớp.

Mục tiêu của phương pháp là phân loại dữ liệu bằng cách tìm hướng phân biệt nhất (most discriminative direction). LDA là tìm các hướng mà theo đó các lớp được phân chia tốt nhất từ

đó làm giảm số chiều dữ liệu trong khi giữ lại nhưng thông tin phân biệt các lớp nhiều nhất có thể.



Nhược điểm:

- LDA chỉ tạo ra nhiều nhất  $c - 1$  vector chiều ứng với  $c - 1$  giá trị riêng tìm được ( $c$  là số lớp của dữ liệu). Do đó nếu quá trình phân loại dữ liệu mà  $c - 1$  vector chiều này vẫn chưa đáp ứng được yêu cầu phân loại thì ta phải sử dụng thêm các phương pháp khác.
- Nếu dữ liệu cần phân loại có phân phối không phải là dạng Gauss (non-Gaussian distribution, tức là các phân phối không thể được biểu diễn bằng các hàm số), LDA sẽ không thể tìm được được các vector chiều tối ưu.
- Khi áp dụng LDA với các chương trình có kích cỡ mẫu nhỏ sẽ thường cho kết quả không tốt.



## Chương II: Nhận dạng khuôn mặt dựa vào phương pháp biểu diễn thưa

### 2.1 Giới thiệu

Bài toán biểu diễn thưa xuất hiện nhiều trong các bài toán thực tế thuộc nhiều lĩnh vực khác nhau của đời sống, đặc biệt trong các lĩnh vực xử lý tín hiệu, xử lý hình ảnh, học máy và thị giác máy tính, ví dụ như các bài toán: khử nhiễu ảnh, xử lý ảnh mờ, khôi phục ảnh, phân lớp ảnh và phân vùng ảnh.

Những kết quả nghiên cứu đầu tiên về bài toán biểu diễn thưa xuất hiện trong bài báo của Stephane Mallat và Zhifeng Zang [7] vào năm 1993, trong đó các tác giả giới thiệu khái niệm từ điển thay thế cho các phép biến đổi dựa trên các cơ sở truyền thống (cơ sở Fourier, cơ sở Wavelet, ...). Trong công trình này, tác giả nhận định với một từ điển cho trước, bài toán biểu diễn thưa hay việc tìm biểu diễn hay xấp xỉ thưa nhất của một tín hiệu dựa trên từ điển sẽ linh hoạt hơn trường hợp tìm biểu diễn thưa dựa trên một cơ sở truyền thống, tương ứng.

Vào năm 1998, Scott Shaobing Chen, David Donoho và Michael Saunders (Chen và cộng sự) đã giới thiệu một kỹ thuật theo đuổi khác sử dụng kỹ thuật nói lỏng bằng cách thay thế bài toán biểu diễn thưa với tiêu chuẩn chọn nghiệm dựa trên chuẩn 0 bằng bài toán nói lỏng là bài toán tối ưu lồi với tiêu chuẩn chọn nghiệm dựa trên chuẩn 1. Trong một số trường hợp, các tác giả đã chỉ ra rằng nghiệm của bài toán biểu diễn thưa và nghiệm của bài toán nói lỏng là trùng nhau. Khi đó, rất nhiều các phương pháp tối ưu lồi đã biết có thể được sử dụng để giải bài toán biểu diễn thưa.

Từ những điểm nổi bật trên, chúng tôi tìm hiểu và áp dụng bài toán biểu diễn thưa để giải quyết bài toán nhận dạng khuôn mặt với phương pháp chọn nghiệm dựa trên chuẩn 1.

### 2.2 Ý tưởng

Bài toán được xây dựng bằng một tập gồm  $k$  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 training. 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 ma trận, chúng tôi sẽ tạm gọi là tập training samples  $A$  gồm  $k$  lớp và mỗi lớp gồm  $p$  vector đặc trưng khuôn mặt, khi đó lớp  $A_i = [v_{i,1}, v_{i,2}, v_{i,3}, \dots, v_{i,p}] \in \mathbb{R}^{m \times p}$ .

Từ mẫu training của lớp  $A_i$ , với mỗi ảnh vector đặc trưng của ảnh test  $y \in \mathbb{R}^m$  là khuôn mặt của đối tượng thuộc lớp thứ  $i$  sẽ xấp xỉ nằm trong mẫu tuyến tính sau:

$$y = \alpha_{i,1}v_{i,1} + \alpha_{i,2}v_{i,2} + \alpha_{i,3}v_{i,3} + \dots + \alpha_{i,p}v_{i,p} \quad (1)$$

Tuy nhiên lớp  $i$  là chưa biết, do vậy ta cần định nghĩa một ma trận  $A$  gồm tất cả  $k$  lớp training mỗi lớp có  $p$  phần tử.

$$A = [A_1, A_2, A_3, \dots, A_k] = [v_{1,1}, v_{1,2}, \dots, v_{1,p}, \dots, v_{k,1}, v_{k,2}, \dots, v_{k,p}] \in \mathbb{R}^{m \times n} \text{ (n là số cột của A)}$$

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

$$y = Ax \in \mathbb{R}^m \quad (2)$$

Trong đó:  $x = [0, \dots, 0, \alpha_{i,1}, \alpha_{i,2}, \dots, \alpha_{i,p_i}, 0, \dots, 0]^T \in \mathbb{R}^n$  là một vector hệ số mà tất cả là 0 ngoại trừ những hệ số tương ứng của lớp  $A_i$

Vấn đề đặt ra ở đây là nếu  $m < n$  thì phương trình có vô số nghiệm và  $x_0$  cần tìm không phải là duy nhất. Thông thường, vấn đề này được giải quyết bằng tiêu chuẩn chọn nghiệm  $l_2$ : (giải bài toán xấp xỉ bằng chuẩn 2)

$$\widehat{x^2} = \arg \min \|x\|_2 \quad (P2)$$

$$\text{Trong đó: } \|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

Tuy nhiên bài toán (P2) thường mang đến số nghiệm khác 0 khá lớn và không tối ưu khi ứng dụng để giải phương trình (2). Có nghĩa là số phần tử khác 0 càng lớn sẽ tương đương với nhiều bức ảnh training thuộc các lớp khác nhau. Để giải quyết khó khăn này chúng ta sử dụng tiêu chuẩn chọn nghiệm  $l_0$ :

$$\widehat{x_0} = \arg \min \|x\|_0 \quad (3)$$

$$y = Ax \in \mathbb{R}^m$$

Tức là giải bài toán  $y = Ax$  với số phần tử khác 0 là nhỏ nhất.

### 2.3 Bài toán biểu diễn thưa với chuẩn $l_1$

Do (3) là bài toán NP-hard nên ta cần đưa về các bài toán xấp xỉ khác. Những phát triển gần đây trong lý thuyết biểu diễn thưa và lấy mẫu nén chỉ ra rằng nếu  $x$  được tìm thấy là đủ thưa bài toán  $l_0$  ta đưa về bài toán xấp xỉ  $l_1$

$$(l_1): \widehat{x_1} = \arg \min \|x\|_1 \quad (4)$$

$$\text{Trong đó: } \|x\|_1 = |x_1| + |x_2| + \dots + |x_n|$$

Theo các nghiên cứu thì việc giải bài toán (2) bằng phương pháp xấp xỉ  $l_1$  cho nghiệm tương đối tốt trong thời gian hợp lý khi biết  $x$  rất thưa.

#### a. Phân lớp dựa vào biểu diễn thưa

Ta đưa vào một vector đặc trưng của ảnh  $y$  tương ứng với một đối tượng trong tập mẫu training, trước tiên tìm biểu diễn thưa của  $x$

$$\text{Với } \widehat{x_1} = \arg \min \|x\|_1$$

$$\|Ax - y\|_2 \leq \epsilon$$

Ý tưởng chính ở đây là những phần tử khác 0 của  $x_1$  sẽ liên kết với những cột thuộc lớp  $i$  trong  $A$ , và từ đó ta có thể dễ dàng chỉ định ảnh test  $y$  thuộc lớp nào. Tuy nhiên do các yếu tố gây nhiễu hay lỗi mô hình hóa dẫn tới việc một phần nhỏ phần tử khác 0 liên kết với nhiều hơn một lớp. Dựa trên các mẫu training, 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 biểu diễn tuyến tính của  $y$  thông qua tập mẫu training.

Gọi  $\delta_i: \mathbb{R}^n \rightarrow \mathbb{R}^n$  là những hệ số đặc trưng của lớp  $i$ . Với  $x \in \mathbb{R}^n$ ,  $\delta_i(x) \in \mathbb{R}^n$  là một vector mà những hệ số khác 0 của  $x$  tương ứng với lớp  $i$ . Với những hệ số của  $\delta_i$  của lớp  $i$  ta có thể tính gần đúng  $y$  với  $\hat{y}_i = A \delta_i(\hat{x}_1)$ . Tiếp theo chúng ta có thể chỉ định  $y$  thuộc lớp nào có sự khác biệt nhỏ nhất giữa  $y$  và  $y_1$ :

$$\min_i r_i(y) = \|y - A \delta_i(\hat{x}_1)\|_2$$

### Thuật toán 1:

1: Đầu vào: ma trận tập các vector đặc trưng của mẫu training và vector đặc trưng của ảnh test  $y$

$A = [A_1, A_2, \dots, A_k] \in \mathbb{R}^{m \times n}$  ứng với  $k$  lớp, ảnh test  $y \in \mathbb{R}^m$  (Và hệ số  $\epsilon > 0$ )

2: Chuẩn hóa các cột của  $A$  theo chuẩn  $l_2$

3: Giải bài toán tối ưu xấp xỉ bằng chuẩn  $l_1$

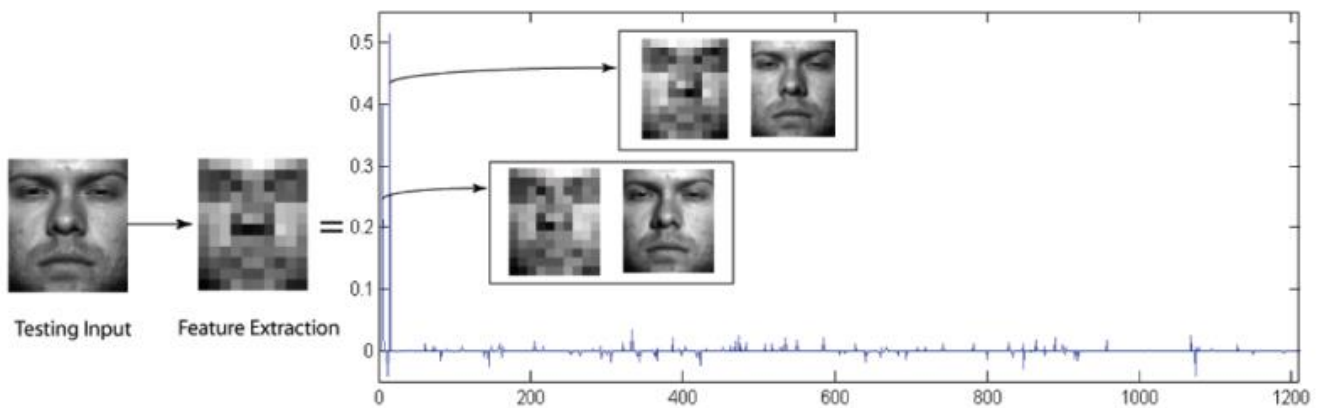
$$\hat{x}_1 = \arg \min \|x\|_1 \text{ trong đó } y = Ax \quad (5)$$

Hoặc:

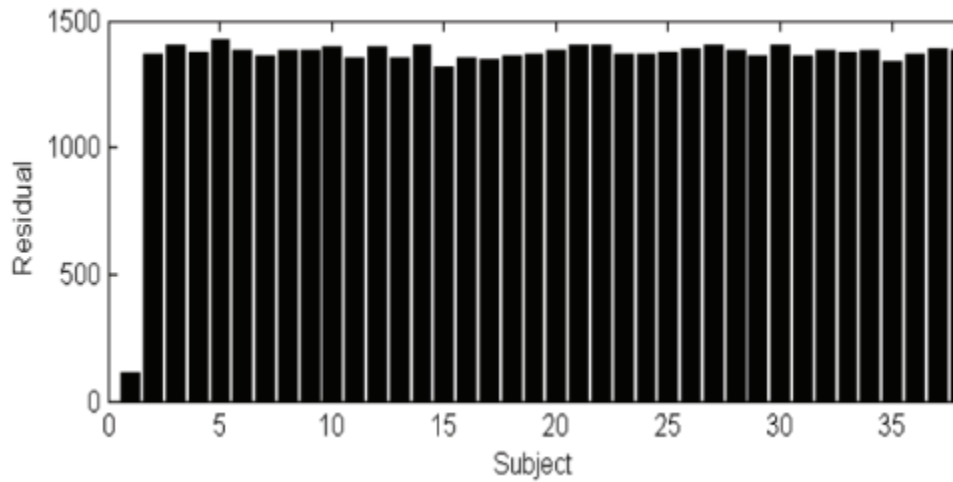
$$\hat{x}_1 = \arg \min \|x\|_1 \text{ trong đó } \|Ax - y\| \leq \epsilon$$

4: Tính số dư  $r_i(y) = \|y - A \delta_i(\hat{x}_1)\|_2$  với  $i = 1, \dots, k$ .

5: Xác định  $y = \arg \min_i r_i(y)$ .



(a)

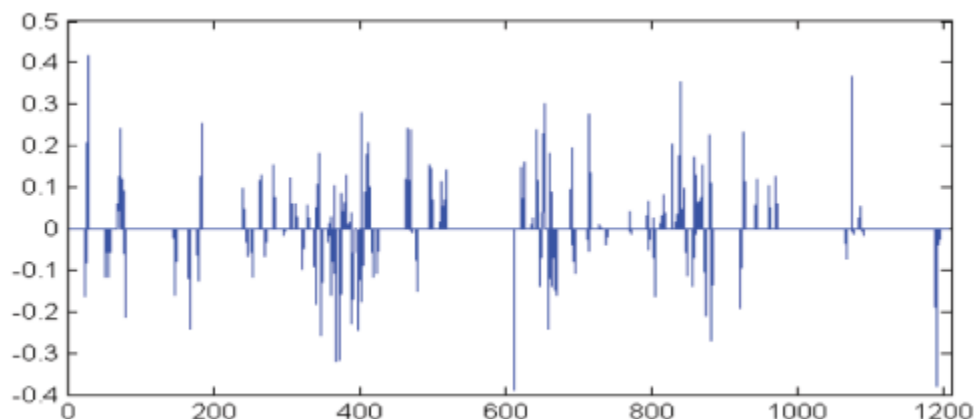


(b)

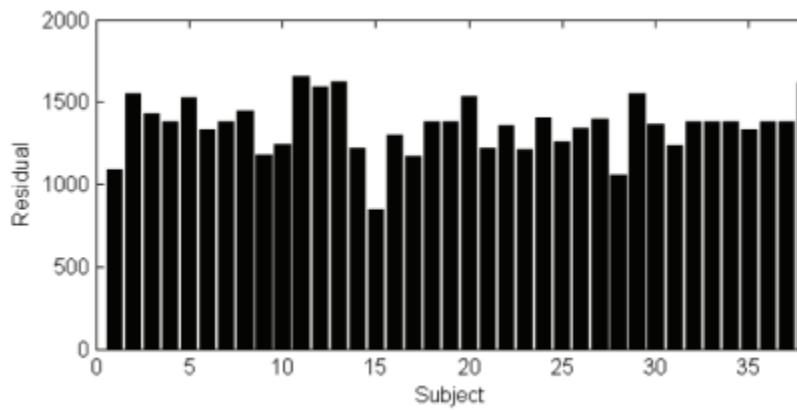
(Hình 1)

Hình 1: Nhận dạng ảnh  $12 \times 10$  pixels. Ảnh  $y$  thuộc về đối tượng thứ nhất. Giá trị của  $x$  được tính bằng thuật toán 1 được thể hiện ở hình (a). Độ sai lệch  $r_i$  của ảnh test được thể hiện ở hình b. Tỷ lệ giữa 2 sai số nhỏ nhất là 1:8.6

Ví dụ 1: (So sánh giữa chuẩn  $l_1$  với chuẩn  $l_2$ ). Để minh họa cho ‘thuật toán 1’, chúng ta lựa chọn ngẫu nhiên một nửa trong số 2412 ảnh và xem chúng là tập training cho một nửa còn lại. Ở ví dụ này, chúng ta giảm số chiều từ  $192 \times 168$  xuống  $12 \times 10$ . 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  $120 \times 1206$  và phương trình  $y = Ax$  có vô số nghiệm. Hình 1a cho thấy những đặc trưng và những ảnh gốc tương ứng với 2 hệ số lớn nhất. Hai hệ số lớn nhất là ứng với 2 ảnh mẫu training từ đối tượng thứ nhất. Hình 1b chỉ ra độ sai khác  $\delta_i(\hat{x}_1)$ ,  $i = 1, \dots, 38$ . Với thuật toán 1 đạt được độ chính xác 92.1 % với Extended Yale B database



(a)



(b)

(Hình 2)

Hình 2: Hình a cho thấy những hệ số từ phương pháp chuẩn  $l^2$  cùng sử dụng ảnh test như ở hình 1, 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

## b. Xác định tính hợp lệ dựa vào biểu diễn thưa (xác định tính hợp lệ của ảnh test)

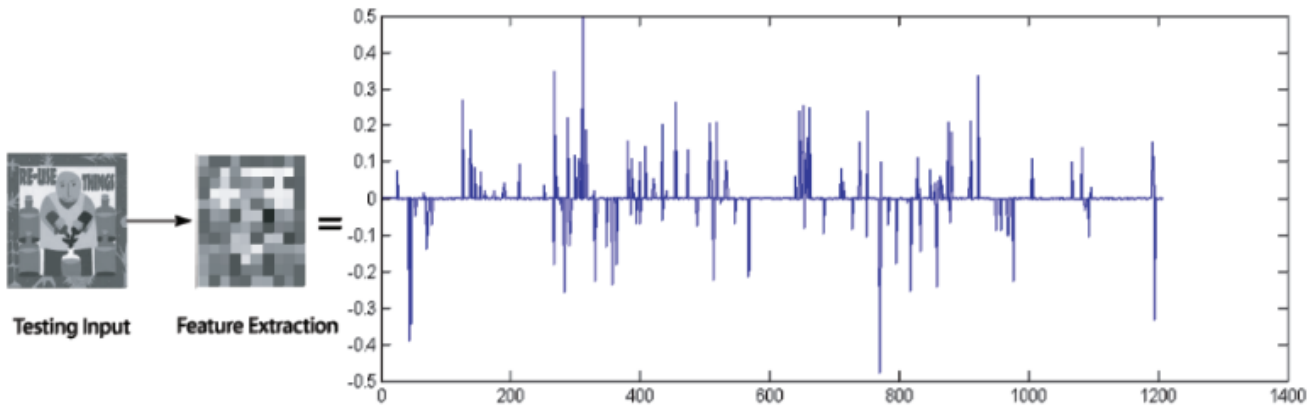
Trước khi phân lớp một ảnh test chúng ta cần xem xét nó có phải là một ảnh hợp lệ hay không nghĩa là kiểm tra xem nó có phải là ảnh của một đối tượng nào đấy trong cơ sở dữ liệu không? Khả năng để phát hiện ảnh là không hợp lệ cũng là một phần quan trọng trong hệ thống nhận dạng hiện nay. Với một hệ 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 cơ sở dữ liệu 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 Nearest Neighbor hay Nearest Subspace thường sử dụng hệ số sai khác  $r_i(y)$  để xác định tính hợp lệ và hơn 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  $r_i(y)$  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 giữa các phép thử với các lớp.

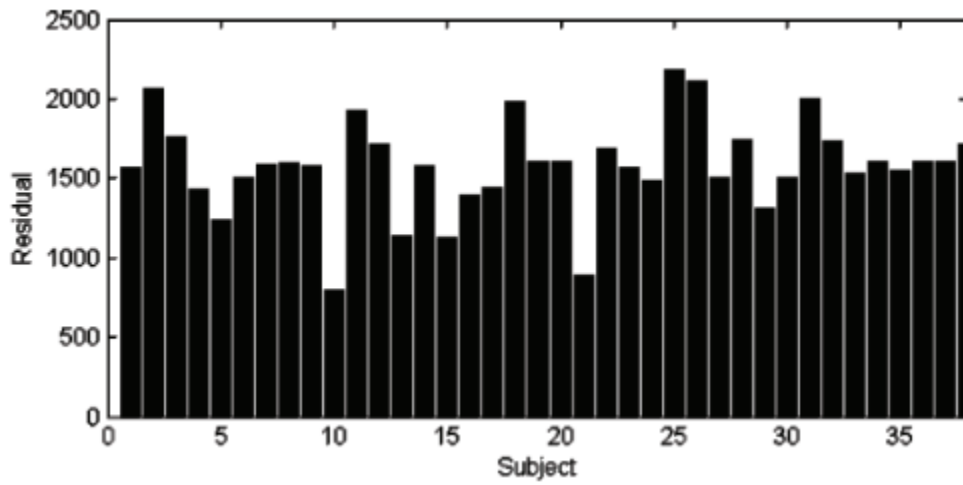
Trong mô hình biểu diễn thưa, hệ số  $\hat{x}_1$  được tính tổng thể trên toàn bộ các ảnh training. Với một ý nghĩa nào đó, nó có thể khai thác sự phân bố của lớp training để xem xét sự hợp lệ của bức ảnh. Các nghiên cứu cho rằng hệ số của  $\hat{x}$  mang đến sự đánh giá tốt hơn cho việc xác định tính hợp lệ của ảnh hơn là việc sử dụng số dư.

Ví dụ 2: Sự tập trung của các hệ số khác 0

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



(a)



(b)

Hình 3: a) Hệ số biểu diễn thưa của một bức ảnh không hợp lệ dựa trên tập training A.

b) Số dư  $r$  của bức ảnh qua  $\delta_i(\hat{x})$  qua các hệ số biểu diễn thưa dựa trên thuật toán chuẩn  $l_1$ . Tỷ lệ giữa 2 số dư  $r$  nhỏ nhất là 1:1.2

So với hệ số của một hình ảnh hợp lệ như với hình 1, những hệ số này không tập trung vào bất kỳ một đối tượng cụ thể nào mà trải rộng trên toàn tập training. Do vậy sự phân bố các hệ số của  $\hat{x}$  mang đến một phần thông tin quan trọng để xác định bức ảnh có hợp lệ hay không? Một bức ảnh hợp lệ sẽ có những hệ số thưa mà những phần tử khác 0 tập trung chủ yếu vào một đối tượng nào đó trong khi một bức ảnh không hợp lệ những hệ số này sẽ trải rộng trên toàn bộ tập training.

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

**Định nghĩa 1:** Sparsity Concentration Index (SCI)

$$SCI(x) = \frac{k \cdot \max_i \|\delta_i(x)\|_1 / \|x\|_1 - 1}{k - 1} \in [0, 1]$$

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 các lớp. Vì thế ta đặt ra một ngưỡng  $t \in [0, 1]$  nào đó và chấp nhận ảnh đó là hợp lệ hay không.

$$SCI(x) \geq t$$

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

Thay vì dựa vào sự thống kê của  $r_i(y)$  để xác định tính hợp lệ của bức ảnh. Bằng phương pháp tiếp cận này ta đã tách thành hai công việc riêng biệt: số dư  $r$  để nhận dạng và hệ số  $x$  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 một đối tượng là tương đương với nhiều lớp. Ví dụ trong hệ thống nhận dạng khi một khuôn mặt khá giống với một vài đối tượng và có hệ số  $r$  khá nhỏ tương đương với các lớp khác nhau. 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 đối tượng của tập training hay không.

## 2.4 Trích chọn đặc trưng

Nhiều phương pháp trích chọn đặc trưng được nghiên cứu để tìm ra cách biểu diễn tốt nhất cho mỗi khuôn mặt với số chiều thấp hơn. Một số phương pháp trích chọn đặc trưng điển hình như: Eigen-faces, Fisherfaces, Laplacianfaces. Một số phương pháp khác trích chọn những vùng 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 trích chọn đặc trưng phức tạp hơn bao gồm nonlinear và kernel features.

Lợi ích của việc trích chọn đặc trưng là giảm thông tin dữ liệu và chi phí tính toán nhưng vẫn mang lại hiệu suất 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 từ không gian ảnh sang không gian đặc trưng có thể biểu diễn qua ma trận  $R \in \mathbb{R}^{d \times m}$  với  $d \ll m$

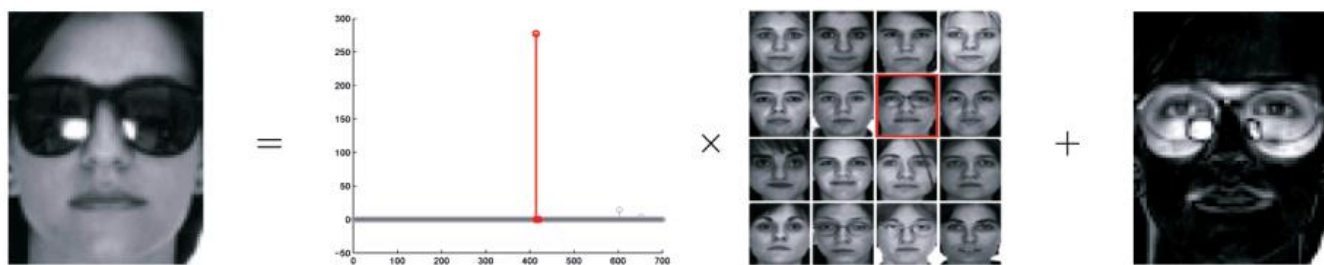
$$\tilde{y} = Ry = RAx_0 \in \mathbb{R}^d$$

Bây giờ ta giải quyết ‘thuật toán 1’ với :

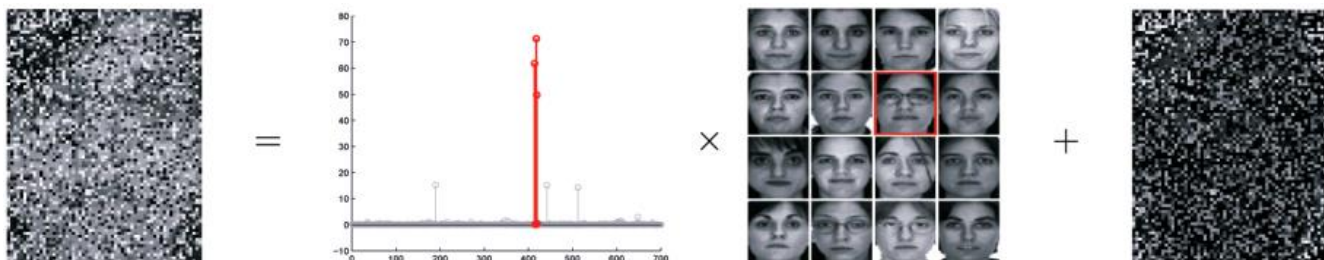
$$\hat{x}_1 = \arg \min ||x_1||$$

$$||RAx - \tilde{y}||_2 \leq \epsilon$$

Với số dư  $\epsilon > 0$  và dĩ nhiên ma trận  $A$  ở ‘thuật toán 1’ bây giờ được thay bằng ma trận  $RA \in \mathbb{R}^{d \times n}$  với  $d$  là số chiều đặc trưng và ảnh test  $y$  được thay bằng ảnh test  $\tilde{y}$ . 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 1’, nó chỉ đòi hỏi ảnh test  $\tilde{y}$  nằm trong hoặc gần với phạm vi của  $RA_i$



(a)



(b)

Sau khi trích chọn đặc trưng hay giảm chiều ta giải bài toán biểu diễn thưa trên ảnh và tập từ điển đã giảm chiều (hình b). Với ảnh test  $\tilde{y}$  bên trái bằng hệ số của  $x_1$  nhân với từ điển cộng với sai số thừa. Những đường màu đỏ là hệ số của  $x_1$  tương ứng với những đối tượng của ảnh trong tập từ điển.

## 2.5 Giải bài toán chuẩn 1 với phương pháp quy hoạch tuyến tính

Trong đề tài khóa luận này tôi trình bày phương pháp quy hoạch tuyến tính với thuật toán đơn hình để giải bài toán (1).

$$y = Ax$$

$$\hat{x}_1 = \arg \min \|x\|_1$$

### 2.5.1 PHÂN LOẠI DẠNG BÀI TOÁN QUY HOẠCH TUYẾN TÍNH

#### a. Dạng tổng quát

Bài toán QHTT dạng tổng quát với  $n$  ẩn là bài toán có dạng:

$$(1) f(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \min(\max)$$

$$(2) a_{i,1}x_1 + a_{i,2}x_2 + \dots + a_{i,n}x_n \begin{cases} \geq \\ \leq \\ = \end{cases} b_i, i = 1, 2, \dots, m$$

$$(3) x_j \begin{cases} \geq 0 \\ \leq 0 \\ \text{tùy ý} \end{cases}, j = 1, 2, \dots, n$$

Trong đó:

- (1) hàm mục tiêu



- (2) là hệ số ràng buộc chính
- (3) là ràng buộc dấu
- (2) (3) được gọi là hệ ràng buộc chung của bài toán

Khi đó:

- Mỗi vector  $x = (x_1, x_2, \dots, x_n)$  thỏa mãn (2) và (3) được gọi là một phương án (PA) của bài toán.
- Mỗi phương án  $x$  thỏa mãn (1), nghĩa là tại đó hàm mục tiêu đạt giá trị nhỏ nhất (lớn nhất) trên tập các phương án được gọi là một phương án tối ưu (PATU) của bài toán.
- Giải một bài toán QHTT là đi tìm một phương án tối ưu của nó hoặc chỉ ra rằng bài toán vô nghiệm, nghĩa là bài toán không có PATU.

### b. Dạng chính tắc của bài toán QHTT

- (1)  $f(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \min(\max)$
- (2)  $a_{i,1}x_1 + a_{i,2}x_2 + \dots + a_{i,n}x_n = b_i, i = 1, 2, \dots, m$
- (3)  $x_j \geq 0, j = 1, 2, \dots, n$

**Nhận xét.** Bài toán QHTT dạng chính tắc là bài toán QHTT dạng tổng quát trong đó:

- Các ràng buộc chính đều là phương trình.
- Các ẩn đều không âm.

### c. Dạng chuẩn

- (1)  $f(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \min(\max)$
- (2)  $a_{i,1}x_1 + a_{i,2}x_2 + \dots + a_{i,n}x_n = b_i, i = 1, 2, \dots, m$
- (3)  $x_j \geq 0, j = 1, 2, \dots, n$

Trong đó:

- Các hệ số tự do đều không âm.
- Trong ma trận hệ số tự do có đủ m vector cột đơn vị:  $e_1, e_2, \dots, e_m$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \dots, e_m = \begin{pmatrix} 0 \\ 0 \\ \dots \\ m \end{pmatrix}$$

Khi đó:

- Các ẩn ứng với các vector cột đơn vị được gọi là các ẩn cơ bản. Cụ thể ẩn ứng với vector cột đơn vị  $k$  là ẩn cơ bản thứ  $k$ .
- Một phương án mà các ẩn cơ bản đều bằng 0 được gọi là phương án cơ bản.
- Một phương án cơ bản có đủ m thành phần dương được gọi là không suy biến. Ngược lại một phương án cơ bản có ít hơn m thành phần dương được gọi là suy biến.

## 2.5.2 Thuật toán đơn hình giải bài toán min

**Bước lặp thứ nhất** (bảng đơn hình thứ nhất)

## 1. Lập bảng đơn hình xuất phát

Vẽ bảng đơn hình và ghi vào đó các thành phần sau của bài toán dạng chuẩn

- Dòng 1. Ghi các ẩn của bài toán (kể cả ẩn phụ).
- Dòng 2. Ghi các hệ số của các ẩn trong hàm mục tiêu.
- Cột 2. Ghi các ẩn cơ bản của bài toán theo thứ tự từ ẩn cơ bản thứ nhất đến ẩn cơ bản cuối cùng, ta gọi cột này là cột ẩn cơ bản.
- Cột 1: Ghi tương ứng các hệ số của các ẩn cơ bản trong hàm mục tiêu, ta gọi cột này là cột hệ số cơ bản.
- Cột 3. Ghi các số hạng tự do của hệ ràng buộc chính theo thứ tự từ trên xuống dưới, ta gọi cột này là cột phương án.
- Cột 4. Ghi ma trận điều kiện A của bài toán.

Tính hệ số ước lượng  $\Delta_j$  của các ẩn  $x_j$  ( $j = 1, 2, \dots, n$ ) và ghi tương ứng vào dòng dưới cột 4, với  $\Delta_j$  được tính theo công thức sau:

$$\Delta_j = (\text{cot } 1) * (A_j) - (hsx_j) \quad (hsx_j: \text{hệ số của ẩn } x_j \text{ trong hàm mục tiêu})$$

Nếu  $x_j$  là ẩn cơ bản thì  $\Delta_j = 0$

Tính hệ số  $f_0 = (\text{cot } 1) * (\text{cot } 3)$  ghi dưới cột 3.

## 2. Xác định phương án cơ bản xuất phát

Với bảng đơn hình vừa lập được thì phương án cơ bản xuất phát  $x^0$  của bài toán được xác định như sau: Cho các ẩn ở cột 2 nhận giá trị tương ứng ở cột 3, các ẩn còn lại nhận giá trị 0. Trị số của hàm mục tiêu tại phương án cơ bản xuất phát  $x^0$  là  $f_0 = f(x^0)$ .

## 3. Đánh giá tính tối ưu của phương án cơ bản xuất phát

- Điều kiện tối ưu: Nếu hệ số ước lượng  $\Delta_j \leq 0 \quad \forall j$  thuật toán kết thúc
- Dấu hiệu của bài toán không có PATU:  $\exists \Delta_k > 0$  và  $a_{ik} \leq 0 \quad \forall i$ , thuật toán kết thúc
- Nếu không xảy ra 2 bước trên thì thuật toán tiếp tục lặp trong bước lặp thứ 2

**Bước lặp thứ hai** (Bảng đơn hình thứ hai)

### 1. Tìm ẩn đưa vào

Ẩn được chọn đưa vào: Ẩn ứng với  $\Delta_k > 0$  lớn nhất.

### 2. Tìm ẩn đưa ra

Thực hiện phép chia lần lượt các số của cột phương án cho các số dương của cột chủ yếu và ghi các thương số  $\lambda_i$  đó vào cột cuối cùng.

Xác định  $\lambda_r = \min\{\lambda_i\}$ . Khi đó  $x_r$  là ẩn mà ta đưa ra khỏi hệ ẩn cơ bản. Dòng có chứa  $x_r$  được gọi là dòng chủ yếu. Số dương nằm trên dòng chủ yếu và cột chủ yếu được gọi là hệ số chủ yếu.

### 3. Lập bản đơn hình thứ 2

- Cột 2: Thay ẩn đưa ra bằng ẩn đưa vào, các ẩn cơ bản còn lại giữ nguyên. Dòng có ẩn đưa vào gọi là dòng chuẩn.
- Cột 1: Thay hệ số của ẩn đưa ra bằng hệ số của ẩn đưa vào, các hệ số của các ẩn cơ bản còn lại giữ nguyên. Các thành phần còn lại được xác định theo dòng như sau:
  - Dòng chuẩn = Dòng chủ yếu chia cho hệ số chủ yếu.
  - Dòng thứ i = Dòng thứ i (cũ) –  $a_{iv}$ .dòng chuẩn. ( $a_{iv}$ : số nằm trên giao của dòng i và cột chủ yếu). Các hệ số ước lượng và trị số của hàm mục tiêu trong bảng thứ hai được tính và ghi như bảng thứ nhất.

#### 4. Xác định và đánh giá phương án cơ bản thứ hai (như bước lặp thứ nhất)

### 2.5.3 Áp dụng thuật toán đơn hình để giải bài toán phân lớp

Ta quay về bài toán tìm x với chuẩn 1 để phân lớp đối tượng

$$\hat{x}_1 = \arg \min \|x\|_1$$

$$y = Ax$$

Xem y là một vector m chiều có các hệ số là  $[y_1, y_2, \dots, y_m]^T$ , từ điển A có dạng :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \text{ khi đó hệ số } x \text{ cần tìm là } [x_1, x_2, \dots, x_n]^T$$

Ta cần giải bài toán quy hoạch tuyến tính sau:

$$\text{Hàm mục tiêu: } f: x_1 + x_2 + \dots + x_n \rightarrow \min$$

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = y_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = y_2$$

...

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = y_m$$

Ví dụ: Giả sử từ điển  $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 2 & 1 & 0 \end{bmatrix}$  và vector đặc trưng  $y = [6, 2, 7]^T$ . Ta có bài toán quy hoạch

tuyến tính sau:

$$\text{Hàm mục tiêu: } f: x_1 + x_2 + x_3 + x_4 \rightarrow \min$$

$$\begin{cases} x_1 + x_2 + x_3 = 6 \\ x_2 - x_4 = 2 \\ 2x_2 + x_3 = 7 \end{cases}$$

$$x_i \geq 0, \forall i \in [1..4]$$

Đưa bài toán về dạng chuẩn tắc bằng cách thêm 2 biến không âm  $x_5, x_6$ :

$$f: x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \rightarrow \min$$

$$\begin{cases} x_1 + x_2 + x_3 = 6 \\ x_2 - x_4 + x_5 = 2 \\ 2x_2 + x_3 + x_6 = 7 \end{cases}$$

Sau khi thực hiện thuật toán đơn hình ta giải được  $x$  như sau:

Hệ số cơ bản	Ấn cơ bản		x1	x2	x3	x4	x5	x6	$\lambda$
			1	1	1	1	1	1	1
1	x1	6	1	1	1	0	0	0	6
1	x5	2	0	[1]	0	-1	1	0	2
1	x6	7	0	2	1	0	0	1	3.5
		15	0	3	1	-2	0	0	
1	x1	4	1	0	1	1	-1	0	4
1	x2	2	0	1	0	-1	1	0	
1	x6	3	0	0	[1]	2	-2	1	3
		9	0	0	1	1	-3	0	
1	x1	1	1	0	0	-1	1	-1	
1	x2	2	0	1	0	-1	1	0	
1	x3	3	0	0	1	2	-2	1	
		6	0	0	0	-1	-1	-1	

Bước lặp kết thúc khi tất cả  $\Delta_j \leq 0$ ,  $x = [1, 2, 3, 0]^T$  và  $f_{\min} = 6$

## Chương III: Thực nghiệm bài toán

### 3.1 Mô tả bài toán

Chương trình cần xây dựng tập dữ liệu hay còn gọi là tập từ điển bằng cách lưu trữ các vector đặc trưng của các đối tượng, từ đó với một bức ảnh khuôn mặt đầu vào bất kì, chương trình cần nhận dạng được bức ảnh đấy thuộc đối tượng nào trong tập dữ liệu.

Bức ảnh dùng để test có thể được chụp trong các trường hợp khác nhau như: chụp trong phòng, ngoài trời, chụp với các điều kiện chiếu sáng và chất lượng khác nhau, khuôn mặt bị che khuất một phần hay biểu cảm khác nhau.

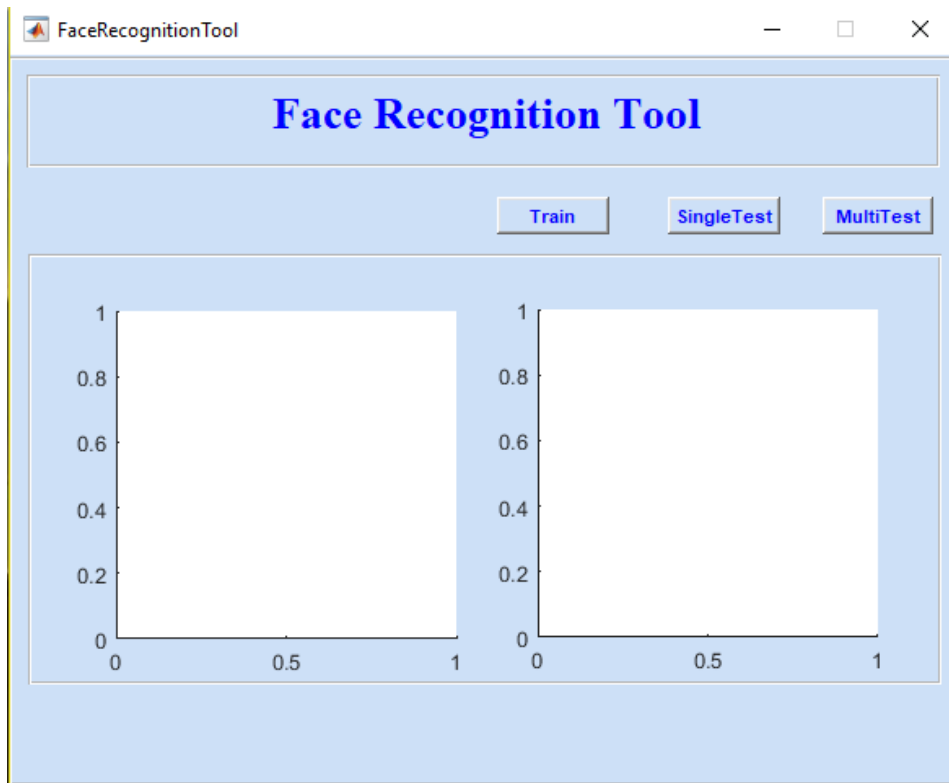
Từ các kết quả thử nghiệm để đưa ra nhận xét, thời gian và tính hiệu quả của chương trình

### 3.2 Phân tích, thiết kế chương trình

Công nghệ và cấu hình phần cứng sử dụng: Matlab2016, Window 10, Ram 6GB, Chip Intel core i5 CPU 1.7GHz.

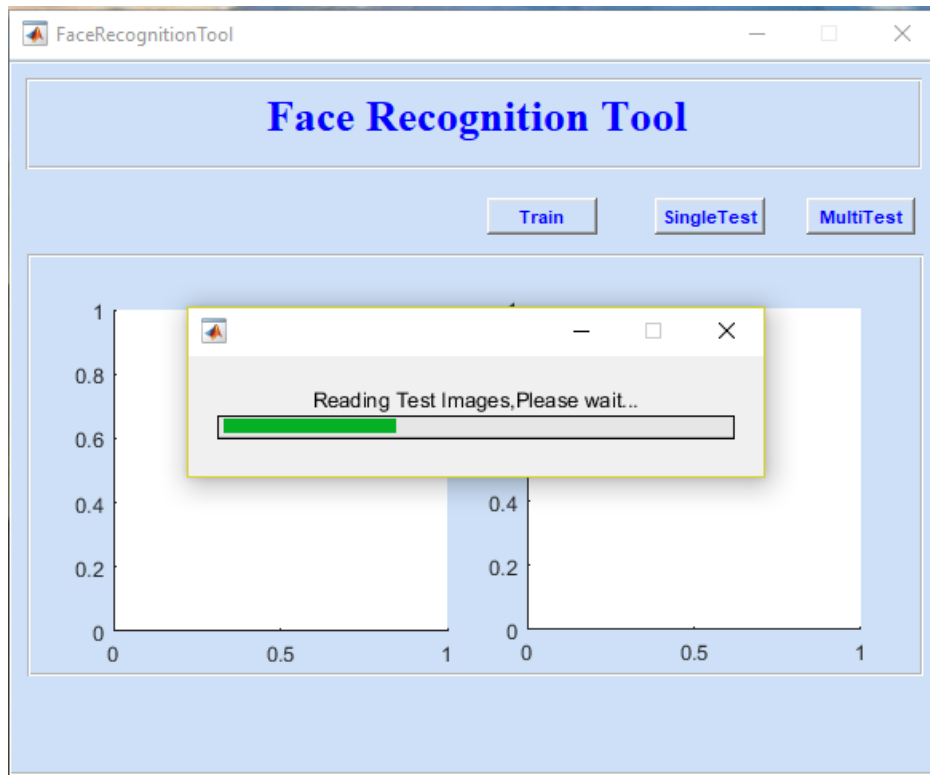
Tập dữ liệu: Gồm 42 đối tượng

- Bộ train: 210 bức ảnh ứng với 42 đối tượng
- Bộ test: 216 bức ảnh ứng với 42 đối tượng không trùng với bộ train



### 3.2.1 Xây dựng từ điển

**Mô tả:** Khi người dùng ấn vào nút train chương trình cho phép người dùng lựa chọn đường dẫn thư mục chứa tập dữ liệu training (.\\FaceDatabase\\Database1\\Train\_Data) sau đó thực hiện xây dựng tập từ điển.

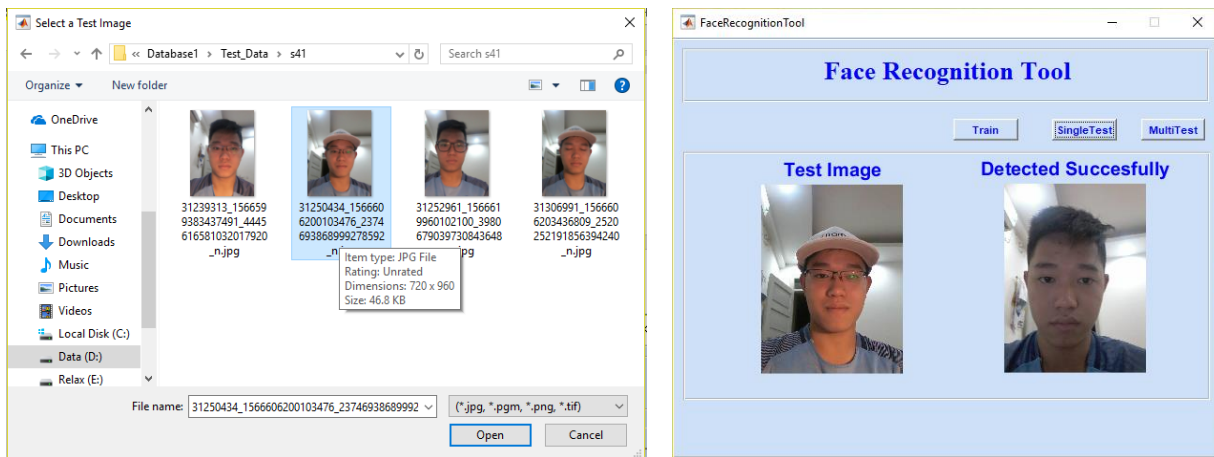


**Chức năng:** Chương trình đọc toàn bộ ảnh trong thư mục, với mỗi ảnh sẽ thực hiện lần lượt các quy trình sau:

- Biến đổi thành ảnh xám (nếu ảnh là ảnh màu)
- Giảm số chiều của ảnh
- Lưu trữ các ảnh như là một vector đặc trưng trong tập từ điển

### 3.2.2 Nhận dạng 1 ảnh

**Mô tả:** Khi người dùng ấn vào nút SingleTest chương trình cho phép người dùng lựa chọn đường dẫn đến ảnh cần test sau đó thực hiện nhận dạng ảnh test dựa trên tập từ điển đã được lưu trữ trong cơ sở dữ liệu.

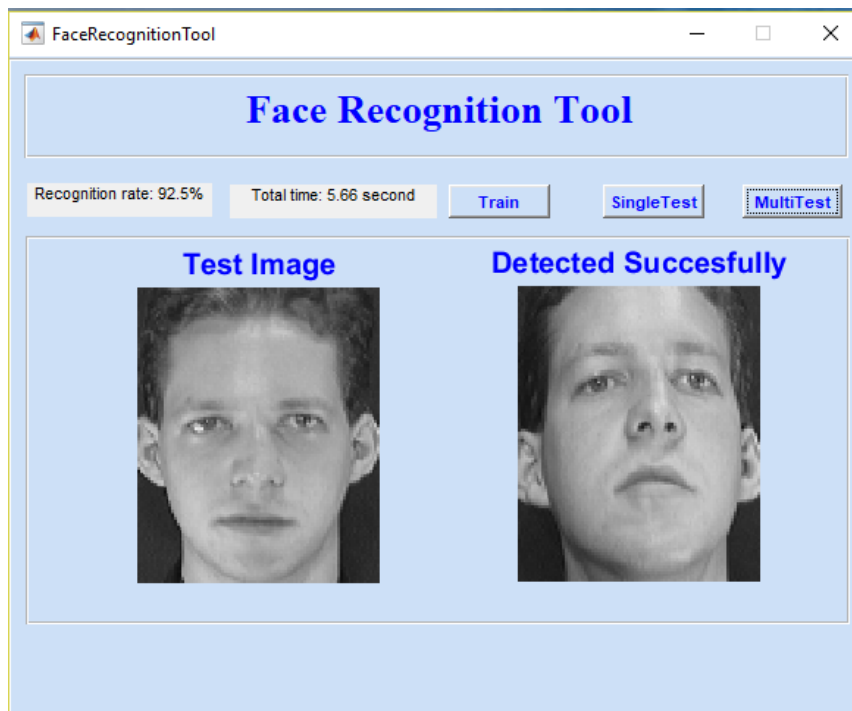


**Chức năng:** Chương trình đọc ảnh test được chọn và thực hiện lần lượt các chức năng sau:

- Biến đổi thành ảnh test thành ảnh xám (nếu ảnh là ảnh màu)
- Giảm số chiều của ảnh tương ứng với số chiều của các vector đặc trưng trong từ điển
- Giải bài toán biểu diễn thưa tìm  $x$  với thuật toán linear programming với vector đặc trưng của ảnh test và tập từ điển
- Đưa ra kết quả nhận dạng dựa vào kết quả của  $x$

### 3.2.3 Đánh giá kết quả nhận dạng trên toàn bộ tập test

**Mô tả:** Khi người dùng ấn nút MultiTest chương trình cho phép lựa chọn thư mục để test các ảnh mong muốn (.\\FaceDatabase\\Database1\\Test\_Data) sau đó thực hiện nhận dạng tất cả các ảnh trong thư mục được chọn và từ đó đưa ra kết quả nhận dạng về độ chính xác và thời gian thực hiện chương trình.



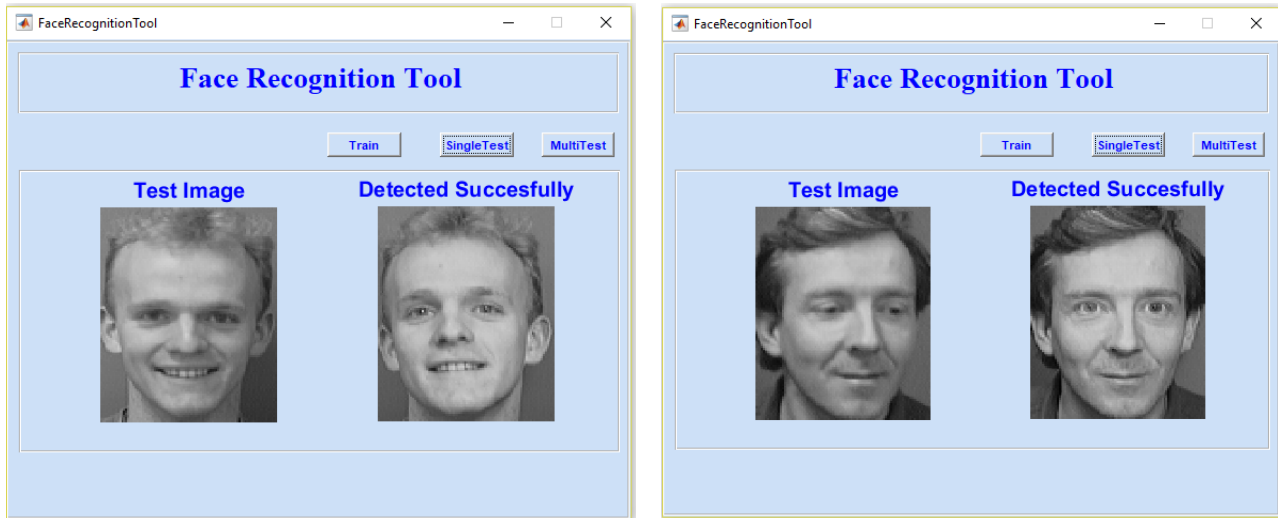
**Chức năng:** Chương trình đọc toàn bộ ảnh trong thư mục, thay vì thực hiện từng ảnh một như SingleTest chương trình sẽ nhận dạng toàn bộ ảnh một cách tự động từ đó đưa ra kết quả về độ chính xác và thời gian để nhận dạng tất cả các ảnh.

### 3.3 Kết quả

#### 3.3.1 Nhận dạng với ảnh xám

**Mục đích:** Kiểm tra tính chính xác của chương trình khi nhận dạng đối tượng với dữ liệu đưa vào là ảnh xám.

**Kết quả:**

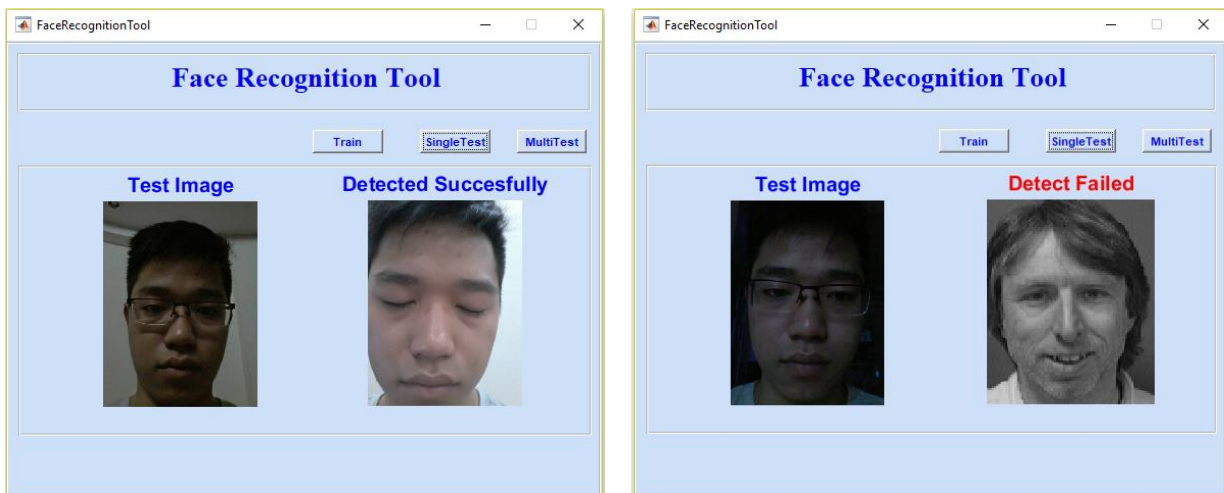


**Nhận xét:** Chương trình cho kết quả chính xác và thời gian xử lý nhanh hơn so với ảnh màu.

#### 3.3.2 Nhận dạng với độ chiếu sáng thấp

**Mục đích:** Kiểm tra tính chính xác của chương trình khi nhận dạng trên cùng một đối tượng được chụp ở vị trí như nhau (khuôn mặt chính giữa camera) nhưng với điều kiện chiếu sáng khác nhau (ảnh bên phải có độ chiếu sáng thấp hơn so với ảnh bên trái).

**Kết quả:**



**Nhận xét:** Độ chính xác của chương trình ảnh hưởng bởi điều kiện chiếu sáng của bức ảnh. Với những bức ảnh được chụp trong điều kiện thiếu ánh sáng sẽ làm cho kết quả nhận dạng không

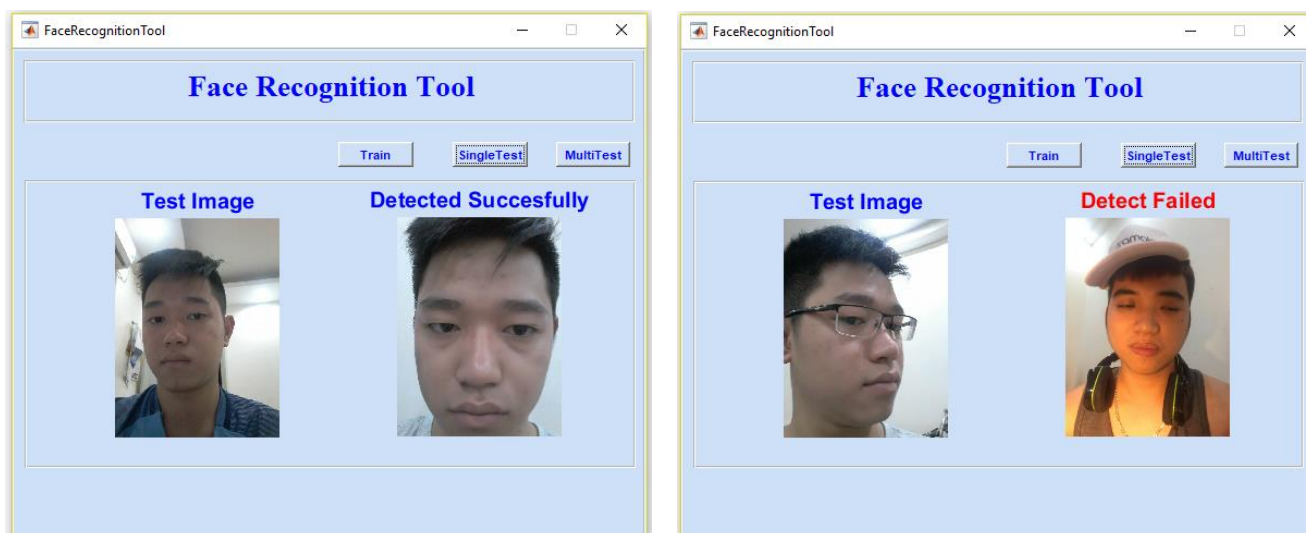


chính xác như mong muốn. Tuy nhiên những bức ảnh được chụp trong một mức ngưỡng chiếu sáng vừa đủ sẽ cho ta kết quả nhận dạng chính xác.

### 3.3.3 Nhận dạng với góc quay khuôn mặt khác nhau

**Mục đích:** Kiểm tra tính chính xác của chương trình khi nhận dạng trên cùng một đối tượng được chụp ở vị trí như nhau (khuôn mặt chính giữa camera) nhưng với góc quay khuôn mặt khác nhau từ đó tìm ra góc quay khuôn mặt phù hợp để không ảnh hưởng đến tính chính xác của chương trình.

**Kết quả:**

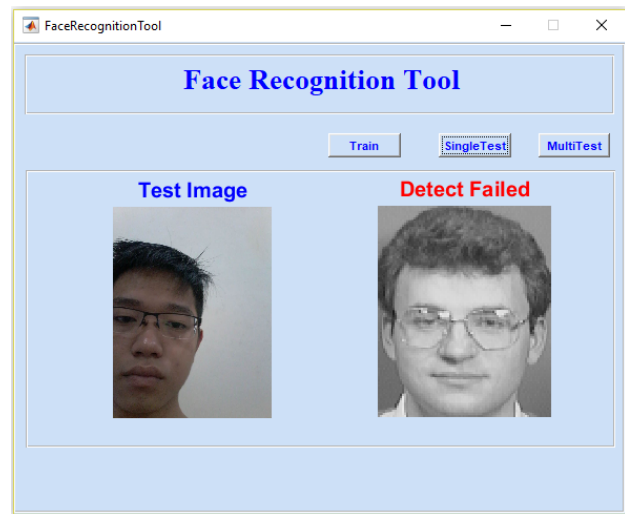
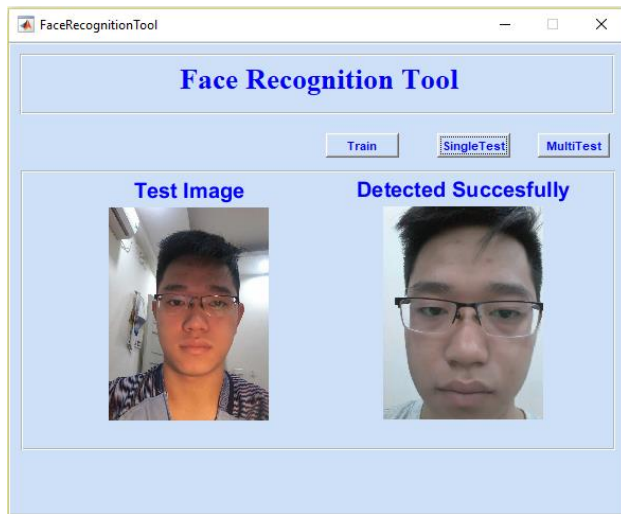


**Nhận xét:** Với những khuôn mặt có góc quay lớn hơn khoảng  $45^0$  thì kết quả nhận dạng sẽ không chính xác

### 3.3.4 Nhận dạng với vị trí khuôn mặt khác nhau

**Mục đích:** Kiểm tra tính chính xác của chương trình khi nhận dạng với vị trí các khuôn mặt khác nhau, khuôn mặt không nằm chính giữa bức ảnh thì kết quả còn chính xác nữa hay không.

**Kết quả:**

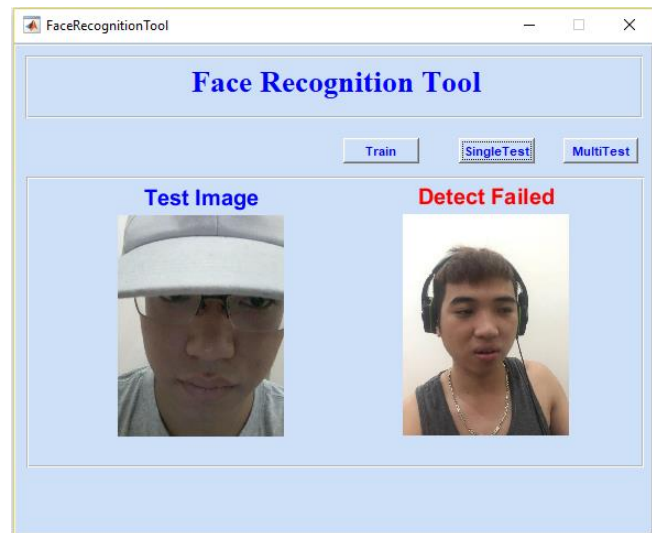
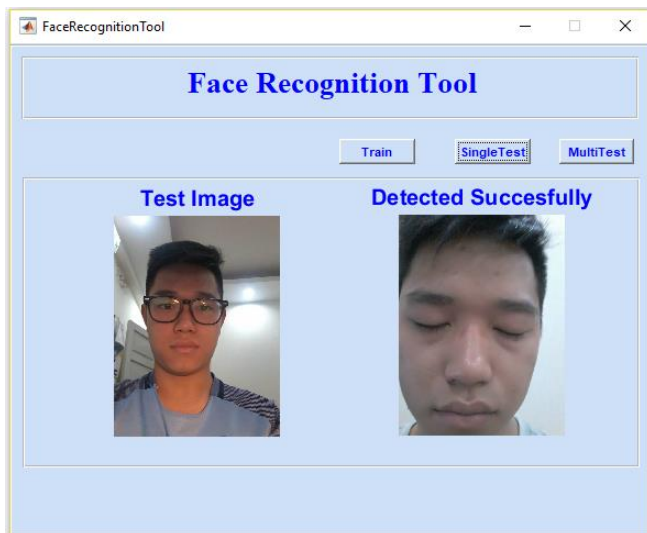


**Nhận xét:** Chương trình cho kết quả tốt khi khuôn mặt nằm chính giữa hay gần chính giữa bức ảnh, tuy nhiên với các bức ảnh có khuôn mặt nằm ở ngoài  $\frac{1}{2}$  bức ảnh sẽ không còn chính xác nữa.

### 3.3.5 Nhận dạng với khuôn mặt bị che

**Mục đích:** Kiểm tra kết quả nhận dạng khi khuôn mặt bị che một phần do bức ảnh người test đội mũ, đeo kính, có râu, tóc che một phần khuôn mặt, ...

**Kết quả:**

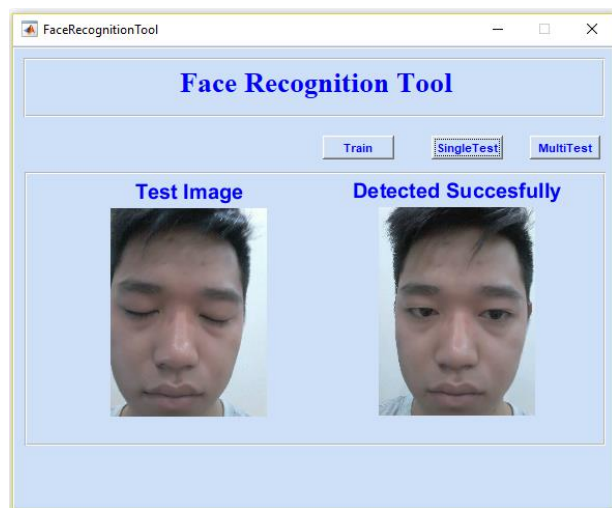


**Nhận xét:** Khi khuôn mặt bị che quá  $\frac{1}{3}$  thì kết quả nhận dạng không còn chính xác.

### 3.3.6 Nhận dạng với khuôn mặt có các biểu cảm hay góc nghiêng khác nhau

**Mục đích:** Kiểm tra kết quả nhận dạng với các khuôn mặt có biểu cảm trạng thái khác nhau hay góc nghiêng đầu khác nhau

**Kết quả:**

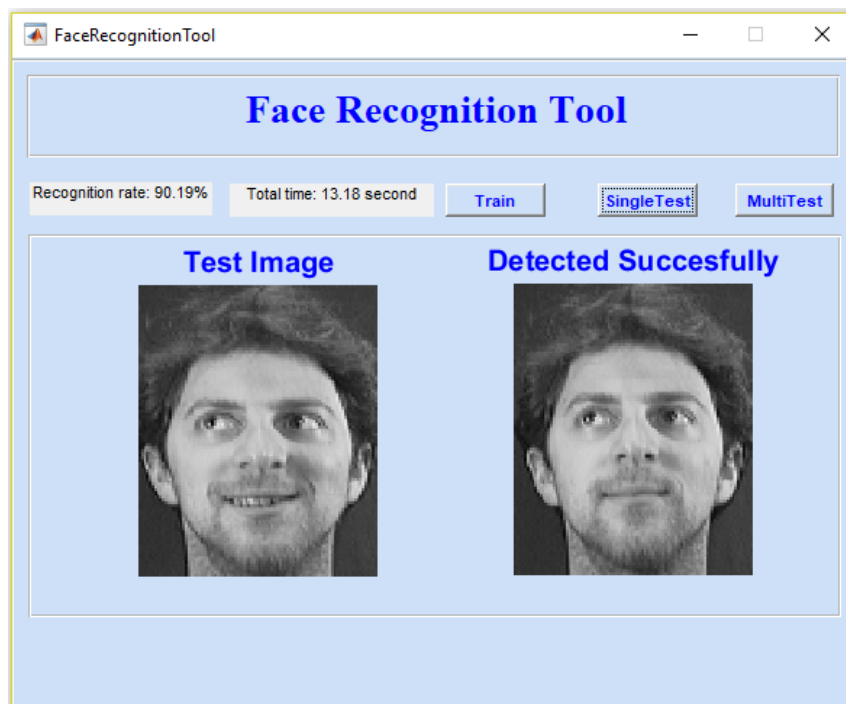


**Nhận xét:** Chương trình cho kết quả chính xác khi khuôn mặt có biểu cảm khác nhau và góc nghiêng đầu không quá  $35^0$ .

### 3.3.7 Thời gian và độ chính xác trên tập ảnh test

**Mục đích:** Tính thời gian và độ chính xác của chương trình khi nhận dạng tự động tập tất cả các ảnh test.

**Kết quả:**



**Nhận xét:** Chương trình nhận diện tự động trên tập gồm 216 ảnh test bao gồm cả ảnh xám và ảnh màu cho kết quả:

- Độ chính xác: 90.2%
- Tổng thời gian nhận dạng: 13.18s
- Thời gian trung bình nhận dạng 1 ảnh: 0.06s

## **Chương IV: Kết luận**

### **4.1 Các kết quả đạt được**

Bài toán nhận dạng khuôn mặt là một trong những lĩnh vực đang được nghiên cứu và phát triển mạnh mẽ hiện nay, tuy có nhiều hướng tiếp cận và phương pháp để tăng độ chính xác của các hệ thống nhận dạng song mỗi phương pháp đều có ưu và nhược điểm nhất định.

Trong thực tế các phương pháp hay thuật toán không chỉ được sử dụng riêng lẻ mà nó được sử dụng kết hợp với nhau hay có những biến đổi thích hợp để đảm bảo được yêu cầu của bài toán và giải quyết được những khó khăn trong vấn đề nhận dạng khuôn mặt như: góc chụp, vị trí khuôn mặt, hướng, điều kiện chiếu sáng, điều kiện môi trường hay khuôn mặt bị che mất một phần nào đấy do các yếu tố khách quan.

Qua quá trình tìm hiểu, hiểu được các quy trình chính trong chương trình xử lý ảnh nói chung và bài toán nhận dạng nói riêng, hiểu được cơ bản các hướng tiếp cận và một số phương pháp trong bài toán nhận dạng nổi bật trong đó là ứng dụng “phương pháp biểu diễn thưa trong bài toán nhận dạng”

Với các kết quả thử nghiệm khác nhau cho thấy sự ảnh hưởng của các yếu tố khách quan đến kết quả của chương trình nhận dạng khuôn mặt như: Vị trí khuôn mặt trong ảnh, điều kiện chiếu sáng, chất lượng ảnh, góc quay khuôn mặt, góc nghiêng đầu, .... Tuy nhiên trong điều kiện cho phép hoặc nằm trong mức ngưỡng thì các kết quả thử nghiệm đạt kết quả chính xác khá cao và thời gian trung bình cho từng bức ảnh khá thấp cụ thể là chưa đến 0.1 giây.

### **4.2 Hướng phát triển tiếp theo**

Tiếp tục nghiên cứu và áp dụng các phương pháp khác nhau để đưa ra sự so sánh đối chiếu về hiệu quả và thời gian thực hiện chương trình. Áp dụng một số thuật toán tích chọn đặc trưng, thêm phần xử lý để xác định khuôn mặt nhằm nâng cao kết quả nhận dạng. Tuy nhiên do thời gian thực hiện luận văn không nhiều và kiến thức còn hạn chế nên còn có những sai sót, em rất mong nhận được sự góp ý của quý thầy cô và bạn đọc.

## **Tài liệu tham khảo**

### **Tiếng Việt**

1. Phạm Thế Bảo, Nguyễn Thành Nhựt, Cao Minh Thịnh, Trần Anh Tuấn, Phan Phúc Doãn (2015), “Tổng quan các phương pháp xác định khuôn mặt người”.

### **Tiếng Anh**

2. Ion Marqu'es (2010), “Face Recognition Algorithms”
3. John Wright, Allen Y. Yang, Arvind Ganesh, S. Shankar Sastry, Fellow, and Yi Ma (2009), “Robust Face Recognition via Sparse Representation”.
4. Michael Elad (2010), “Sparse and Redundant Representations”
5. Rafael C. Gonzalez and Richard E. Woods (2002), “Digital Image Processing”
6. Santa Rosa (2016), “Linear Discriminant Analysis”
7. S.G. Mallat and Zhifeng Zhang (1993), "Matching pursuits with timefrequency dictionaries”