

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

-----o0o-----



## **ĐỒ ÁN TỐT NGHIỆP**

HẢI PHÒNG 2013

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

-----o0o-----

**TÌM HIỂU ĐẶC TRƯNG LỖI 3D VÀ BÀI TOÁN  
PHÁT HIỆN MẶT NGƯỜI TRONG ẢNH**

**ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY**

Ngành: Công nghệ Thông tin

HẢI PHÒNG - 2013

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

-----o0o-----

**TÌM HIỂU ĐẶC TRƯNG LỜI 3D VÀ BÀI TOÁN  
PHÁT HIỆN MẶT NGƯỜI TRONG ẢNH**

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công nghệ Thông tin

Sinh viên thực hiện: Phạm Thành Huân

Giáo viên hướng dẫn: PGS.TS. Đỗ Năng Toàn

Mã số sinh viên: 1351010019

HẢI PHÒNG - 2013

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**  
**TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG**

**CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  
***Độc lập - Tự do - Hạnh phúc***

-----o0o-----

## **NHIỆM VỤ THIẾT KẾ TỐT NGHIỆP**

Sinh viên: Phạm Thành Huân

Mã SV: 1351010019

Lớp: CT1301

Ngành: Công nghệ Thông tin

Tên đề tài: Tìm hiểu đặc trưng lời 3D và bài toán phát hiện mặt người trong ảnh.

## **NHIỆM VỤ ĐỀ TÀI**

1. Nội dung và các yêu cầu cần giải quyết trong nhiệm vụ đề tài tốt nghiệp

a. Nội dung

b. Các yêu cầu cần giải quyết

## **CÁN BỘ HƯỚNG DẪN ĐỀ TÀI TỐT NGHIỆP**

**Người hướng dẫn thứ nhất:**

Họ và tên: Đỗ Năng Toàn

Học hàm, học vị: Phó Giáo Sư - Tiến Sĩ

Cơ quan công tác: Viện khoa học & Công nghệ Việt Nam

Nội dung hướng dẫn:

**Người hướng dẫn thứ hai:**

Họ và tên: .....

Học hàm, học vị: .....

Cơ quan công tác: .....

Nội dung hướng dẫn: .....

.....  
.....

Đề tài tốt nghiệp được giao ngày    tháng    năm 2013

Yêu cầu phải hoàn thành trước ngày    tháng    năm 2013

Đã nhận nhiệm vụ: Đ.T.T.N

Sinh viên

Đã nhận nhiệm vụ: Đ.T.T.N

Cán bộ hướng dẫn Đ.T.T.N

PGS.TS. Đỗ Năng Toàn

Hải Phòng, ngày .....tháng.....năm 2013

HIỆU TRƯỞNG

***GS.TS.NGŨT Trần Hữu Nghị***

## PHÂN NHẬN XÉT TÓM TẮT CỦA CÁN BỘ HƯỚNG DẪN

1. Tinh thần thái độ của sinh viên trong quá trình làm đề tài tốt nghiệp:

.....

.....

.....

.....

.....

.....

2. Đánh giá chất lượng của đề tài tốt nghiệp (so với nội dung yêu cầu đã đề ra trong nhiệm vụ đề tài tốt nghiệp)

.....

.....

.....

.....

.....

.....

3. Cho điểm của cán bộ hướng dẫn:

( Điểm ghi bằng số và chữ )

.....

.....

Ngày.....tháng.....năm 2013

Cán bộ hướng dẫn chính

( Ký, ghi rõ họ tên )

**PHẦN NHẬN XÉT ĐÁNH GIÁ CỦA CÁN BỘ CHẤM PHẢN BIỆN ĐỀ  
TÀI TỐT NGHIỆP**

**1. Đánh giá chất lượng đề tài tốt nghiệp (về các mặt như cơ sở lý luận, thuyết minh chương trình, giá trị thực tế, ...)**

**2. Cho điểm của cán bộ phản biện**

*( Điểm ghi bằng số và chữ )*

.....  
.....

Ngày.....tháng.....năm 2013

Cán bộ chấm phản biện

*( Ký, ghi rõ họ tên )*



## MỤC LỤC

	Trang
<b>LỜI CẢM ƠN .....</b>	1
<b>DANH MỤC CHỮ VIẾT TẮT .....</b>	2
<b>PHẦN MỞ ĐẦU .....</b>	3
<b>CHƯƠNG 1: KHÁI QUÁT VỀ XỬ LÝ ẢNH VÀ BÀI TOÁN PHÁT HIỆN MẶT NGƯỜI.....</b>	5
<b>1.1. Khái quát về xử lý ảnh .....</b>	5
1.1.1. Một số khái niệm cơ bản.....	5
1.1.2. Một số vấn đề trong xử lý ảnh .....	6
1.1.2.1. Các hệ thống xử lý ảnh .....	6
1.1.2.2. Các hình thái của ảnh.....	8
1.1.3. Một số ứng dụng trong xử lý ảnh .....	9
<b>1.2. Bài toán nhận dạng mặt người.....</b>	10
1.2.1. Bài toán .....	10
1.2.2. Những khó khăn của nhận dạng khuôn mặt .....	11
1.2.3. Tầm quan trọng của bài toán nhận diện mặt người .....	14
1.2.4. Các ứng dụng đặc trưng của bài toán nhận diện mặt người.....	14
1.2.5. Xây dựng hệ thống nhận diện mặt người đặc trưng .....	15
1.2.6. Một số phương pháp nhận diện mặt người .....	16
1.2.6.1. Dựa trên tri thức .....	16
1.2.6.2. Hướng tiếp cận dựa trên đặc trưng không thay đổi .....	18
1.2.6.3. Hướng tiếp cận dựa trên so khớp mẫu .....	20
1.2.6.4. Hướng tiếp cận dựa trên diện mạo.....	21

1.2.7. Pháp phát hiện mặt người dựa trên đặc trưng lời.....	22
<b>CHƯƠNG 2: PHÁT HIỆN MẶT NGƯỜI DỰA VÀO ĐẶC TRƯNG LỜI.....</b>	<b>24</b>
<b>2.1. Đặc trưng lời .....</b>	<b>24</b>
2.1.1. Điểm lời .....	25
2.1.2. Dò tìm và lấy vùng lời .....	25
2.1.3. Dò tìm và phát hiện vùng lời ở nhiều mức khác nhau.....	28
<b>2.2. Phát hiện mặt người.....</b>	<b>31</b>
2.2.1. Xây dựng cấu trúc cây .....	31
2.2.2. Xây dựng hàm tính độ tương đồng giữa hai cây.....	34
2.2.2.1. Độ tương đồng giữa hai nút trên cây .....	34
2.2.2.2. Độ tương đồng giữa hai cây .....	36
2.2.2.3. Không gian cây và khoảng cách giữa hai cây .....	37
<b>Chương 3: CÀI ĐẶT THỬ NGHIỆM.....</b>	<b>39</b>
<b>3.1. Bài toán .....</b>	<b>39</b>
<b>3.2. Phân tích bài toán.....</b>	<b>39</b>
<b>3.3. Một số kết quả chương trình.....</b>	<b>39</b>
<b>PHẦN KẾT LUẬN.....</b>	<b>44</b>
<b>TÀI LIỆU THAM KHẢO .....</b>	<b>45</b>

## LỜI CẢM ƠN

Trước tiên em xin được bày tỏ sự trân trọng và lòng biết ơn đối với thầy giáo PGS.TS. Đỗ Năng Toàn giảng viên – Viện Khoa Công nghệ thông tin. Trong suốt thời gian học và làm đồ án tốt nghiệp, thầy đã dành rất nhiều thời gian quý báu để tận tình chỉ bảo, hướng dẫn, định hướng cho em thực hiện đồ án.

Tác giả xin được cảm ơn các thầy cô giáo Trường Đại học Dân lập Hải phòng đã giảng dạy trong quá trình học tập, thực hành, làm bài tập, giúp em hiểu thấu đáo hơn các nội dung học tập và những hạn chế cần khắc phục trong việc học tập, nghiên cứu và thực hiện bản đồ án này.

Em xin cảm ơn các bạn bè và nhất là các thành viên trong gia đình đã tạo mọi điều kiện tốt nhất, động viên, cổ vũ trong suốt quá trình học tập và đồ án tốt nghiệp.

*Hải Phòng, ngày tháng năm 2013.*

Sinh viên

Phạm Thành Huân

**DANH MỤC CHỮ VIẾT TẮT**

AD	Analog to Digital
Ppi	Pixel per inch
Dpi	Dot per inch
BMP	Bit map
GIF	Graphics Interchanger Format do hang ComputerServer Incorporated (Mỹ) đề xuất 1990.
JPEG	Joint Photograp Expert Group : tên của nhóm nghiên cứu các chuẩn nén cho ảnh, thành lập 1982. Tên cũ là IOS. JPEG chính thức thành lập năm 1986.
PEL	Picture Elenment
JPG	Joint Photographic Experts Group
1D	Một chiều (1 Dimention)
2D	Hai chiều (2 Dimentions)
3D	Ba chiều (3 Dimentions)
CNN	Mạng nơ ron tế bào (Cellular Neural Network)
S – N	South – North (Nam – Bắc)
E – W	East – West (Đông – Tây )
NW – SE	North West – South East (Tây Bắc – Đông Nam)
NE – SW	North East – South West (Đông Bắc – Tây Nam)

## PHẦN MỞ ĐẦU

Công nghệ thông tin ngày càng phát triển và có vai trò hết sức quan trọng không thể thiếu trong cuộc sống hiện đại. Con người ngày càng tạo ra những cỗ máy thông minh có khả năng tự nhận biết và xử lý được các công việc một cách tự động, phục vụ cho lợi ích của con người.

Trong những năm gần đây, một trong những bài toán nhận được nhiều sự quan tâm và tốn nhiều công sức nhất của lĩnh vực công nghệ thông tin, đó chính là bài toán nhận dạng. Tuy mới xuất hiện chưa lâu nhưng nó đã rất được quan tâm vì tính ứng dụng thực tế của bài toán cũng như sự phức tạp của nó.

Bài toán nhận dạng có rất nhiều lĩnh vực như: nhận dạng vật chất (nước, lửa, đất, đá, gỗ...) nhận dạng chữ viết, nhận dạng giọng nói, nhận dạng hình dáng, nhận dạng khuôn mặt.. trong đó phổ biến và được ứng dụng nhiều hơn cả là bài toán nhận diện khuôn mặt.

Để nhận dạng được khuôn mặt, bước đầu tiên để nhận dạng là phát hiện ra khuôn mặt, điều này thực sự là quan trọng và hết sức khó khăn. Cho đến tận bây giờ, các nhà nghiên cứu vẫn chưa đạt được sự ưng ý trong việc giải quyết các khó khăn của bài toán và cho kết quả hoàn toàn đúng. Tuy nhiên, những gì đã đạt được cũng đủ để chúng ta áp dụng rộng rãi và đem lại những lợi ích to lớn trong cuộc sống.

Với sự hấp dẫn của bài toán và những thách thức còn đang ở phía trước, với niềm đam mê công nghệ hiện đại và những ứng dụng thực tế tuyệt vời của nó, với khát khao khám phá và chinh phục những chi thức mới mẻ.. em đã chọn đề tài nghiên cứu: **PHÁT HIỆN MẶT NGƯỜI TRONG ẢNH** làm đề tài nghiên cứu và bảo vệ luận văn tốt nghiệp của mình.

*Về lý thuyết :*

- Tìm hiểu khái quát về xử lý ảnh bài toán phát hiện mặt người.
- Phát hiện khuôn mặt dựa vào đặc trưng lời.

*Về thực tiễn :*

- Cài đặt thử nghiệm chương trình đã tìm hiểu được.

Cấu trúc chính của đồ án gồm 3 chương :

Chương 1: Khái quát về xử lý ảnh và bài toán phát hiện mặt người

Trình bày khái quát về xử lý ảnh và bài toán nhận dạng mặt người.

Chương 2: Phát hiện khuôn mặt dựa vào đặc trưng lồi

Trình bày đặc trưng lồi và phát hiện mặt người..

Chương 3: Cài đặt thử nghiệm

Chương trình ứng dụng và một số kết quả thu được.

# CHƯƠNG 1. KHÁI QUÁT VỀ XỬ LÝ ẢNH VÀ BÀI TOÁN PHÁT HIỆN MẶT NGƯỜI

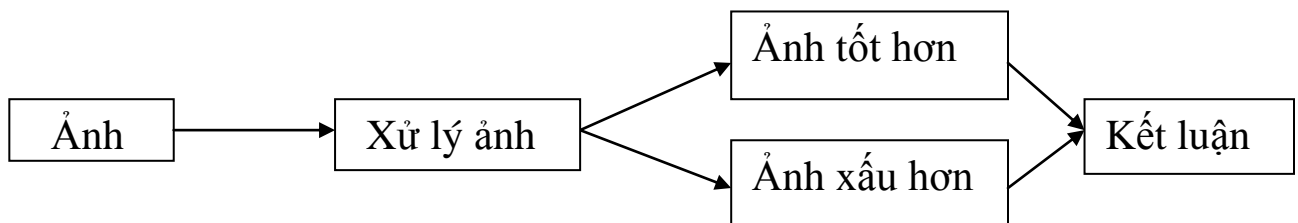
## 1.1. Khái quát về xử lý ảnh

### 1.1.1. Một số khái niệm cơ bản

Xử lý ảnh là một trong những mảng quan trọng nhất trong kỹ thuật thị giác máy tính, là tiền đề cho nhiều nghiên cứu thuộc lĩnh vực này. Hai nhiệm vụ cơ bản của quá trình xử lý ảnh là nâng cao chất lượng thông tin hình ảnh và xử lý số liệu cung cấp cho các quá trình khác trong đó có việc ứng dụng thị giác vào điều khiển.

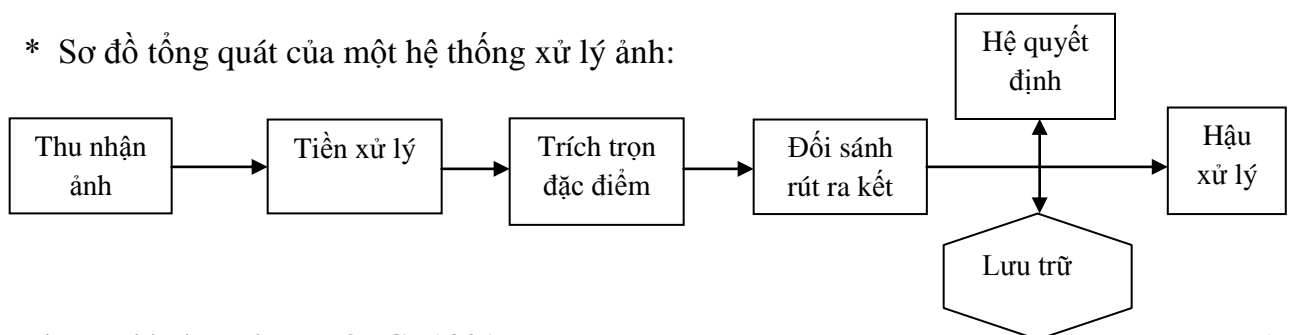
Quá trình bắt đầu từ việc thu nhận ảnh nguồn (từ các thiết bị thu nhận ảnh dạng số hoặc tương tự) gửi đến máy tính. Dữ liệu ảnh được lưu trữ ở định dạng phù hợp với quá trình xử lý. Người lập trình sẽ tác động các thuật toán tương ứng lên dữ liệu ảnh nhằm thay đổi cấu trúc ảnh phù hợp với các ứng dụng khác nhau.

Quá trình xử lý nhận dạng ảnh được xem như là quá trình thao tác ảnh đầu vào nhằm cho ra kết quả mong muốn. Kết quả đầu ra của một quá trình xử lý ảnh có thể 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 nào đó của đối tượng trong không gian và nó có thể xem như một hàm n biến  $P(c, c_1, c_2, \dots)$ . 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:



## 1.1.2. Một số vấn đề trong xử lý ảnh

### 1.1.2.1. Các hệ thống xử lý ảnh

#### \* Thu nhận ảnh

Đây là bước đầu tiên trong quá trình xử lý ảnh. Để thực hiện điều này ta cần có các thiết bị thu nhận ảnh bao gồm camera, scanner các thiết bị thu nhận này có thể cho ảnh đen trắng.

Các thiết bị thu nhận ảnh có 2 loại chính ứng với 2 loại ảnh thông dụng Raster và Vector. Các thiết bị thu nhận ảnh thông thường Raster là camera, còn các thiết bị thu nhận ảnh Vector là sensor hoặc bộ số hóa (digitalizer) hoặc được chuyển đổi từ ảnh Raster.

Các thiết bị thu ảnh thông thường gồm camera cộng với bộ chuyển đổi tương tự số AD (Analog to Digital) hoặc scanner chuyên dụng. Các thiết bị thu nhận ảnh này có thể cho ảnh đen trắng hoặc ảnh màu. Đầu ra của scanner là ảnh ma trận số mà ta quen gọi là bản đồ ảnh (ảnh Bitmap). Bộ số hoá (digitalizer) sẽ tạo ảnh vector có hướng. Nhìn chung, các hệ thống thu nhận ảnh thực hiện hai quá trình:

- Cảm biến : biến đổi năng lượng quang học thành năng lượng điện.
- Tổng hợp năng lượng điện thành ảnh

#### \* Tiền xử lý

Ở bước này, ảnh sẽ được cải thiện về độ tương phản, khử nhiễu, khôi phục ảnh, nắn chỉnh hình học... Với mục đích làm cho chất lượng ảnh trở nên tốt hơn nữa, chuẩn bị cho các bước xử lý phức tạp kế tiếp sau đó.

**Khử nhiễu:** Nhiễu được chia thành hai loại: nhiễu hệ thống và nhiễu ngẫu nhiên. Đặc trưng của nhiễu hệ thống là tính tuần hoàn. Do vậy, có thể khử nhiễu này bằng việc sử dụng phép biến đổi Fourier và loại bỏ các đỉnh điểm. Đối với nhiễu ngẫu nhiên, trường hợp đơn giản là các vết bản tương ứng với các điểm sáng hay tối, có thể khử nhiễu bằng phương pháp nội suy, lọc trung vị, lọc trung bình.

**Chỉnh mức xám:** Đây là kỹ thuật nhằm chỉnh sửa tính không đồng đều của thiết bị thu nhận hoặc độ tương phản giữa các vùng ảnh.



**Chỉnh tán xạ:** Ảnh thu nhận từ các thiết bị quang học hay điện tử có thể bị mờ, nhòe. Phương pháp biến đổi Fourier dựa trên tích chập của ảnh với hàm tán xạ cho phép giải quyết việc hiệu chỉnh này.

**Nấn chỉnh hình học:** Những biến dạng hình học thường do các thiết bị điện tử và quang học gây ra. Do đó phương pháp hiệu chỉnh hình ảnh dựa trên mô hình được mô tả dưới dạng phương trình biến đổi ảnh biến dạng  $f(x,y)$  thành ảnh lý tưởng  $f(x',y')$  như sau:

$$\begin{cases} x' = h_x(x,y) \\ y' = h_y(x,y) \end{cases} \quad (1.1)$$

Trong đó  $h_x, h_y$  là các phương trình tuyến tính (biến dạng do phối cảnh) hay bậc hai (biến dạng do ống kính camera).

#### \* Trích chọn đặc điểm

Vì lượng thông tin chứa trong ảnh là rất lớn, trong khi đó đa số ứng dụng chỉ cần một số thông tin đặc trưng nào đó, cần có bước trích chọn đặc điểm để giảm lượng thông tin không lờ ăy. Việc trích chọn hiệu quả đặc điểm 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.

Các đặc điểm của đối tượng được trích chọn tùy theo mục đích nhận dạng. Có thể nêu ra một số đặc điểm của ảnh sau đây:

**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,...)

**Đặc điểm biên và đường biên:** Đặc trưng cho đường biên của đối tượng và do rất hữu ích trong việc trích chọn các đặc tính bất biến được dung khi nhận dạng đối tượng. Các đặc điểm này có thể được trích chọn nhờ toán tử Gradien, toán tử Laplace, toán tử chéo không (zero crossing)..

#### \* Đối sánh, nhận dạng

Nhận dạng tự động (automatic recognition), mô tả đối tượng, phân loại và phân nhóm các mẫu là những vấn đề quan trọng trong thị giác máy, được ứng dụng trong nhiều ngành khoa học khác nhau. Ví dụ mẫu có thể là ảnh của vân tay, ảnh của một vật nào đó được chụp, một chữ viết, khuôn mặt người hoặc một ký đồ tín hiệu tiếng nói. Khi biết một mẫu nào đó, để nhận dạng hoặc phân loại mẫu đó.

Hệ thống nhận dạng tự động bao gồm ba khâu tương ứng với ba giai đoạn chủ yếu sau đây:

- Thu nhận dữ liệu và tiền xử lý.
- Biểu diễn dữ liệu.
- Nhận dạng, ra quyết định.

Bốn cách tiếp cận khác nhau trong lý thuyết nhận dạng là:

- Đối sánh mẫu dựa trên các đặc trưng được trích chọn.
- Phân loại thống kê.
- Đối sánh cấu trúc.
- Phân loại dựa trên mạng nơ-ron nhân tạo.

Trong các ứng dụng rõ ràng là không thể chỉ dùng có một cách tiếp cận đơn lẻ để phân loại “tối ưu” do vậy cần sử dụng cùng một lúc nhiều phương pháp và cách tiếp cận khác nhau. Do vậy, các phương thức phân loại tổ hợp hay được sử dụng khi nhận dạng và nay đã có những kết quả có triển vọng dựa trên thiết kế các hệ thống lai (hybrid system) bao gồm nhiều mô hình kết hợp.

#### **1.1.2.2. Các hình thái của ảnh**

\* Chuyển ảnh màu thành ảnh xám

Đơn vị tế bào của ảnh số là pixel. Tùy theo mỗi định dạng là ảnh màu hay ảnh xám mà từng pixel có thông số khác nhau. Đối với ảnh màu từng pixel sẽ mang thông tin của ba màu cơ bản tạo ra bản màu khả kiến là Đỏ (R), Xanh lá (G) và Xanh biển (B) [Thomas 1892]. Trong mỗi pixel của ảnh màu, ba màu cơ bản R, G và B được bố trí sát nhau và có cường độ sáng khác nhau. Thông thường, mỗi màu cơ bản được biểu diễn bằng tám bit tương ứng 256 mức độ màu khác nhau. Như vậy mỗi pixel chúng ta sẽ có  $28 \times 3 = 224$  màu (khoảng 16.78 triệu màu). Đối với ảnh xám, thông thường mỗi pixel

mang thông tin của 256 mức xám (tương ứng với tám bit) như vậy ảnh xám hoàn toàn có thể tái hiện đầy đủ cấu trúc của một ảnh màu tương ứng thông qua tám mặt phẳng bit theo độ xám.

\* Lược đồ xám của ảnh (Histogram)

Lược đồ xám của một ảnh số có các mức xám trong khoảng  $[0, L-1]$  là một hàm rời rạc  $p(r_k) = n_k/n$ . Trong đó  $n_k$  là số pixel có mức xám thứ  $r_k$ ,  $n$  là tổng số pixel trong ảnh và  $k=0, 1, 2, \dots, L-1$ . Do đó  $P(r_k)$  cho một xấp xỉ xác suất xảy ra mức xám  $r_k$ . Vẽ hàm này với tất cả các giá trị của  $k$  sẽ biểu diễn khái quát sự xuất hiện các mức xám của một ảnh. Chúng ta cũng có thể thể hiện lược đồ mức xám của ảnh thông qua tần suất xuất hiện mỗi mức xám qua hệ tọa độ vuông góc  $xOy$ . Trong đó, trục hoành biểu diễn số mức xám từ 0 đến  $N$  (số bit của ảnh xám). Trục tung biểu diễn số pixel của mỗi mức xám.

### 1.1.3. Một số ứng dụng trong xử lý ảnh

Như đã nói ở trên, các kỹ thuật xử lý ảnh trước đây chủ yếu được sử dụng để nâng cao chất lượng hình ảnh, chính xác hơn là tạo cảm giác về sự gia tăng chất lượng ảnh quang học trong mắt người quan sát. Thời gian gần đây, phạm vi ứng dụng xử lý ảnh mở rộng không ngừng, có thể nói hiện không có lĩnh vực khoa học nào không sử dụng các thành tựu của công nghệ xử lý ảnh số.

Trong y học các thuật toán xử lý ảnh cho phép biến đổi hình ảnh được tạo ra từ nguồn bức xạ X-ray hay nguồn bức xạ siêu âm thành hình ảnh quang học trên bề mặt film x-quang hoặc trực tiếp trên bề mặt màn hình hiển thị. Hình ảnh các cơ quan chức năng của con người sau đó có thể được xử lý tiếp để nâng cao độ tương phản, lọc, tách các thành phần cần thiết (chụp cắt lớp) hoặc tạo ra hình ảnh trong không gian ba chiều (siêu âm 3 chiều).

Trong lĩnh vực địa chất, hình ảnh nhận được từ vệ tinh có thể được phân tích để xác định cấu trúc bề mặt trái đất. Kỹ thuật làm nổi đường biên (image enhancement) và khôi phục hình ảnh (image restoration) cho phép nâng cao chất lượng ảnh vệ tinh và tạo ra các bản đồ địa hình 3-D với độ chính xác cao.

Trong ngành khí tượng học, ảnh nhận được từ hệ thống vệ tinh theo dõi thời tiết cũng được xử lý, nâng cao chất lượng và ghép hình để tạo ra ảnh bề mặt trái đất trên

một vùng rộng lớn, qua đó có thể thực hiện việc dự báo thời tiết một cách chính xác hơn.

Xử lý ảnh còn được sử dụng rộng rãi trong lĩnh vực hình sự và các hệ thống bảo mật hoặc kiểm soát truy cập: quá trình xử lý ảnh với mục đích nhận dạng vân tay hay khuôn mặt cho phép phát hiện nhanh các đối tượng nghi vấn cũng như nâng cao hiệu quả hệ thống bảo mật cá nhân cũng như kiểm soát ra vào. Ngoài ra, có thể kể đến các ứng dụng quan trọng khác của kỹ thuật xử lý ảnh tĩnh cũng như ảnh động trong đời sống như tự động nhận dạng, nhận dạng mục tiêu quân sự, máy nhìn công nghiệp trong các hệ thống điều khiển tự động, nén ảnh tĩnh, ảnh động để lưu và truyền trong mạng viễn thông. v. v. v...

## **1.2. Bài toán nhận dạng mặt người**

### **1.2.1. Bài toán nhận dạng mặt người**

Nhận diện mặt người là bài toán xuất hiện cách đây chưa lâu, ra đời chỉ vài chục năm nhưng đã có rất nhiều sự quan tâm, nghiên cứu về nó. Bài toán nhận dạng mặt người có rất nhiều ứng dụng thực tế trong cuộc sống, giải trí và an ninh bảo mật, chính vì thế bài toán nhận diện mặt người luôn nhận được nhiều sự quan tâm, có sức hấp dẫn, lôi cuốn mọi người ngày càng hoàn thiện nó.

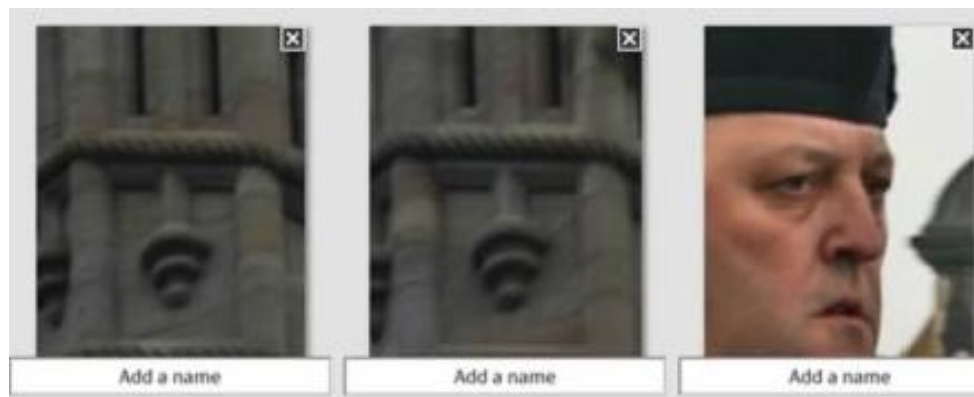
Các nghiên cứu ban đầu về bài toán nhận diện mặt người khá đơn giản với ảnh đen trắng và chụp thẳng và thường chỉ có một đến hai khuôn mặt trong một bức ảnh, về sau mở rộng ra với nhiều điều kiện phức tạp hơn như: ảnh màu, ảnh có nhiều khuôn mặt, ánh sáng trong ảnh không tốt hay điều kiện một trường xung quanh phức tạp. Bài toán nhận dạng mặt người cũng có nhiều hướng nghiên cứu khác nhau, nhưng cơ bản nhất có thể kể ra hai hướng cơ bản chính sau:

- Nhận diện một khuôn mặt trong một bức ảnh, so sánh với các mẫu có sẵn trong thư viện để tìm ra một người nào đó. Ứng dụng của nó có thể dùng lĩnh vực an ninh, truy tìm tội phạm.
- Xác định vị trí của nhiều khuôn mặt trong một bức ảnh, trong một bức ảnh có nhiều khuôn mặt với nhiều kích thước và sáng tối khác nhau, bài toán sẽ tìm và chỉ ra vị trí các khuôn mặt có trong bức ảnh. Ứng dụng của nó có thể dùng trong lĩnh vực giải trí.

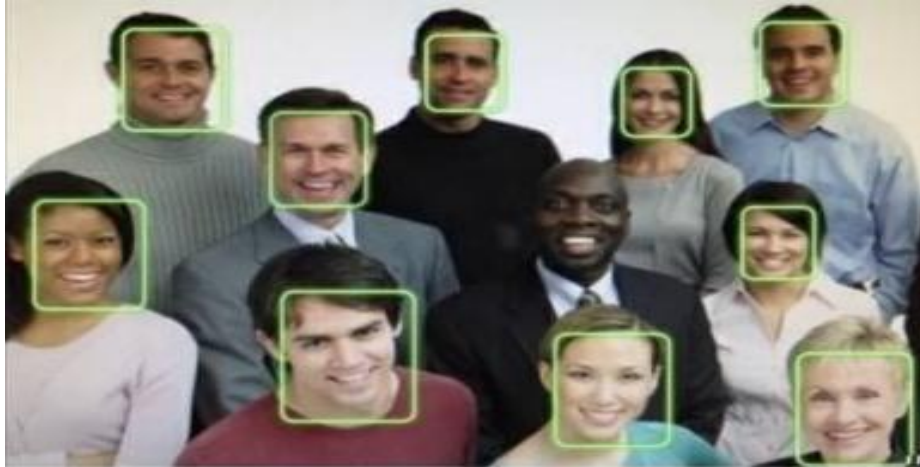
### 1.2.2. Những khó khăn của nhận dạng khuôn mặt

Bài toán nhận diện mặt người trong ảnh là một trong những bài toán phức tạp và gặp nhiều khó khăn nhất trong các bài toán nhận diện. Có thể kể đến các khó khăn thường gặp nhất trong bài toán nhận diện khuôn mặt người như:

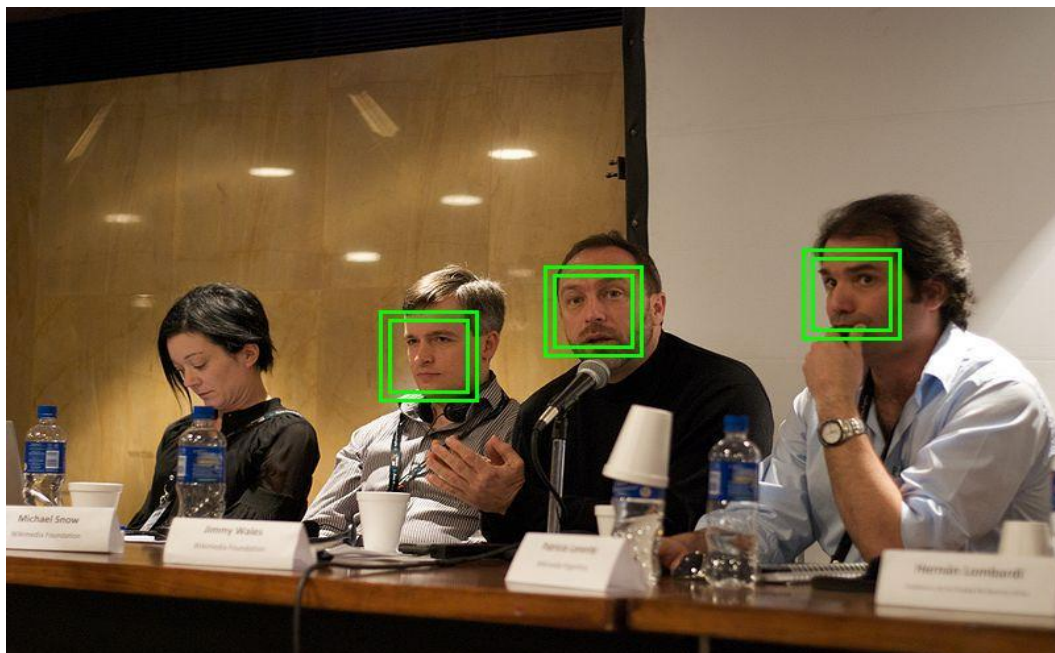
**Nền ảnh phức tạp:** nền của ảnh phức tạp là một trong những khó khăn nhất trong bài toán nhận diện khuôn mặt người trong ảnh, khuôn mặt người sẽ dễ bị nhầm lẫn với nhiều khung cảnh phức tạp xung quanh và ảnh hưởng rất nhiều đến quá trình phân tích và rút trích các đặc trưng của khuôn mặt trong ảnh, có thể dẫn đến không nhận ra khuôn mặt hoặc là nhận nhầm các khung cảnh xung quanh thành khuôn mặt người



**Màu sắc của da mặt:** màu sắc của da mặt quá tối hoặc gần với màu sắc của khung cảnh môi trường xung quanh cũng là một khó khăn với bài toán nhận diện mặt người. Nếu màu sắc của da người quá tối thì thuật toán sẽ gặp khó khăn trong việc nhận diện các đặc trưng và có thể không tìm ra được khuôn mặt người



**Góc chụp khuôn mặt:** khuôn mặt đang nhìn thẳng nhưng góc chụp của máy ảnh lại lệch nhiều so với hướng nhìn thẳng của khuôn mặt hoặc lúc chụp ảnh mặt người quay nghiêng sang một bên đó nhiều đều là những khó khăn rất lớn trong bài toán nhận diện mặt người. Hầu hết các ứng dụng hiện tại đều chỉ áp dụng cho bức ảnh có khuôn mặt nhìn thẳng hoặc gần như nhìn thẳng mà bỏ qua những khuôn mặt bị nghiêng quá nhiều, khi đó một số bộ phận trên khuôn mặt sẽ bị che khuất hoặc là biến dạng, gây khó khăn cho quá trình tìm kiếm và rút trích các đặc trưng



**Sự thể hiện cảm xúc quá nhiều và phụ kiện đi kèm:** sự thể hiện cảm xúc quá nhiều sẽ làm biến đổi khuôn mặt so với bình thường khá nhiều, gây ra không ít khó khăn cho việc nhận diện khuôn mặt một cách đơn giản, tuy nhiên nó cũng là thách thức để phát triển nên các hệ thống thông minh nhận biết cảm xúc của con người. ngoài ra việc dùng nhiều các phụ kiện đi kèm trang trí trên khuôn mặt cũng là một khó khăn đối với việc nhận biết khuôn mặt, ví dụ một người xăm đầy mặt và đeo nhiều khuyên trên mặt, tóc tai bù xù...



**Điều kiện ánh sáng và môi trường xung quanh:** điều kiện của môi trường lúc chụp ảnh có ảnh hưởng rất nhiều đến quá trình phân tích và nhận biết các đặc trưng của khuôn mặt. Trong đó ánh sáng là một phần rất quan trọng để nhận biết được khuôn mặt dễ dàng, những bức ảnh chụp trong điều kiện thiếu hoặc thừa ánh sáng quá nhiều đều gây khó khăn rất nhiều, ngoài ra các điều kiện về thời tiết ( mưa to, tuyết rơi...) đều gây ảnh hưởng tới quá trình phân tích và nhận biết khuôn mặt người.

Ngoài ra còn các yếu tố khác như: mặt người bị che khuất bởi các đối tượng khác có trong ảnh, các nét mặt khác nhau như: vui, buồn, ngạc nhiên...



### 1.2.3. Tầm quan trọng của bài toán nhận diện mặt người

Nếu như công nghệ nhận dạng giọng nói phù hợp với các ứng dụng call-center và nhược điểm của nó là tiếng ồn, không phù hợp với nơi công cộng đông người thì công nghệ nhận dạng chữ ký cũng gây nhiều phiền phức cho người sử dụng vì khó duy trì được chữ ký giống nhau ngay trong cùng một thời điểm. Trong khi đó, công nghệ nhận dạng vân tay hiện đã được dùng khá phổ biến, nhưng cũng có nhược điểm là bị ảnh hưởng bởi độ ẩm da và đặc biệt sẽ không chính xác cao với người có tay hay tiếp xúc với hoá chất. Bên cạnh đó, nhận dạng bằng vân tay còn có những hạn chế về việc giả mạo, thậm chí kẻ gian có thể sử dụng tay của người khác để xác thực. Còn công nghệ nhận dạng mống mắt là một công nghệ có nhiều ưu điểm về độ chính xác cũng như khả năng chống giả mạo. Việc xác thực, nhận dạng trong công nghệ này dựa trên sơ đồ mạch máu trong võng mạc mắt. Sơ đồ này rất ít thay đổi kể từ lúc mới sinh cho tới khi già và đặc biệt là khi chết, hệ thống sơ đồ mạch máu này sẽ biến mất. Vì thế, không thể dùng người chết để xác thực cho việc truy cập trái phép. Tuy nhiên, công nghệ này lại bị tác động bởi nhiều yếu tố khác như độ rộng của mắt, lông mi, kính đeo và khó triển khai phổ biến trên diện rộng do độ phức tạp của các thiết bị.

Trong khi đó, mặc dù độ chính xác không là phải ưu điểm nổi bật của công nghệ nhận dạng mặt người song không giống với những phương pháp khác, nhận dạng mặt người là công nghệ nhận dạng không cần đến sự tiếp xúc trực tiếp giữa đối tượng và thiết bị thu nhận. Thay vào đó, công nghệ này mang tính theo dõi, giám sát và rất thuận tiện cho những ứng dụng bảo vệ mục tiêu, chống khủng bố tại những điểm công cộng đông người. Đây cũng là ưu điểm nổi trội của nhận dạng mặt người mà các công nghệ nhận dạng khác khó có thể có được.

### 1.2.4. Các ứng dụng đặc trưng của bài toán nhận diện mặt người

Bài toán phát hiện mặt người trong ảnh là một bài toán phức tạp nhưng lại hấp dẫn rất nhiều người nghiên cứu về nó, có lẽ bởi tính ứng dụng thực tế của bài toán là rất lớn và đa dạng trong cuộc sống hiện đại. Có thể kể ra một số ứng dụng thông dụng nhất của bài toán nhận diện mặt người:

- **Hệ thống giao tiếp thông minh giữa người và máy:** con người có thể xây dựng những hệ thống thông giao tiếp giữa người và máy tính thông qua việc nhận diện khuôn mặt, biểu cảm trên khuôn mặt người để dự



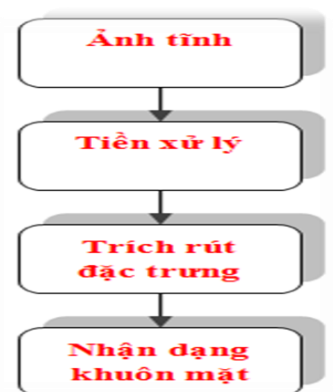
đoán, nhận biết trạng thái tâm lí hiện thời của người đó. Một ngôi nhà thông minh trong tương lai có thể nhận biết được chủ nhân của nó thông qua nhận biết khuôn mặt, dáng người, giao tiếp qua giọng nói, vẫy tay...

- **Nhận dạng tội phạm:** hệ thống có thể nhận diện ra một khuôn mặt ngay tức thì và đối chiếu với hàng triệu bản ghi có sẵn trong cơ sở dữ liệu để chỉ ra đó có thể là một tội phạm đang truy tìm hay không, hoặc đó có thể là một nhân vật nào đó đặc biệt cần quan tâm..vv..
- **Giải trí:** trong hầu hết các máy ảnh hiện đại ngày nay đều có chức năng tự động nhận diện mặt người để có thể lấy độ nét, điều chỉnh ánh sáng cho phù hợp với khung cảnh xung quanh. Trên một số trang web cũng đã áp dụng công nghệ tự động nhận diện mặt người và so sánh với kho dữ liệu khổng lồ của mình để đưa ra những lời chào, dịch vụ thông minh nhất cho người sử dụng..
- **Hệ thống quan sát,** theo dõi và bảo vệ: các hệ thống camera sẽ xác định đâu là con người và theo dõi người đó...
- **Điều khiển ra vào các cơ quan,** văn phòng: cho phép nhân viên ra vào các khu vực quan trọng mà không cần phải đăng nhập hay dùng thẻ. Nếu kết hợp với sử dụng vẫy tay hay hốc mắt thì sẽ đem lại kết quả chính xác cao.
- **Tổ chức tìm kiếm** liên quan đến con người thông qua khuôn mặt trên nhiều hệ cơ sở dữ liệu lớn
- **Phân tích cảm xúc** của khuôn mặt người..v..v..

### 1.2.5. Xây dựng hệ thống nhận diện mặt người đặc trưng

Một hệ thống nhận diện khuôn mặt người có thể được xây dựng qua rất nhiều công đoạn khác nhau và rất phức tạp, nhưng ta có thể khái quát chung mọi hệ thống nhận dạng khuôn mặt người gồm có 3 bước cơ bản sau:

- **Tiền xử lý:** hệ thống nhận vào một ảnh tĩnh, sau đó có thể xử lý ảnh cho chất lượng tốt hơn, như chỉnh lại độ sáng,



giảm độ nhiễu .. điều này giúp cho quá trình rút trích các đặc trưng của khuôn mặt được dễ dàng hơn rất nhiều. Quá trình tiền xử lý thường khá đơn giản và nhanh gọn nên không cần dùng các thuật toán phức tạp và mất nhiều thời gian.

- Trích rút các đặc trưng: Trích rút đặc trưng là kỹ thuật sử dụng các thuật toán để lấy ra những thông tin mang những đặc điểm riêng biệt của một người
- Nhận diện khuôn mặt: sau khi đã rút trích các đặc trưng, sẽ đưa vào khối nhận dạng để phân lớp đối tượng.

### 1.2.6. Một số phương pháp nhận diện mặt người

Hiện tại có khá nhiều phương pháp nhằm xác định khuôn mặt người trong ảnh, từ ảnh có chất lượng không tốt cho đến tốt, không màu(đen trắng) đến có màu.. Nhưng có thể phân chia các phương pháp này thành bốn hướng tiếp cận chính sau:

- Hướng tiếp cận dựa trên tri thức: Mã hóa các hiểu biết của con người về khuôn mặt thành các luật. các luật này thường mô tả quan hệ của các đặc trưng của khuôn mặt
- Hướng tiếp cận dựa trên diện mạo: hướng tiếp cận này dùng một tập ảnh huấn luyện cho trước, rồi sau đó hệ thống sẽ xác định khuôn mặt người, do đó nó còn có tên là: hướng tiếp cận dựa trên phương pháp học.
- Hướng tiếp cận dựa trên so sánh khớp mẫu: Dùng các mẫu chuẩn của khuôn mặt người để mô tả cho khuôn mặt hay các đặc trưng của khuôn mặt.
- Hướng tiếp cận dựa trên các đặc trưng không thay đổi: 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, điều kiện ánh sáng hay vị trí đặt máy ảnh thay đổi

#### 1.2.6.1. Dựa trên tri thức

Trong hướng tiếp cận này, các luật sẽ phụ thuộc rất lớn vào tri thức của những tác giả nghiên cứu về bài toán xác định khuôn mặt người. Đây là hướng tiếp cận dạng top-down. Dễ dàng xây dựng các luật cơ bản để mô tả các đặc trưng của khuôn mặt và các quan hệ tương ứng. Ví dụ, một khuôn mặt thường có hai mắt đối xứng nhau qua trục thẳng đứng ở giữa khuôn mặt và có một mũi, một miệng. Các quan hệ của các đặc trưng có thể được mô tả như quan hệ về khoảng cách và vị trí. Thông thường sẽ trích đặc trưng của khuôn mặt trước tiên để có được các ứng viên, sau đó các ứng viên này

sẽ được xác định thông qua các luật để biết ứng viên nào là khuôn mặt và ứng viên nào không phải khuôn mặt. Thường áp dụng quá trình xác định để giảm số lượng xác định sai.

Một vấn đề khá phức tạp khi dùng hướng tiếp cận này là làm sao chuyển từ tri thức con người sang các luật một cách hiệu quả. Nếu các luật này quá chi tiết thì khi xác định có thể xác định thiếu các khuôn mặt có trong ảnh, vì những khuôn mặt này không thể thỏa mãn tất cả các luật đưa ra. Nhưng các luật tổng quát quá thì có thể chúng ta sẽ xác định lầm một vùng nào đó không phải là khuôn mặt mà lại xác định là khuôn mặt. Và cũng khó khăn mở rộng yêu cầu từ bài toán để xác định các khuôn mặt có nhiều tư thế khác nhau.

Hai tác giả Yang và Huang dùng một phương thức theo hướng tiếp cận như sau để xác các khuôn mặt. Hệ thống của hai tác giả này bao gồm ba mức luật. Ở mức cao nhất, dùng một khung cửa sổ quét trên ảnh và thông qua một tập luật để tìm các ứng viên có thể là khuôn mặt. Ở mức kế tiếp, hai ông dùng một tập luật để mô tả tổng quát hình dáng khuôn mặt. Còn ở mức cuối cùng lại dùng một tập luật khác để xem xét ở mức chi tiết các đặc trưng khuôn mặt. Một hệ thống đa độ phân giải có thứ tự được dùng để xác định. Các luật ở mức cao nhất để tìm ứng viên như: “vùng trung tâm khuôn mặt có bốn phần với một mức độ đều cơ bản”, “phân xung quanh bên trên của một khuôn mặt có một mức độ đều cơ bản”, và “mức độ khác nhau giữa các giá trị xám trung bình của phần trung tâm và phần bao bên trên là đáng kể”. Độ phân giải thấp nhất của ảnh dùng để tìm ứng viên khuôn mặt mà còn tìm ở các mức phân giải tốt hơn. Ở mức hai, xem xét biểu đồ histogram của các ứng viên để loại bớt ứng viên nào không phải là khuôn mặt, đồng thời dò ra cạnh bao xung quanh ứng viên. Ở mức cuối cùng, những ứng viên nào còn lại sẽ được xem xét các đặc trưng của khuôn mặt về mắt và miệng. Hai ông đã dùng một chiến lược “từ thô đến mịn” hay “làm rõ dần” để giảm số lượng tính toán trong xử lý. Mặc dù tỷ lệ chính xác chưa cao, nhưng đây là tiền đề cho nhiều nghiên cứu sau này

Hai tác giả Kotropoulos và Pitas đưa một phương pháp tương tự dùng trên độ phân giải thấp. Hai ông dùng phương pháp chiếu để xác định các đặc trưng khuôn mặt, Kanade đã thành công với phương pháp chiếu để xác định biên của khuôn mặt. Với

$I(x,y)$  là giá trị xám của một điểm trong ảnh có kích thước  $m \times n$  ở tại vị trí  $(x,y)$ , các hàm để chiếu ảnh theo phương ngang và thẳng đứng được định nghĩa như sau:

$$\begin{cases} HI(x) = \sum_{y=1}^n I(x, y) \\ VI(y) = \sum_{x=1}^m I(x, y) \end{cases} \quad (1.2)$$

Dựa trên biểu đồ hình chiếu ngang, có hai cực tiểu địa phương khi hai ông xét quá trình thay đổi độ dốc của HI, đó chính là cạnh bên trái và phải của hai bên đầu. Tương tự với hình chiếu dọc VI, các cực tiểu địa phương cũng cho ta biết vị trí miệng, đỉnh mũi, và hai mắt. Các đặc trưng này đủ để xác định khuôn mặt.

#### 1.2.6.2. Hướng tiếp cận dựa trên đặc trưng không thay đổi.

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

#### ***Các đặc trưng khuôn mặt :***

Sirohey đưa một phương pháp xác định khuôn mặt từ một ảnh có hình nền phức tạp. Phương pháp dựa trên cạnh (dùng phương pháp Candy và heuristics để loại bỏ các cạnh để còn lại duy nhất một đường bao xung quanh khuôn mặt. Một hình ellipse dùng để bao khuôn mặt, tách biệt vùng đầu và hình nền. Tỷ lệ chính xác của Phạm Thành Huân – Lớp CT1301

thuật toán là 80%. Cũng dùng phương pháp cạnh như Sirohey, Chetverikov và Lerch dùng một phương pháp dựa trên blob và streak (hình dạng giọt nước và sọc xen kẽ), để xác định theo hướng các cạnh. Hai ông dùng hai blob tối và ba blob sáng để mô tả hai mắt, hai bên gò má, và mũi. Mô hình này dùng các streak để mô tả hình dáng ngoài của khuôn mặt, lông mày, và môi. Dùng ảnh có độ phân giải thấp theo biến đổi Laplace để xác định khuôn mặt thông qua blob.

Leung trình bày một mô hình xác suất để xác định khuôn mặt ở trong ảnh có hình nền phức tạp trên cơ sở một bộ xác định đặc trưng cục bộ và so khớp đồ thị ngẫu nhiên. Dùng năm đặc trưng (hai mắt, hai lỗ mũi, phần nối giữa mũi và miệng) để mô tả một khuôn mặt. Luôn tính quan hệ khoảng cách với các đặc trưng cặp (như mắt trái, mắt phải), dùng phân bố Gauss để mô hình hóa. Một mẫu khuôn mặt được đưa ra thông qua trung bình tương ứng cho một tập đa hướng, đa tỷ lệ của bộ lọc đạo hàm Gauss. Từ một ảnh, các đặc trưng ứng viên được xác định bằng cách so khớp từng điểm ảnh khi lọc tương ứng với vector mẫu (tương tự mỗi tương quan), chọn hai ứng viên đặc trưng đứng đầu để tìm kiếm cho các đặc trưng khác của khuôn mặt. Giống như xây dựng một đồ thị quan hệ mỗi node của đồ thị tương ứng như các đặc trưng của một khuôn mặt, đưa xác suất vào để xác định. Tỷ lệ xác định chính xác là 86%.

### ***Kết cấu khuôn mặt:***

Khuôn mặt con người có những kết cấu riêng biệt mà có thể dùng để phân loại so với các đối tượng khác. Augusteijn và Skufca cho rằng hình dạng của khuôn mặt dùng làm kết cấu phân loại, gọi là kết cấu giống khuôn mặt (face-like texture). Tính kết cấu qua các đặc trưng thống kê thứ tự thứ hai (SGLD) trên vùng có kích thước  $16 \times 16$  điểm ảnh. Có ba loại đặc trưng được xem xét: màu da, tóc, và những thứ khác. Hai ông dùng mạng neural về mối tương quan cascade cho phân loại có giám sát các kết cấu và một ánh xạ đặc trưng tự tổ chức Kohonen để gom nhóm các lớp kết cấu khác nhau. Hai tác giả đề xuất dùng phương pháp bầu cử khi không quyết định được kết cấu đưa vào là kết cấu của da hay kết cấu của tóc.

Manian và Ross dùng biến đổi wavelet để xây dựng tập dữ liệu kết cấu của khuôn mặt trong ảnh xám thông qua nhiều độ phân giải khác nhau kết hợp xác suất thông kê để xác định khuôn mặt người. Mỗi mẫu sẽ có chín đặc trưng. Tỷ lệ chính xác là 87%, tỷ lệ xác định sai là 18%.

***Sắc màu của da:***

Thông thường các ảnh màu không xác định trực tiếp trên toàn bộ dữ liệu ảnh mà các tác giả dùng tính chất sắc màu của da người (khuôn mặt người) để chọn ra được các ứng viên có thể là khuôn mặt người (lúc này dữ liệu đã thu hẹp đáng kể) để xác định khuôn mặt người. Tôi sẽ trình bày chi tiết về mô hình hóa màu da người ở một bài sau.

***Đặc trưng***

Gần đây có nhiều nghiên cứu sử dụng các đặc trưng toàn cục như: màu da người, kích thước, và hình dáng để tìm các ứng viên khuôn mặt, rồi sau đó sẽ xác định ứng viên nào là khuôn mặt thông qua dùng các đặc trưng cục bộ (chi tiết) như: mắt, lông mày, mũi, miệng, và tóc. Tùy mỗi tác giả sẽ sử dụng tập đặc trưng khác nhau.

**1.2.6.3. Hướng tiếp cận dựa trên so khớp mẫu**

Trong so khớp mẫu, các mẫu chuẩn của khuôn mặt (thường là khuôn mặt được chụp thẳng) sẽ được xác định trước hoặc xác định các tham số thông qua một hàm. Từ một ảnh đưa vào, tính các giá trị tương quan so với các mẫu chuẩn về đường viền khuôn mặt, mắt, mũi và miệng. Thông qua các giá trị tương quan này mà các tác giả quyết định có hay không có tồn tại khuôn mặt trong ảnh. Hướng tiếp cận này có lợi thế là rất dễ cài đặt, nhưng không hiệu quả khi tỷ lệ, tư thế, và hình dáng thay đổi. Nhiều độ phân giải, đa tỷ lệ, các mẫu con, và các mẫu biến dạng được xem xét thành bất biến về tỷ lệ và hình dáng.

***Hướng tiếp cận này có thể đi theo phương pháp sau.***

***Xác định mẫu trước***

Tác giả Sakai đã cố gắng thử xác định khuôn mặt người chụp thẳng trong ảnh. Ông dùng vài mẫu con về mắt, mũi, miệng, và đường viền khuôn mặt để mô hình hóa một khuôn mặt. Mỗi mẫu con được định nghĩa trong giới hạn của các đoạn thẳng. Các đường thẳng trong ảnh được trích bằng phương pháp xem xét thay đổi gradient nhiều nhất và so khớp các mẫu con. Đầu tiên tìm các ứng viên thông qua mối tương quan giữa các ảnh con và các mẫu về đường viền. Sau đó, so khớp với các mẫu con khác. Hay nói một cách khác, giai đoạn đầu xem như là giai đoạn sơ chế để tìm ứng viên,

giai đoạn thứ hai là giai đoạn tinh chế để xác định có tồn tại hay không một khuôn mặt người. Ý tưởng này được duy trì cho đến các nghiên cứu sau này.

Còn tác giả Craw đưa ra một phương pháp xác định khuôn mặt người dựa vào các mẫu về hình dáng của các ảnh được chụp thẳng (dùng vẽ bề ngoài của hình dáng khuôn mặt). Đầu tiên dùng phép lọc Sobel để tìm các cạnh. Các cạnh này sẽ được nhóm lại theo một số ràng buộc. Sau đó, tìm đường viền của đầu, quá trình tương tự được lặp đi lặp lại với mỗi tỷ lệ khác nhau để xác định các đặc trưng khác như: mắt, lông mày, và môi. Sau đó Craw mô tả một phương thức xác định dùng một tập có 40 mẫu để tìm các đặc trưng khuôn mặt và điều khiển chiến lược dò tìm.

### *Các mẫu bị biến dạng*

Tác giả Yuille dùng các mẫu biến dạng để mô hình hóa các đặc trưng của khuôn mặt, mô hình này có khả năng linh hoạt cho các đặc trưng khuôn mặt. Trong hướng tiếp cận này, các đặc trưng khuôn mặt được mô tả bằng các mẫu được tham số hóa. Một hàm năng lượng (giá trị) được định nghĩa để liên kết các cạnh, đỉnh, và thung lũng trong ảnh để tương ứng với các tham số trong mẫu. Mô hình này tốt nhất khi tối thiểu hàm năng lượng qua các tham số, Mặc dù kết quả tốt với mẫu biến dạng trong theo vết đối tượng trên đặc trưng không mô hình theo lưới, một hạn chế của hướng tiếp cận này là các mẫu biến dạng phải được khởi tạo trong phạm vi gần các đối tượng để xác định.

Hai tác giả Huang và Su dùng lý thuyết dòng chảy để xác định đường viền khuôn mặt dựa trên đặc tính hình học. Hai ông dùng lý thuyết tập đồng mức (Level Set) để loang từ các khởi động ban đầu để có được các khuôn mặt người.

### **1.2.6.4. Hướng tiếp cận dựa trên diện mạo**

Trái ngược với các phương pháp so khớp mẫu với các mẫu đã được định nghĩa trước bởi những chuyên gia, các mẫu trong hướng tiếp cận này được học từ các ảnh mẫu. Một cách tổng quát, các phương pháp theo hướng tiếp cận này áp dụng các kỹ thuật theo hướng xác suất thống kê và máy học để tìm những đặc tính liên quan của khuôn mặt và không phải là khuôn mặt. Các đặc tính đã được học ở trong hình thái các mô hình phân bố hay các hàm biệt số nên dùng có thể dùng các đặc tính này để xác

định khuôn mặt người. Đồng thời, bài toán giảm số chiều thường được quan tâm để tăng hiệu quả tính toán cũng như hiệu quả xác định.

Có nhiều phương pháp áp dụng xác suất thống kê để giả quyết. Một ảnh hay một vector đặc trưng xuất phát từ một ảnh được xem như một biến ngẫu nhiên  $x$ , và biến ngẫu nhiên có đặc tính là khuôn mặt hay không phải khuôn mặt bởi công thức tính theo các hàm mật độ phân lớp theo điều kiện.

$P(x | \text{khuôn mặt})$  và  $P(x | \sim \text{khuôn mặt})$

Có thể dùng phân loại Bayes hoặc khả năng cực đại để phân loại một ứng viên là khuôn mặt hay không phải là khuôn mặt. Không thể cài đặt trực tiếp phân loại Bayes bởi vì số chiều của  $x$  khá cao, bởi vì  $P(x | \text{khuôn mặt})$  và  $P(x | \text{không phải khuôn mặt})$  là đa phương thức, và chưa thể hiểu nếu xây dựng các dạng tham số hóa một cách tự nhiên cho  $P(x | \text{khuôn mặt})$  và  $P(x | \text{không phải khuôn mặt})$ . Có khá nhiều nghiên cứu theo hướng tiếp cận này quan tâm xấp xỉ có tham số hay không có tham số cho  $P(x | \text{khuôn mặt})$  và  $P(x | \text{không phải khuôn mặt})$ .

Các tiếp cận khác trong hướng tiếp cận dựa trên diện mạo là tìm một hàm biệt số (như: mặt phẳng quyết định, siêu phẳng để tách dữ liệu, hàm ngưỡng) để phân biệt hai lớp dữ liệu: khuôn mặt và không phải khuôn mặt. Bình thường, các mẫu ảnh được chiếu vào không gian có số chiều thấp hơn, rồi sau đó dùng một hàm biệt số (dựa trên các độ đo khoảng cách) để phân loại, hoặc xây dựng mặt quyết định phi tuyến bằng mạng neural đa tầng. Hoặc dùng SVM (Support Vector Machine) và các phương thức kernel, chiếu hoàn toàn các mẫu vào không gian có số chiều cao hơn để dữ liệu bị rời rạc hoàn toàn và ta có thể dùng một mặt phẳng quyết định phân loại các mẫu khuôn mặt và không phải khuôn mặt.

### 1.2.7. Pháp phát hiện mặt người dựa trên đặc trưng lồi

Năm 1983, James L. Crowley đã đưa ra khái niệm ridge và peak.

Ridge là các điểm lồi trên ảnh. Tập các điểm ridge trên ảnh sẽ tạo thành những đường xương sống và các đường này sẽ tạo thành các chòm tia gọi là peak (chùm). James L. Crowley sử dụng phép hiệu của lọc Low Pass để rút ra các điểm ridge (lồi) trên ảnh, và sau đó một thuật toán duyệt để kết chúng lại với nhau thành các đặc trưng ridge và peak. Tiếp sau đó, một số các phương pháp được đề xuất để rút trích các thông



tin về ridge. Hầu hết các phương pháp này dùng một bộ lọc để tăng cường thông tin về cạnh trên ảnh sau đó dò tìm quỹ tích các điểm cực trị. Quỹ tích các điểm cực trị này được xem là các đặc trưng lồi và lõm. Trong đặc trưng lồi và lõm được dò tìm trên nhiều độ phân giải khác nhau, cho kết quả đáng kể. Tuy nhiên, chất lượng của các đặc trưng lồi lõm này còn phụ thuộc khá nhiều vào điều kiện chiếu sáng cũng như các mức phân giải được lựa chọn trước. Để giải quyết sự phụ thuộc này, Lindeberg đã đề xuất một phương pháp tự động chọn độ phân giải tốt nhất. Gần đây Hải Trần đưa ra một cách tiếp cận sử dụng Laplacian để tăng cường thông tin cạnh sau đó dò tìm ridge và peak dưới nhiều mức khác nhau. Trong hướng tiếp cận này, các đặc trưng ridge và peak được biểu diễn dưới dạng các đồ thị quan hệ hoặc các cây cấp bậc với các tầng là các ridge và peak được dò tìm tại các mức khác nhau.

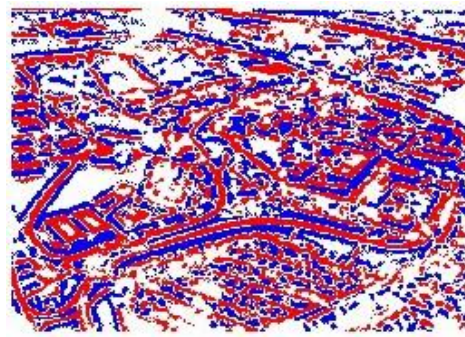
## CHƯƠNG 2. PHÁT HIỆN MẶT NGƯỜI DỰA VÀO ĐẶC TRƯNG LỖI

### 2.1. Đặc trưng lỗi

Ta có thể hiểu một cách đơn giản rằng: điểm lỗi là những điểm cao nhất... ví dụ là đỉnh của những dãy núi là điểm lỗi. Trong ảnh vệ tinh, lỗi là đường xương sống của dãy núi. Trong ảnh y tế, lỗi có thể là các đường mạch máu hay các cơ quan nhô cao. Trong ảnh khuôn mặt, mũi và các phần nhô của hai má được coi như là lỗi.

Tác giả Hải Trần sử dụng một bộ lọc Laplacian-of-Gaussian, sau đó áp dụng một thuật toán để dò tìm quỹ tích của các điểm cực trị. Hình 2.1 biểu diễn các đặc trưng lỗi, các điểm lỗi được tô màu đỏ. Chúng ta lưu ý rằng các điểm lỗi rút trích được chỉ biểu diễn được các vùng lỗi nhỏ, chi tiết, chẳng hạn như: chân mày, mũi, và miệng.

Trong những phần tiếp theo, chúng ta sẽ dò tìm các đặc trưng lỗi tại các mức khác nhau. Ví dụ như, tại mức thấp nhất, cần dò tìm các đặc trưng lỗi có kích thước nhỏ và chi tiết như mắt, mũi, v.v... Trong những mức độ cao hơn kế tiếp, sẽ dò tìm các vùng lỗi lớn hơn. Đến mức độ lớn nhất, toàn bộ khuôn mặt có thể trở thành một vùng lỗi. Nói cách khác, ở mức lớn sẽ dò được vùng lỗi kích thước lớn và ít chi tiết hơn. Cũng giống như khi nhìn một người từ rất xa, ta chỉ có thể nhìn được dáng, gần hơn một tí ta có thể nhìn được mặt nhưng không nhìn được các chi tiết mắt, mũi, miệng trên khuôn mặt, và gần hơn nữa ta mới có thể nhìn chi tiết mắt mũi, miệng. Việc này tương ứng với cách tiếp cận phân tích nhiều mức khác nhau.



*Hình 2.1 Dò tìm thông tin lỗi và lõm*

**2.1.1. Điểm lồi**

Gọi  $I(x, y)$  là một hàm hai biến biểu diễn một ảnh  $I$ . Hàm số này thiết lập một mặt  $(x, y, I(x, y))$ . Mục tiêu của chúng ta là dò tìm các điểm có khả năng là điểm lồi trên mặt này. Một cách chính xác theo toán học, điểm lồi là những điểm mà tại đó cả hai đạo hàm:

$$\frac{\partial^2 I}{\partial x^2} \text{ và } \frac{\partial^2 I}{\partial y^2} \tag{2.1}$$

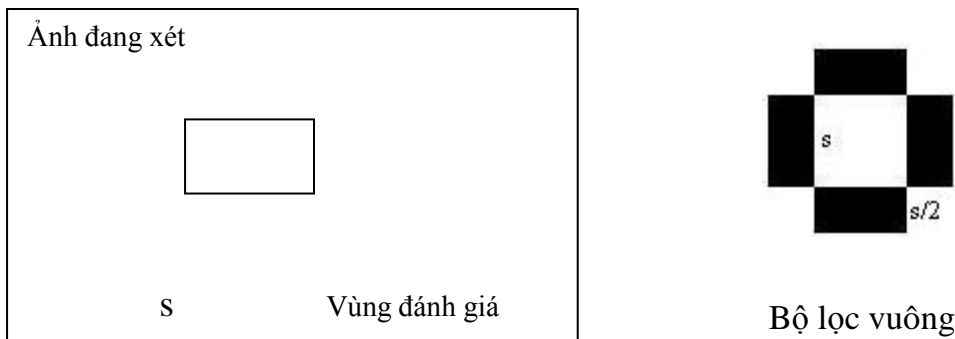
đạt cực trị. Trên thực tế, do ảnh là một hàm số không liên tục, hàm rời rạc do sự số hóa của máy tính, vì vậy sẽ rất khó khăn nếu chúng ta chỉ dò tìm các điểm cực trị. Và chúng ta sẽ bị mất một số thông tin của các điểm ảnh lồi cũng như những điểm lân cận chúng. Vì vậy ta đưa ra các công thức sau đây nhằm giảm bớt các điều kiện ràng buộc đối với điểm lồi :

$$R_P = \{(x,y) \mid |LoG(x,y)| \geq \delta \wedge LoG(x,y) > 0\} \tag{2.2}$$

Trong đó  $R_P$  là tập các điểm lồi. Ký hiệu  $LoG(x, y)$  là giá trị của Laplacian-of-Gaussian tại điểm  $(x, y)$ . Mục đích là rút trích các vùng lồi chứ không chỉ rút trích điểm lồi.

**2.1.2. Dò tìm và lấy vùng lồi**

Để đánh giá một vùng có phải là lồi hoặc không lồi, chúng ta đề xuất một bộ lọc như sau:



**Hình 2.2 Dò tìm vùng trên ảnh**

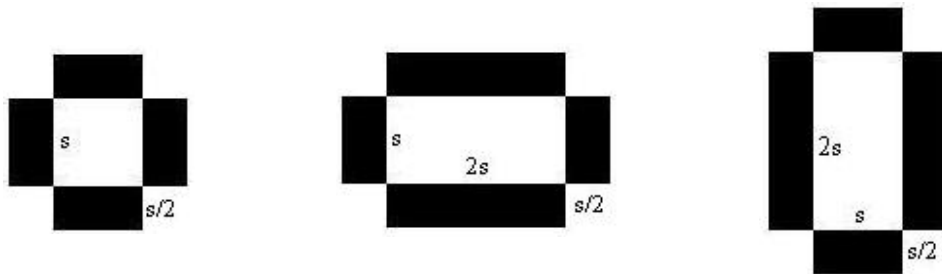
Ứng với vùng đang xét – một hình vuông  $s \times s$ , chúng ta tính giá trị của bộ lọc trên vùng này theo công thức 2.2. Đặt  $S_1, S_2, S_3$ , và  $S_4$  là các hình chữ nhật đen xung quanh hình vuông  $R$  giá trị lọc  $F(R)$  được định nghĩa như sau:

$$F(R) = 2sum(R) - \sum_{i=1}^4 sum(S_i) \tag{2.3}$$

Trong công thức (2.3), hàm  $sum(X)$  là hàm tổng giá trị độ sáng trong vùng  $X$ . Dựa trên giá trị  $F(R)$  vừa tính được, chúng ta đánh giá vùng  $R$  theo công thức sau:

$$R = \begin{cases} ridge & \Leftrightarrow |F(R)| \geq \delta \wedge F(R) > 0 \\ valley & \Leftrightarrow |F(R)| \geq \delta \wedge F(R) < 0 \\ normal & otherwise \end{cases} \tag{2.4}$$

Công thức 2.4 là công thức mở rộng của (2.1) và (2.2) nhằm để dò tìm các vùng lồi trên ảnh. Sự khác biệt chủ yếu ở chỗ ta sử dụng bộ lọc hiệu trong (2.4) thay vì sử dụng Laplacian-of-Gaussian trong (2.1) và (2.2). Chúng ta để ý rằng, bằng cách thay kích thước  $S$  của vùng đang xét, hay nói cách khác là kích thước của lọc, chúng ta có thể dò tìm các vùng lồi ứng với nhiều kích cỡ, mức độ khác nhau. Ngoài ra, để dò tìm và rút trích các vùng lồi có hình dạng khác nhau, sử dụng một tập các lọc hình học có hình dạng thay vì chỉ sử dụng một bộ lọc vuông.



Lọc hình vuông

Lọc chữ nhật ngang

Lọc chữ nhật đứng

**Hình 2.3 Tập các bộ lọc**

Do đó ta có thể biến đổi công thức (2.3) thành công thức sau:

$$F_S(R) = 2sum(R) - \sum_{i=1}^4 sum(S_i) \quad (2.5)$$

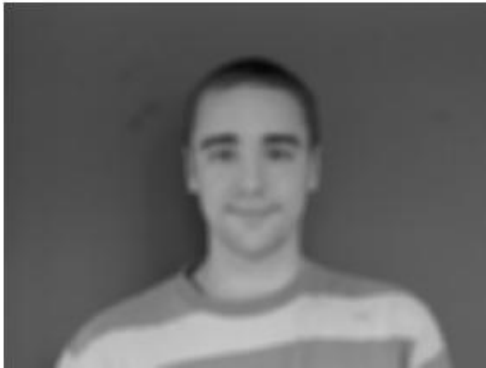
$$F_R(R) = 3sum(R) - 2\sum_{i=1}^4 sum(S_i) \quad (2.6)$$

$F_S(R)$  được sử dụng để tính giá trị lọc vuông còn  $F_R(R)$  được sử dụng để tính giá trị các lọc hình chữ nhật. Chúng ta dễ dàng nhận thấy rằng các công thức tính lọc đều tương tự với công thức của một Laplacian, nhưng có một chút khác biệt là sử dụng vùng thay vì từng điểm đơn. Trong công thức (2.5) và (2.6), hệ số đứng trước hàm  $sum(R)$  có được là do tỷ lệ giữa diện tích vùng ở tâm (vùng đang xét) và tổng diện tích của vùng lân cận. Trong hình (2.3a), tổng diện tích của các vùng đen xung quanh gấp đôi diện tích phần ở giữa, vì vậy chúng ta có hệ số 2. Trong hình (2.3b) và (2.3c), tổng diện tích các vùng đen gấp rưỡi diện tích phần hình chữ nhật ở giữa. Một cách chính xác chúng ta phải dùng hệ số 1.5 trước hàm  $sum(R)$ , nhưng việc tính toán trên số thực rất chậm, do đó để tăng tốc độ tính toán ta dùng 2 hệ số nguyên là 3 và 2 thay vì chỉ dùng 1 hệ số 1.5.

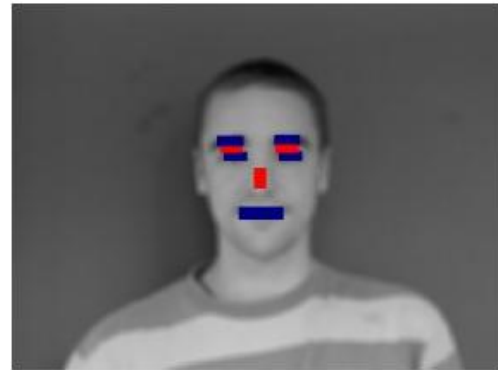
Một thông số quan trọng khác trong công thức (2.4) là ngưỡng  $\partial$ . Ngưỡng  $\partial$  phải phụ thuộc tuyến tính vào diện tích của vùng đang xét. Chúng ta sử dụng  $\partial_1=KS^2$  đối với lọc hình vuông có kích thước S và  $\partial_2=4KS^2$  đối với các lọc hình chữ nhật có kích thước S. K là một thông số thực nghiệm. Ngưỡng  $\partial_1$  là tích của K và S<sup>2</sup> – diện tích của lọc vuông. Tương tự, trong 2.6, chúng ta đã gấp đôi giá trị của  $F_R(R)$  để tránh chuyện tính toán trên số thực. Vì thế, chúng ta cũng phải gấp đôi ngưỡng  $\partial_2$  để tăng độ chính xác của việc đánh giá.

### 2.1.3. Dò tìm và phát hiện vũng lỗi ở nhiều mức khác nhau

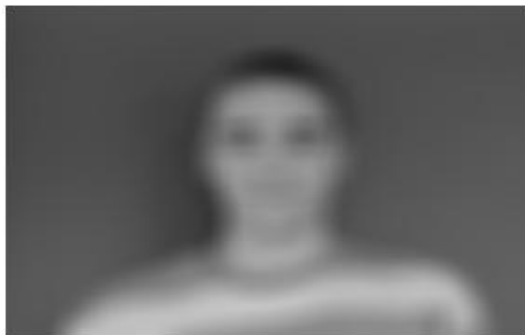
Sau đây là một phương pháp sử dụng các lọc hiệu ở nhiều mức khác nhau để dò tìm các vùng lỗi. Các lọc được giới thiệu trong phần trước sẽ được sử dụng với nhiều kích thước khác nhau vì vậy chúng sẽ dò tìm được các vùng lỗi ở nhiều mức khác nhau.



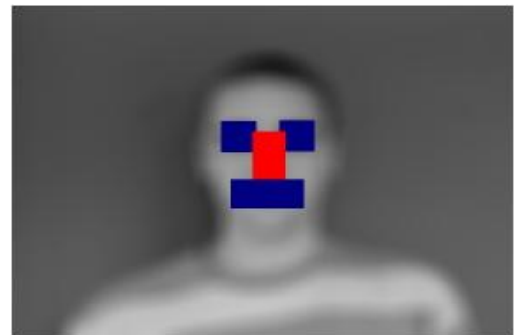
$\sigma=10$ , ảnh sau khi smoth với  
Gaussian  $10\sqrt{2}$



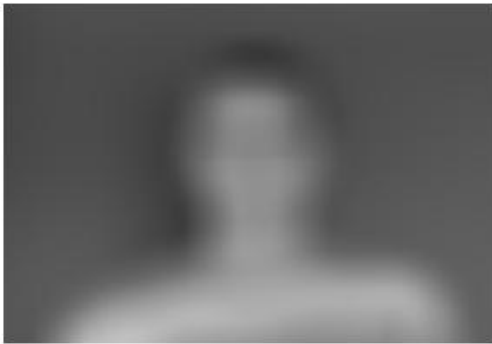
Các vùng lỗi dò tìm với kích  
thước lọc 10



$\sigma=20$ , ảnh sau khi smoth với  
Gaussian  $20\sqrt{2}$

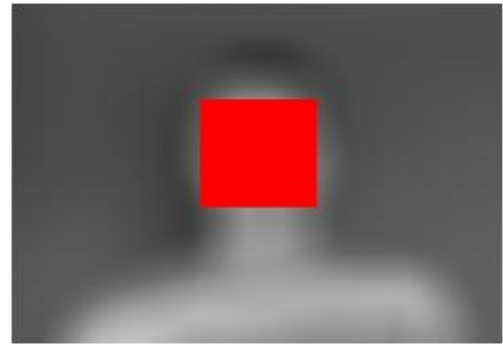


Các vùng lỗi dò tìm với kích  
thước lọc 20



$\sigma=60$ , ảnh sau khi smooth với

Gaussian  $60\sqrt{2}$



Các vùng lỗi dò tìm với kích

thước lọc 60

**Hình 2.4 Dò tìm vùng lỗi ở nhiều mức khác nhau**

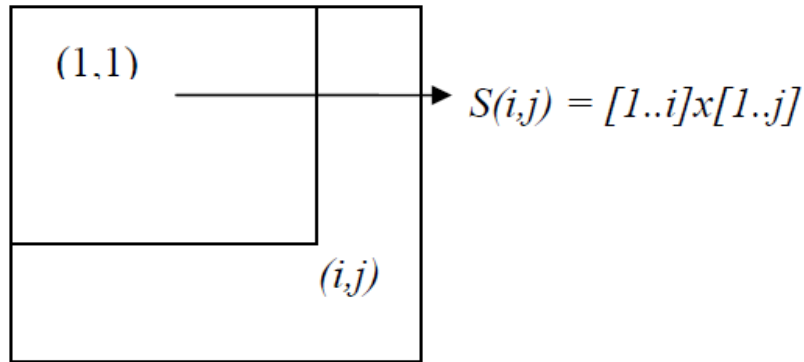
Đặt  $\partial_1, \partial_2, \dots, \partial_n$  là các mức độ được định nghĩa trước và có độ lớn tăng dần.  $\partial_1$  là mức độ thấp nhất. Tại mức độ này, chúng ta có thể dò tìm các vùng lỗi, có kích thước nhỏ và chi tiết.  $\partial_n$  là mức độ cao nhất mà tại đó chúng ta có thể dò tìm các vùng lỗi lớn hơn, ít chi tiết hơn. Tại mỗi mức, trước tiên áp dụng một lọc Gaussian với kích thước  $\partial_i\sqrt{2}$  sau đó chúng ta sử dụng 3 bộ lọc hiệu như hình 2.3 với kích thước  $S = \partial_i$ . Trong hình 2.4, các hình vuông và hình chữ nhật màu đỏ là các vùng lỗi. Trong thực tế, có nhiều vùng lỗi khác cũng được dò tìm được, tuy nhiên trong hình 2.4, chỉ tập trung vào các vùng lỗi dò tìm được trong vùng ảnh quan tâm – vùng ảnh khuôn mặt. Trong phần tiếp theo, sẽ trình bày phương pháp cấu trúc các vùng lỗi rút trích được này thành các cây cấp bậc.

#### 2.1.4. Tối ưu tốc độ dò tìm

Trong vùng đang xét, chúng ta phải thường xuyên tính tổng độ sáng điểm ảnh của một vùng hình chữ nhật, một số hình trong chúng có thể chồng lên nhau. Việc này làm cho tính toán chậm đi rất nhiều do phải tính nhiều lần tổng độ sáng cho một vùng. Để giảm độ phức tạp của tính toán này, chúng tôi sử dụng phương pháp tương tự như ảnh tích phân (integral image) và kỹ thuật lập trình quy hoạch động. Phương pháp này được trình bày như sau:

Cho trước một ảnh mức xám  $I(x, y)$ , chúng ta định nghĩa một ảnh tích phân  $S(x, y)$  sao cho  $S(x, y)$  là tổng mức sáng của toàn bộ điểm ảnh trong hình chữ nhật  $R(1, 1, x, y)$ . Chúng ta có thể dễ dàng tính được  $S(x, y)$  theo công thức truy hồi:

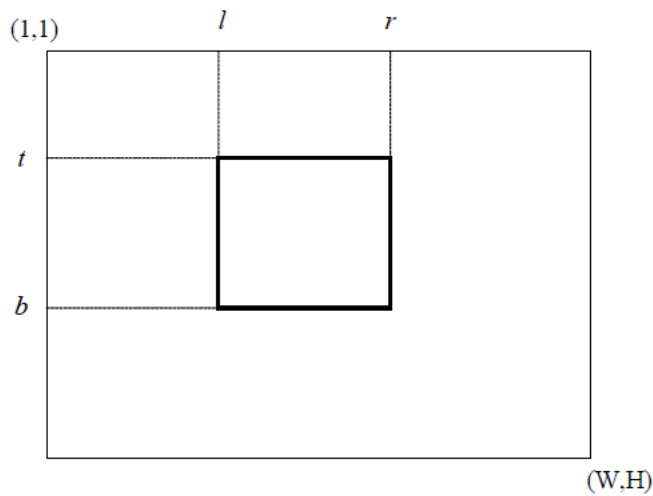
$$S(x, y) = \begin{cases} 0 & x = 0 \vee y = 0 \\ S(x-1, y) + S(x, y-1) - S(x-1, y-1) + I(x, y) & x \neq 0 \wedge y \neq 0 \end{cases} \quad (2.7)$$



**Hình 2.5 Ảnh tích phân**

Từ ảnh  $S(x, y)$ , ta dễ dàng tính được tổng độ sáng của một hình chữ nhật bất kỳ  $R(l, r, t, b)$  theo công thức sau:

$$S_R = S(r, b) - S(r, t) - S(l, b) + S(l, t) \quad (2.8)$$



**Hình 2.6: Tính tổng độ sáng cho hình chữ nhật  $R(l, r, t, b)$**

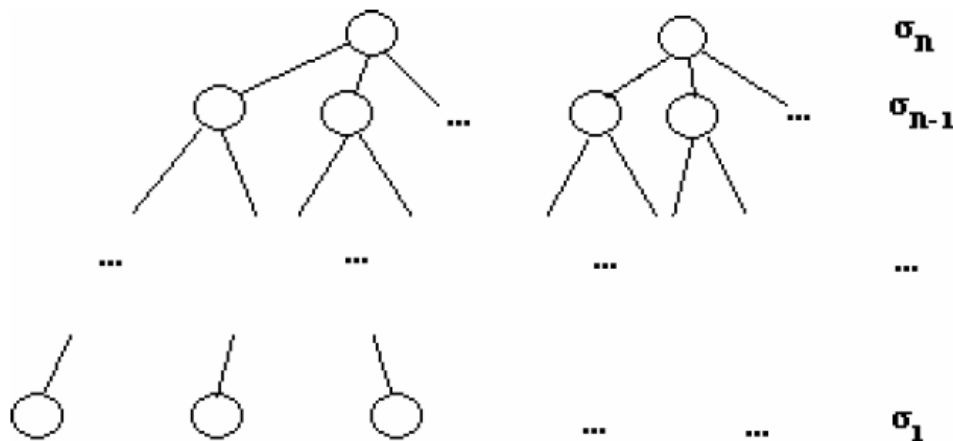


Phương pháp này cải tiến đáng kể tốc độ tính toán bởi vì chúng ta chỉ cần tính một lần ảnh tích phân và sử dụng nó trong tất cả các cửa sổ hình chữ nhật con và đặc biệt độ phức tạp tính toán không phụ thuộc vào kích thước của vùng đang xét.

## 2.2. Phát hiện mặt người

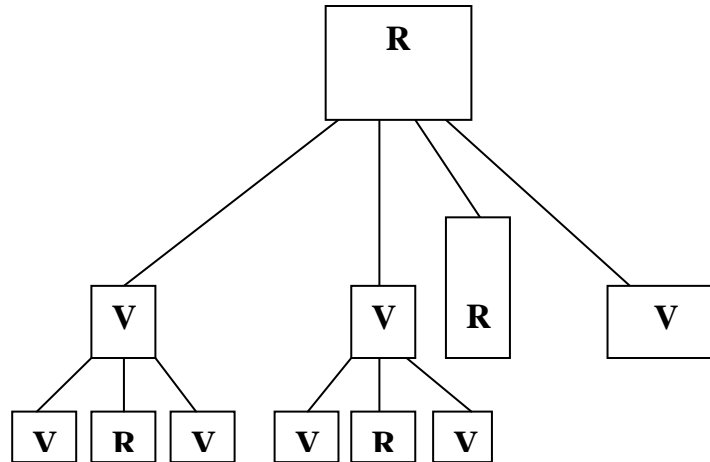
### 2.2.1. Xây dựng cấu trúc cây

Từ các vùng lỗi rút trích được, chúng ta xây dựng các cây quan hệ cấp bậc. Những vùng được rút trích ở mức cao nhất  $\sigma_n$  sẽ trở thành những nút gốc của các cây sẽ tạo. Những vùng rút trích được ở mức thấp nhất  $\sigma_1$  sẽ được chèn vào cây như là những nút lá (xem hình 2.7)



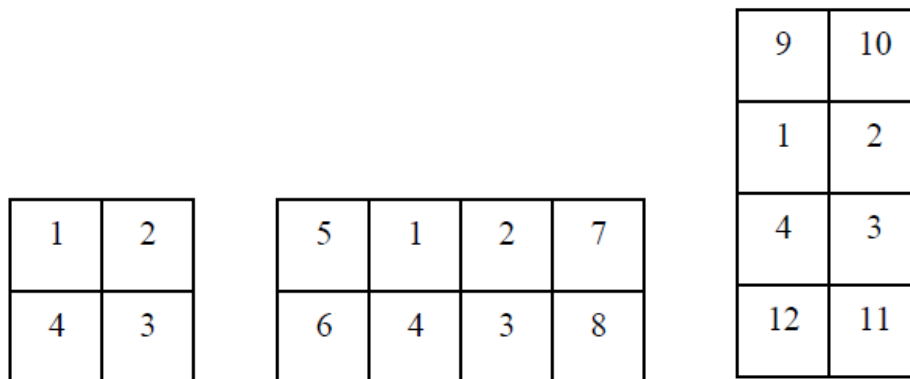
**Hình 2.7 Tạo cây cấp bậc**

Công việc chính của việc tạo cây là kết nối các vùng rút trích được tại một mức đến các vùng của mức kế tiếp. Nếu vùng R tại mức  $\sigma_1$  bị che khuất hơn hai phần ba diện tích của nó bởi vùng R' của mức  $\sigma_{i+1}$ , R sẽ được gán nhãn là nút con của R'. Bằng cách này chúng ta có thể dễ dàng tạo cây cấp bậc từ các vùng đã rút trích được. Các vùng rút trích được trong hình được xây dựng thành cây như sau :



**Hình 2.8 Một cây rút trích từ khuôn mặt**

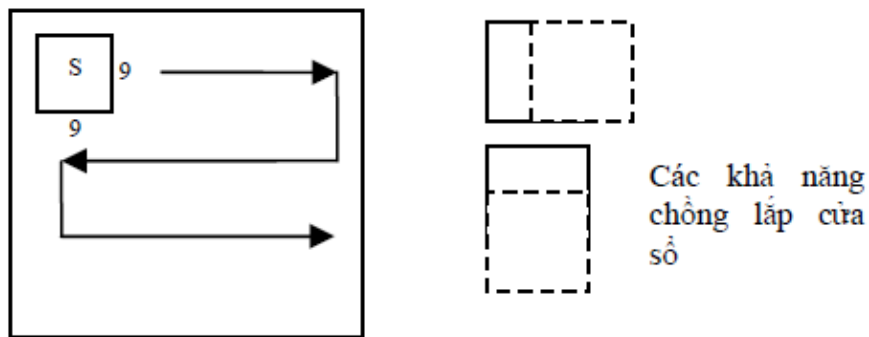
Đối với mỗi nút trong cây, chúng ta cần lưu các thông tin của nút này như: loại (lồi hay lõm), hình dạng (vuông, chữ nhật ngang, chữ nhật đứng), vị trí tương đối của nó so với nút cha của nó (có giá trị từ 0 đến 12, xem hình 2.9), và thông tin về độ sáng của nó.



**Hình 2.9: vị trí vùng tương đối của nút con**

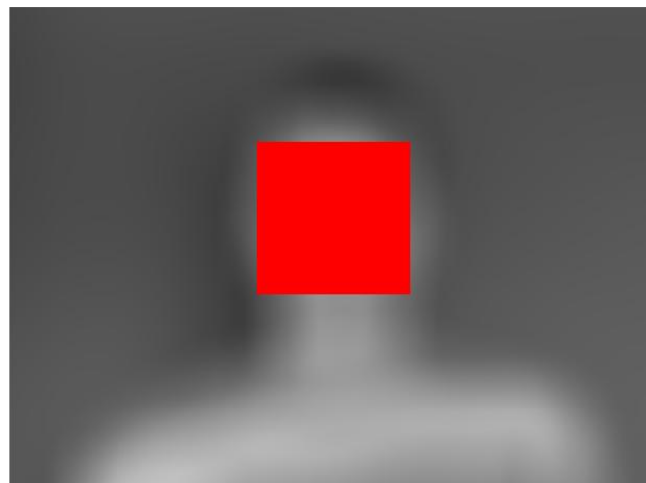
Trong hình biểu diễn các vị trí tương đối của nút con trên nút cha lần lượt có dạng vuông, chữ nhật ngang, và chữ nhật đứng. Việc xét một nút con có vị trí tương đối như thế nào đối với nút cha khá đơn giản. Nếu tâm của nút con (tâm của hình vuông hay hình chữ nhật) rơi vào vùng nào trong hình thì chỉ số của vùng đó chính là giá trị vị trí tương đối của nút con trong nút cha. Lưu ý rằng đối với các nút gốc giá trị này sẽ là 0. Bên cạnh đó thông tin về độ sáng cũng rất quan trọng trong các bài toán phát hiện và nhận dạng khuôn mặt về sau. Để lưu trữ thông tin này trên cây, ta sử dụng

phương pháp phân tích thành phần chính (Principal Component Analysis – PCA) để giảm số chiều của ảnh mức xám trên vùng ảnh của nút đó. Bằng cách này ta giảm được đáng kể thông tin phải lưu trữ về độ sáng của ảnh. Thay vì phải lưu trữ toàn bộ điểm sáng của các vùng lỗi (các nút), chúng ta chỉ cần lưu một vector PCA chứa nội dung của các điểm sáng này. Cách này giúp cho giảm thông tin phải lưu trữ, xử lý mà cũng không mất mát quá nhiều thông tin (xem hình 2.10)

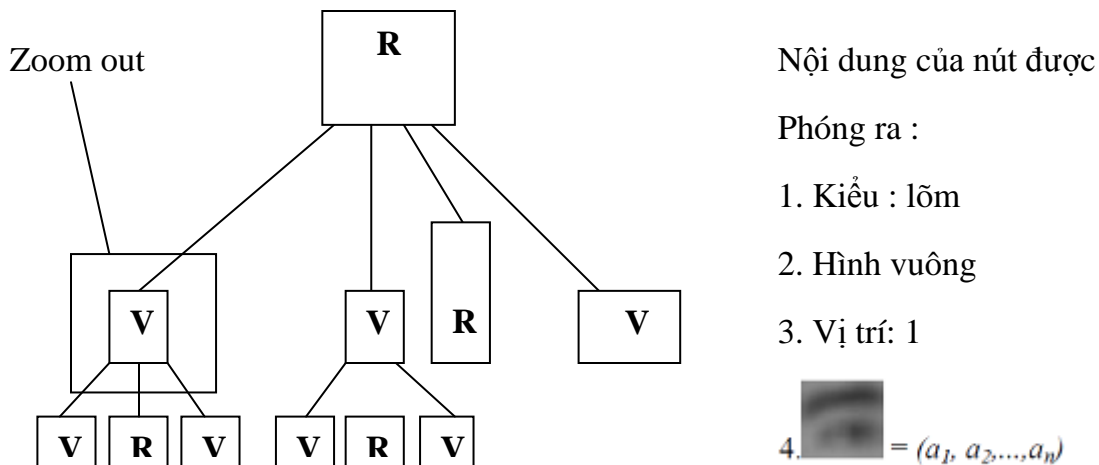


**Hình 2.10: cách tính vecto đại diện độ sáng cho nút**

Cụ thể ta thực hiện như sau: dùng một cửa sổ con S kích thước 9x9 quét theo đường zic-zac trên ảnh xám của vùng lỗi. Các cửa sổ này có thể chồng lấp lên nhau hai phần ba kích thước của nó. Tại mỗi điểm, ta lưu lại trung bình mức sáng của các điểm ảnh trong cửa sổ S. Như vậy ta lưu được một vector tương ứng cho biểu diễn mức sáng của điểm ảnh tại nút đó.



**Vị trí của nút gốc trên ảnh**



**Hình 2.11: cây rút trích được và các thông tin trên một nút**

Bên cạnh đó, ta sử dụng một số luật để tỉa bớt nhánh cho các cây. Nếu một nút cùng kiểu so với nút cha của nó thì nó được xem là một nút lặp. Trong trường hợp này, nút lặp sẽ được tỉa khỏi cây. Nếu nó là một nút lá, chúng ta dễ dàng xóa nó khỏi cây. Ngược lại, nếu nó không phải là nút lá, tất cả các nút con của nó sẽ được gán trở thành con của nút cha nó. Ví dụ chúng ta có hai vùng lỗi hình chữ nhật đứng ở mức  $\sigma_2$  và  $\sigma_3$ . Hai vùng này thật sự chỉ là một. Trong thực tế, do hiệu ứng của phép lọc Gaussian tại các mức khác nhau, chúng được rút trích ra từ ảnh với các kích thước khác nhau vì thế vùng lỗi ở mức  $\sigma_3$  đã được xóa khỏi cây.

**2.2.2. Xây dựng hàm tính độ tương đồng giữa hai cây**

Trong phần này ta cùng tìm hiểu một cách tính độ tương đồng giữa hai cây theo cách biểu diễn như phần trước. Việc xây dựng hàm tương đồng giữa hai cây là rất cần thiết trong tất cả các xử lý liên quan đến đặc trưng này. Nó có tính chất quyết định ảnh hưởng đến độ chính xác của phát hiện và nhận dạng khuôn mặt.

**2.2.2.1. Độ tương đồng giữa hai nút trên cây**

Để xây dựng hàm tính độ tương đồng giữa hai cây trước hết ta cần xây dựng hàm tính độ tương đồng của hai nút trên hai cây. Với điều kiện rằng hai nút thuộc hai cây khác nhau này cùng mức  $\sigma_i$  nào đó.

Đặt  $N_1$  và  $N_2$  là hai nút cùng cấp trên hai cây  $T_1$  và  $T_2$ . Khi đó  $N_1$  và  $N_2$  có những thuộc tính sau:

$$N_1, N_2 \in Node = \{(t,s,p,i)\} \quad (2.9)$$

Trong đó:

$t \in \{\text{ridge}\}$  : là loại của nút (lồi).

$s = \{\text{square, horizontal rectangle, vertical rectangle}\}$  là hình dạng của nút (vuông, hình chữ nhật ngang, hình chữ nhật đứng).

$p \in [0..12]$  : là vị trí tương đối của nút trong nút cha của nó.

$i \in (a_1, a_2, \dots, a_n)$  : là vector chứa thông tin độ sáng của nút. Trong đó  $n$  thay đổi tùy theo cấp của nút. Ở mức  $\sigma_i$  càng cao, kích thước của vùng lồi lõm tương ứng với nút càng lớn, số chiều  $n$  của vector này càng cao.

Dựa trên các thuộc tính của node định nghĩa trong công thức 3.9, ta xây dựng công thức tính độ tương đồng giữa 2 nút như sau:

$$NS(N_1, N_2) = \alpha_1 E(N_1.t, N_2.t) + \beta_1 E(N_1.s, N_2.s) + \gamma_1 E(N_1.p, N_2.p) + (1 - \alpha_1 - \beta_1 - \gamma_1) EV(N_1.i, N_2.i) \quad (2.10)$$

Trong công thức 2.10,  $\alpha_1$ ,  $\beta_1$ , và  $\gamma_1$  là các thông số thực nghiệm, hàm  $E(x,y)$  dùng để xét sự giống nhau giữa 2 thuộc tính được định nghĩa như sau:

$$E(x, y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases} \quad (2.11)$$

Bên cạnh đó ký hiệu  $EV(x,y)$  là hàm tính độ tương đồng giữa 2 vector được chuẩn hóa về đoạn  $[0,1]$ . Trong đó  $x$  và  $y$  là hai vector độ sáng có  $n$  chiều  $(a_1, a_2, \dots, a_n)$ .

$$EV(x, y) = 1 - \frac{\sqrt{\sum_{i=1}^n (x.a_i - y.a_i)^2}}{255\sqrt{n}} \quad (2.12)$$

Do mỗi thành phần  $a_i$  trong vector độ sáng có giá trị là độ sáng ( $a_i \in [0,255]$ ) nên dễ dàng chứng minh được  $EV(x,y) \in [0,1]$ .  $EV(x,y)$  đạt giá trị lớn nhất là 1 khi hai vector  $x$  và  $y$  trùng nhau. Và  $EV(x,y)$  đạt giá trị nhỏ nhất là 0 khi hai vector có khoảng cách Euclid xa nhau nhất. Và từ đó ta cũng dễ dàng suy ra được giá trị của NS ( $N1, N2$ )  $\in [0,1]$  với mọi cặp nút  $N1$  và  $N2$ .

### 2.2.2.2. Độ tương đồng giữa hai cây

Trên cơ sở hàm tính độ tương đồng giữa hai nút, ta xây dựng lên hàm tính độ tương đồng giữa hai cây. Gọi  $T1$  và  $T2$  và hai cây đang cần tính độ tương đồng.  $T1$  và  $T2$  có những thuộc tính sau:

$$T1, T2 \in Tree = \{(r, n, c1, c2, \dots, cn)\} \quad (2.13)$$

Trong đó:

$r$  : là nút gốc của cây  $r \in Node$ .

$n$  : là số nhánh con của cây  $n \in N$ .

$c_i$  : là cây (nhánh) con thứ  $i$  của cây  $c_i \in Tree$  và  $i \in [1..n]$ .  $c_i$  có thể là một nút là hoặc cũng có thể là một cây khác.

Từ công thức (2.13), ta xây dựng công thức tính độ tương đồng của 2 cây như sau:

$$TS(T_1, T_2) = \alpha_2 NS(T_1.r, T_2.r) + (1 - \alpha_2) \frac{\sum_{i=1}^{T_1.n} \sum_{j=1}^{T_2.n} TS(T_1.c_i, T_2.c_j)}{T_1.n T_2.n} \quad (2.14)$$

Trong đó  $\alpha_2$  là thông số thực nghiệm. Các thông số thực nghiệm  $\alpha_1$ ,  $\beta_1$ , và  $\gamma_1$  trong công thức (2.10) và  $\alpha_2$  trong (2.14) sẽ được thảo luận chi tiết trong phần thực nghiệm nhằm chọn thông số thực nghiệm tốt nhất. Theo công thức (2.14) ta cũng dễ dàng nhận thấy  $TS(T_1, T_2) \in [0, 1]$  với mọi cặp cây  $T_1$  và  $T_2$ . Giá trị của hàm này càng lớn, hai cây này càng giống nhau càng nhiều hơn.

Ngoài ra, trong phần thử nghiệm có thử nghiệm đặc trưng lỗi mà không sử dụng đến thông tin độ sáng. Khi đó công thức (2.10) tính độ tương đồng giữa hai nút được đơn giản hóa như sau:

### 2.2.2.3. Không gian cây và khoảng cách giữa hai cây

$$NS(N_1, N_2) = \alpha_1 E(N_1.t, N_2.t) + \beta_1 E(N_1.s, N_2.s) + (1 - \alpha_1 - \beta_1) E(N_1.p, N_2.p)$$

Trong một số trường hợp, hàm tính độ tương đồng giữa hai cây không đủ để tính toán trong một số trường hợp. Để đưa về bài toán không gian, ta xây dựng một “không gian cây” (tree-space), không gian này sẽ là cơ sở để cho các bài toán phát hiện và nhận dạng khuôn mặt về sau.

Trong không gian cây này, mỗi cây được xem là một điểm trong không gian. Trên cơ sở đó ta xây dựng hàm tính khoảng cách giữa hai cây như sau:

$$\|T_1, T_2\| = 1 - TS(T_1, T_2) \tag{2.15}$$

Dễ dàng nhận thấy, khoảng cách giữa hai cây là một giá trị thuộc đoạn  $[0, 1]$ . Và khoảng cách tiến dần về bằng không khi hai cây càng giống nhau.

Trên cơ sở các công thức tính độ tương đồng, khoảng cách này, ta xây dựng một mô hình thống kê để dò tìm và phát hiện khuôn mặt. Mô hình có thể dùng cho nhận dạng khuôn mặt cũng như mở rộng cho bài toán phân loại đối tượng.

## CHƯƠNG 3. CHƯƠNG TRÌNH THỬ NGHIỆM

### 3.1. Bài toán

Nội dung : Tìm hiểu đặc trưng lời 3D và bài toán phát hiện mặt người trong ảnh.

- Đầu vào : Đưa ảnh đầu vào để phân tích
- Đầu ra : Xác định vùng ảnh chứa mặt người

Trích lọc khuôn mặt

### 3.2. Phân tích bài toán

Chương trình thử nghiệm được xây dựng bằng ngôn ngữ C# sử dụng bộ Visual Studio 2008 với các hàm API hỗ trợ phát hiện khuôn mặt trong thư viện Luxand FaceSDK [25]. Chương trình gồm :

- + Phát hiện, đánh dấu
- + Lọc vùng khuôn mặt

Hoạt động của chương trình :

**Bước 1:** Đưa một ảnh có định dạng JPG (hoặc PNG, BMP, GIF, JPEG)

**Bước 2 :** Xác định vùng ảnh chứa mặt người và trích chọn khuôn mặt.

**Bước 3:** Hiện thị ảnh kết quả

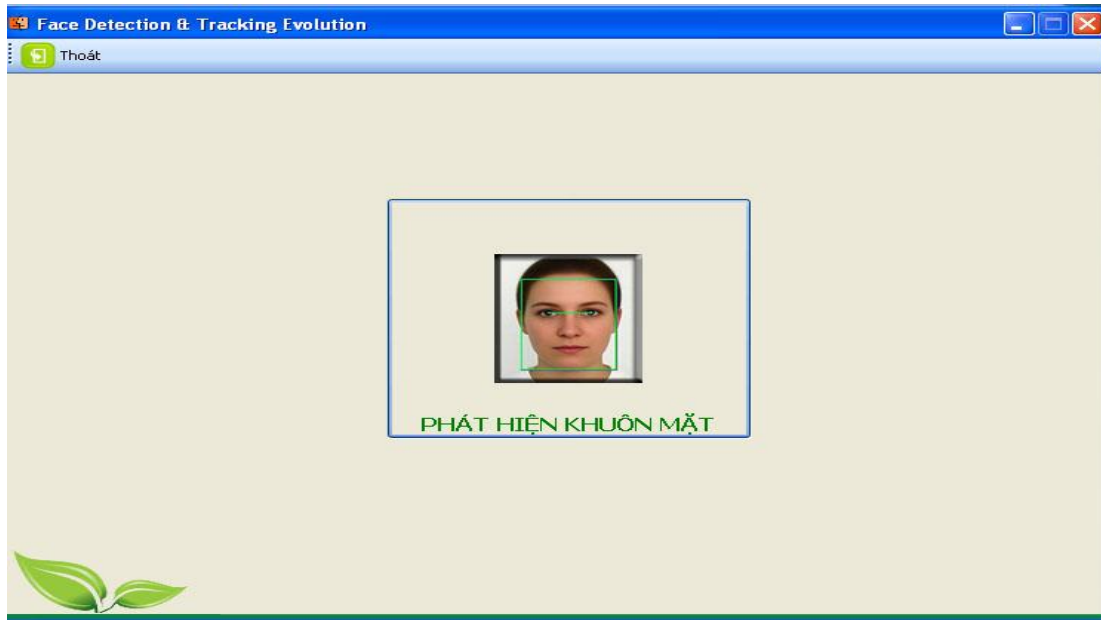
### 3.3. Một số kết quả chương trình

Một số modul chính của chương trình

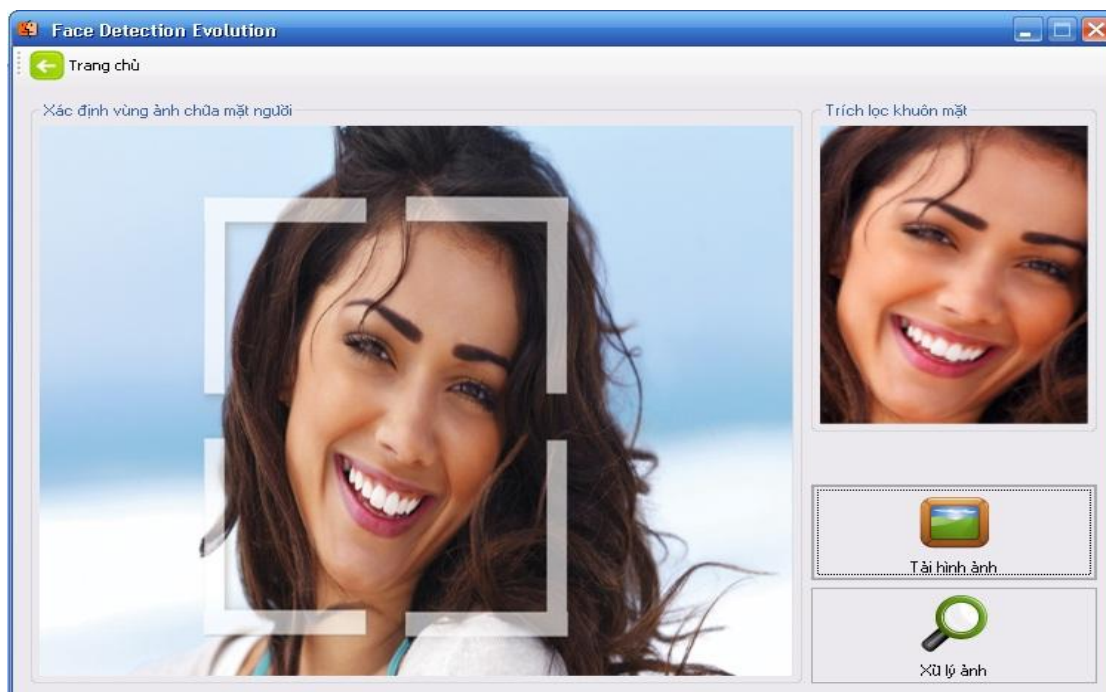
1. Đưa một số hình ảnh vào
2. Xác định vùng ảnh chứa mặt người
3. Hiện thị kết quả

Một số giao diện chính của chương trình:

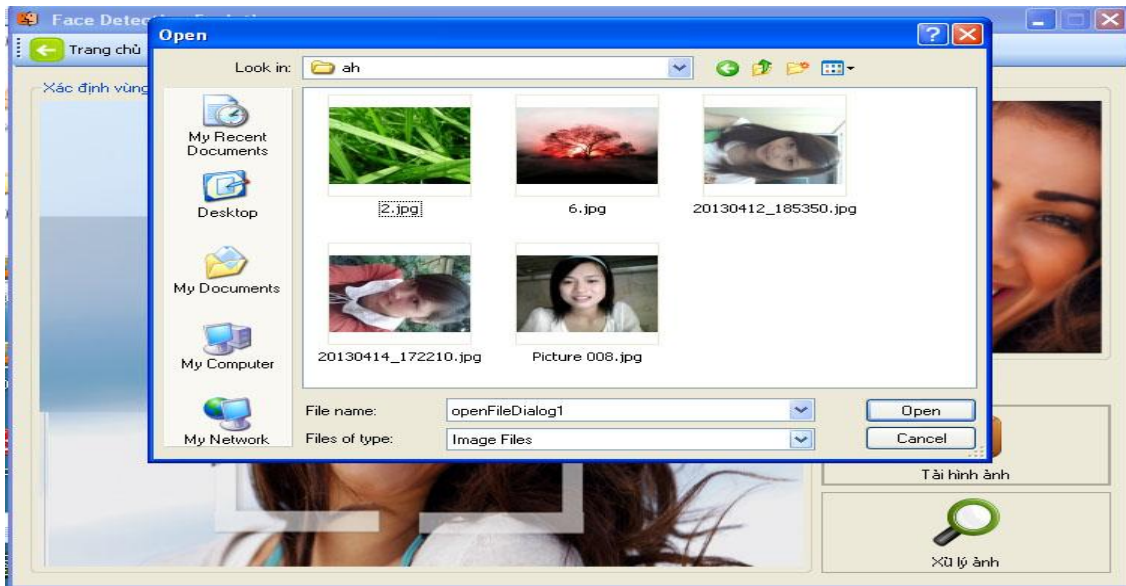




**Hình 3.1** Giao diện chính của chương trình

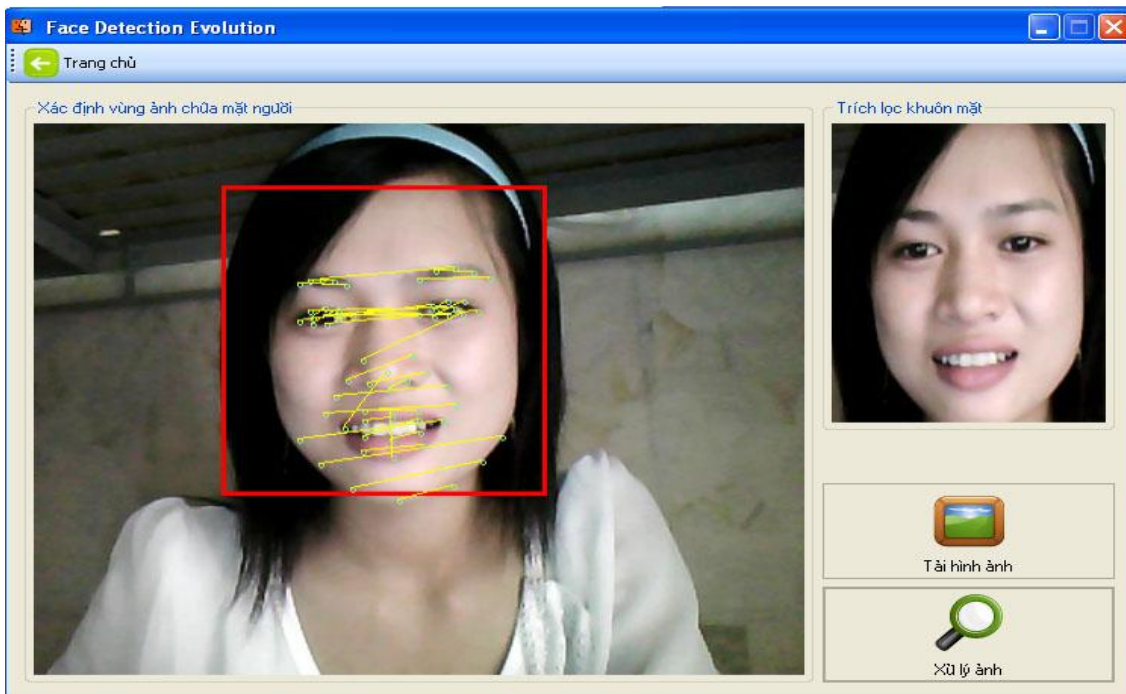


**Hình 3.2** Giao diện phát hiện khuôn mặt

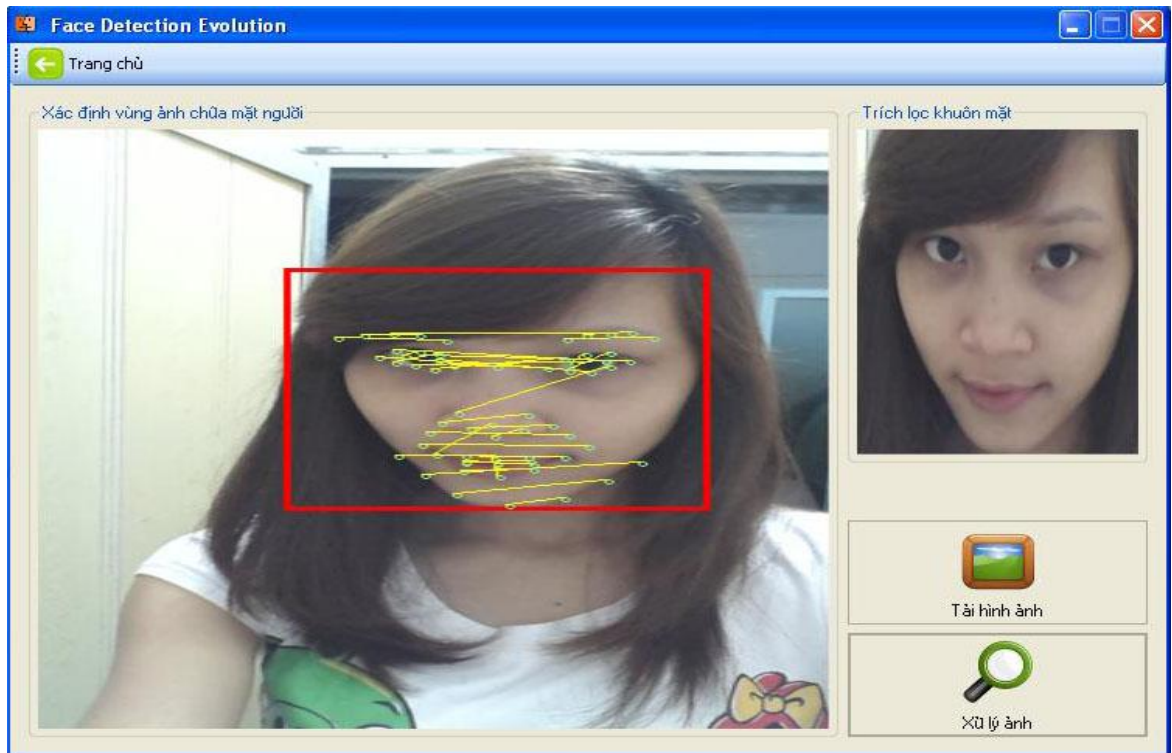


**Hình 3.3** Giao diện modul chọn ảnh đầu vào

Một số kết quả thu được:



**Hình 3.4** Ảnh kết quả qua khuôn mặt biểu cảm

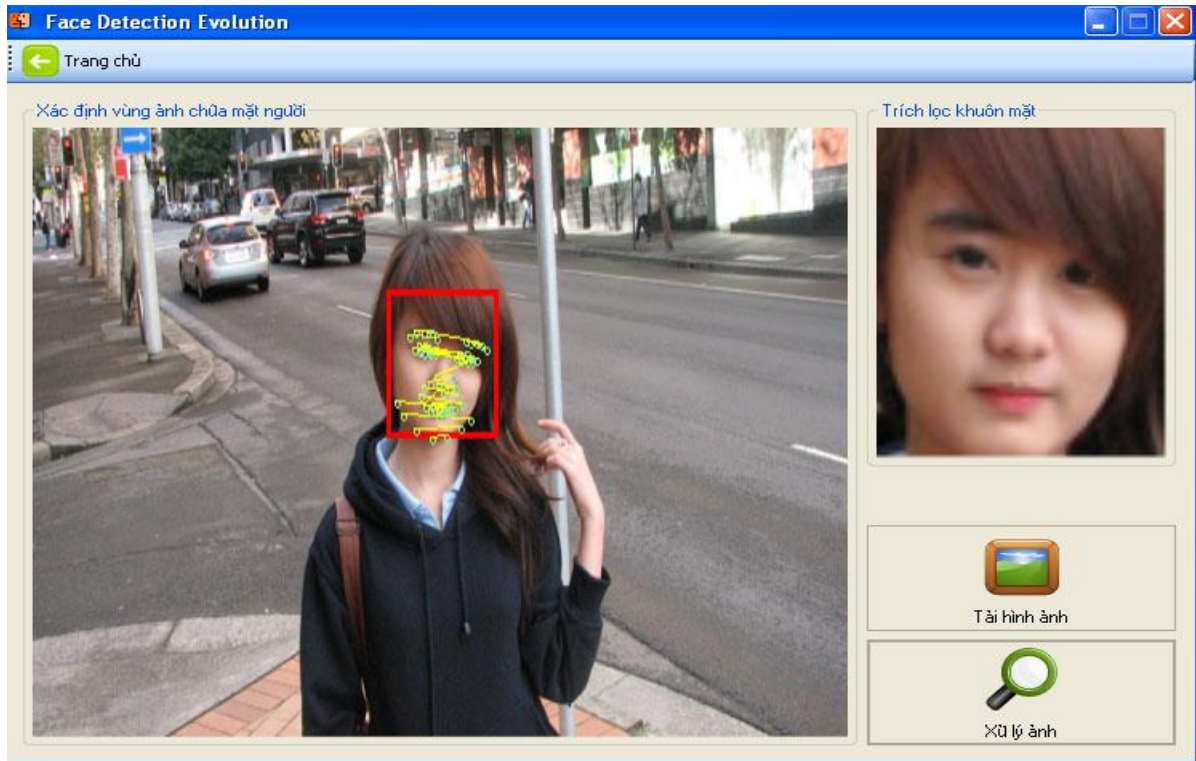


*Hình 3.5 Ảnh kết quả qua ảnh chụp trong nhà*

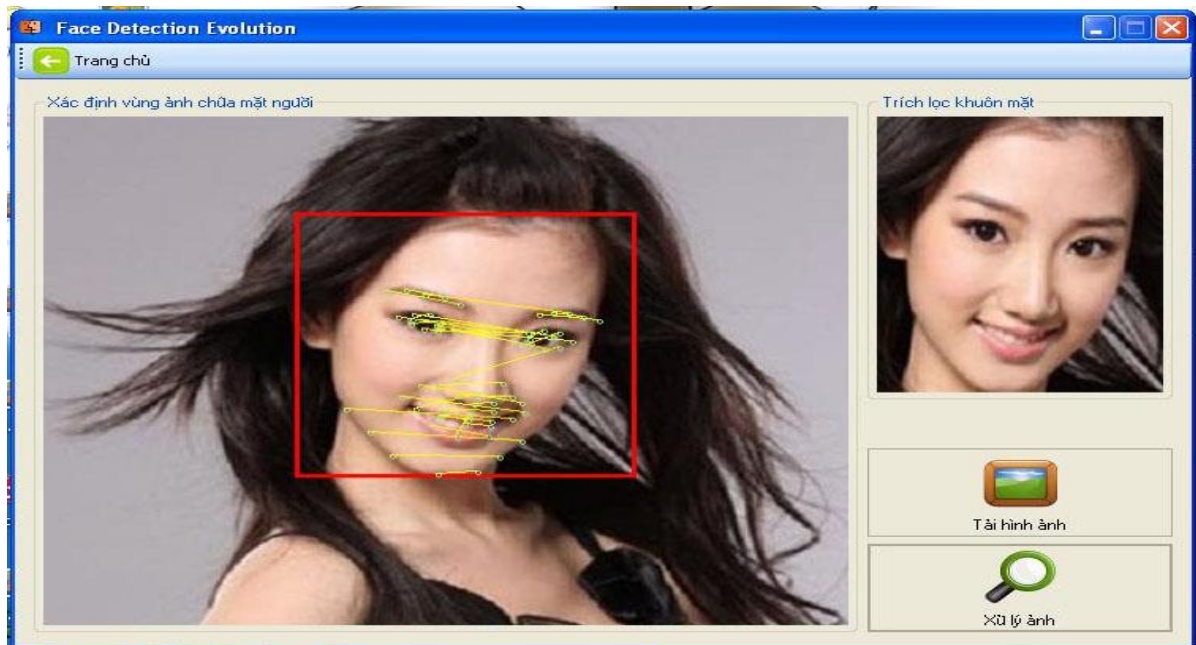


*Hình 3.6 Ảnh kết quả qua ảnh chụp thiếu sáng*





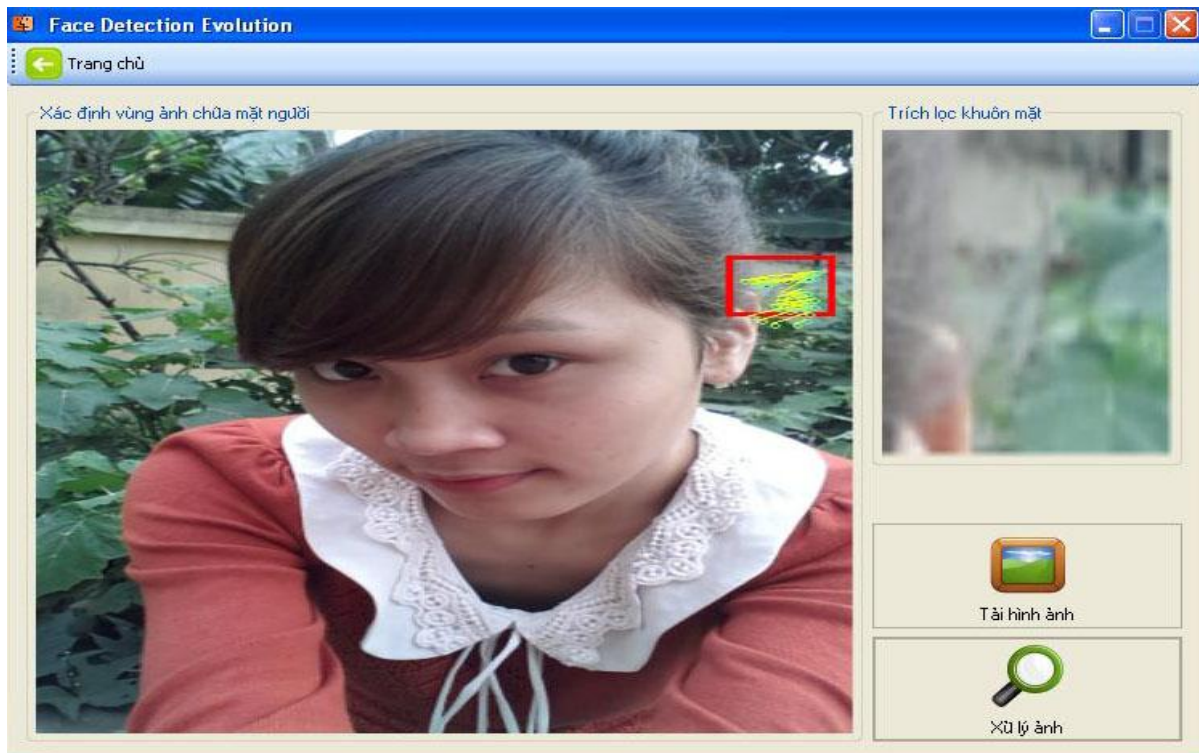
*Hình 3.7 Ảnh kết quả qua ảnh chụp ngoài trời*



*Hình 3.8 Ảnh kết quả mặt nghiêng*



*Hình 3.9 Ảnh kết quả phát hiện khuôn mặt trong ảnh đặt biệt*



*Hình 3.10 Ảnh kết quả phát hiện nhầm khi dò tìm khuôn mặt*

## PHẦN KẾT LUẬN

### 1. Kết quả đạt được

Trong quá trình nghiên cứu và thực hiện đồ án dưới sự định hướng dẫn của thầy hướng dẫn, em đã tìm hiểu được một cách tổng quan về xử lý ảnh và bài toán phát hiện mặt người dựa vào đặc trưng lồi 3D. Dựa vào những tài liệu tìm được em đã tiến hành cài đặt chương trình thử nghiệm, xây dựng thuật toán xác định vùng ảnh chứa khuôn mặt. Chương trình thử nghiệm cho kết quả khả quan với khả năng mô phỏng khuôn mặt bằng một khung lưới theo định dạng 3D.

### 2. Tồn tại

Trong một số trường hợp, việc phát hiện khuôn mặt vẫn xảy ra nhầm lẫn các vùng khác (không phải là khuôn mặt) trong ảnh là khuôn mặt. Điều này xảy ra do ảnh thu nhận có chứa nhiều vùng ảnh có lưới *Ridge* có kết cấu tương tự với kết cấu *Ridge* của khuôn mặt, do đó chương trình xem nó như là một thành phần biểu diễn khuôn mặt.

### 3. Hướng phát triển

Hướng tới việc xây dựng khung lưới định dạng 3D thực thụ với việc gán trực tọa độ Oxyz lên vùng *Ridge* dò tìm được nhằm tăng cường khả năng nhận biết các khu vực có khả năng là mắt mũi hay miệng và các vùng khác như: chân mày, cằm...

Bên cạnh đó để tăng cường khả năng xử lý, em mở rộng việc sử dụng các bộ lọc tương thích để tăng cường quá trình tiền xử lý ảnh để loại bớt các vùng nhiễu không cần thiết. Em tin rằng bài toán này có thể mở rộng áp dụng cho bài toán nhận dạng đối tượng ở mức tổng quát.

## TÀI LIỆU THAM KHẢO

### [A]. Tài liệu Tiếng Anh

- [1]. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, “Detecting Faces in Images: A Survey”, IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI), Vol. 24, No 1, pp. 34-47, Jan 2002.
- [2]. T. Kanade, “Picture Processing Computer Complex and Recognition of Human Faces”, PhD thesis, Kyoto Univ., 1973.
- [3]. G. Yang, T. S. Huang, “Human Face Detection in Complex Background”, Pattern Recognition, Vol. 27, No. 1, pp. 53-63, 1994.
- [4]. C. Kotropoulos, I. Pitas, “Rule-based Face Detection in Frontal Views”, Proc. Int’l Conf. Acoustics, Speech and Signal Processing, Vol. 4, pp. 2637-2540, 1997.
- [5]. T. K. Leung, M.C. Burl, P. Perona, “Finding Faces in Cluttered Scenes Using Random Labeled Graph Matching”, Proc. 5th IEEE Conf. Computer Vision (ICCV’95), pp. 637-644, 1995.
- [6]. K. C. Yow, R. Cipolla, “Feature-Based Human Face Detection”, Image and Vision Computing, Vol. 15, No. 9, pp. 713-735, 1997.
- [7]. I. Craw, D. Tock, A. Bennett, “Finding Face Features”, Proc. 2nd European Conf. Computer Vision (ECCV’92), Vol. 2, pp. 92-96, 1992.

### [B]. Tài liệu Tiếng Việt

- [8]. PGS.TS Đỗ Năng Toàn, TS. Phạm Việt Bình, “Giáo trình Xử Lý Ảnh”, Đại học Thái Nguyên, 2007.
- [9]. Phạm Thế Bảo, Nguyễn Thành Nhựt, Cao Minh Thịnh, Trần Anh Tuấn, Phan Phúc Doãn, “Tổng quan các phương pháp xác định khuôn mặt người”, Tạp chí Công nghệ thông tin & Truyền thông, 2007.

### [C]. Cơ sở dữ liệu ảnh khuôn mặt và thư viện lập trình

- [10]. Luxand FaceSDK 4.0, Trial Version, Luxand.Inc, 2011.  
<http://www.luxand.com>.