

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



ISO 9001:2008

ĐỒ ÁN TỐT NGHIỆP

NGÀNH: CÔNG NGHỆ THÔNG TIN

HẢI PHÒNG - 2013

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

**TÌM HIỂU ĐẶC TRƯNG LỖM 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 : Nguyễn Thị Thơm
Giáo viên hướng dẫn : PGS. TS. Đỗ Năng Toàn
Mã số sinh viên : 1351010031

HẢI PHÒNG - 2013

MỤC LỤC

MỤC LỤC	1
DANH MỤC HÌNH VẼ	3
LỜI CẢM ƠN	4
LỜI MỞ ĐẦU	5
CHƯƠNG 1. KHÁI QUÁT VỀ XỬ LÝ ẢNH VÀ BÀI TOÁN PHÁT HIỆN MẶT NGƯỜI	6
1.1. Khái quát về xử lý ảnh	6
1.1.1. Một số khái niệm cơ bản.....	6
1.1.2. Một số vấn đề trong xử lý ảnh.....	7
1.1.2.1. Các hệ thống xử lý ảnh	7
1.1.2.2. Các hình thái của ảnh.....	8
1.1.3. Một số ứng dụng trong xử lý ảnh.....	9
1.2. Bài toán nhận dạng mặt người.....	9
1.2.1. Bài toán nhận dạng mặt người.....	9
1.2.2. Những khó khăn của nhận dạng khuôn mặt	10
1.2.3. Tầm quan trọng của bài toán nhận diện mặt người	11
1.2.4. Các ứng dụng đặc trưng của bài toán nhận diện mặt người.....	12
1.2.5. Xây dựng hệ thống nhận diện mặt người đặc trưng.....	13
1.2.6. Một số phương pháp nhận diện mặt người.....	13
1.2.6.1. Dựa trên tri thức.....	14
1.2.6.2. Hướng tiếp cận dựa trên đặc trưng không thay đổi.	15
1.2.6.3. Hướng tiếp cận dựa trên so khớp mẫu	18
1.2.6.4. Hướng tiếp cận dựa trên diện mạo.....	19
1.3. Phát phát hiện mặt người dựa trên đặc trưng lồi	20
CHƯƠNG 2 : TÌM HIỂU VỀ CÁC ĐẶC TRƯNG KHUÔN MẶT	21
2.1. Đặc trưng lồi	21
2.2. Rút trích các đặc trưng lồi	21

2.2.1.Điểm lõm.....	21
2.2.2. Dò tìm và lấy vùng lõm.....	22
2.2.3.Dò và phát hiện vùng lõm ở nhiều mức khác nhau.....	24
2.2.3.1.Tối ưu tốc độ dò tìm.....	25
2.2.4.Phát hiện mặt người.....	27
2.2.4.1.Xây dựng cấu trúc cây lõm	27
2.2.4.2.Xây dựng hàm tính độ tương đồng giữa hai cây.....	29
2.2.4.2.1. Độ tương đồng giữa hai nút trên cây	30
2.2.4.2.2. Không gian cây và khoảng cách giữa hai cây	32
2.3.1. Gán nhãn.....	32
2.3.2. Thống kê.....	33
2.3.3. Đánh giá dùng cho phát hiện khuôn mặt.....	34
2.3.4. Hậu xử lí.....	35
CHƯƠNG 3 CHƯƠNG TRÌNH THỬ NGHIỆM.....	36
3.1.Bài toán.....	36
3.2.Phân tích thiết kế	36
3.3.Một số kết quả chương trình.....	36
3.3.1 Tập ảnh thử nghiệm.....	36
3.3.2 Kết quả chương trình.....	37
3.3.3 Kết quả thực nghiệm	41
PHẦN KẾT LUẬN.....	43
TÀI LIỆU THAM KHẢO	44

DANH MỤC HÌNH VẼ

- Hình 1.1 Quá trình xử lý ảnh
- Hình 1.2 Các bước cơ bản trong một hệ thống xử lý ảnh
- Hình 1.3 Nền ảnh phức tạp
- Hình 1.4 Hệ thống nhận diện mặt người đặc trưng
- Hình 1.5 Kết cấu khuôn mặt
- Hình 1.6 Màu sắc da mặt
- Hình 2.1. Đồ tìm thông tin lốm
- Hình 2.2 Đồ tìm vùng lốm trên ảnh
- Hình 2.3 Tập các bộ lọc
- Hình 2.4 Đồ tìm vùng lốm ở nhiều mức khác nhau
- Hình 2.5 Ảnh tích phân
- Hình 2.6 Tính tổng độ sáng cho hình chữ nhật $R(l,t,r,b)$
- Hình 2.7 Tạo cây cấp bậc
- Hình 2.8 Một cây rút trích từ khuôn mặt
- Hình 2.9 Vị trí vùng tương đối của nút con
- Hình 2.10 Cách tính vectơ đại diện độ sáng cho nút
- Hình 2.11 Cây rút trích được và các thông tin trên một nút
- Hình 2.12 Mô hình phát hiện khuôn mặt
- Hình 3.1 Các ảnh thử nghiệm
- Hình 3.2 Giao diện chính của chương trình
- Hình 3.3 Giao diện phát hiện khuôn mặt
- Hình 3.4 Phát hiện đúng ra khuôn mặt
- Hình 3.5 Phát hiện đúng ra khuôn mặt
- Hình 3.6 Chỉ phát hiện được một khuôn mặt
- Hình 3.7 Không phát hiện được ra khuôn mặt

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.

Em 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 5 tháng 6 năm 2013.

Sinh viên

Nguyễn Thị Thơm

LỜI 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.. tôi đã chọn đề tài nghiên cứu: TÌM HIỂU ĐẶC TRƯNG LỖM 3D VÀ BÀI TOÁN PHÁT HIỆN MẶT NGƯỜI TRONG ẢNH làm đề tài nghiên cứu và bảo vệ luận văn tốt nghiệp đại học của mình.

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

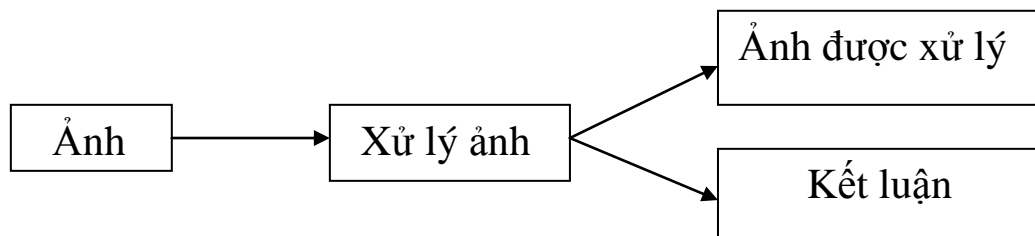
1.1. Khái quát về xử lý ảnh

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.

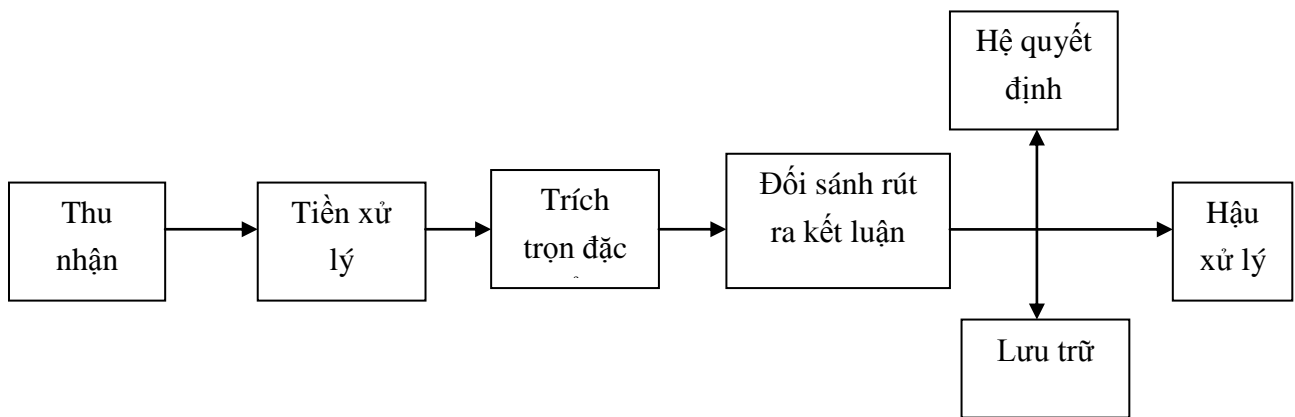
Quá trình xử lý nhận dạng ảnh được xem như là quá trình thao tác ảnh đầu vào nhằm cho ra kết quả mong muốn. Kết quả đầu ra của một quá trình xử lý ảnh có thể là một ảnh “đã được xử lý” hoặc một kết luận.



Hình 1.1. Quá trình xử lý ảnh

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:



Hình 1.2. Các bước cơ bản trong 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

a. Tiền xử lý

Tiền xử lý là giai đoạn đầu tiên trong xử lý ảnh số. Tùy thuộc vào quá trình xử lý tiếp theo trong giai đoạn này sẽ thực hiện các công đoạn khác nhau như: nâng cấp, khôi phục ảnh, nắn chỉnh hình học, khử nhiễu v.v..

b. Trích chọn đặc điểm

Các đặc điểm của đối tượng được trích chọn tùy theo mục đích nhận dạng trong quá trình xử lý ảnh. Trích chọn hiệu quả các 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.

c. 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

a. 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.

b. 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.

1.2. Bài toán nhận dạng mặt người

1.2.1. Bài toán nhận dạng mặt người

Hệ thống nhận dạng mặt người là một hệ thống nhận vào là một ảnh hoặc một đoạn video (một chuỗi các ảnh). Qua xử lý tính toán hệ thống xác định được vị trí mặt người trong ảnh (nếu có) và xác định là người nào trong số những người hệ thống đã được biết (qua quá trình học) hoặc là người lạ.

1.2.2. Những khó khăn của nhận dạng khuôn mặt

a. Tư thế góc chụp : Ảnh chụp khuôn mặt có thể thay đổi rất nhiều bởi vì góc chụp giữa camera và khuôn mặt. Chẳng hạn như : chụp thẳng, chụp xéo bên trái 45^0 hay xéo bên phải 45^0 , chụp từ trên xuống, chụp từ dưới lên, v.v...). Với các tư thế khác nhau, các thành phần trên khuôn mặt như mắt, mũi, miệng có thể bị khuất một phần hoặc thậm chí khuất hết. Khuôn mặt đang nhìn thẳng nhưng góc chụp của máy ảnh lại lệch nhiều so với hướng nhìn thẳng của khuôn mặt hoặc là lúc chụp ảnh mặt người quay nghiêng sang một bên nào đó nhiều đều là những khó khăn rất lớn trong bài toán nhận diện mặt người.

b. Sự xuất hiện hoặc thiếu một số thành phần của khuôn mặt: Các đặc trưng như râu mép, râu hàm, mắt kính, v.v... có thể xuất hiện hoặc không. Vấn đề này làm cho bài toán càng khó khăn hơn rất nhiều.

c. Sự biểu cảm của khuôn mặt: Biểu cảm của khuôn mặt người có thể làm ảnh hưởng đáng kể lên các thông số của khuôn mặt. Chẳng hạn, cùng một khuôn mặt một người, nhưng có thể sẽ rất khác khi họ cười hoặc sợ hãi....

d. Sự che khuất: Khuôn mặt có thể bị che khuất bởi các đối tượng khác hoặc các khuôn mặt khác.

e. Hướng của ảnh: Các ảnh của khuôn mặt có thể biến đổi rất nhiều với các góc quay khác nhau của trục camera. Chẳng hạn chụp với trục máy ảnh nghiêng làm cho khuôn mặt bị nghiêng so với trục của ảnh.

f. Điều kiện của ảnh: Ảnh được chụp trong các điều kiện khác nhau về: chiếu sáng, về tính chất camera (máy kỹ thuật số, máy hồng ngoại, v.v...) ảnh hưởng rất nhiều đến chất lượng ảnh khuôn mặt.

g. 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.



Hình 1.3 Nền ảnh phức tạp

h. 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.

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 móng 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.

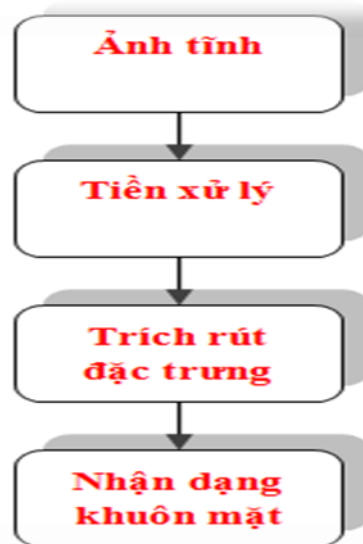
1.2.4. Các ứng dụng đặc trưng của bài toán nhận diện mặt người

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, vân 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 vân 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:



Hình 1.4 Hệ thống nhận diện mặt người đặc trưng

- 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 :

- 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”, “phần

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:

$$\begin{cases} HI(x) = \sum_{y=1}^n I(x, y) \\ VI(y) = \sum_{x=1}^m I(x, y) \end{cases} \quad (2.1)$$

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.

a.Các đặc trưng khuôn mặt :

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

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 mô 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%.

b.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 16×16 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%.



Hình 1.5 Kết cấu khuôn mặt

c.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.



Hình 1.6 Màu sắc da mặt

d. Đa đặc trưng

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

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:

a. 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.

b. 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 các 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 giải 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.

1.3. Phương pháp phát hiện mặt người dựa trên đặc trưng lõm

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 và valley (lõm). 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.

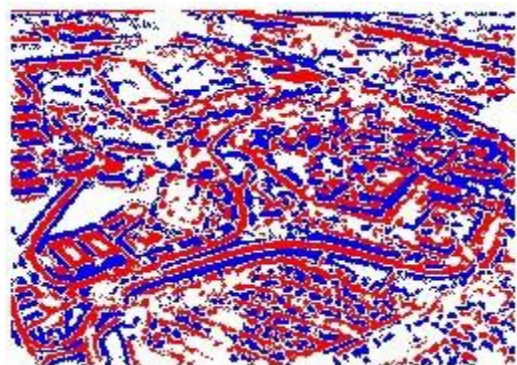
CHƯƠNG 2 : TÌM HIỂU VỀ CÁC ĐẶC TRƯNG KHUÔN MẶT

2.1. Đặc trưng lõm

Ta có thể hiểu một cách đơn giản rằng: lõm là những điểm thấp nhất, ví dụ như lõm là những thung lũng. Trong ảnh vệ tinh, lõm là đường khe chạy giữa hai dãy núi. Trong ảnh y tế, lõm là các thành phần trũng xuống. Còn trong ảnh khuôn mặt, các phần hốc mắt được coi là lõm.

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 lõm tương tự như kết quả của Hải Trần. Các điểm lõm được tô màu xanh. Chúng ta lưu ý rằng các điểm lõm rút trích được chỉ biểu diễn được các vùng lõm nhỏ, chi tiết, chẳng hạn như: mắt, chân mày, mũi, và miệng.

Trong những phần tiếp theo, chúng ta sẽ dò tìm các đặc trưng lõm tại các mức khác nhau. Ví dụ như, tại mức thấp nhất, cần dò tìm các đặc trưng lõm 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 lõm lớn 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 lõm

2.2. Rút trích các đặc trưng lõm

2.2.1. Điểm lõm

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 lõm trên mặt này. Một cách chính xác theo toán học, điểm lõm là những điểm mà tại đó cả hai đạo hàm:

$$\frac{\partial^2 I}{\partial x^2} \text{ và } \frac{\partial^2 I}{\partial y^2} \tag{2.2}$$

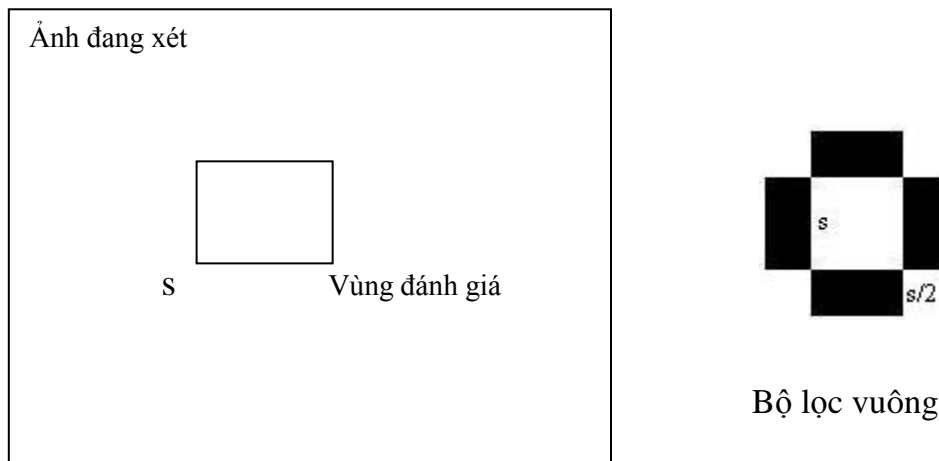
đạt cực trị. Trên thực tế, do ảnh là một hàm số không liên tục, hàm rời rạc do sự số hóa của máy tính, vì vậy sẽ rất khó khăn nếu chúng ta chỉ dò tìm các điểm cực trị. Và chúng ta sẽ bị mất một số thông tin của các điểm ảnh lõm cũng như những điểm lân cận chúng. Vì vậy ta đưa ra các công thức sau đây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

$$V_P = \{(x,y) \mid |LoG(x,y)| \geq \delta \quad LoG(x,y) < 0 \tag{2.3}$$

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

2.2.2. Dò tìm và lấy vùng lõm

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



Hình 2.2 Dò tìm vùng lõm trên ảnh

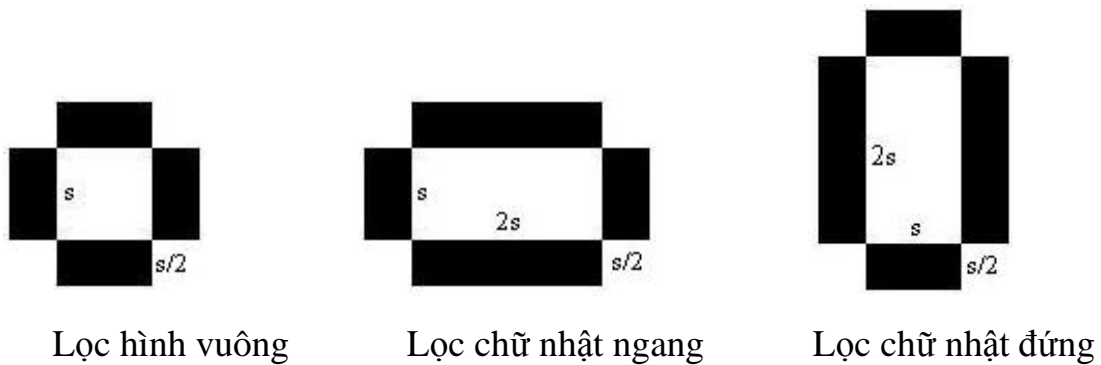
Ứng với vùng đang xét – một hình vuông $s \times s$, chúng ta tính giá trị của bộ lọc trên vùng này theo công thức (2.3). Đặt $S_1, S_2, S_3,$ và S_4 là các hình chữ nhật đen xung quanh hình vuông R giá trị lọc $F(R)$ được định nghĩa như sau:

$$F(R) = 2sum(R) - \sum_{i=1}^4 sum(S_i) \tag{2.4}$$

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:

$$R = \begin{cases} ridge & \Leftrightarrow |F(R)| \geq \delta \wedge F(R) > 0 \\ valley & \Leftrightarrow |F(R)| \geq \delta \wedge F(R) < 0 \\ normal & otherwise \end{cases} \tag{2.5}$$

Công thức (2.5) là công thức mở rộng của (2.2) và (2.3) nhằm để dò tìm các vùng lõm trên ảnh. Sự khác biệt chủ yếu ở chỗ ta sử dụng bộ lọc hiệu trong (2.5) thay vì sử dụng Laplacian-of-Gaussian trong (2.2) và (2.3). 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 lõm ứng với nhiều kích cỡ, mức độ khác nhau. Ngoài ra, để dò tìm và rút trích các vùng lõm 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.



Hình 2.3 Tập các bộ lọc

Do đó ta có thể biến đổi công thức (2.4) thành công thức sau:

$$F_S(R) = 2sum(R) - \sum_{i=1}^4 sum(S_i) \tag{2.6}$$

$$F_R(R) = 3sum(R) - 2\sum_{i=1}^4 sum(S_i) \tag{2.7}$$

$F_S(R)$ 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 (2.6) và (2.7), hệ số đứng trước hàm $\text{sum}(R)$ 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 $\text{sum}(R)$, 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 (2.5) là ngưỡng ∂ . Ngưỡng ∂ 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 ∂_1 là tích của K và S^2 – diện tích của lọc vuông. Tương tự, trong (2.7), chúng ta đã gấp đôi giá trị của $F_R(R)$ để tránh chuyển tính toán trên số thực. Vì thế, chúng ta cũng phải gấp đôi ngưỡng ∂_2 để tăng độ chính xác của việc đánh giá.

2.2.3. Dò và phát hiện vùng lõm ở nhiều mức khác nhau

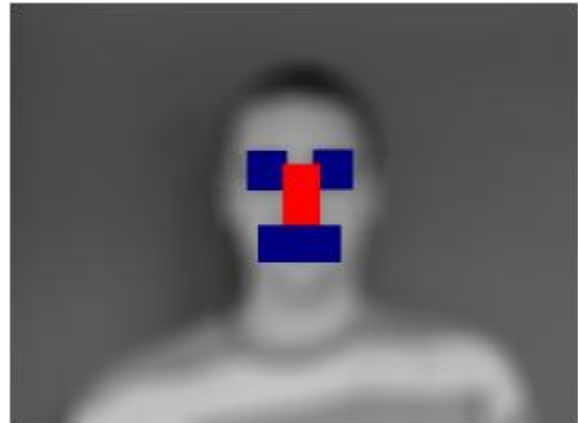
Sau đây là một phương pháp sử dụng các lọc hiệu ở nhiều mức khác nhau để dò tìm các vùng lõm. Các lọc được giới thiệu trong phần trước sẽ được sử dụng với nhiều kích thước khác nhau vì vậy chúng sẽ dò tìm được các vùng lõm ở nhiều mức khác nhau.



$\sigma=10$, ảnh sau khi smoth với

Các vùng lõm dò với kích thước lọc 1

Gaussian $10\sqrt{2}$



$\sigma=20$, ảnh sau khi smoth với

Các vùng lõm dò tìm với kích

Gaussian $20\sqrt{2}$

thước lọc 20

Hình 2.4 Dò tìm vùng lõm 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. ∂_1 là mức độ thấp nhất. Tại mức độ này, chúng ta có thể dò tìm các vùng lõm, có kích thước nhỏ và chi tiết. ∂_n là mức độ cao nhất mà tại đó chúng ta có thể dò tìm các vùng lõm 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 xanh là các vùng lõm. Trong thực tế, có nhiều vùng lõm khác cũng được dò tìm được, tuy nhiên trong hình (2.4), chỉ tập trung vào các vùng lõm 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 lõm rút trích được này thành các cây cấp bậc.

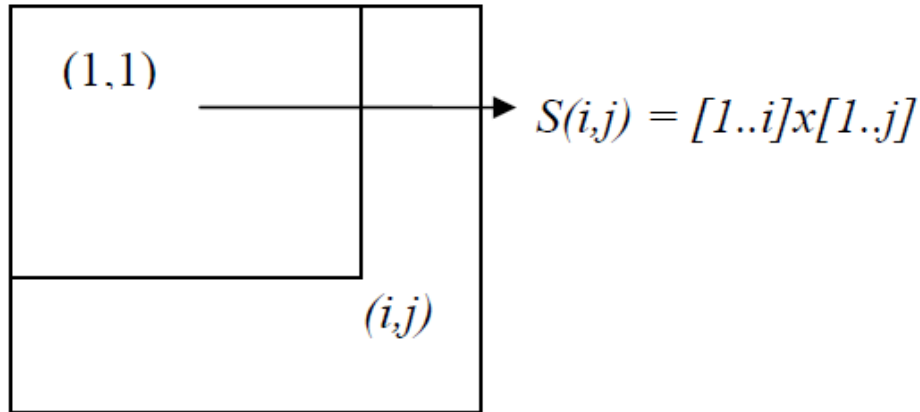
2.2.3.1. Tối ưu tốc độ dò tìm

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, 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:

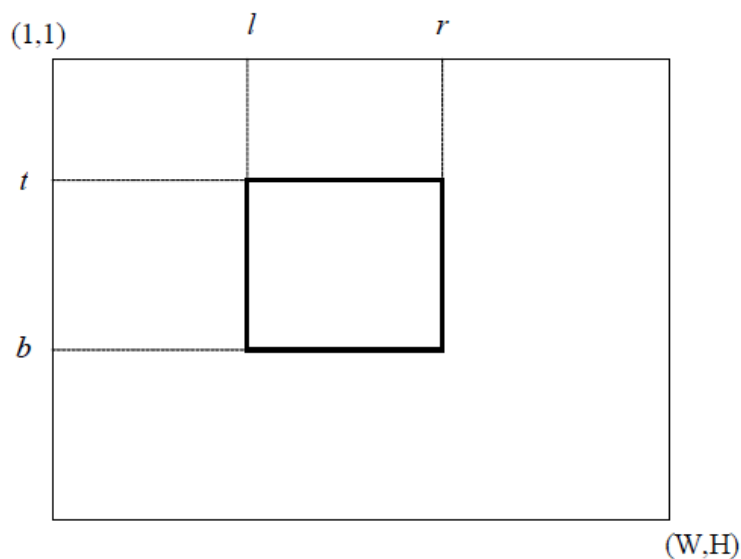
$$S(x, y) = \begin{cases} 0 & x = 0 \vee y = 0 \\ S(x-1, y) + S(x, y-1) - S(x-1, y-1) + I(x, y) & x \neq 0 \wedge y \neq 0 \end{cases} \quad (2.8)$$



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:

$$S_R = S(r, b) - S(r, t) - S(l, b) + S(l, t) \quad (2.9)$$



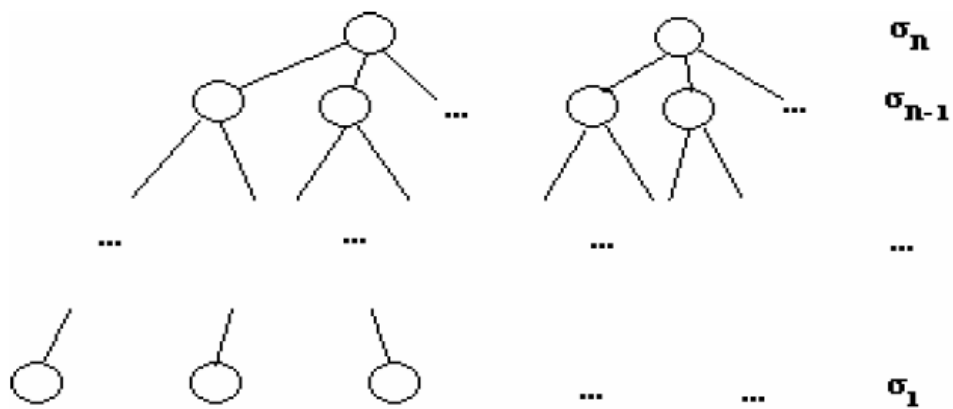
Hình 2.6: Tính tổng độ sáng cho hình chữ nhật $R(l, t, r, b)$

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.

2.2.4. Phát hiện mặt người

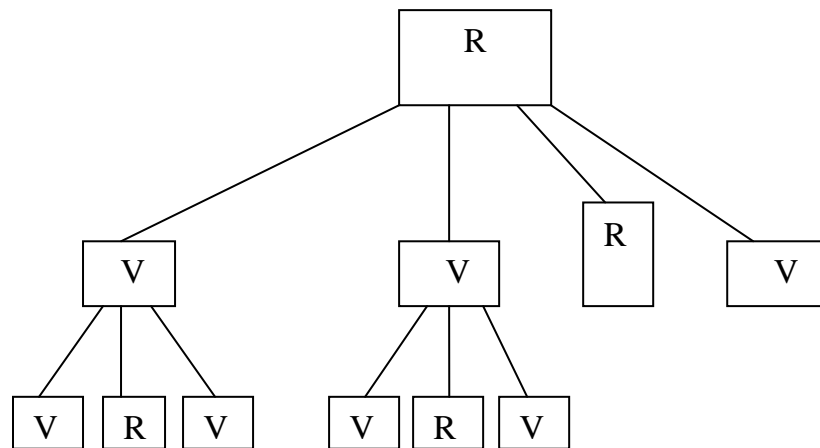
2.2.4.1. Xây dựng cấu trúc cây lổm

Từ các vùng lổm rút trích được, chúng ta xây dựng các cây quan hệ cấp bậc. Những vùng được rút trích ở mức cao nhất σ_n sẽ trở thành những nút gốc của các cây sẽ tạo. Những vùng rút trích được ở mức thấp nhất σ_1 sẽ được chèn vào cây như là những nút lá.



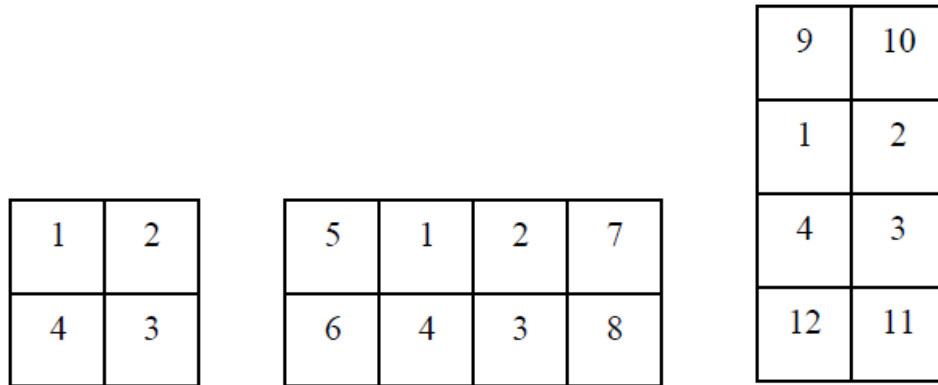
Hình 2.7 Tạo cây cấp bậc

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 σ_1 bị che khuất hơn hai phần ba diện tích của nó bởi vùng R' của mức σ_{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 :



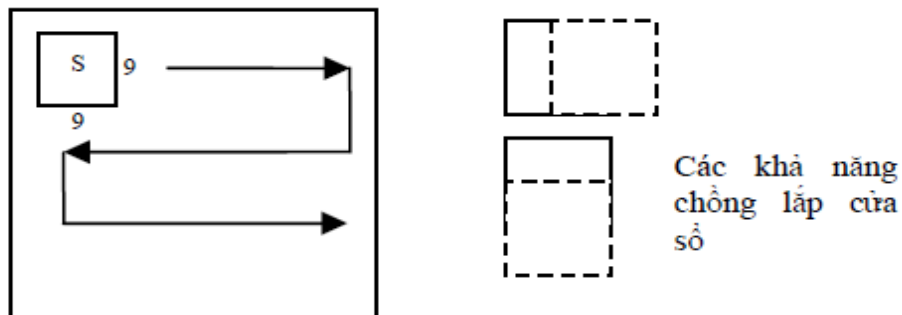
Hình 2.8 Một cây rút trích từ khuôn mặt

Đối với mỗi nút trong cây, chúng ta cần lưu các thông tin của nút này như: loại (lỗm), 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.



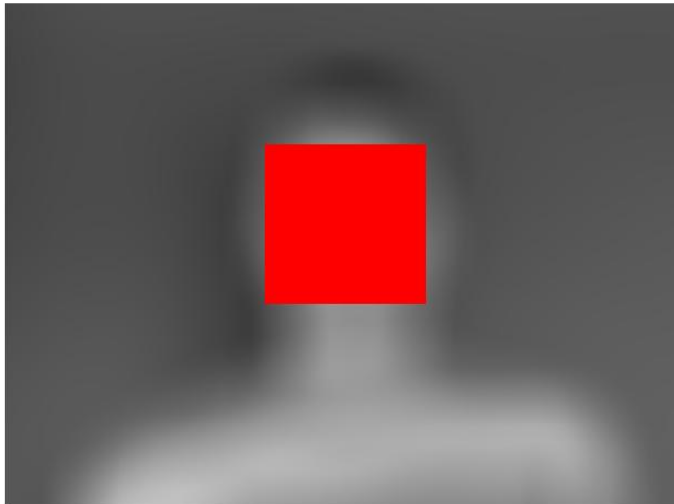
Hình 2.9: vị trí vùng tương đối của nút con

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 lỗm (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.

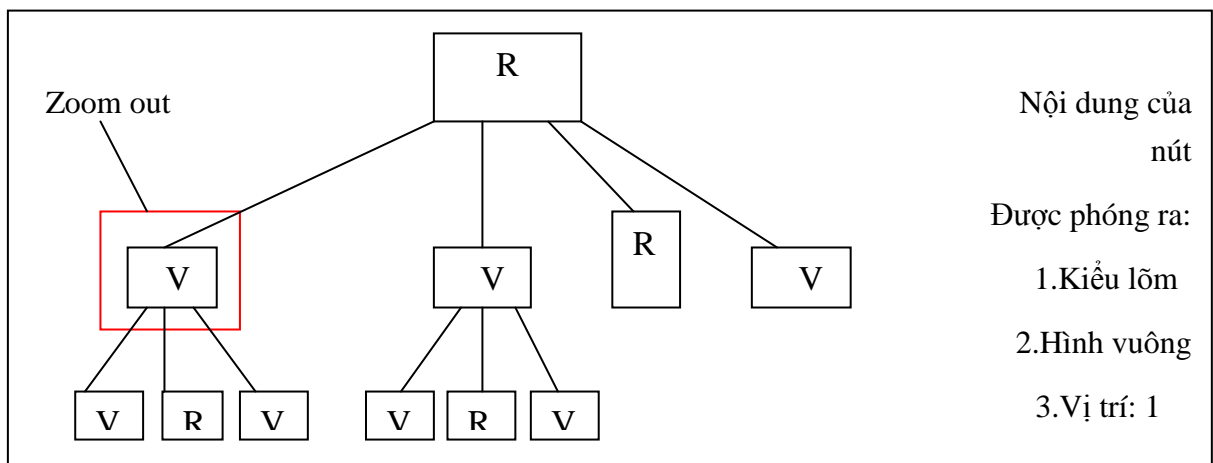


Hình 2.10: cách tính vectơ đại diện độ sáng cho nút

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 lõm. 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 tia bớt nhánh cho các cây. Nếu một nút cùng kiểu (lõm) so với nút cha của nó thì nó được xem là một nút lặp. Trong trường hợp này, nút lặp sẽ được tia khỏi cây. Nếu nó là một nút lá, chúng ta dễ dàng xóa nó khỏi cây.

2.2.4.2. 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 lõm 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.

2.2.4.2.1. 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 σ_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:

$$N_1, N_2 \in Node = \{(t,s,p,i)\} \tag{2.10}$$

Trong đó:

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

$s = \{\text{square, horizontal rectangle, vertical rectangle}\}$ 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 σ_i càng cao, kích thước của vùng 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:

$$NS(N_1, N_2) = \alpha_1 E(N_1.t, N_2.t) + \beta_1 E(N_1.s, N_2.s) + \gamma_1 E(N_1.p, N_2.p) + (1 - \alpha_1 - \beta_1 - \gamma_1) EV(N_1.i, N_2.i) \tag{2.11}$$

Trong công thức (2.11), α_1 , β_1 , và γ_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:

$$E(x, y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases} \tag{2.12}$$

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) .

$$EV(x, y) = 1 - \frac{\sqrt{\sum_{i=1}^n (x.a_i - y.a_i)^2}}{255\sqrt{n}} \tag{2.13}$$

Do mỗi thành phần ai 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 (N_1, N_2) $\in [0,1]$ với mọi cặp nút N_1 và N_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 T_1 và T_2 và hai cây đang cần tính độ tương đồng. T_1 và T_2 có những thuộc tính sau:

$$T_1, T_2 \in Tree = \{(r, n, c_1, c_2, \dots, c_n)\} \tag{2.14}$$

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:

$$TS(T_1, T_2) = \alpha_2 NS(T_1.r, T_2.r) + (1 - \alpha_2) \frac{\sum_{i=1}^{T_1.n} \sum_{j=1}^{T_2.n} TS(T_1.c_i, T_2.c_j)}{T_1.n T_2.n} \tag{2.15}$$

Trong đó α_2 là thông số thực nghiệm. Các thông số thực nghiệm α_1, β_1 , và γ_1 trong công thức (2.10) và α_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 (2.15) 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 chúng tôi có thử nghiệm đặc trưng lỗi lờm 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:

$$NS(N_1, N_2) = \alpha_1 E(N_1.t, N_2.t) + \beta_1 E(N_1.s, N_2.s) + (1 - \alpha_1 - \beta_1) E(N_1.p, N_2.p) \tag{2.16}$$

2.2.4.2.2. Không gian cây và khoảng cách giữa hai cây

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:

$$\|T_1, T_2\| = 1 - TS(T_1, T_2) \tag{2.17}$$

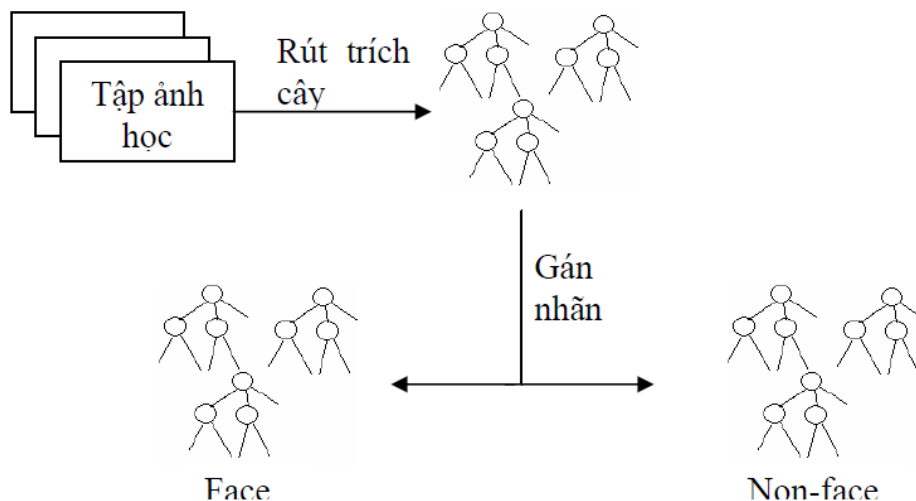
Để dàng nhận thấy, khoảng cách giữa hai cây là một giá trị thuộc đoạn [0,1]. Và khoảng cách tiến dần về bằng không khi hai cây càng giống nhau.

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. Phần mô hình này sẽ được trình bày trong chương sau. Các phần thảo luận về khả năng mở rộng của mô hình này cho bài toán nhận dạng khuôn mặt và phân loại đối tượng sẽ được trình bày trong chương kết luận và hướng phát triển.

2.3. Phát hiện mặt người trong ảnh dựa trên đặc trưng lôm

2.3.1. Gán nhãn

Từ tập ảnh học, ta dùng phương pháp trình bày trong chương 3 để rút trích các cây lôm trên ảnh. Các cây này được gán nhãn thủ công. Mỗi cây bất kỳ được gán nhãn là mặt (face) hay không phải khuôn mặt (non-face). Như vậy, từ một tập ảnh học ta rút trích được hai tập cây. Tập cây biểu diễn khuôn mặt và tập cây biểu diễn đối tượng không phải là khuôn mặt.



2.3.2. Thống kê

Mục đích của bước này là xấp xỉ được một cấu trúc cây khuôn mặt. Trong bước gán nhãn, tôi tiến hành rút trích cây và gán nhãn. Kết quả thu được một số lượng cây thuộc lớp “khuôn mặt”. Các cây này được dùng làm dữ liệu học thống kê. Mô hình thống kê được sử dụng như sau:

Đặt $S(T_x)$ là tổng khoảng cách của một cây T_x bất kỳ đến tất cả các cây còn lại trong tập học.

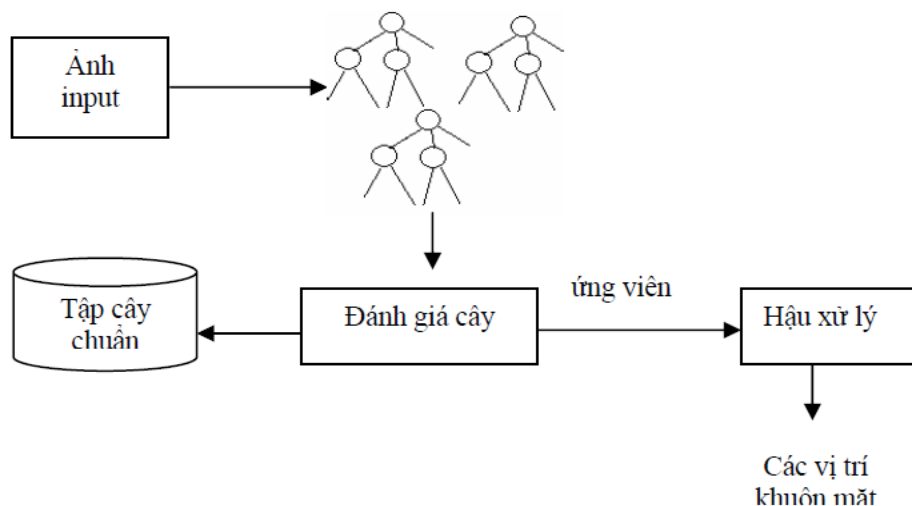
$$S(T_x) = \sum_{i=1}^{i=n \wedge i \neq x} TS(T_x, T_i) \tag{2.18}$$

Ta sẽ có tập học trở thành

$$Training = \{(T_i, S(T_i)), \forall T_i \in FaceSet\} \tag{2.19}$$

Trong đó FaceSet là tập các cây được gán nhãn là “khuôn mặt”. Việc học thống kê theo mô hình của này khá đơn giản. Trước tiên, chọn trong tập Training ra k cây có giá trị S(T) bé nhất. k cây này sẽ là đại diện tiêu biểu cho một lớp khuôn mặt.

Từ kết quả k cây rút trích được, tạm gọi là tập cây chuẩn, ta có thể dùng chúng để phát hiện khuôn mặt theo mô hình như sau:



Hình 2.12 Mô hình phát hiện khuôn mặt

Trong mô hình này, trước tiên từ ảnh đầu vào ta dùng phương pháp rút trích lõm ta sẽ nhận được một danh sách các cây lõm. Ứng với mỗi cây ta dựa vào tập

cây chuẩn để kiểm tra xem liệu đây có phải là cây biểu diễn “khuôn mặt” hay không. Nếu thỏa điều kiện đánh giá nó sẽ trở thành ứng viên khuôn mặt. Trong bước cuối cùng, hậu xử lý, các ứng viên sẽ được kiểm tra ràng buộc. Và sau cùng là khử trùng lặp khuôn mặt do có một số cây có một phần của nút gốc chồng lên nhau.

2.3.3. Đánh giá dùng cho phát hiện khuôn mặt

Dựa trên tập cây chuẩn, một cây bất kỳ được đánh giá có khả năng là khuôn mặt hay không dựa trên trung bình khoảng cách của cây đang xét đến tất cả các cây trong tập cây chuẩn.

$$\overline{D}(t) = \frac{\sum_{i=1}^k \|t, T_i\|}{k} \quad (2.20)$$

Trong đó:

$\overline{D}(t)$: là trung bình khoảng cách của cây t đến các cây trong tập chuẩn

t : là cây đang xét

T_i : là các cây trong tập k cây chuẩn

Căn cứ vào giá trị \overline{D} , cây t được đánh giá là ứng viên mặt hay không dựa theo công thức sau:

$$t \in FaceCandidate \Leftrightarrow \overline{D}(t) \leq \delta_1 \quad (2.21)$$

Ngưỡng δ_1 là ngưỡng thực nghiệm, ngưỡng này sẽ được lấy khoảng 0.2 đến 0.4. Ngoài cách tính trung bình khoảng cách, ta có thể đánh giá ứng viên thông qua việc tính trung bình độ tương đồng của cây ứng viên với tập cây chuẩn.

$$\overline{S}(t) = \frac{\sum_{i=1}^k TS(t, T_i)}{k} \quad (2.22)$$

Trong đó:

$\overline{S}(t)$: là trung bình độ tương đồng của cây t so với các cây trong tập chuẩn

t : là cây đang xét

T_i : là các cây trong tập k cây chuẩn

Tương tự như trường hợp trên, giá trị $\overline{S}(t)$ sẽ được dùng để đánh xem cây t có khả năng là khuôn mặt hay không.

$$t \in \text{FaceCandidate} \Leftrightarrow \overline{S}(t) \geq \delta_2 \quad (2.23)$$

Để thấy ngưỡng δ_2 là số gần với một, có thể nói $\delta_2 \cong 1 - \delta_1$. Ngưỡng này sẽ được trong lấy khoảng từ 0.6 đến 0.8. Tuy nhiên, về bản chất hai cách này hoàn toàn giống nhau.

2.3.4. Hậu xử lí

Các kết quả thu được cho thấy rằng có một vùng khuôn mặt được phát hiện có một phần chồng lấp lên nhau. Có thể phát hiện ra nhiều cây ứng viên được đánh giá là “khuôn mặt”, nhưng thật sự chúng là sự chồng lấp của nhiều cây lên một vị trí khuôn mặt. Vì vậy, chúng ta cần một thuật toán tốt để giảm những phát hiện các ứng viên trùng này. Trước hết ta sử dụng một đánh giá đơn giản để liên kết các vùng trùng lấp. Nếu hai ứng viên khuôn mặt có vùng diện tích bị chồng lấp lên nhau quá một phần ba diện tích của nó, chúng được xem là các mặt trùng và được liên kết lại thành một vùng lớn hơn. Vùng diện tích xem xét đó chính là vùng diện tích của nút gốc của cây ứng viên.

Sau khi gom nhóm được vùng, thu được một hay nhiều vùng chứa khuôn mặt. Tuy nhiên, do sai số của kích thước của cây và kích thước khuôn mặt, nên luôn có được vùng khuôn mặt phát hiện được lớn hơn khuôn mặt thực tế. Áp dụng tiếp kỹ thuật xác định vùng lõm với mức nhỏ hơn nhằm để phát hiện chính xác vùng. ta chỉ áp dụng thuật toán dò tìm đặc trưng lõm trên vùng bao khuôn mặt, sau đó liên kết các vùng lõm cùng loại lại với nhau.

Vùng khuôn mặt thật sự sẽ là vùng liên thông lớn nhất trong đó. Với cách như vậy, chúng ta sẽ thu nhỏ, và xác định chính xác hơn vị trí khuôn mặt.

CHƯƠNG 3 CHƯƠNG TRÌNH THỬ NGHIỆM

3.1. Bài toán

- Nội dung : Tìm hiểu đặc trưng lồi 3D và bài toán phát hiện mặt người trong ảnh
- + Đầu vào : Ảnh có chứa mặt người.
- + Đầu ra : Ảnh vào có thêm khoanh vùng mặt phát hiện được.

3.2. Phân tích thiết kế

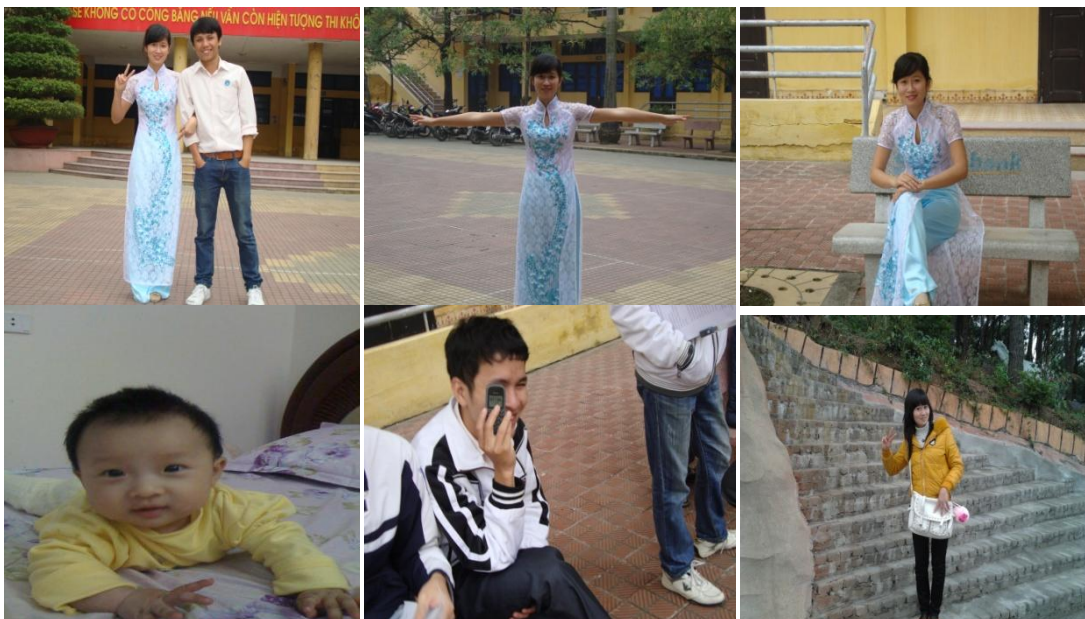
Bài toán : Tìm hiểu đặc trưng lồi 3D và Bài toán phát hiện mặt người trong ảnh chương trình thực hiện trên open CV

Phát hiện mặt người trong ảnh có nhiều ứng dụng trong cuộc sống, tuy nhiên, trong khóa luận này, tôi chỉ xây dựng một chương trình nhỏ để minh họa cho các lý thuyết ở trên. Cụ thể đây là một chương trình phát hiện mặt người trong ảnh được viết trên nền visual C# của Microsoft, sử dụng thư viện mã nguồn mở OpenCV của Intel.

Dựa trên thư viện Luxand để thực hiện chương trình: Phát hiện, đánh dấu và lọc vùng khuôn mặt.

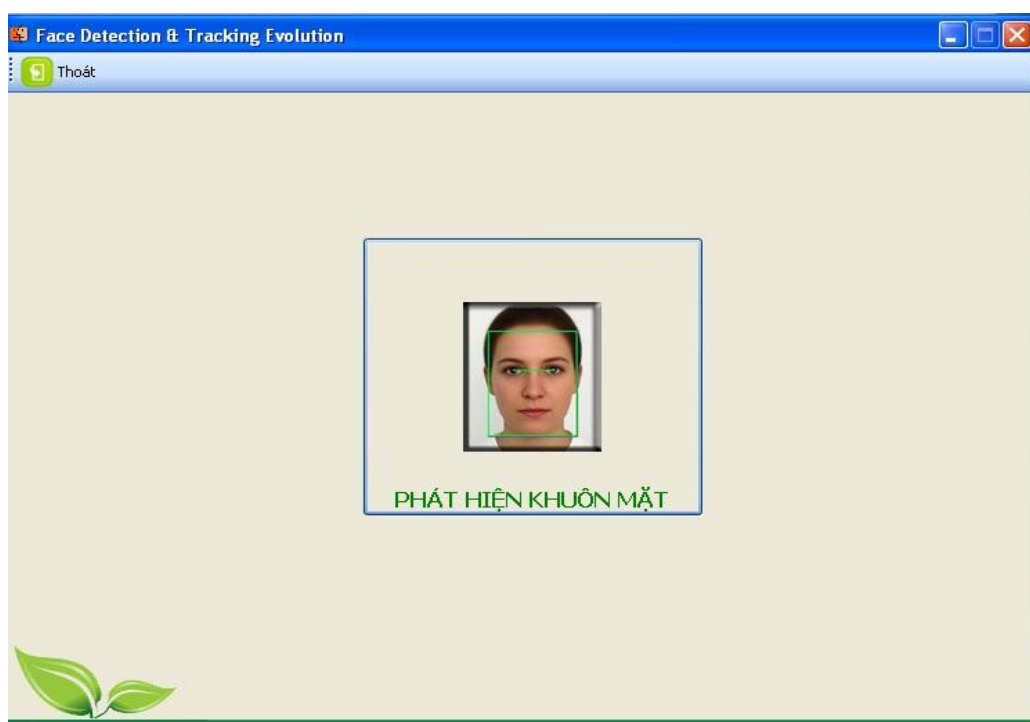
3.3. Một số kết quả chương trình

3.3.1 Tập ảnh thử nghiệm



Hình 3.1 Các ảnh thử nghiệm

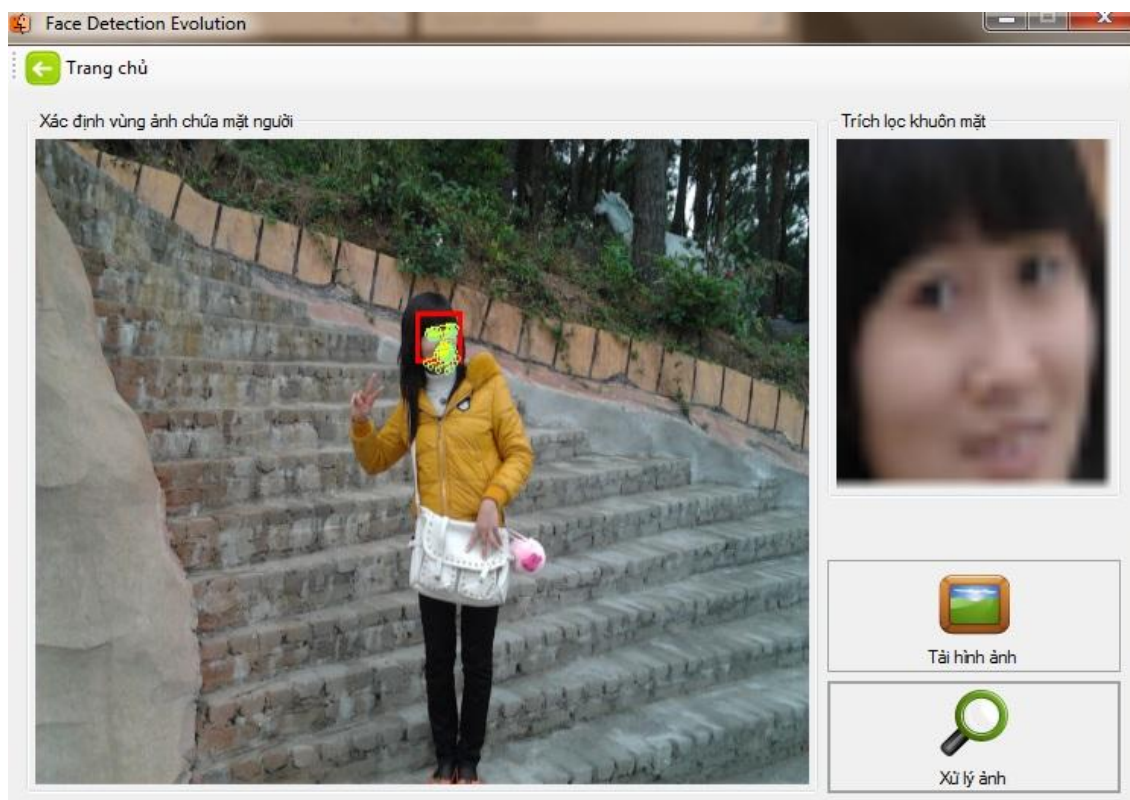
3.3.2 Kết quả chương trình



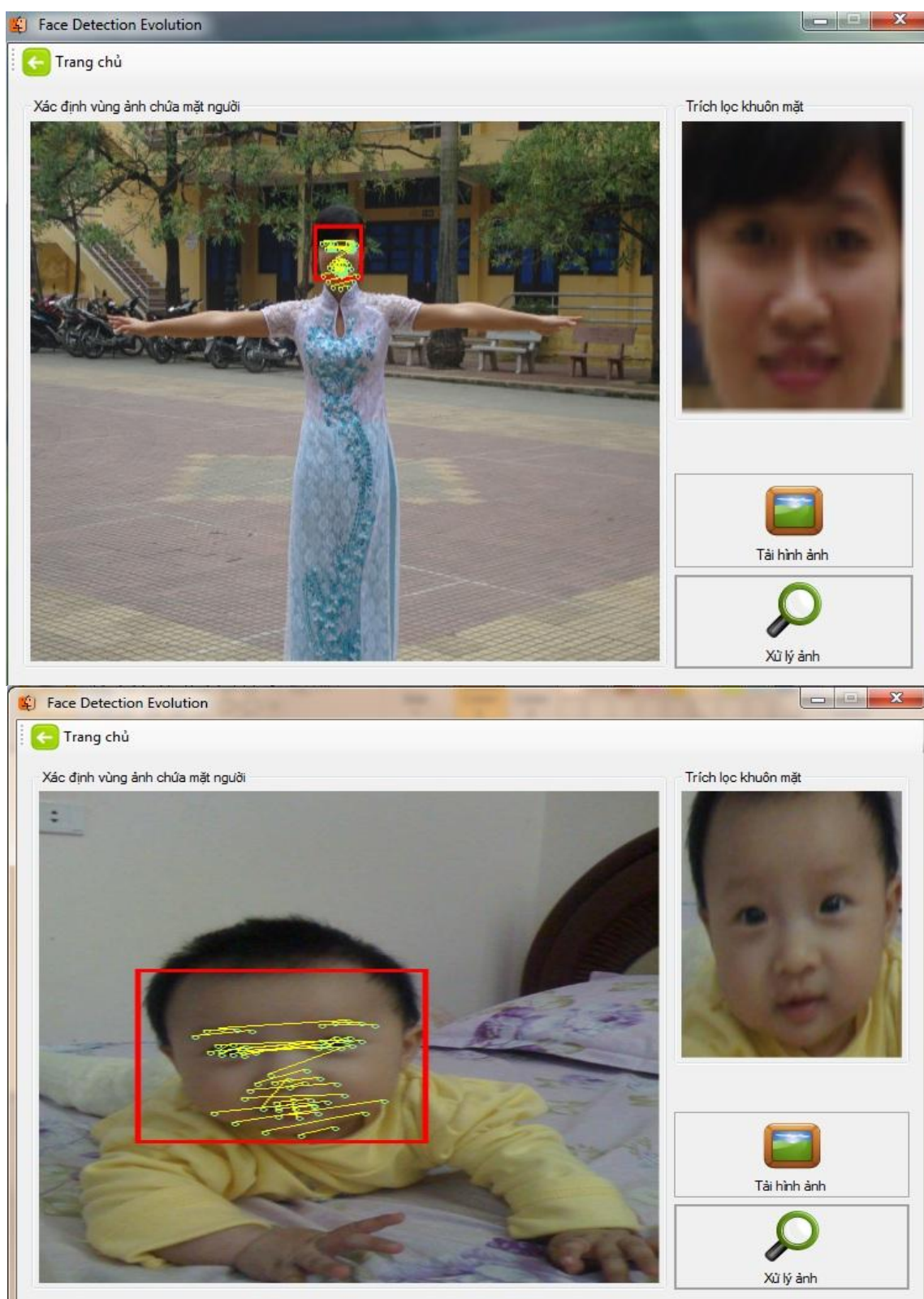
Hình 3.2 Giao diện chính của chương trình



Hình 3.3 Giao diện phát hiện khuôn mặt



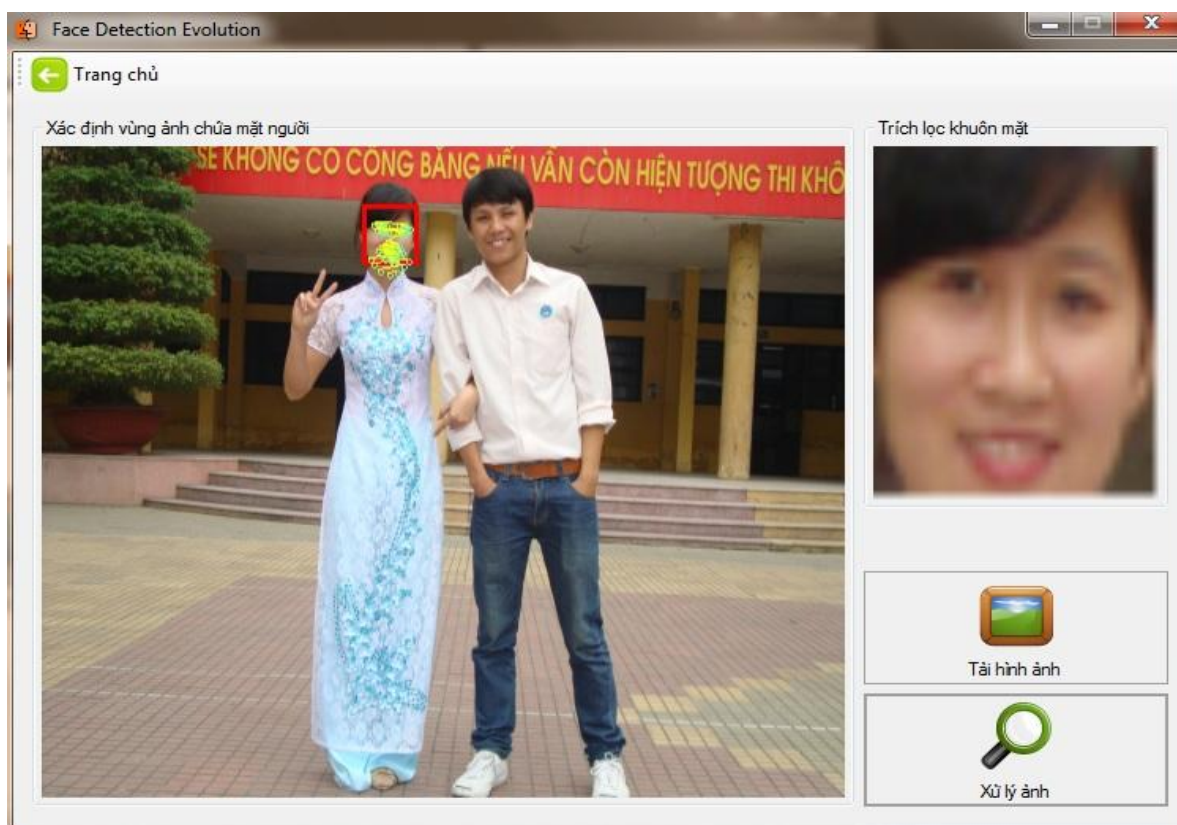
Hình 3.4 Phát hiện đúng ra khuôn mặt



Hình 3.5 Phát hiện đúng ra khuôn mặt

3.3.3 Kết quả thực nghiệm

Tập ảnh thử nghiệm	Kết quả	Đúng	Sai	Tỉ lệ chính xác
ảnh đầu vào	20	17	3	85 %



Hình 3.6 Chỉ phát hiện được một khuôn mặt



Hình 3.7 Không phát hiện được ra khuôn mặt

PHẦN KẾT LUẬN

Qua luận văn này, tôi đã đóng góp được một hướng tiếp cận mới cho bài toán phát hiện khuôn mặt. Bên cạnh đó, hứa hẹn việc mở ra một loại đặc trưng mới có thể phát triển cho xác định vị trí của các thành phần trên khuôn mặt như mắt, mũi, miệng, v.v... Ngoài ra có thể phát triển rộng hơn để trở thành bài toán nhận dạng khuôn mặt.

Tôi đã tìm hiểu được bài toán phát hiện mặt người dựa trên đặc trưng lõm 3D để nhận dạng được khuôn mặt và xác định được các ứng dụng đặc trưng của bài toán phát hiện mặt người. Từ đó ta xây dựng được hệ thống nhận diện người .

Chương trình phát hiện mặt người trong ảnh đã nhân dạng được ra khuôn mặt và các chi tiết trên khuôn mặt. Nhằm đưa ra được những ứng dụng thực tế trong cuộc sống hiện đại:

- Phân tích cảm xúc của khuôn mặt người.
- Tổ chức tìm kiếm liên quan đến con người thông qua khuôn mặt trên nhiều hệ cơ sở dữ liệu lớn.
- Điều khiển ra vào các cơ quan, văn phòng.
- Hệ thống quan sát, theo dõi và bảo vệ.
- Giải trí.
- Nhận dạng tội phạm.
- Hệ thống giao tiếp thông minh giữa người và máy.

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.
- [8]. A. Lanitis, C. J. Taylor, T. F. Cootes, “An Automatic Face Identification System Using Flexible Appearance Models”, Image and Vision Computing, Vol. 13, No. 5, pp. 393-401, 1995.

[B]. Tài liệu Tiếng Việt

- [9]. PGS.TS Đỗ Năng Toàn, TS. Phạm Việt Bình, “Giáo trình Xử Lý Ảnh”, Đại học Thái Nguyên, 2007.
- [10]. 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.
- [11]. Luxand FaceSDK 4.0, Trial Version, Luxand.Inc, 2011.

<http://www.luxand.com>.