|  |
| --- |
| BAN CƠ YẾU CHÍNH PHỦ  **HỌC VIỆN KỸ THUẬT MẬT MÃ**  ¯¯¯¯¯¯¯¯¯¯¯¯¯¯¯¯  Logo HvKTMM |
| ĐỒ ÁN TỐT NGHIỆP  **NGHIÊN CỨU, XÂY DỰNG ỨNG DỤNG NHẬN DIỆN KHUÔN MẶT TRÊN ANDROID** |
| Ngành: An toàn thông tin  Mã số: |
| Hà Nội, 2020 |
| BAN CƠ YẾU CHÍNH PHỦ  **HỌC VIỆN KỸ THUẬT MẬT MÃ**  ¯¯¯¯¯¯¯¯¯¯¯¯¯¯¯¯  Logo HvKTMM |
| ĐỒ ÁN TỐT NGHIỆP  **NGHIÊN CỨU, XÂY DỰNG ỨNG DỤNG NHẬN DIỆN KHUÔN MẶT TRÊN ANDROID** |
| Ngành: An toàn thông tin  Mã số:  *Sinh viên thực hiện*:  Mã SV: AT120626 Họ và tên: **Phạm Quang Huy**    *Người hướng dẫn 1*:  **TS. Phạm Duy Trung**  Khoa An toàn thông tin – Học viện Kỹ thuật mật mã  *Người hướng dẫn 2*:  **TS. Nguyễn Nhất Hải**  Viện Công nghệ thông tin và Truyền Thông – Đại học Bách Khoa Hà Nội |
| Hà Nội, 2020 |

Mục lục

MỤC LỤC

[Danh mục kí hiệu và viết tắt](#_Toc37365136)

[Danh mục hình vẽ](#_Toc37365137)

[Danh mục bảng](#_Toc37365138)

[Lời cảm ơn](#_Toc37365139)

[Lời nói đầu](#_Toc37365140)

[CHƯƠNG 1. CƠ SỞ LÝ THUYẾT VÀ BÀI TOÁN TÌM KIẾM KHUÔN MẶT 1](#_Toc37365141)

[1.1. Khái quát về xử lý ảnh 1](#_Toc37365142)

[1.1.1. Khái niệm xử lý ảnh 1](#_Toc37365143)

[1.1.2. Một số vấn đề cơ bản trong xử lý ảnh 1](#_Toc37365144)

[1.2. Bài toán phát hiện khuôn mặt 2](#_Toc37365145)

[1.2.1. Một số đặc trưng về khuôn mặt 2](#_Toc37365146)

[1.2.2. Tìm hiểu về bài toán phát hiện khuôn mặt 3](#_Toc37365147)

[1.2.3. Các phương pháp chính phát hiện mặt người 4](#_Toc37365148)

[1.2.4. Những khó khăn trong việc phát hiện khuôn mặt người 5](#_Toc37365149)

[1.2.5. Tìm hiểu về phương pháp phát hiện khuôn mặt của Viola và Johns 5](#_Toc37365150)

[TÀI LIỆU THAM KHẢO 9](#_Toc37365151)

Danh mục kí hiệu và viết tắt

Danh mục hình vẽ

Danh mục bảng

Lời cảm ơn

Trước tiên em xin được bày tỏ sự trân trọng và lòng biết ơn đối với thầy giáo TS. Phạm Duy Trung – Khoa An toàn thông tin và thầy giáo TS. Nguyễn Nhất Hải - Viện Công nghệ thông tin và Truyền Thông Đại học Bách Khoa Hà Nội. Trong suốt thời gian làm đồ án tốt nghiệp, các 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 Học viện Kỹ thuật Mật Mã 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.

|  |  |
| --- | --- |
|  | **SINH VIÊN THỰC HIỆN ĐỒ ÁN**  Phạm Quang Huy |
|  |  |

Lời nói đầu

Ngày nay, cùng với những tiến bộ vuợt bậc của khoa học kỹ thuật nói chung, bộ môn khoa học xử lý ảnh đã và đang thu được những thành tựu lớn lao và chứng tỏ vài trò không thể thiếu với những ứng dụng sâu rộng trong khoa học kỹ thuật cũng như đời sống xã hội. Một bộ phận của khoa học xử lý ảnh là lĩnh vực thị giác máy tính hiện đang thu hút rất nhiều sự quan tâm của các nhà nghiên cứu xử lý ảnh với mục tiêu xây dựng nên một thế giới trong đó hệ thống thị giác kỳ diệu của con người có thể được mô phỏng bởi các hệ thống máy tính, đem lại khả năng cảm nhận bằng thị giác cho các hệ thống về môi trường xung quanh.

Mơ ước về một hệ thống máy tính có thể hoà nhập vào thế giới con người với đầy đủ các giác quan trong đó thị giác đóng vai trò quan trọng đang dần dần được hiện thực hoá với những đóng góp nghiên cứu của các nhà khoa học trên phạm vi toàn thế giới. Một trong những bài toán nhận được nhiều sự quan tâm của lĩnh vực thị giác máy tính là bài toán nhận dạng khuôn mặt.

Đồng thời trong những năm gần đây, điện thoại di động dần trở thành một phần không thể thiếu trong cuộc sống của con người hiện đại, đồng thời xu hướng phát triển các ứng dụng trên các nền tảng di động đang ngày càng phổ biến. Hiện nay, các thế hệ điện thoại thông minh mới có tốc độ xử lý và khả năng lưu trữ ngày càng lớn. Nhận dạng khuôn mặt tren thiết bị di động vì thế sẽ có tiềm năng để phát triển để trợ giúp con người trong nhiều tác vụ khác nhau.

Vì vậy, tôi đã chọn đề tài tốt nghiệp của mình là “*Nghiên cứu, xây dựng ứng dụng nhận diện khuôn mặt trên Android*” nhằm tạo ra một ứng dụng có thể nhận dạng một người nào đó thông qua khuôn mặt của họ. Việc tạo ra một sản phẩm thực tế sẽ giúp tiết kiệm thời gian và công sức cho các công việc như điểm danh sinh viên, xác thực bảo vệ quyền riêng tư cá nhân, …

Mục tiêu đặt ra khi thực hiện đồ án là:

1. Hiểu được tổng quan về bài toán tìm kiếm khuôn mặt

2. Hiểu được tổng quan về bài toán nhận diện khuôn mặt

3. Xây dựng phần mềm nhận diện khuôn mặt

Sau thời gian khoảng ba tháng thực hiện đồ án, các mục tiêu về cơ bản đã đạt được. Tuy nhiên thị giác máy tính là lĩnh vực khoa học còn mới và phức tạp, thời gian thực hiện đồ án tương đối ngắn nên chắc chắn không tránh khỏi thiếu sót. Rất mong được sự góp ý của các thầy cô, cũng như các bạn học để đồ án này được hoàn thiện hơn.

**SINH VIÊN THỰC HIỆN ĐỒ ÁN**

**Phạm Quang Huy**

CHƯƠNG 1. CƠ SỞ LÝ THUYẾT VÀ BÀI TOÁN TÌM KIẾM KHUÔN MẶT

## 1.1. Khái quát về xử lý ảnh

### 1*.1.1. Khái niệm xử lý ảnh*

Xử lý ảnh là một ngành khoa học còn tương đối mới mẻ so với nhiều ngành khoa học khác, nhất là trên qui mô công nghiệp.

Xử lý ảnh là quá trình thực hiện các thao tác trên ảnh đầu vào cho ra kết quả như mong muốn. Ảnh kết quả có thể khác so với ảnh ban đầu tốt hơn hoặc xấu hơn so với ảnh đầu vào.

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

1. Một số khái niệm cơ bản:

* Ảnh: là một tập hợp hữu hạn các điểm ảnh kề nhau. Ảnh thường được biểu diễn bằng một ma trận 2 chiều, mỗi phần tử của ma trận tương ứng với một điểm ảnh.
* 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.
* Mức xám: là kết quả sự mã hóa tương ứng một cường độ sáng của mỗi điểm ảnh với 1 giá trị số - kết quả của quá trình lượng hóa.
* Biểu diễn ảnh: Trong biểu diễn ảnh người ta thường dùng các phần tử đặc trưng của ảnh là pixel. Việc xử lý ảnh số yêu cầu ảnh phải được mẫu hóa và lượng tử hóa. Một số mô hình được dùng trong biểu diễn ảnh: mô hình toán, mô hình thống kê.

1. Tăng cường ảnh – khôi phục ảnh:

* Tăng cường ảnh là bước quan trọng tạo tiền đề cho xử lý ảnh. Nó gồm các kỹ thuật: lọc độ tương phản, khử nhiễu, nổi màu….
* Khôi phục ảnh là nhằm loại bỏ các suy giảm trong ảnh.

1. Biến đổi ảnh: Thuật ngữ biến đổi ảnh thường được dùng để nói tới một lớp các ma trận đơn vị và các kỹ thuật dùng để biến đổi ảnh. Có nhiều loại biến dạng được dùng như: biến đổi Fourier, sin,cosin ….
2. Nhận dạng ảnh: Nhận dạng ảnh là quá trình liên quan đến các mô tả đối tượng mà người ta muốn đặc tả nó. Người ta đã áp dụng kỹ thuật nhận dạng khá thành công với nhiều đối tượng khác nhau như: nhận dạng vân tay, nhận dạng chữ viết… Có bốn cách tiếp cận khác nhau:

+/ Đối sánh mẫu dựa trên các đặc trưng được trích chọn.

+/ Phân loại thống kê.

+/ Đối sánh cấu trúc.

+/ Phân loại dựa trên mạng nơron nhân tạo.

1. Nén ảnh: Dữ liệu ảnh cũng như các dữ liệu khác cần phải lưu trữ hay truyền đi trên mạng mà lượng thông tin để biểu diễn cho một ảnh là rất lớn. Do đó cần phải giảm lượng thông tin hay nén dữ liệu là một nhu cầu cần thiết. Nén ảnh thường được tiến hành theo cả hai khuynh hướng là nén có bảo toàn và không bảo toàn thông tin.

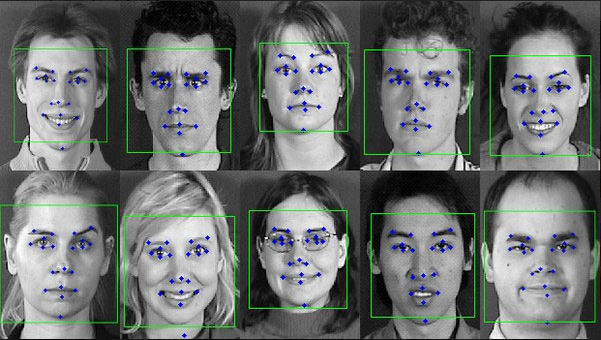
## 1.2. Bài toán phát hiện khuôn mặt

### 1.2.1. Một số đặc trưng về khuôn mặt

Chúng ta nhận diện gương mặt mọi người qua các đặc điểm, hệ thống nhận diện cũng vậy. Mỗi khuôn mặt đều có nhiều điểm mốc, những phần lồi lõm tạo nên các đặc điểm của khuôn mặt. Các hệ thống nhận diện gương mặt định nghĩa những điểm này là những điểm nút. Mỗi mặt người có khoảng 80 điểm nút. Có thể nhận diện một số điểm nút như sau:

* Khoảng cách giữa hai mắt
* Chiều rộng của mũi
* Độ sâu của hốc mắt
* Hình dạng của xương gò má
* Độ dài của xương hàm
* ...

Chuyên gia Michael Sheehan - chuyên gia thuộc Đại học California, Berkeley đã nghiên cứu cùng với các đồng nghiệp của mình đã tiến hành thu thập và phân tích dữ liệu hàng ngàn cá nhân khác nhau. Các chỉ số được tiến hành đo rất cụ thể và chi tiết: khoảng cách hai mắt, chiều cao mũi… cho tới chiều dài bắp chân.



Hình 1 Sự khác nhau của khuôn mặt người

Sau khi phân tích, họ nhận ra rằng khuôn mặt chính là bộ phận đặc biệt nhất. Nếu như đặc điểm ở chân, tay và các bộ phận khác hoàn toàn có thể dự đoán trước thì chúng ta không thể làm điều này với khuôn mặt. Đây là bộ phận duy nhất khác nhau hoàn toàn ở mỗi người, không ai giống ai và có thể coi là chứng minh thư tự nhiên của mỗi cá nhân[1].

### 1.2.2. Tìm hiểu về bài toán phát hiện khuôn mặt

Trong nhiều năm gần đây, có rất nhiều công trình nghiên cứu về bài toán xác định khuôn mặt người từ ảnh đen trắng, xám đến ảnh màu. Ban đầu chỉ là những bài toán đơn giản, mỗi ảnh chỉ có một khuôn mặt nhìn thẳng và đầu luôn phải ở tư thế thằng đứng trong ảnh đen trắng, không đáp ứng được nhu cầu ngày càng cao trong cuộc sống, khoa học ngày nay. Vì thế đã có những cải tiến nghiên cứu về bài toán phát hiện khuôn mặt người trong những môi trường phức tạp hơn, có nhiều khuôn mặt người trong ảnh hơn.

Về cơ bản, một hệ thống nhận dạng mặt người làm việc giống thị giác của con người khi cần nhận dạng (nhận ra) một ai đó khi nhìn vào một bức ảnh. Ở bước đầu tiên hệ thống cần xác định xem có mặt người trong bức ảnh input hay không, sau đó nếu có thì chỉ vùng ảnh đó sẽ được quan tâm, xử lý, bước này gọi là phát hiện khuôn mặt (face detection). Tiếp đến, các đặc điểm là đặc trưng nhất của khuôn mặt (giúp phân biệt ảnh mặt người này với người khác và các các ảnh thuộc về cùng một người) sẽ được trích chọn (extract) để nhận dạng, bước này gọi là trích chọn đặc trưng (feature extraction). Trong bước tiếp theo, gọi là nhận dạng (recognition), hệ thống sẽ đối sánh các đặc trưng vừa nhận được với một kho dữ liệu các đặc trưng của những người đã biết rõ danh tính để xác định xem các đặc trưng đó là của khuôn mặt nào và đưa ra kết luận về danh tính tương ứng với khuôn mặt có tỉ lệ tương đồng lớn nhất. Ở phần này chúng ta sẽ tìm hiểu về quá trình phát hiện khuôn mặt (face detection).

Phát hiện mặt người là một kỹ thuật để phát hiện vị trị và kích thước khuôn mặt người trong các ảnh bất kỳ. Kỹ thuật này chỉ nhận biết về các đặc trưng của khuôn mặt và bỏ qua những đặc trưng khác.

### 1.2.3. Các phương pháp chính phát hiện mặt người

Dựa vào tính chất của các phương pháp xác định mặt người trên ảnh, các phương pháp này được chia thành bốn loại chính, tương ứng với bốn hướng tiếp cận khác nhau. Ngoài ra cũng có rất nhiều nghiên cứu mà phương pháp xác định mặt người không chỉ dựa vào một hướng mà có liên quan đến nhiều hướng.

- Hướng tiếp cận dựa trên tri thức: Dựa vào các thuật toán, mã hoá các đặc trưng và quan hệ giữa các đặc trưng của khuôn mặt thành các luật. Đây là hướng tiếp cận theo kiểu top-down.

- Hướng tiếp cận dựa trên đặc trưng không thay đổi: Xây dựng các thuật toán để tìm các đặc trưng của khuôn mặt mà các đặc trưng này không thay đổi khi tư thế khuôn mặt hay vị trí đặt camera thay đổi.

- Hướng tiếp cận dựa trên so sánh khớp mẫu: Dùng các mẫu chuẩn của khuôn mặt (các mẫu này đã được chọn và lưu trữ) để mô tả các khuôn mặt hay các đặc trưng của khuôn mặt. Phương pháp này có thể dùng để xác định vị trí hay dò tìm khuôn mặt trên ảnh.

- Hướng tiếp cận dựa trên diện mạo: Trái với hướng tiếp cận dựa trên khuôn mẫu, các mô hình hay các mẫu sẽ được học từ một tập ảnh huấn luyện mà thể hiện tính chất tiêu biểu của sự xuất hiện của mặt người trong ảnh. Sau đó hệ thống sẽ xác định mặt người. Phương pháp này còn được biết đến với tên gọi tiếp cận theo hướng học máy.

### 1.2.4. Những khó khăn trong việc phát hiện khuôn mặt người

- Các khuôn mặt trong ảnh có thể có có những hướng nhìn khác nhau như: nhìn thẳng, nhìn nghiêng hay nhìn lên nhìn xuống. Cùng trong một ảnh nhưng sẽ có nhiều hướng nhìn khác nhau của khuôn mặt vì vậy việc phát hiện cũng sẽ gặp khó khăn.

- Trong một khuôn mặt không chỉ có những đặc trưng là khuôn mặt mà còn có một số chi tiết không phải là đặc trưng của khuôn mặt nên việc phát hiện khuôn mặt cũng có thể bị sai.

- Một số khuôn mặt bị che khuất bởi các đối tượng khác cũng gây cản trở cho việc phát hiện khuôn mặt.

- Sự biểu cảm của khuôn mặt có thể làm thay đổi đáng kể các đặc trưng và thông số của khuôn mặt. Như cùng một người nhưng khi cười, tức giận hay sợ hãi cũng dẫn đến sự khác biệt của khuôn mặt.

- Tập các ảnh khuôn mặt trong huấn luyện không thể bao quát được tất cả các biến đổi có thể có trên khuôn mặt của một người trong thế giới thực.

### 1.2.5. Tìm hiểu về phương pháp phát hiện khuôn mặt của Viola và Johns

#### 1.2.5.1. Tổng quan về phương pháp

Phát hiện mặt người là bài toán cơ bản được xây dựng từ nhiều năm nay, có nhiều phương pháp được đưa ra như sử dụng template matching, neuron network…Cho tới nay bài toán này hầu như được giải quyết dựa trên phương pháp sử dụng trình phân loại Haar Cascades của Viola và Johns được cài đặt trong thư viện OpenCv. Phương pháp này được cho là đơn giản và kết quả phát hiện là tương đối cao, lên tới 95%. Các hãng sản xuất máy ảnh như Canon, Samsung… cũng đã tích hợp nó vào trong các sản phẩm của mình.

Trước năm 2000 các kỹ thuật để phát hiện khuôn mặt vẫn chưa đủ tin cậy, chậm và yêu cầu đầu vào thủ công. Đến năm 2001, Viola và Jones đã phát minh ra trình phân loại Haar Cascades, tạo nên sự cách mạng hoá trong phương phát phát hiện khuôn mặt, nó có thể phát hiện các đối tượng ngay trong thời gian thực với độ chính xác 95% sau khi được Lienhart và Maydt cải thiện hơn năm 2002 [2].

Đây là một cách tiếp cận dựa trên máy học, chức năng xếp tầng được đào tạo từ rất nhiều hình ảnh tích cực và tiêu cực. Sau đó, nó được sử dụng để phát hiện các đối tượng trong các hình ảnh khác [3]. Khi training, chương trình sẽ rút ra đặc trưng từ các ảnh có mặt người (tích cực) rồi so sánh với các ảnh không có mặt người (tiêu cực), nếu đặc trưng đó được tìm thấy trong ảnh tiêu cực thì đặc trưng đó không phù hợp.

Về tổng quan bài toán phát hiện khuôn mặt cũng chính là bài toán phân loại 2 lớp mặt và không phải mặt. Trình phân loại Haar Cascades là bộ phân loại tầng với mỗi tầng là một bộ phân loại Adaboost. Ở các tầng sẽ sử dụng các đặc trưng Haar-like đặt vào các vùng ảnh để tính toán các giá trị đặc trưng, từ những giá trị đặc trưng này đưa vào bộ phân loại Adaboost sẽ xác định ảnh có khuôn mặt hay không.

#### 1.2.5.2. Tiền xử lý ảnh

Phương pháp Haar Cascades được thực hiện trên ảnh xám (gray image). Mỗi điểm ảnh (pixel) sẽ có giá trị mức xám từ 0 đến 255 (không gian màu 8 bit). Như vậy phương pháp Haar Cascades sẽ không khai thác những đặc điểm về màu sắc khuôn mặt để nhận dạng song vẫn rất hiệu quả. Ảnh màu sẽ được chuyển về ảnh xám để nhận dạng.

Sau khi được chuyển thành ảnh xám, ảnh lại được tiếp tục chuyển thành “ảnh tích hợp” và các đặc trưng Haar-like sẽ làm việc trực tiếp trên ảnh tích hợp.

#### 1.2.5.3. Đặc trưng Haar-like

a) Đối tượng nhận dạng

Trên ảnh, vùng khuôn mặt là tập hợp các điểm ảnh có những mối quan hệ khác biệt so với các vùng ảnh khác, nhưng mối quan hệ này tạo lên các đặc trưng riêng của khuôn mặt. Tất cả khuôn mặt người đều có chung những đặc điểm sau khi đã chuyển qua ảnh xám, ví dụ như:

- Vùng hai mắt sẽ tối hơn vùng má và vùng trán, tức màu xám của vùng này cao hơn so với hai vùng còn lại.

- Vùng sống mũi sẽ sáng hơn vùng mắt.

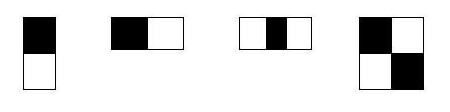
- …

Và còn rất nhiều những đặc điểm khác của khuôn mặt và các đặc trưng Haar-like dựa vào các đặc điểm này để nhận dạng.

Về tổng quát, các đặc trưng Haar-like không chỉ được sử dụng để nhận dạng khuôn mặt mà có thể dùng để nhận dạng bất kì một đối tượng nào trên ảnh (thân người, tay, chân, ô tô, đồ vật, ….). Bởi vì cũng giống như khuôn mặt, mỗi đối tượng có những đặc điểm riêng biệt đặc trưng bởi các vùng điểm ảnh, công việc của đặc trưng Haar-like là tính toán các giá trị tương quan giữa các vùng ảnh đó.

b) Các đặc trưng Haar-like

Đặc trưng Haar-like do Viola và Jones công bố gồm 4 đặc trưng cơ bản để xác định khuôn mặt người [4]. Mỗi đặc trưng Haar-like là một miền hình chữ nhật được chia thành 2,3 hoặc 4 hình chữ nhật nhỏ phân biệt quy ước bằng màu trắng và đen:



Hình 2 Bốn đặc trưng Haar-like

Để sử dụng các đặc trưng này vào việc xác định khuôn mặt người, 4 đặc trưng Haar-Like cơ bản được mở rộng ra và được chia làm 3 tập đặc trưng như sau:

* Đặc trưng cạnh (edge feature) https://images.viblo.asia/47709099-37f8-48f2-be1b-4a5dc6a41c5b.png
* Đặc trưng đường (line feature) 
* Đặc trưng xung quanh tâm (center-surround features) https://images.viblo.asia/988200ee-7f89-4d58-bdb3-52479e62dc35.png

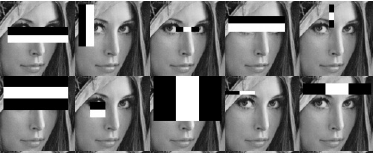
Giá giá trị của một đặc trưng Haar-Like là sự chênh lệch giữa tổng các giá trị xám của các pixel trong vùng đen với tổng các giá trị xám của các pixel trong vùng trắng như trong công thức sau:

f(x) = Tổng vùng đen (các mức xám) – Tổng vùng trắng (các mức xám)

Vậy khi được đặt lên một vùng ảnh, đặc trưng Haar-like sẽ tính toán và đưa ra giá trị đặc trưng f(x) của vùng đó.

#### 1.2.5.4. Cách áp dụng đặc trưng Harr-like để phát hiện khuôn mặt trong ảnh

Để phát hiện khuôn mặt, hệ thống sẽ cho một cửa sổ con (sub-window) quét lên toàn bộ ảnh đầu vào. Sau mỗi lần quét hết toàn bộ ảnh đầu vào thì cửa sổ con sẽ tăng kích thước lên. Như vậy sẽ có rất nhiều ảnh con ứng với từng cửa sổ con, các đặc trưng Haar-like sẽ được đặt lên các cửa sổ con này từ đó tính ra các giá trị đặc trưng. Sau đó các giá trị này được bộ phân loại xác nhận xem khung hình đó có phải khuôn mặt hay không.



Hình 3 Các đặc trưng Haar-like được đặt lên tại khung hình

Như hình vẽ trên đây, tại mỗi khung hình các đặc trưng Haar-like được đặt lên ở các vị trí khác nhau vào sau đó tính các giá trị đặc trưng. Ứng với mỗi đặc trưng trên, một bộ phân loại yếu (weak classifier) được định nghĩa như sau:

=

Trong đó:

- x: Cửa sổ con đang xét.

- : Giá trị đặc trưng Haar-like thứ k.

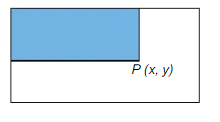
- : Ngưỡng (là giá trị đã được rút ra sau quá trình huấn luyện).

Ta hiểu công thức trên đơn giản như sau: Đặc trưng Haar-like có k vị trí đặt trong cửa sổ con x, khi giá trị của đặc trưng Haar-like tại vị trí thứ k vượt qua một ngưỡng thì bộ phân lớp sẽ kết luật cửa sổ con x là khuôn mặt, còn không qua ngưỡng đó thì không là khuôn mặt.

Sẽ có những vị trí trên ảnh không phải là khuôn mặt nhưng lại cho ra giá trị đặc trưng vượt ngưỡng và bộ phân lớp yếu sẽ kết luận đấy là khuôn mặt, nhưng ta không chỉ dùng một đặc trưng Haar-like mà dùng rất nhiều đặc trưng ở những vị trí và kích thước khác nhau nên sẽ có được những kết luận chính xác.

#### 1.2.5.5. Ảnh tích hợp (Integral image)

Để tính các giá trị của đặc trưng Haar-like, ta phải tính tổng của các vùng pixel trên ảnh. Mà số lượng các đặc trưng Haar-like rất nhiều nên để tính toán đòi hỏi khối lượng tính toán rất lớn. Vì vậy ảnh tích hợp được đưa ra để nhằm tính toán nhanh chóng các đặc trưng, giảm thời gian xử lý. Giá trị của ảnh tích hợp tại ví trí (x, y) là tổng các điểm ảnh thuộc hình chữ nhật xác định bởi góc trên bên trái (0,0) và góc phải dưới (x,y).



Hình 4 Ảnh tích hợp

Ảnh tích hợp được định nghĩa theo công thức:

P(x,y) =

Trên thực tế để chuyển một ảnh thành ảnh tích hợp ta dùng công thức truy hồi sau:

s(x, y) = x(x, y-1) + i(x, y) với s(x, -1) = 0

P(x, y) = P(x-1, y) + s(x, y) P(-1, y) = 0

|  |  |  |
| --- | --- | --- |
| 1 | 2 | 3 |
| 2 | 4 | 6 |
| 3 | 6 | 9 |

Ví dụ chuyển một ảnh 3x3 có giá trị xám như bên dưới thành ảnh tích hợp

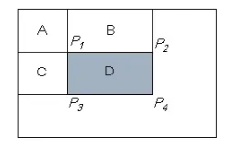
|  |  |  |
| --- | --- | --- |
| 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

==>

Ảnh ban đầu Ảnh tích hợp

Sau khi chuyển ảnh cần nhận dạng thành ảnh tích hợp, việc tính toán giá trị các đặc trưng Haar-like sẽ rất đơn giản.

Để tính giá trị đặc trưng Haar-like, ta phải tính được tổng giá trị điểm ảnh trong một vùng chữ nhật trên ảnh. Ví dụ như vùng D trong hình vẽ :



Hình 5 Vùng giá trị điểm ảnh cần tính

Với A, B, C, D: là tổng các giá trị điểm ảnh trong từng vùng.

P1, P2, P3, P4: là giá trị của ảnh tích hớp tại 4 đỉnh của D.

Nếu như là ảnh xám bình thường thì để tính D ta phải tính tổng tất cả các giá trị điểm ảnh trong D, miền D càng lớn thì số phép cộng càng nhiều. Nhưng với ảnh tích hợp dù miền D có kích thước như thế nào thì D cũng chỉ cần tính thông qua 4 giá trị tại 4 đỉnh.

Ta có: P1 = A, P2 = A+B, P3 = A+C, P4 = A+B+C+D

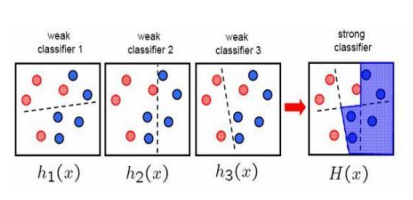
Suy ra :

D = (A+B+C+D) – (A+B) – (A+C) + A

= P4 – P2 – P3 + P1

Để chọn các đăc trưng Haar-like dùng cho việc thiết lập ngưỡng, Viola và Jones sử dụng một phương pháp máy học được gọi là Adaboost [5]. Adaboost sẽ kết hợp các bộ phân loại yếu để tạo thành một bộ phân loại mạnh.

Adaboost sử dụng thêm khái niệm trọng số (weight) để đánh dấu các mẫu khó nhận dạng. Trong quá trình huấn luyện, cứ mỗi weak classifiers được xây dựng, thuật toán sẽ tiến hành cập nhật lại trọng số để chuẩn bị cho việc xây dựng weak classifier kế tiếp: tăng trọng số của các mẫu bị nhận dạng sai và giảm trọng số của các mẫu được nhận dạng đúng bởi các weak classifier vừa xây dựng. Bằng cách này weak classifier sau có thể tập trung vào các mẫu mà các weak classifier trước nó làm chưa tốt. Sau cùng, các weak classifier sẽ được kết hợp tùy theo mức độ tốt của chúng để tạo nên strong classifier.



Hình 6 Kết hợp các bộ phân loại yếu thành bộ phân loại mạnh

#### 1.2.5.6. Cascade of classifier

Giả sử chúng ta huấn luyện được một bộ phân lớp mạnh để phát hiện khuôn mặt thông qua 10 bộ phân lớp yếu. Nếu ta dùng bộ phân lớp mạnh này đưa vào các cửa sổ con (sub window) chạy khắp một tấm ảnh đưa vào để phát hiện khuôn mặt thì sẽ rất tốn chi phí tính toán và thời gian. Tại mỗi cửa sổ con trên tấm ảnh chúng ta lại phải dùng cả 10 bộ phân lớp yếu. Trong khi đó, ta thấy trong một tấm ảnh thì khuôn mặt chiếm tỉ lệ ít và tại những cửa sổ không phải là khuôn mặt ta có thể loại bỏ mà chỉ cần dùng một bộ phân lớp mạnh khác gồm ít bộ phân lớp yếu hơn.

Viola và Jones giải quyết vấn đề này bằng chuỗi các bộ phân lớp (Cascade of classifier). Ta sẽ có một chuỗi các bộ phân lớp, trong đó mỗi bộ phân lớp được xây dựng bằng thuật toán Adaboost.

Sub

Window

Không là khuôn mặt

Là khuôn mặt

Không là khuôn mặt

Là khuôn mặt

Không là khuôn mặt

Là khuôn mặt

Hình 7Mô hinh Cascade of classifier

Mỗi cửa sổ con sẽ được cho đi qua các bộ phân lớp này:

- Bộ phân lớp đầu tiên sẽ loại bỏ phần lớn các cửa sổ không phải khuôn mặt và cho đi qua các cửa sổ được cho là khuôn mặt. Ở đây, bộ phân lớp này rất đơn giản nên độ phức tạp cũng rất thấp. Tất nhiên, vì nó đơn giản nên trong số các cửa sổ được nhận dạng là khuôn mặt sẽ có một số lượng lớn cửa sổ bị nhận dạng sai

- Những cửa sổ được cho đi qua bởi bộ phân lớp đầu sẽ được xem xét bởi bộ phân lớp sau đó. Nếu bộ phân lớp cho rằng đó không phải là khuôn mặt thì loại bỏ cửa sổ, nếu bộ phân lớp cho rằng đó là khuôn mặt thì ta lại cho đi qua và chuyển đến bộ phân lớp phía sau.

- Những bộ phân lớp càng về sau thì càng phức tạp hơn, đòi hỏi sự tính toán nhiều hơn. Cửa sổ nào không thoả mãn bộ phân lớp sẽ bị loại bỏ. Chỉ những cửa sổ đi qua được tất cả các bộ phân lớp thì ta mới quyết định đó là khuôn mặt.

Tóm lại, chuỗi các bộ phân lớp sẽ xử lý các mẫu (cửa sổ con) đi vào theo nguyên tắc sau: nếu một bộ phân lớp nào đó cho rằng đó không phải là mặt người thì ta loại bỏ ngay, còn nếu bộ phân lớp cho rằng đó là khuôn mặt thì ta chuyển tiếp đến bộ phân lớp sau. Nếu một mẫu trót lọt hết tất cả các bộ phân lớp thì ta mới quyết định đó là khuôn mặt.

TÀI LIỆU THAM KHẢO

[1]. *Nguyên nhân khiến khuôn mặt con người không ai giống ai*, https://khoahoc.tv/nguyen-nhan-khien-khuon-mat-con-nguoi-khong-ai-giong-ai-56493

[2]. Alrashed, H. H. *Face Recognition On Android,* Submitted in partial fulfilment of the requirements for the first year of Master of Information Sciences Massey University 2014

[3]. *Face Detection using Haar Cascades*, https://opencv-python-tutroals.readthedocs.io/en/latest/py\_tutorials/py\_objdetect/py\_face\_detection/py\_face\_detection.html

[4]. Viola and Jones. *Rapid object detection using a boosted cascade of simple feature*, Computer Vision and Pattern Recognition, 2001.

[5]. *Viola–Jones object detection framework*, https://en.wikipedia.org/wiki/Viola%E2%80%93Jones\_object\_detection\_framework