

Mục lục

Phần mở đầu	2
Chương 1: Khái quát về xử lý ảnh và bài toán phát hiện mặt người	4
1.1. Khái quát về xử lý ảnh	4
1.1.1. Xử lý ảnh là gì:	4
1.1.2. Một số vấn đề cơ bản trong xử lý ảnh:	4
1.2. Bài toán phát hiện mặt người trong ảnh.	5
1.2.1. Bài toán:	5
1.2.2. Những khó khăn phát hiện khuôn mặt người trong ảnh.	7
1.2.3. Một số lĩnh vực ứng dụng phát hiện khuôn mặt người trong ảnh	7
Chương 2: Đặc trưng haar trong phát hiện mặt người	10
2.1. Đặc trưng haar	10
2.2. Thuật toán Adaboost	12
Chương 3: Xây dựng ứng dụng	19
3.1. Bài toán:	19
3.2. Phân tích và lựa chọn công cụ	19
3.2.1. Phân tích bài toán.	19
3.2.2. Sơ lược về OpenCV	20
3.2.2.1. OpenCV là gì:	20
3.2.2.2. Cấu trúc OpenCV	21
3.3. Một số kết quả của chương trình	22
Phần kết luận	25
Tài liệu tham khảo	26

Phân mở đầu

Công nghệ thông tin đang được ứng dụng trong mọi lĩnh vực của cuộc sống. Với một hệ thống máy tính, chúng ta có thể làm được rất nhiều việc, tiết kiệm thời gian và công sức. Đặc biệt là công việc nhận dạng người. Ngày xưa, muốn tìm kiếm một người tình nghi trong siêu thị, sân bay hay bất kỳ đâu ở những nơi công cộng, các nhân viên an ninh phải tìm kiếm trên từng màn hình theo dõi. Ngày nay, công việc đó đã được làm tự động nhờ các hệ thống nhận dạng mặt người. Phát hiện mặt người trong ảnh là một phần quan trọng của hệ thống nhận dạng mặt người, giải quyết tốt việc phát hiện mặt người sẽ giúp tiết kiệm thời gian và nâng cao độ chính xác của việc nhận dạng khuôn mặt.

Bài toán phát hiện mặt người được nghiên cứu từ những năm 1990. Trong những 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 thẳng đứng trong ảnh đen trắng, không đáp ứng được nhu cầu ngày càng cao trong cuộc sống, khoa học ngày nay.

Vì thế 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, và có nhiều tư thế thay đổi trong ảnh.

Trong đồ án này mục tiêu chính của em là tìm hiểu về đặc trưng haar và áp dụng vào bài toán phát hiện mặt người trong ảnh.

Nội dung đề án bao gồm:

Chương 1: Khái quát về xử lý ảnh và bài toán phát hiện mặt người: Giới thiệu về xử lý ảnh, tổng quan về bài toán phát hiện mặt người trong ảnh, những khó khăn và một số lĩnh vực ứng dụng phát hiện mặt người.

Chương 2: Đặc trưng Haar trong phát hiện mặt người: Tìm hiểu về các đặc trưng Haar và thuật toán Adaboost để phát hiện mặt người.

Chương 3: Xây dựng ứng dụng: Xây dựng một chương trình đơn giản phát hiện mặt người trong ảnh dựa vào đặc trưng Haar.

Phần kết luận.

Tài liệu tham khảo.

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. Xử lý ảnh là gì:

Xử lý ảnh là một 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:

a) Một số khái niệm 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.

b) 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.

c) 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.

d) 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

e) 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.

f) 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 mặt người trong ảnh.

1.2.1. Bài toán:

a) Giới thiệu về bài toán phát hiện mặt người trong ảnh:

Trong những 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 thẳng 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, và có nhiều tư thế thay đổi trong ảnh.

Xác định khuôn mặt người là một kỹ thuật để xác định vị trí và kích thước khuôn mặt người trong các ảnh bất kỳ. Kỹ thuật này nhận biết về các đặc trưng của khuôn mặt và bỏ qua những thứ khác.

b) Bài toán phát hiện mặt người là gì:

Phát hiện mặt người là một kỹ thuật để phát hiện vị trí 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.

c) 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 hóa 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 trên khuôn mặt 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 theo tiếp cận theo các phương pháp học máy.

1.2.2. Những khó khăn phát hiện khuôn mặt người trong ảnh.

- Khuôn mặt trong ảnh có thể 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 nhận dạng sẽ rất 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 nhận dạng khuôn mặt cũng có thể sai.

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

- Sự biểu cảm của khuôn mặt: 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.

- Giới hạn về số ảnh cần thiết cho tập huấn luyện, tập các ảnh khuôn mặt huấn luyện không thể bao quát được tất cả các biến đổi có thể có trên khuôn mặt của một người cần nhận dạng trong thế giới thực.

1.2.3. Một số lĩnh vực ứng dụng phát hiện khuôn mặt người trong ảnh.

- ✚ Nhận dạng đối tượng giúp cơ quan chức năng quản lý:

Dựa vào ảnh một người, nhận dạng xem người đây có phải là tội phạm hay không bằng cách so sánh với các ảnh tội phạm đang được lưu trữ.

✚ Hệ thống camera quan sát, theo dõi và bảo vệ:

Sử dụng camera để quan sát, theo dõi và phát hiện tội phạm trong đám đông. Ví dụ như người đó có lấy trộm đồ, xâm nhập bất hợp pháp hay không.

✚ Giúp con người bảo mật:

Các ứng dụng về bảo mật rất đa dạng, một trong số đó là công nghệ nhận dạng mặt người. Ví dụ như công nghệ nhận dạng mặt người của laptop, để vào được laptop người dùng phải ngồi trước máy tính và phải sử dụng webcam để chụp ảnh khuôn mặt mình và cho máy “học” thuộc các đặc điểm của khuôn mặt giúp cho quá trình đăng nhập sau này.

✚ Lưu trữ khuôn mặt:

Xác định mặt người có thể được ứng dụng trong các trạm rút tiền tự động để lưu trữ khuôn mặt của người rút tiền để tránh trường hợp chủ thẻ bị mất thẻ khi đó sẽ giúp nhân viên ngân hàng xử lý dễ dàng hơn.

✚ Một số ứng dụng khác của bài toán nhận dạng mặt người:

- Trong lĩnh vực an ninh sân bay, xuất nhập cảnh: để nhận dạng được tội phạm quốc gia, tội phạm quốc tế trà trộn chung với hành khách. Hay tội phạm trong nước tìm cách trốn ra nước ngoài bằng đường hàng không.

- Điều khiển vào ra ở công ty, văn phòng... Trong mỗi công ty hay một tổ chức mỗi nhân viên sẽ có một quyền hạn nhất định trong phạm vi làm việc của mình. Để ngăn cản sự xâm nhập người ta sử dụng hệ thống nhận dạng như vân tay hay mắt... để hạn chế quyền hạn của nhân viên.

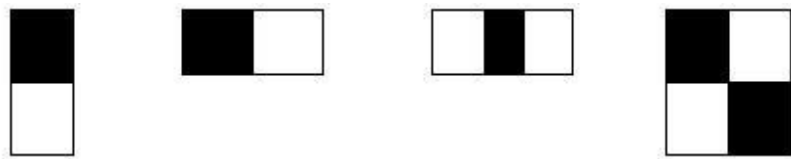
- Tìm kiếm và tổ chức dữ liệu liên quan đến con người thông qua khuôn mặt trên nhiều hệ cơ sở dữ liệu lưu trữ lớn như internet, các hãng truyền hình....
- Trong giao thông kiểm tra và phát hiện trạng thái người lái xe có ngủ gật, mất tập trung hay không và thông báo về cho cơ quan chức năng biết khi cần thiết.

Chương 2: Đặc trưng haar trong phát hiện mặt người.

2.1. Đặc trưng haar

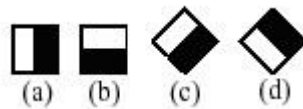
Khuôn mặt được đặc trưng bởi tập hợp các pixel trong vùng khuôn mặt mà các pixel này tạo lên những điểm khác biệt so với các vùng pixel khác. Tuy nhiên với một ảnh đầu vào, việc sử dụng các pixel riêng lẻ lại không hiệu quả. Vì vậy những nhà nghiên cứu đã đưa ra tư tưởng kết hợp các vùng pixel với nhau tạo đặc trưng có khả năng phân loại tốt các vùng của khuôn mặt. Trong số đó đặc trưng haarlike đã được ứng dụng.

Đặc trưng haarlike được tạo thành bằng việc kết hợp các hình chữ nhật đen, trắng với nhau theo một trật tự, một kích thước nào đó.

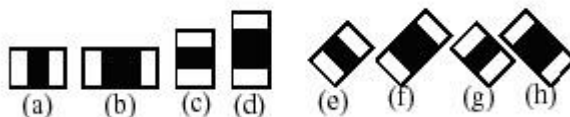


Có 4 đặc trưng haarlike cơ bản được mở rộng, và được chia làm 3 tập đặc trưng sau:

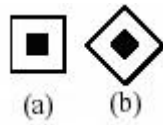
➤ Đặc trưng cạnh



➤ Đặc trưng đường

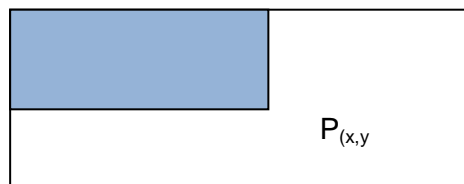


➤ Đặc trưng xung quanh tâm



Để tính giá trị các đặc trưng haarlike, ta tính sự chênh lệch giữa tổng của các pixel của các vùng đen và các vùng trắng theo công thức sau: $f(x) = Tổng_{vùng\ đen}(pixel) - Tổng_{vùng\ trắng}(pixel)$

Để tính các giá trị của đặc trưng haarlike, ta phải tính tổng của các vùng pixel trên ảnh. Nhưng để tính toán các giá trị của đặc trưng haarlike cho tất cả các vị trí trên ảnh đòi hỏi chi phí tính toán khá lớn. Do đó những nhà nghiên cứu đã đưa ra một khái niệm gọi là Integral Image để tính toán nhanh cho các feature cơ bản. Integral Image là một mảng 2 chiều với kích thước bằng với kích thước của ảnh cần tính các đặc trưng haarlike, với mỗi phần tử của mảng này được tính bằng cách tính tổng của điểm ảnh phía trên và bên trái của nó. Bắt đầu từ vị trí trên, bên trái đến vị trí dưới, phải của ảnh, việc tính toán này chỉ dựa trên phép cộng số nguyên đơn giản.



Cách tính Integral Image của ảnh:

$$P(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y')$$

Sau khi đã tính được Integral Image của ảnh, việc tính tổng điểm ảnh của một vùng bất kì nào đó trên ảnh được thực hiện như sau:

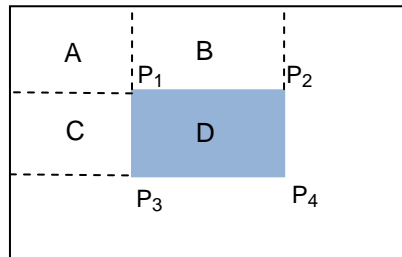
Giả sử ta cần tính tổng điểm ảnh của vùng D như trong hình 2.1:

$$D = A + B + C + D - (A + B) - (A + C) + A$$

Với $A + B + C + D$ chính là giá trị tại điểm P4 trên Integral Image, $A + B$ là giá trị tại điểm P2, $A + C$ là giá trị tại điểm P3, và A là giá trị tại điểm P1.

Vậy ta có thể tính lại biểu thức D ở trên như sau:

$$D = (x_4, y_4) - (x_2, y_2) - (x_3, y_3) + (x_1, y_1)$$



Hình 2.1: Ví dụ cách tính nhanh tổng các điểm ảnh của vùng D trên ảnh

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

2.2. Thuật toán Adaboost

Adaboost là một cách trong hướng tiếp cận dựa trên diện mạo, Viola và Jones dùng AdaBoost kết hợp cascade để xác định khuôn mặt người [17] với các đặc trưng dạng Haar wavelet-like. Tốc độ xử lý khá nhanh và tỷ lệ chính xác hơn 80% trên ảnh xám.

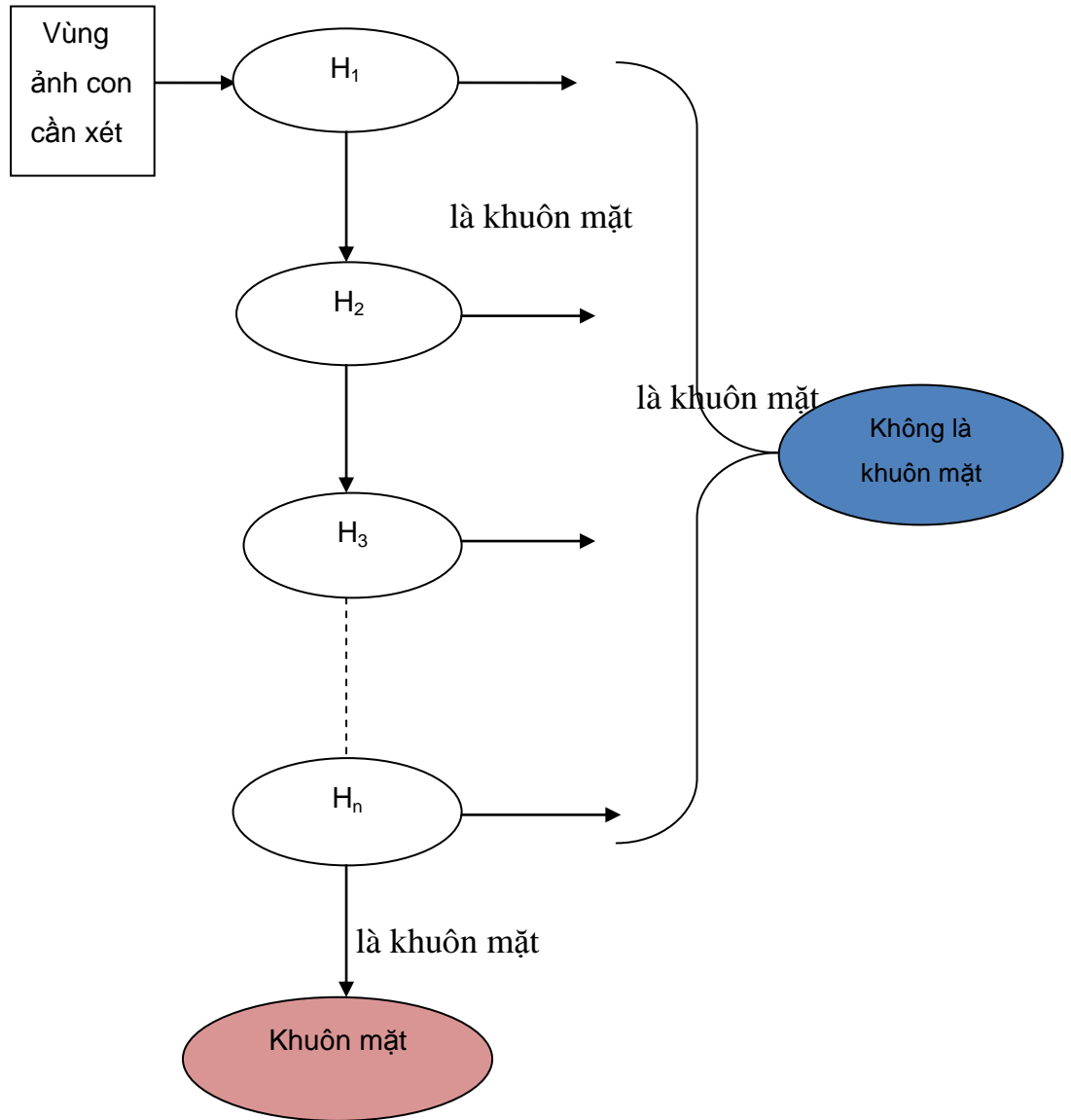
Adaboost là một bộ phân loại mạnh phi tuyến phức dựa trên tiếp cận boosting được Freund và Schapire đưa ra vào năm 1995. Adaboost cũng hoạt động dựa trên nguyên tắc kết hợp tuyến tính

các weak classifiers để hình thành một strong classifier.

Là một cải tiến của của tiếp cận boosting. 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.

Ví dụ: để biết một ảnh có phải là ảnh mặt người không, ta hỏi N người (tương đương với N classifier xây dựng từ N vòng lặp của boosting), đánh giá của mỗi người (tương đương với một weak classifier) chỉ cần tốt hơn ngẫu nhiên một chút (tỉ lệ sai dưới 50%). Sau đó ta sẽ đánh trọng số cho đánh giá của từng người, người nào có khả năng đánh giá tốt các mẫu khó thì mức độ quan trọng của người đó trong kết luận cuối cùng sẽ cao hơn những người chỉ đánh giá tốt được các mẫu dễ. Việc cập nhật lại trọng số của các mẫu sau mỗi vòng lặp boosting chính là để đánh giá độ khó của các mẫu (mẫu có nhiều người đánh giá sai là mẫu càng khó).

Viola và Jones dùng Adaboost kết hợp các bộ phân loại yếu sử dụng các đặc trưng haar-like theo mô hình phân tầng (cascade) như sau:



Hình 2.2 : Mô hình phân tầng kết hợp các bộ phân loại yếu để xác định khuôn mặt.

Các weak classifier $h_k(x)$ là các bộ phân loại yếu được biểu diễn như sau:

$$h_k(x) = \begin{cases} 1 & \text{nếu } p_k f_k(x) < p_k \theta_k \\ 0 & \text{nếu ngược lại} \end{cases}$$

Trong đó:

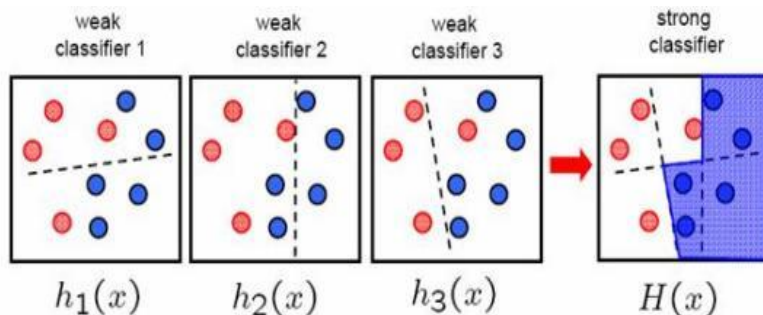
- x : cửa sổ con cần xét
- θ_k : ngưỡng ($\theta = \text{teta}$)
- f_k : giá trị của đặc trưng Haar-like
- p_k : hệ số quyết định chiều của phương trình

Nếu giá trị vector đặc trưng của mẫu cho bởi hàm lượng giác của bộ phân loại vượt qua một ngưỡng cho trước thì mẫu object (đối tượng cần nhận dạng), ngược lại mẫu là background (không phải đối tượng).

Adaboost sẽ kết hợp các bộ phân loại yếu thành bộ phân loại mạnh như sau:

$$H(x) = \text{sign}(a_1 h_1(x) + a_2 h_2(x) + \dots + a_n h_n(x) = \alpha)$$

Với $a_n \geq 0$ là hệ số chuẩn hóa cho các bộ phân loại yếu



Hình 2.3: Kết hợp các bộ phân loại yếu thành bộ phân loại mạnh

Mỗi bộ phân loại yếu sẽ quyết định kết quả cho một đặc trưng haarlike, được xác định ngưỡng đủ nhỏ sao cho có thể vượt được tất cả các bộ dữ liệu mẫu trong tập dữ liệu huấn luyện (số lượng ảnh khuôn mặt trong tập huấn luyện lớn). Trong quá trình xác định khuôn mặt người, mỗi vùng ảnh con sẽ được kiểm tra với các đặc trưng trong chuỗi các đặc trưng haarlike, nếu có một đặc trưng haarlike nào cho ra kết quả là khuôn mặt người thì các đặc trưng khác không cần xét nữa. Thứ tự xét các đặc trưng trong chuỗi các đặc trưng haarlike sẽ được dựa vào trọng số (weight) của đặc trưng đó do Adaboost quyết định dựa vào số lần và thứ tự xuất hiện của các đặc trưng haarlike.

Thuật toán Adaboost:

- ✚ Cho một tập huấn luyện gồm N mẫu có đánh dấu $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Với $x_k \in X = (x_{1k}, x_{2k}, \dots, x_{mk})$ là vector đặc trưng và $y_k \in (-1, 1)$ là nhãn của mẫu (1 ứng với object, -1 ứng với background).
- ✚ Khởi tạo trọng số ban đầu cho tất cả các mẫu: với m là số mẫu đúng (ứng với object và $y= 1$) và l là số mẫu sai (ứng với background và $y= -1$).

$$w_{1,k} = \frac{1}{2m}, \frac{1}{2l}$$

- ✚ Xây dựng T weak classifier

Lặp $t = 1 \dots T$

- Với mỗi đặc trưng trong vector đặc trưng, xây dựng một weak classifier h_j với ngưỡng θ_j và lỗi ϵ_j .

$$\epsilon_j = \sum_k^n w_{t,k} |h_j(x_k) - y_k|$$

- Chọn ra h_j với ϵ_j nhỏ nhất, ta được h_t :

$$h_t : X \rightarrow \{1, -1\}$$

- Cập nhật lại trọng số:

$$w_{t+1,k} = \frac{w_{t,k}}{Z_t} \times \begin{cases} e^{-\alpha_t}, & h_t(x_k) = y_k \\ e^{\alpha_t}, & h_t(x_k) \neq y_k \end{cases}$$

- Trong đó: $\alpha_t = \frac{1}{2} \ln \left(\frac{1-\epsilon_j}{\epsilon_j} \right)$

Z_t : hệ số dùng để đưa W_{t+1} về đoạn $[0, 1]$ (normalization factor)

- ✚ Strong classifier được xây dựng:

$$H(x) = \text{dấu}(\sum_{t=1}^T \alpha_t h_t(x))$$

Quá trình huấn luyện bộ phân loại được thực hiện bằng một vòng lặp mà ở mỗi bước lặp, thuật toán sẽ chọn ra weak classifier h_t thực hiện việc phân loại với lỗi ϵ_t nhỏ nhất (do đó sẽ là bộ phân loại tốt nhất) để bổ sung vào strong classifier. Mỗi khi chọn được một bộ phân loại h_t Adaboost sẽ tính giá trị α_t theo công thức ở trên. α_t cũng được chọn trên nguyên tắc làm giảm thiểu giá trị lỗi ϵ_t .

Hệ số α_t nói lên mức độ quan trọng của h_t .

- Trong công thức của bộ phân loại $H(x)$:

$$H(x) = \text{dấu} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$$

Ta thấy tất cả bộ phân loại h_t đều có đóng góp vào kết quả phân loại của $H(x)$, và mức độ đóng góp của chúng phụ thuộc vào giá trị α_t tương ứng: h_t với α_t càng lớn thì nó càng có vai trò quan trọng trong $H(x)$.

- Trong công thức tính α_t :

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_j}{\varepsilon_j} \right)$$

Dễ thấy giá trị α_t tỉ lệ nghịch với ε_j . Bởi vì ht được chọn với tiêu chí đạt ε_j nhỏ nhất, do đó nó sẽ bảo đảm giá trị α_t lớn nhất. Công thức này do Freund và Schapire đưa ra.

Sau khi tính được giá trị α_t , AdaBoost tiến hành cập nhật lại trọng số của các mẫu: tăng trọng số các mẫu mà ht phân loại sai, giảm trọng số các mẫu mà ht phân loại đúng. Bằng cách này, trọng số các mẫu phản ánh được độ khó nhận dạng của mẫu đó và ht+1 sẽ ưu tiên học cách phân loại các mẫu này.

Vòng lặp xây dựng strong classifier sẽ dừng lại sau T lần lặp. Trong thực tế cài đặt, người ta ít sử dụng giá trị T vì không có công thức nào đảm bảo tính được giá trị T tối ưu cho quá trình huấn luyện. Thay vào đó người ta sử dụng giá trị max false positive hay max false alarm (tỉ lệ nhận dạng sai tối đa các mẫu background). Tỉ lệ này của bộ phân loại cần xây dựng không được phép vượt quá giá trị này. Khi đó, qua các lần lặp, false alarm của strong classifier $H_t(x)$ xây dựng được (tại lần thứ t) sẽ giảm dần, và vòng lặp kết thúc khi tỉ lệ này thấp hơn max false alarm.

Chương 3: Xây dựng ứng dụng

3.1. Bài toán:

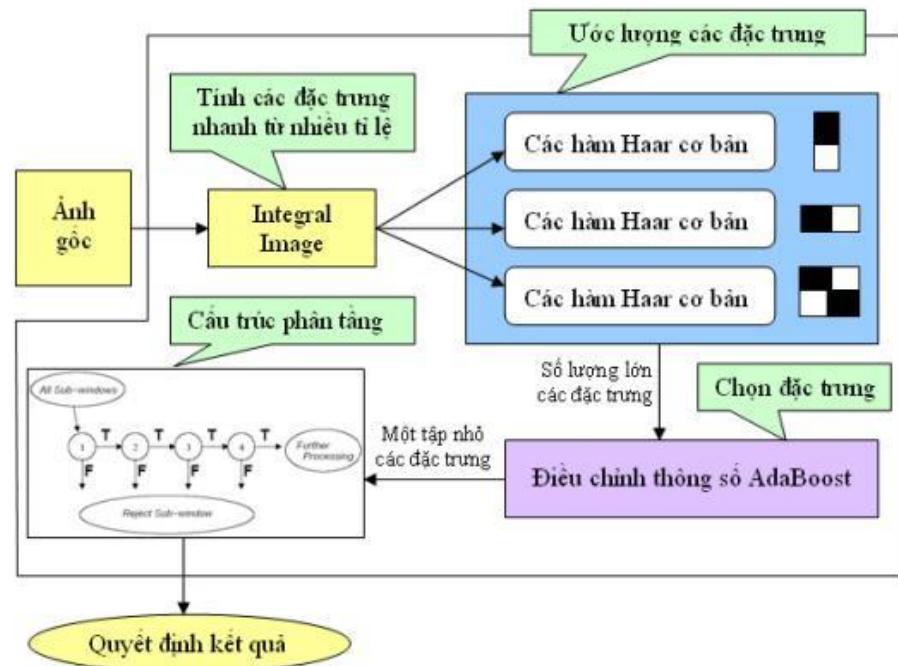
Đưa ảnh gốc ban đầu vào hệ thống nhận dạng khuôn mặt sau đó hệ thống sẽ sử dụng các đặc trưng haar để phát hiện khuôn mặt.

3.2. Phân tích và lựa chọn công cụ.

3.2.1. Phân tích bài toán.

Đầu vào: ảnh gốc.

Đầu ra: ảnh đã được phát hiện khuôn mặt.



Hình 3.1: Hệ thống phát hiện khuôn mặt người

Từ ảnh gốc ban đầu ta sẽ tính Integral image, là mảng hai chiều với phần tử (x,y) sẽ được tính bằng tổng của các phần tử (x',y') với $x' < x$ và $y' < y$, làm như vậy để tính nhanh tổng của các giá trị mức xám của một vùng hình chữ nhật bất kỳ trên ảnh gốc. Các vùng ảnh con này sẽ được đưa qua các hàm Haar cơ bản để ước lượng đặc trưng, kết quả ước lượng sẽ được đưa qua bộ

điều chỉnh Adaboost để loại bỏ nhanh các đặc trưng không có khả năng là các đặc trưng của khuôn mặt người. Chỉ có một tập nhỏ các đặc trưng mà bộ điều chỉnh Adaboost cho là có khả năng là đặc trưng của khuôn mặt người mới được chuyển sang cho bộ quyết định kết quả (là tập các bộ phân loại yếu). Bộ quyết định sẽ tổng hợp kết quả là khuôn mặt người nếu kết quả của các bộ phân loại yếu trả về là khuôn mặt người.

Mỗi bộ phân loại yếu sẽ quyết định kết quả cho một đặc trưng haarlike, được xác định ngưỡng đủ nhỏ sao cho có thể vượt được tất cả các bộ dữ liệu mẫu trong tập dữ liệu huấn luyện (số lượng ảnh khuôn mặt trong tập huấn luyện).

3.2.2. Sơ lược về OpenCV

3.2.2.1. OpenCV là gì:

OpenCV viết tắt của từ Open Source Computer Vision Library.

- OpenCV là một thư viện mã nguồn mở phục vụ cho việc nghiên cứu hay phát triển về thị giác máy tính.
- Tối ưu hóa và xử lý ứng dụng trong thời gian thực.
- Giúp cho việc xây dựng các ứng dụng xử lý ảnh, thị giác máy tính ... một cách nhanh hơn.
- OpenCV có hơn 500 hàm khác nhau, được chia làm nhiều phần phục vụ các công việc như: xử lý ảnh, an ninh, camera quan sát, nhận diện, robot...

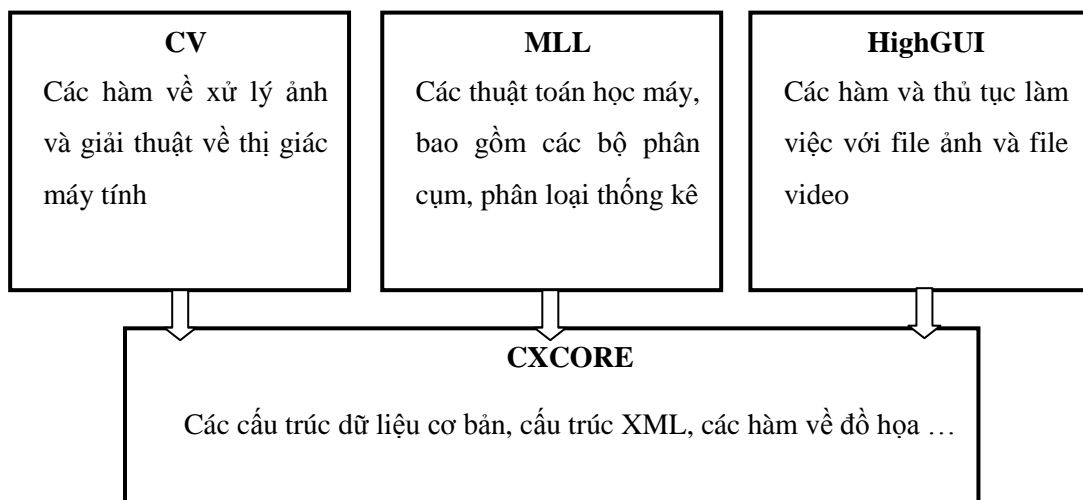
Thư viện được viết bằng ngôn ngữ C và C++ có thể chạy trên các hệ điều hành Linux, Window và MacOSX. OpenCV được thiết kế để nâng cao hiệu suất tính toán và nhấn mạnh đến hệ thống thời gian thực. OpenCV đưa ra một hệ thống đơn giản, dễ sử dụng giúp mọi người nhanh chóng xây dựng các ứng dụng

trong thị giác máy, kể cả các hệ thống kiểm tra trong nhà máy, bức ảnh trong lĩnh vực y học, bảo mật, robot học... Nó chứa các lập trình xử lý ảnh đơn giản, kể cả thực thi các hàm bậc cao như dò tìm khuôn mặt, theo dõi khuôn mặt, nhận dạng khuôn mặt...

OpenCV được giới thiệu vào tháng 1/1999, OpenCV đã được sử dụng trong rất nhiều ứng dụng, các sản phẩm và các nghiên cứu như: trong lĩnh vực hàng không, sử dụng giảm nhiễu trong y học, phân tích đối tượng, an ninh, hệ thống dò tìm, theo dõi tự động và hệ thống bảo mật...., ngoài ra nó còn được sử dụng trong nhận dạng âm thanh. OpenCV còn là một chìa khóa quan trọng trong các robot sử dụng thị giác như Stanford, Asimo.

3.2.2.2. Cấu trúc OpenCV.

Cấu trúc của opencv được chia làm 5 phần chính, 4 trong số đó được chia ra như trong hình sau:



Hình 3.2: Cấu trúc cơ bản của OpenCV

CV (computer vision) là thành phần chứa những xử lý ảnh cơ sở và thuật toán thị giác máy tính mức cao.

MLL(machine learning library) là thư viện machine learning, cái này bao gồm rất nhiều lớp thống kê và gộp công cụ xử lý.

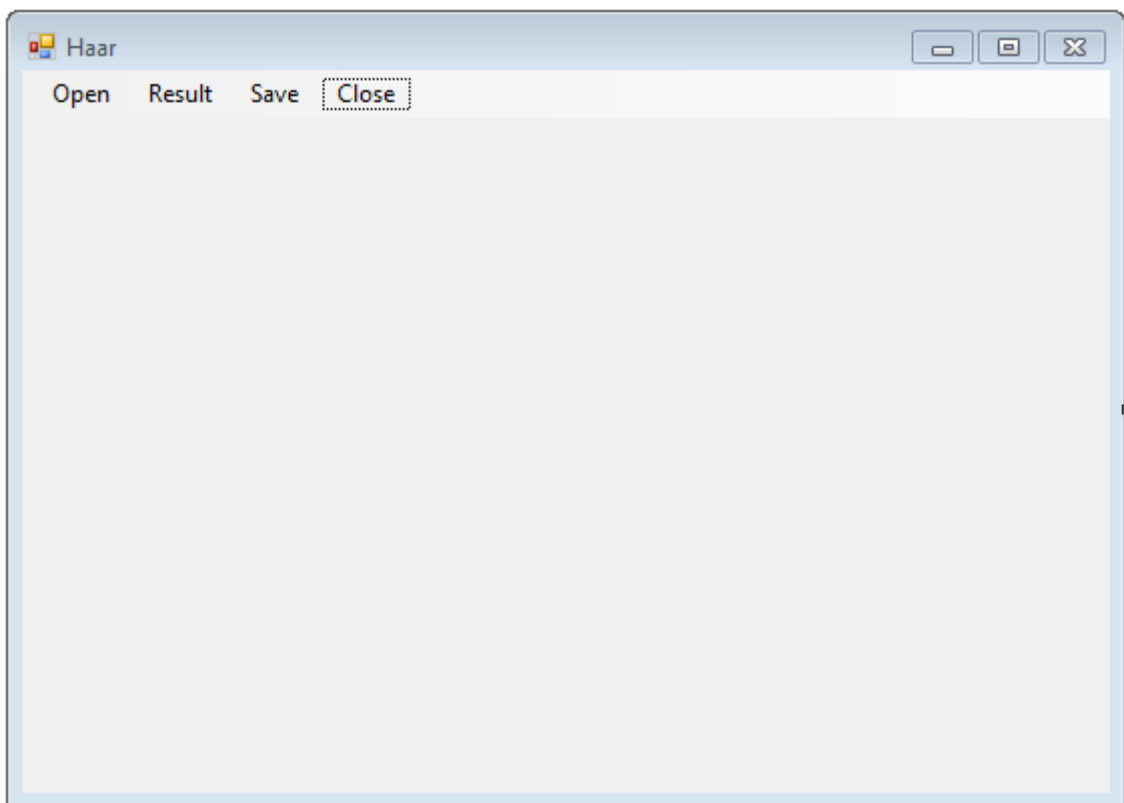
HighGUI chứa các thủ tục vào ra và các hàm dùng cho việc lưu trữ và tải ảnh video.

CXCore chứa cấu trúc và nội dung dữ liệu cơ sở.

3.3. Một số kết quả của chương trình.

Chức năng chính của chương trình bao gồm:

- Đọc file ảnh.
- Phát hiện khuôn mặt trong ảnh.
- Lưu lại ảnh đã được phát hiện.



Hình 3.3: Giao diện chính

Đầu tiên chương trình sẽ nhận đầu vào là file ảnh. Sau khi đã có ảnh đầu vào thì chuyển sang giai đoạn xử lý ảnh đó là phát hiện các khuôn mặt có trong ảnh.



Hình 3.4: Kích chọn nút Open trên thanh Menu

Sau khi đã mở ảnh để phát hiện được khuôn mặt người trong ảnh trên thanh menu ta chọn Result. Khi đó chương trình sẽ gọi tới hàm đặc trưng haar đã được xây dựng sẵn trong thư viện OpenCV trích chọn ra các đặc trưng là mặt người



Hình 3.5: Kích chọn Result trên thanh menu

Lưu hình ảnh đã được phát hiện mặt người trong ảnh ta kích chọn Save trên thanh công cụ.

Phần kết luận.

Qua tìm hiểu bài toán phát hiện mặt người trong ảnh em đã thấy được sự quan trọng của xử lý ảnh cũng như sự phát triển và tầm quan trọng của bài toán phát hiện mặt người trong đời sống.

Trong quá trình xây dựng chương trình em đã tìm hiểu về thư viện mã nguồn mở OpenCV. Qua đó biết cách sử dụng các hàm được xây dựng sẵn trong thư viện OpenCV.

Về chương trình em mới chỉ thực hiện phát hiện mặt người trong ảnh, kết quả đạt được như mong muốn, thời gian phát hiện nhanh đối với mặt người chụp thẳng, chất lượng ảnh tốt, không có vật cản...., còn trong các ảnh như chụp hơi nghiêng so với máy chụp 100 hay ảnh có màu sắc tối thì chương trình vẫn chưa phát hiện ra khuôn mặt trong ảnh đó. Em sẽ tìm hiểu thêm và cố gắng khắc phục.

Tài liệu tham khảo

a) Tài liệu tiếng việt:

- Nhập môn xử lý ảnh số (Lương Mạnh Bá , Nguyễn Thanh Thúy).
- Giáo trình Xử lý ảnh (TS. Đỗ Năng Toàn, TS Phạm Việt Bình).

b) Tài liệu Internet:

- www.comvisap.com
- <http://www.ieev.org/2010/03/adaboost-haar-features-face-detection.html>

.....