

XÂY DỰNG CƠ SỞ DỮ LIỆU KHUÔN MẶT TỰ ĐỘNG

PHỤC VỤ NHẬN DẠNG

*Huỳnh Hữu Hưng, Trần Thanh Điệp, Tạ Thị Ái Nhi**

TÓM TẮT

Phát hiện khuôn mặt người là một kỹ thuật dùng để xác định vị trí và kích thước của khuôn mặt người trong một ảnh bất kỳ. Kỹ thuật này nhận biết các đặc trưng của các khuôn mặt và bỏ qua những phần khác trong ảnh như tòa nhà, cây cối, xe cộ,...Hiện nay, nó được ứng dụng trong nhiều lĩnh vực với mục đích bảo mật. Bài báo này sẽ giới thiệu một phương pháp phát hiện khuôn mặt dựa trên hướng tiếp cận theo diện mạo sử dụng bộ phân loại mạnh AdaBoost. Dựa trên phương pháp này, chúng ta có thể xây dựng được cơ sở dữ liệu (CSDL) khuôn mặt một cách tự động phục vụ cho việc nhận dạng.

Từ khóa: Xây dựng cơ sở dữ liệu khuôn mặt người, phát hiện khuôn mặt người, nhận dạng khuôn mặt người

1. Đặt vấn đề

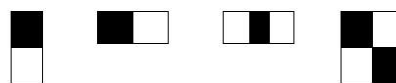
Hiện nay bài toán nhận dạng được ứng dụng khá rộng rãi trong việc xác minh tội phạm, trong các ứng dụng bảo mật của máy tính hoặc trong việc nhận dạng khuôn mặt của người rút tiền tại các trạm rút tiền tự động ATM,...

Trong phạm vi bài báo này sẽ giới thiệu một phương pháp để xây dựng một cơ sở dữ liệu khuôn mặt từ ảnh tĩnh một cách tự động phục vụ nhận dạng dựa trên thuật toán phát hiện khuôn mặt sử dụng bộ phân loại mạnh AdaBoost. Bên cạnh đó, nó còn tạo tiền đề cho việc trích lọc tất cả các khuôn mặt người có trong một video khi kết hợp cơ sở dữ liệu này với các thuật toán theo dõi đối tượng và giúp xây dựng chức năng đếm tổng số các khuôn mặt xuất hiện trong đoạn video.

Xây dựng CSDL khuôn mặt tự động từ ảnh

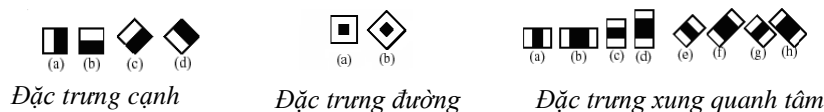
1.1. Đặc trưng Haar-like

Đặc trưng Haar-like do Viola và Jones 3.[1] công bố bao gồm 4 đặc trưng cơ bản để xác định khuôn mặt người. Mỗi đặc trưng Haar-like là sự kết hợp của hai hoặc ba hoặc bốn hình chữ nhật "trắng" và "đen" như trong Hình 1.



Hình 1. Các đặc trưng Haar-like cơ bản

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 **Error! Reference source not found.**



Hình 2. Các đặc trưng mở rộng

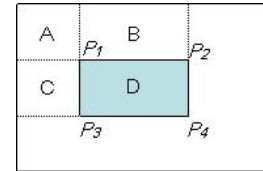
Dùng các đặc trưng trên, ta có thể tính được giá trị của đặc trưng Haar-like là sự chênh lệch giữa tổng các pixel của các vùng đen và các vùng trắng như trong công thức (1) **Error! Reference source not found.**

$$f(x) = \text{Tổng vùng đen (các mức xám của pixel)} - \text{Tổng vùng trắng (các mức xám của pixel)} \quad (1)$$

Để tính các giá trị của đặc trưng Haar-like, ta phải tính tổng của các vùng pixel trên ảnh. Nhưng để tính toán các giá trị của các đặc trưng Haar-like cho tất cả các vị trí trên ảnh đòi hỏi chi phí tính toán khá lớn, không đáp ứng được cho các ứng dụng thời gian thực. Do đó, Viola và Jones đưa ra một khái niệm gọi là Integral Image [1], đó là một mảng 2 chiều với kích thước bằng với kích thước của ảnh cần tính các đặc trưng Haar-like, với mỗi phần tử của mảng này được tính bằng cách tính tổng của điểm ảnh phía trên (dòng-1) và bên trái (cột-1) của nó. Bắt đầu từ vị trí trên, bên trái đến vị trí dưới, bên phải của ảnh, việc tính toán này đơn thuần chỉ dựa trên phép cộng số nguyên đơn giản, do đó tốc độ thực hiện rất nhanh.

Sau khi đã tính được Integral Image, việc tính tổng các giá trị mức xám của một vùng bất kỳ nào đó trên ảnh thực hiện rất đơn giản theo cách sau:

Giả sử ta cần tính tổng các giá trị mức xám của vùng D như trong Hình 3, ta có thể tính như sau: $D = A + B + C + D - (A+B) - (A+C) + A$



Hình 3. Ví dụ tính mức xám

Với $A + B + C + D$ chính là giá trị tại điểm P_4 trên Integral Image, tương tự như vậy $A+B$ là giá trị tại điểm P_2 , $A+C$ là giá trị tại điểm P_3 , và A là giá trị tại điểm P_1 . Vậy ta có thể viết lại biểu thức tính D ở trên như công thức (2)

$$D = \underbrace{(x_4, y_4)}_{A+B+C+D} - \underbrace{(x_2, y_2)}_{A+B} - \underbrace{(x_3, y_3)}_{A+C} + \underbrace{(x_1, y_1)}_A \quad (2)$$

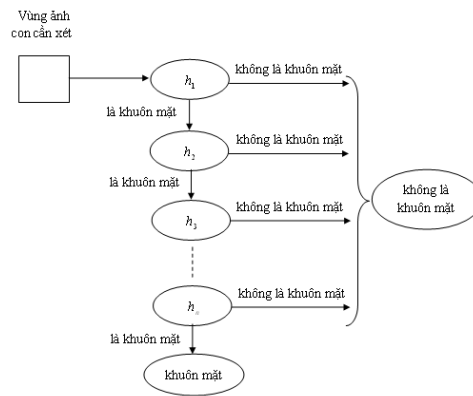
Tiếp theo, để chọn các đặc trưng Haar-like dùng cho việc thiết lập ngưỡng, Viola và Jones sử dụng một phương pháp máy học được gọi là AdaBoost. AdaBoost sẽ kết hợp các bộ phân loại yếu để tạo thành một bộ phân loại mạnh.

1.2. AdaBoost

AdaBoost là một bộ phân loại mạnh phi tuyến phức dựa trên hướng tiếp cận boosting được Freund và Schapire đưa ra vào năm 1995 [2]. Adaboost cũng hoạt động trên nguyên tắc kết hợp tuyến tính các bộ phân loại yếu để hình thành một bộ phân loại mạnh.

Trong quá trình huấn luyện, cứ mỗi bộ phân loại yếu được xây dựng, thuật toán sẽ tiến hành cập nhật lại trọng số để chuẩn bị cho việc xây dựng bộ phân loại yếu kế tiếp bằng cách tăng trọng số của các mẫu bị nhận dạng sai và giảm trọng số của các mẫu được nhận dạng đúng bởi bộ phân loại yếu vừa xây dựng.

Điều này giúp bộ phân loại yếu sau có thể tập trung vào các mẫu mà các bộ phân loại yếu trước nó làm chưa tốt. Sau cùng, các bộ phân loại yếu sẽ được kết hợp tùy theo mức độ tốt của chúng để tạo nên bộ phân loại mạnh. 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 Hình 4.



Hình 4. Mô hình phân tầng kết hợp các bộ phân loại yếu để xác định khuôn mặt

Trong đó, h_k là các bộ phân loại yếu, được biểu diễn bởi công thức (3) 3.[1]

x : Cửa sổ con cần xét

θ_k : Ngưỡng

f_k : Giá trị của đặc trưng Haar-like

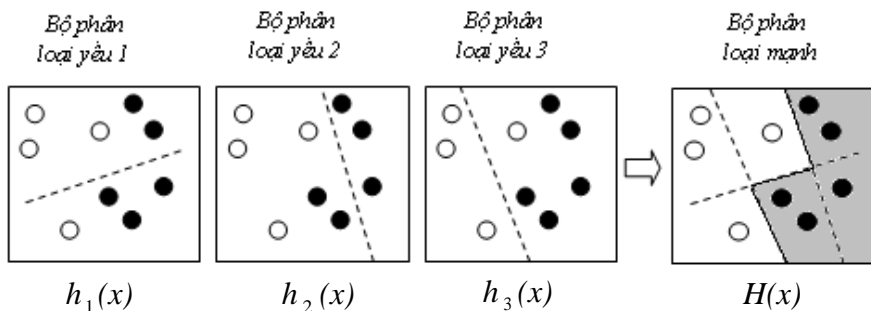
p_k : Hệ số quyết định chiều của phương trình

$$h_k = \begin{cases} 1 & \text{nếu } p_k f_k(x) < p_k \theta_k \\ 0 & \text{ngược lại} \end{cases} \quad (3)$$

AdaBoost sẽ kết hợp các bộ phân loại yếu thành bộ phân loại mạnh như Hình 5.

Giá trị của bộ phân loại mạnh $H(x) = \text{sign}(a_1 h_1(x) + a_2 h_2(x) + \dots + a_n h_n(x))$ ($a = \text{alpha}$)

Với: $a_i \geq 0$ là hệ số chuẩn hoá cho các bộ phân loại yếu

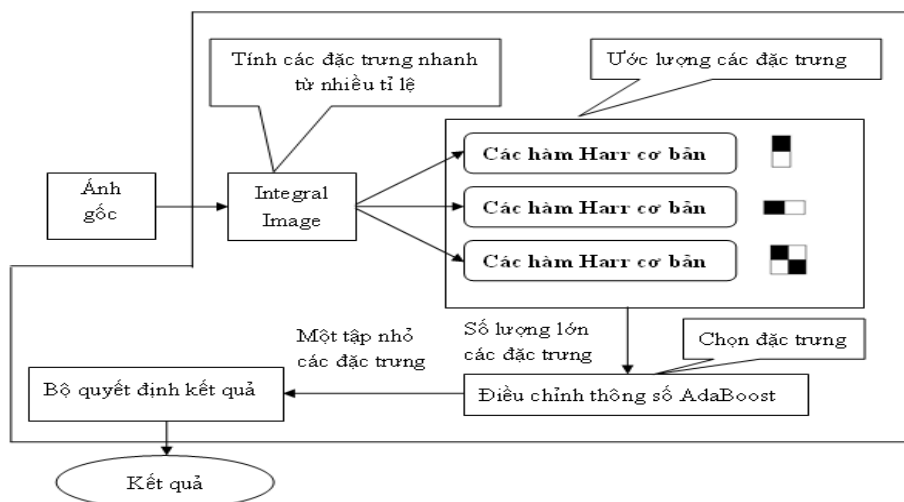


Hình 5. Kết hợp các bộ phân loại yếu thành bộ phân loại mạnh

1.3. Hệ thống phát hiện và xác định vị trí khuôn mặt trên ảnh

Trong Hình 6, từ ảnh gốc ban đầu, ta sẽ được tính Integral Image, là mảng 2 chiều với phần tử (x, y) sẽ được tính bằng tổng của các phần tử (x', y') với $x' < x$ và $y' < y$,

mục đích là 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 đặc trưng 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 sẽ được chuyển sang cho bộ quyết định kết quả (là tập các bộ phân loại yếu có cấu trúc như Hình 5). Bộ quyết định sử dụng mô hình phân tầng như Hình 4 sẽ tổng hợp kết quả dựa trên kết quả trả về của các bộ phân loại yếu.



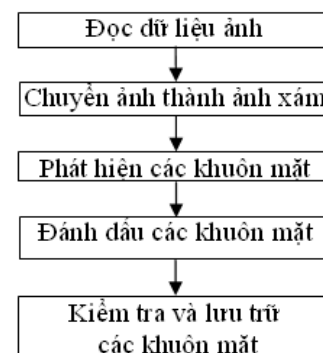
Hình 6. Hệ thống phát hiện và xác định vị trí khuôn mặt người trên ảnh

Chú ý rằng 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 Haar-like, nếu có một đặc trưng Haar-like 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 Haar-like sẽ được dựa vào trọng số (weight) của đặc trưng đó và do AdaBoost quyết định dựa vào số lần và thứ tự xuất hiện của các đặc trưng Haar-like.

1.4. Thuật toán xây dựng CSDL khuôn mặt tự động từ ảnh

Thuật toán của việc xây dựng CSDL từ ảnh được mô tả như Hình 7. Với đầu vào là các tập tin ảnh với định dạng (Joint Photographic Group) hoặc JPEG (Joint Photographic Experts Group), hệ thống sẽ đọc dữ liệu ảnh, chuyển thành ảnh xám và tiến hành phát hiện khuôn mặt có trong ảnh.

Việc phát hiện khuôn mặt được thực hiện bằng thuật toán Adaboost thông qua hệ thống được trình bày như Hình 6 bằng việc sử dụng hàm `cvHaarDetectObjects()` [3]. Hàm này thực hiện việc phát hiện đối tượng dựa trên các đặc trưng



Hình 7. Thuật toán xây dựng CSDL khuôn mặt tự động

haar-like nhờ vào một bộ Cascade 3.[1] được truyền vào cho hàm. Bộ Cascade được xây dựng theo dạng cây (tree-node) và đã được huấn luyện từ trước. Tuy nhiên, việc huấn luyện và tạo bộ dữ liệu mẫu này là một công việc khó và tốn nhiều thời gian vì thế trong phạm vi bài toán này sẽ lấy các mẫu khuôn mặt đã được training sẵn từ Internet để làm dữ liệu mẫu cho chương trình. Sau đó, ta tiến hành so sánh kết quả lấy được với bộ dữ liệu sẵn có và đánh dấu các khuôn mặt tìm được trong ảnh. Cuối cùng, chương trình sẽ thực hiện việc kiểm tra và lưu tất cả các khuôn mặt trong ảnh để tạo thành cơ sở dữ liệu gồm các khuôn mặt phục vụ cho việc nhận dạng tự động.

2. Kết quả và đánh giá

Thử nghiệm chương trình trên một tập ảnh tĩnh gồm 30 ảnh định dạng JPG và JPEG với nhiều kích thước khác nhau, thì độ chính xác thu được của việc nhận dạng khoảng 92.6% và chương trình có thể đem lại một tập dữ liệu các ảnh khuôn mặt khá phong phú với nhiều kích thước phục vụ tốt cho quá trình nhận dạng người tự động thông qua khuôn mặt.



Hình 8. Kết quả của chương trình trên một ảnh mẫu

3. Kết luận

Từ các kết quả thu được trong quá trình thử nghiệm, ta nhận thấy chương trình làm việc khá tốt khi khuôn mặt người nhìn thấy rõ, không bị che khuất và độ sáng của ảnh ổn định. Với cơ sở dữ liệu các khuôn mặt thu được này có thể giúp ích rất nhiều trong việc xây dựng được các chương trình nhận dạng khuôn mặt. Hơn nữa, nó còn là nền tảng cho việc xây dựng các chức năng như theo vết nhiều khuôn mặt cùng lúc, đếm số lượng người trong video.

TÀI LIỆU THAM KHẢO

- [1] P. Viola, M. J. Jones (2004), *Robust real-time face detection*, International Journal of Computer Vision.
- [2] Y. Freund, R. Schapire (1997), *A decision-theoretic generalization of on-line learning and an application to boosting*, Journal of Computer and System Sciences.

- [3] *OpenCV Reference Manual v2.1*
- [4] Rainer Lienhart, Jochen Maydt, *An Extended Set of Haar-like Features for Rapid Object Detection*, Intel Labs, Intel Corporation, Santa Clara, CA 95052, USA.

BUILDING AUTOMATICALLY FACE DATABASE FOR FACE RECOGNITION

Huynh Huu Hung¹, Tran Thanh Diep², Ta Thi Ai Chi³

¹*University of Technology, The University of Danang*

²*Center for Software Development, The University of Danang*

³*Faculty of Information Technology, University of Technology, The University of Danang*

ABSTRACT

Face Detection is a technology that is used to determine the location and size of people's face in any image. This technology detects features of a face and skips other parts of the image such as buildings, trees, transportation, etc. Currently, it is applied in many fields for security purpose. This paper will introduce a face detection method based on face approach using the AdaBoost strong classifier. Based on this method, we can build face database automatically to support face recognition.

Key words: building face database, face recognition, face detection

* Huỳnh Hữu Hưng, Trường Đại học Bách khoa, Đại học Đà Nẵng

Trần Thanh Điệp, Trung Tâm Phát Triển Phần Mềm, Đại học Đà Nẵng

Tạ Thị Ái Nhi, Lớp Cao học KHMT-K12, Khoa Công nghệ Thông tin, Trường Đại học Bách khoa, Đại học Đà Nẵng