

# BÁO CÁO KẾT QUẢ NGHIÊN CỨU

Ngày thực hiện báo cáo : 06/05/2013

<b>TÊN ĐỀ TÀI</b>	Tìm hiểu và áp dụng mô hình AAM để nhận dạng và truy vấn khuôn mặt hai chiều.	
<b>LĨNH VỰC CHUYÊN NGÀNH</b>	Khoa học máy tính	
<b>LOẠI HÌNH NGHIÊN CỨU</b>	Lí thuyết	
<b>NGƯỜI HƯỚNG DẪN</b>	Ths. Âu Bửu Long	
<b>SINH VIÊN THỰC HIỆN 1</b>	Nguyễn Thị Diệu	K35.103.007
<b>SINH VIÊN THỰC HIỆN 2</b>	Huỳnh Thị Yến Phương	K35.103.059

## GIỚI THIỆU

Nhận dạng là một lĩnh vực nghiên cứu thu hút nhiều sự chú ý của các kỹ sư ngành Thị giác máy tính. Nhận dạng nói chung, nhận dạng mặt người nói riêng là bài toán mới chỉ xuất hiện cách đây không lâu - khoảng vài thập niên nhưng đã và đang được nghiên cứu rất rộng rãi. Các nghiên cứu đi từ những bài toán đơn giản như nhận dạng một mặt người trong ảnh đen trắng với một khuôn mặt được chụp thẳng đến những bài toán phức tạp như nhận dạng ảnh màu, có nhiều mặt người trong ảnh với nhiều góc chụp khác nhau.

Bài toán nhận dạng mặt người có rất nhiều ứng dụng trong cuộc sống. Các ứng dụng liên quan đến nhận dạng mặt người có thể kể như: xác minh tội phạm, ứng dụng trong ATM để lưu trữ khuôn mặt của người rút tiền, quản lý nhân viên và an ninh sân bay, đăng nhập máy tính cá nhân sử dụng khuôn mặt,... Đó chính là lý do mà bài toán này thu hút được nhiều sự nghiên cứu trong thời gian dài.

Xây dựng hệ thống nhận dạng mặt người không hề đơn giản, nhìn chung các hệ thống nhận dạng khuôn mặt đều trải qua nhiều bước. Đầu tiên, hệ thống cần phát hiện khuôn mặt (*face detection*) tức là phần ảnh mặt trong dữ liệu đầu vào và cắt lấy phần mặt để nhận dạng. Bước thứ hai là tiền xử lý ảnh (*preprocessing*) bao gồm các bước căn chỉnh ảnh và chuẩn hóa ánh sáng. Tiếp theo là bước trích chọn đặc trưng. Ở bước này, một phương pháp trích chọn đặc trưng phù hợp sẽ được sử dụng với ảnh mặt để trích chọn đặc trưng của khuôn mặt và được biểu diễn dưới dạng vector gọi là vector đặc trưng (*feature vector*). Một số phương pháp trích chọn đặc trưng mang lại hiệu quả cao có thể kể đến như phương pháp phân tích thành phần chính (*Principal Component Analysis – PCA*) [3], mẫu nhị phân cục bộ (*Local Binary Pattern – LBP*) hay Gabor wavelets, Gradient,... Bước tiếp theo là nhận dạng khuôn mặt (*recognition*) tức là xác định danh tính hay nhãn của ảnh đó, ảnh đầu vào là của ai.

Có rất nhiều hệ thống nhận dạng mặt người đã được xây dựng dựa trên các phương pháp khác

nhau như Eigenface, Active Shape Model (ASM), Active Appearance Model (AAM) [4], phương pháp dò tìm Nơron, Support Vector Machine (SVM),...

Mô hình AAM (*Active Appearance Models*) là một trong những mô hình thống kê kinh điển được dùng nhiều trong việc nhận dạng hoặc theo vết các đối tượng thuộc cùng một họ, trong đó mảng được ứng dụng nhiều nhất vẫn là khuôn mặt người. Mô hình này được đề xuất bởi Cootes, Edwards và Taylor vào năm 1998.

Đề tài sẽ trình bày các vấn đề liên quan tới kiến thức nền tảng xây dựng nên khóa luận. Với mục tiêu chính là tìm hiểu mô hình Active Appearance Model và ứng dụng trong nhận dạng khuôn mặt hai chiều. Bên cạnh đó, đề tài cũng sẽ tìm hiểu các mô hình truy vấn ảnh và áp dụng để đưa ra những khuôn mặt gần giống với ảnh đầu vào nhất.

#### GHI CHÚ

## MỤC TIÊU NGHIÊN CỨU CỦA ĐỀ TÀI

Nhận dạng ngày càng có nhiều ứng dụng trong các hoạt động của con người. Bên cạnh đó, với dữ liệu ảnh khổng lồ như ngày nay thì vấn đề truy vấn ảnh để tìm kiếm ảnh tương đồng rất cần thiết và tiện lợi. Do đó, luận văn sẽ tìm hiểu mô hình AAM và áp dụng mô hình AAM trong nhận dạng và truy vấn khuôn mặt hai chiều.

Mục tiêu chính của đề tài là tìm hiểu mô hình AAM và xây dựng một hệ thống nhận dạng khuôn mặt người. Từ ảnh gốc ban đầu, hệ thống sẽ tiến hành phát hiện khuôn mặt. Kết quả của quá trình phát hiện khuôn mặt là ảnh đó có phải là mặt người hay không. Nếu là mặt người, thì sẽ giới hạn khuôn mặt trong khung hình chữ nhật. Tiếp theo, áp dụng mô hình AAM để tiến hành nhận dạng: khởi tạo hình ảnh trung bình – đã được huấn luyện "offline" lên trọng tâm hình chữ nhật chứa khuôn mặt và tính toán vector đặc trưng. Vector đặc trưng này được sử dụng để tính khoảng cách từ ảnh đầu vào với các ảnh trong tập huấn luyện. Cuối cùng, hệ thống truy vấn ảnh sẽ đưa ra những ảnh gần giống với khuôn mặt đầu vào nhất.

AAM là một mô hình thống kê được đề xuất bởi Cootes, Taylor và Edwards vào năm 1998. Ý tưởng của mô hình này là xây dựng hai mô hình con: mô hình xử lý hình dạng và mô hình xử lý vân. Sau đó sẽ kết hợp hai mô hình này lại với nhau bằng cách sử dụng phương pháp phân tích thành phần chính PCA để rút trích đặc trưng. Đặc trưng này được biểu diễn dưới dạng vector gọi là vector đặc trưng. Vector đặc trưng thể hiện được hầu hết những thông tin trên khuôn mặt và được dùng để nhận dạng khuôn mặt. Ưu điểm của mô hình AAM này là tốc độ nhanh, sử dụng giải thuật PCA để phân tích thành phần chính để giảm số chiều của mẫu dữ liệu nhưng vẫn giữ lại những đặc trưng cơ bản giúp quá trình nhận dạng nhanh hơn. Trong quá trình nhận dạng, hệ thống sẽ tiến hành tìm kiếm ảnh giống với ảnh đầu vào để đưa ra kết quả. Cần thiết phải sử dụng một hệ thống truy vấn để loại bỏ những hình ảnh sai, hoặc không giống, tiết kiệm thời gian tìm kiếm và độ chính

xác cao. Để xác định mức độ giống nhau giữa hai ảnh cần so sánh thì cần phải có một giá trị để biểu thị cho sự giống nhau này và giá trị này là độ đo. Do đó đề tài sẽ tìm hiểu và chọn một hệ thống truy vấn cùng với một độ đo tối ưu nhất để truy vấn ảnh.

#### GHI CHÚ

## TÌNH HÌNH NGHIÊN CỨU TRONG VÀ NGOÀI NƯỚC

Với sự phát triển của thị giác máy tính, có rất nhiều phương pháp nhận dạng được nghiên cứu và áp dụng. Mục tiêu chung của các phương pháp này là xây dựng một mô hình phù hợp nhất có thể thể hiện chính xác nhất hình ảnh đầu vào. Có khá nhiều nghiên cứu về mô hình nhận dạng kinh điển và gần đây. Tuy nhiên, mô hình thành công nhất có thể kể đến là mô hình thống kê.

Có hai phương pháp nhận dạng phổ biến hiện nay là nhận dạng dựa trên đặc trưng của các phần tử trên khuôn mặt như biến đổi Gabor Wavelet và mạng Neural, SVM (Support Vector Machine),... và nhận dạng dựa trên xét tổng thể toàn khuôn mặt như phương pháp PCA, LDA, LFA [2]. Trong đó, PCA là phương pháp rút trích đặc trưng nhằm giảm số chiều của ảnh tuy đơn giản nhưng mạng lại hiệu quả tốt. Nhận dạng khuôn mặt dùng PCA kết hợp với mạng Nơron là phương pháp mạng lại hiệu quả nhận dạng cao bởi nó phát huy được ưu điểm của PCA và mạng nơron. Bên cạnh đó, nhận dạng dựa trên các phần đặc trưng của các phần tử trên khuôn mặt cũng được thu hút nhiều chuyên gia nghiên cứu.

Nguyễn Thành Thái (2006) đã kết hợp SVM (Support Vector Machine) và mạng Nơron để nhận dạng khuôn mặt người. Mục tiêu của phương pháp là xác định một đối tượng có phải là học sinh của một trường THPT hay không. Phương pháp này có hai giai đoạn là huấn luyện và nhận dạng. Trong giai đoạn huấn luyện: phương pháp PCA được sử dụng để rút trích vector đặc trưng, sau đó đưa vào bộ nhận dạng SVM để xác định lớp cho mẫu. SVM là phương pháp phân lớp nhanh nhưng vẫn có thể phân một đối tượng nhầm lớp nên phương pháp mạng Nơron đã được sử dụng thêm để so sánh.

Ngoài ra, Từ Minh Hiên – Trần Thị Khánh Hòa (2012) cũng sử dụng mạng Nơron và phương pháp phân tích thành phần chính để nhận dạng khuôn mặt. Những đặc trưng sau khi được rút trích sẽ được đưa vào khối nhận dạng để tiến hành phân lớp đối tượng. Các đối tượng được nhận dạng bằng cách sử dụng hai thuật toán: phương pháp đối sánh dùng khoảng cách Euclides, Mahalanobis và dùng mạng Nơron.

Nguyễn Quốc Bình (2011) đã nhận diện khuôn mặt bằng cách sử dụng giải thuật PCA để xây dựng chương trình quản lý nhân viên.

Võ Hoàng Minh, Trần Bình Long, Lê Hoàng Thái, Trần Hành (2011) đã nhận dạng mặt người dùng Polar Cosine Transform và mạng Radial Basis Function.

Mai Hữu Lợi (2011) đã sử dụng thuật toán PCA để nhận dạng khuôn mặt người.

AAM là mô hình thống kê được phát triển dựa trên mô hình ASM (Active Shape Model) [6].

Trước ASM, một số mô hình thống kê đã ra đời như PDM (Point Distribute Model), Eigenfaces (phương pháp mặt riêng), Active Contour Model (Snakes),..

Đã có rất nhiều thuật toán được phát triển cải tiến cho phương pháp AAM trong những năm gần đây [5]. Mô hình AAM được áp dụng cho những bài toán đơn giản như nhận dạng một mặt người trong ảnh đen trắng với một khuôn mặt được chụp thẳng đến những bài toán phức tạp như nhận dạng ảnh màu. Gần đây nhất, áp dụng mô hình AAM để nhận dạng khuôn mặt trong đoạn video đang nhận được nhiều sự quan tâm nghiên cứu của các kỹ sư ngành thị giác máy tính.

Một dự án nghiên cứu của Mircea Ionita và Phil Treasader: phát triển các thuật toán dựa trên mô hình AAM để xác định vị trí và theo dõi đặc điểm khuôn mặt trong thời gian thực trên điện thoại di động.

Tháng 4, năm 2009, K. Zimmermann, J. Matas, and T. Svoboda sử dụng mô hình AAM để truy vết khuôn mặt. Đầu vào của hệ thống theo vết khuôn mặt là đoạn video chứa khuôn mặt cần theo vết. Vì khuôn mặt cần nhận dạng thay đổi liên tục trong video nên hệ thống sẽ chia video thành từng frame và áp dụng AAM để nhận dạng khuôn mặt trong frame đó.

Tháng 9 năm 2009, T. Zhang, D. Tao, X. Li, and J. Yang đề xuất phương pháp căn chỉnh tập huấn luyện mới để giảm số chiều của tập huấn luyện, cải thiện đáng kể thời gian huấn luyện và số chiều của vector hình dạng.

Tháng 11 năm 2009, J. Saragih and R. Goecke công bố một công thức mới để xây dựng tập huấn luyện tốt hơn. Chính nhờ những thông số này mà quá trình khởi tạo hình dạng ban đầu cho ảnh đầu vào tốt hơn rất nhiều, thuận tiện cho quá trình khớp ảnh.

Các phương pháp nhận như Nơron, SVM (Support Vector Machine),... có khả năng nhận dạng tốt. Tuy nhiên, những phương pháp này tốn nhiều thời gian huấn luyện và nhận dạng. So với các phương pháp nhận dạng trước đây như Eigenface, Snakes, ASM,...thì AAM (Active Appearance Model) mô tả đầy đủ thông tin của khuôn mặt hơn do có sự kết hợp giữa mô hình hình dạng và vân. Đối với các phương pháp mới trong những năm gần đây, AAM ít tốn thời gian huấn luyện và nhận dạng hơn. Ngoài ra còn có một ưu điểm nữa là mô hình này có khả năng khử nhiễu. AAM có khả năng khử nhiễu là do đây là một mô hình thống kê. Mặc dù là mô hình thống kê kinh điển nhưng với những ưu điểm trên, đề tài đã sử dụng mô hình AAM để xây dựng một hệ thống nhận dạng khuôn mặt.

Trên cơ sở nghiên cứu về bài toán nhận dạng mặt người, đề tài đã tìm hiểu mô hình AAM và kết hợp với một hệ thống truy vấn ảnh để nhận dạng khuôn mặt. Khi nghiên cứu về vấn đề này, đề tài không những chỉ nghiên cứu về mô hình AAM và thuật toán AASA để khớp ảnh mà còn tìm hiểu thêm các phương pháp nhận dạng khác và các cơ sở lý thuyết liên quan. Bên cạnh đó, trong quá trình xây dựng chương trình ứng dụng, đề tài đã tìm hiểu về thư viện mã nguồn mở OpenCV, qua đó biết cách sử dụng các hàm trong thư viện OpenCV để thực hiện bước phát hiện khuôn mặt đối với một ảnh đầu vào trước khi tiến hành nhận dạng. So với các công trình nghiên cứu trước đây, chúng em đã tìm hiểu và xây dựng thành công mô hình truy vấn đơn giản ảnh từ vector đặc trưng rút trích được từ mô hình AAM. Tuy nhiên, do sự giới hạn về thời gian và chuyên môn và đề tài nghiên cứu tương đối rộng, hệ thống truy vấn ảnh chưa đánh giá được độ đo tương tự tối ưu cho mô hình và chưa áp dụng được mô hình tri thức trong truy vấn ảnh.

## NỘI DUNG VÀ PHẠM VI NGHIÊN CỨU

Đề tài tìm hiểu mô hình AAM và cơ sở lý thuyết liên quan: xử lý ảnh, phương pháp phân tích thành phần chính (PCA), phát hiện khuôn mặt bằng phương pháp sử dụng đặc trưng Haar-like và giải thuật Adaboost, các hệ thống truy vấn ảnh, các độ đo thông dụng. Trên cơ sở đó, xây dựng hệ thống nhận dạng và truy vấn ảnh trên phần mềm Visual Studio 2010, thư viện OpenCV 2.4.0 và Matlab 7.8 R2009a.

Thông thường, một hệ thống nhận dạng khuôn mặt thường trải qua nhiều bước. Bước đầu tiên không thể thiếu trong nhận dạng là phát hiện khuôn mặt. Nếu bước phát hiện khuôn mặt được thực hiện tốt thì sẽ làm tăng hiệu quả nhận dạng của hệ thống. Phương pháp được sử dụng để phát hiện khuôn mặt là phương pháp sử dụng đặc trưng Haar – like. Sau khi ảnh đầu vào được xác định là có khuôn mặt người trong ảnh thì hệ thống sẽ tiến hành nhận dạng. Tuy nhiên, kết quả đầu ra là một ảnh duy nhất, ảnh này có thể không đúng với yêu cầu. Do đó cần phải sử dụng một hệ thống truy vấn ảnh, tính toán khoảng cách giữa các ảnh và đưa ra tập những ảnh gần giống với ảnh nhận dạng nhất trong CSDL ảnh để loại bỏ những hình ảnh sai, hoặc không giống, tiết kiệm thời gian tìm kiếm và độ chính xác cao. Để xác định mức độ giống nhau giữa hai ảnh cần so sánh thì cần phải có một giá trị để biểu thị cho sự giống nhau này và giá trị này là độ đo tương tự.

Các thông tin khuôn mặt được biểu diễn dưới dạng vector. Giả sử tập huấn luyện gồm  $N$  ảnh, chúng ta sẽ có  $N \times N$  vector. Tuy nhiên, trong số những vector đó có những vector không cần thiết, ta có thể sử dụng PCA để loại bỏ những vector này mà không ảnh hưởng đến thông tin khuôn mặt. Quá trình này gọi là rút trích vector đặc trưng. Mô hình AAM được xây dựng dựa trên hai mô hình con là mô hình hình dạng và mô hình vân. Theo AAM, đặc trưng về khuôn mặt được biểu thị bằng vector hình dạng và vector vân. Bằng việc áp dụng phương pháp PCA kết hợp hai vector này lại ta sẽ xây dựng được một vector duy nhất gọi là vector đặc trưng. Sau đó, sử dụng vector đặc trưng này để xây dựng lại khuôn mặt đầu vào dựa trên tập huấn luyện và tìm những ảnh gần giống nhất với ảnh trong CSDL. Để tìm ảnh gần nhất với ảnh trong CSDL, một phương pháp đơn giản mang lại hiệu quả cao được sử dụng đó là tính khoảng cách từ ảnh đầu vào với các ảnh trong CSDL. Khoảng cách giữa các ảnh được so sánh qua độ đo tương tự. Nếu độ đo tương tự giữa hai ảnh càng nhỏ thì hai ảnh đó càng giống nhau và ngược lại.

Tập huấn luyện gồm 238 ảnh: 74 ảnh nam và 45 ảnh nữ, mỗi người được chọn hai trạng thái: trạng thái bình thường và trạng thái cười. Mỗi ảnh có kích thước 768 x 576 pixel. Ảnh này là ảnh mặt người trong các cơ sở dữ liệu mẫu, chụp từ máy ảnh. Các đối tượng được chụp trong điều kiện ánh sáng bình thường, chụp trực diện hoặc nghiêng một góc không đáng kể so với phương thẳng đứng (nhỏ hơn  $10^0$ ). Ảnh không thể hiện cảm xúc quá nhiều, không bị ngược lên hoặc cúi xuống. Mỗi ảnh chỉ chứa duy nhất một khuôn mặt người, tức là ảnh đầu vào chỉ gồm đối tượng nhận dạng. Các thành phần trên khuôn mặt rõ ràng, không bị che khuất. Đề tài không xem xét đến

các ảnh mặt người nhân tạo.

GHI CHÚ

## CÁC GIẢ THUYẾT VÀ CÁCH TIẾP CẬN GIẢI QUYẾT CÁC BÀI TOÁN

### ❖ Mô hình AAM:

Mô hình diện mạo tích cực (Active Appearance Model – AAM) là mô hình thống kê được đề xuất bởi Cootes, Taylor và Edwards vào năm 1998. Mô hình này là sự kết hợp hai mô hình con: mô hình hình dạng (*shape model*) và mô hình vân (*texture model*).

Ý tưởng của mô hình: từ một ảnh đầu vào, sử dụng PCA để rút trích đặc trưng hình dạng khuôn mặt, hình dạng đối tượng được biểu diễn dưới dạng một vector  $x$ . Tương tự như hình dạng, vân của khuôn mặt cũng được sử dụng PCA và biểu diễn dưới dạng vector  $g$ . Sau đó áp dụng phương pháp phân tích thành phần chính (PCA) để kết hợp hai vector này lại thành một vector đặc trưng duy nhất (*feature vector*).

#### - Mô hình hình dạng:

Mỗi khuôn mặt người trong ảnh được biểu diễn bởi  $n$  điểm mốc, tuy nhiên với cách biểu diễn này sẽ gây khó khăn trong việc tính toán, do đó mỗi khuôn mặt sẽ biểu diễn bằng một vector có dạng:

$$x = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]^T$$

Với tập huấn luyện gồm  $N$  ảnh ta sẽ có  $N$  vector biểu diễn hình dạng khuôn mặt khác nhau. Tuy nhiên, mỗi khuôn mặt có hình dạng, vị trí và hướng khác nhau. Do đó, ta cần phải căn chỉnh (align) các khuôn mặt về cùng 1 vị trí bằng phép tịnh tiến (translation), phép biến đổi tỉ lệ (scale) và phép quay (rotation).

Sau khi tiền xử lý thông qua bước trên ta sẽ xây dựng mô hình xử lý hình dạng dựa trên tập huấn luyện: giống hàng tất cả hình dạng trong tập huấn luyện về cùng 1 hệ tọa độ và biểu diễn bằng vector  $x$ . Áp dụng giải thuật PCA (Principal Component Analysis) để chuyển đổi dữ liệu, ta tìm được biểu diễn của  $x$ :

$$x = \bar{x} + P_s b_s$$

Trong đó:  $\bar{x}$  là hình dạng trung bình,  $P_s$  là ma trận hiệp phương sai,  $b_s$  là tham số hình dạng

#### - Mô hình vân

Tương tự mô hình xử lý hình dạng, ta đưa về vân trung bình, sau đó tính toán sai khác giữa vân trung bình so với từng ảnh rồi mô hình hóa sự sai khác này. Áp dụng PCA ta tính được mô hình vân:

$$g = \bar{g} + P_g b_g$$

Trong đó:  $\bar{g}$  là mô hình vân trung bình,  $P_g$  là ma trận hiệp phương sai,  $b_g$  là tham số mô hình vân.

- Mô hình kết hợp:

Sau khi đã xây dựng 2 mô hình hình dạng, mô hình vân và có được các vector tham số  $b_s$  và  $b_g$ . Mô hình AAM sẽ kết hợp 2 tham số này lại để cho ra 1 tham số mô hình duy nhất. Mô hình này có dạng:

$$b = P_c c$$

Trong đó:  $P_c$  là tập các vector riêng của mô hình mới,  $c$  là tham số của mô hình kết hợp.

Do hình dạng và vân thể hiện 2 yếu tố khác nhau của khuôn mặt, nên khi kết hợp lại, một trong 2 tham số cần được gán trọng số.

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{pmatrix}$$

$W_s$  là một ma trận đường chéo cung cấp một trọng số phù hợp giữa các khoảng cách điểm ảnh và cường độ điểm ảnh.

Bằng việc biến đổi tuyến tính vector tham số  $c$ , hình dạng và diện mạo của khuôn mặt sẽ được thay đổi như sau:

$$x = \bar{x} + P_s W_s^{-1} P_{cs} c \quad , \quad g = \bar{g} + P_g P_{cg} c$$

$$\text{Với: } Q = \begin{pmatrix} P_s \\ P_g \end{pmatrix}$$

Tổng quát hóa phương trình trên ta thu được:

$$x = \bar{x} + Q_s c \quad , \quad g = \bar{g} + Q_g c \quad (*)$$

❖ Ý tưởng của hệ thống:

Từ ảnh gốc ban đầu, hệ thống sẽ tiến hành phát hiện khuôn mặt. Kết quả của quá trình phát hiện khuôn mặt là ảnh đó có phải là mặt người hay không. Nếu là mặt người, thì sẽ giới hạn khuôn mặt trong khung hình chữ nhật. Tiếp theo, áp dụng mô hình AAM để tiến hành nhận dạng: khởi tạo hình ảnh trung bình – đã được huấn luyện "offline" lên trọng tâm hình chữ nhật chứa khuôn mặt và tính toán vector đặc trưng. Vector đặc trưng này được sử dụng để tính khoảng cách từ ảnh đầu vào với các ảnh trong tập dữ liệu ảnh huấn luyện. Cuối cùng, hệ thống truy vấn ảnh sẽ đưa ra những ảnh gần giống với khuôn mặt đầu vào nhất.

- Phát hiện khuôn mặt:

Đây là bước đầu tiên cần thực hiện trong một hệ thống nhận dạng khuôn mặt người. Mục đích của bước phát hiện khuôn mặt người này là làm tăng hiệu quả nhận dạng của hệ thống. Phương pháp được sử dụng để phát hiện khuôn mặt là phương pháp dùng đặc trưng Haar – like.

Đặc trưng Haar-like được Papageorgiou đề xuất lần đầu tiên. Viola và Jones đã kết hợp đặc trưng này cùng với các kỹ thuật:

- + Tích phân ảnh (*Integral Image*) để tính toán các đặc trưng ảnh một cách nhanh chóng.
- + Bộ phân lớp nhiều tầng (*Cascade of Classifiers*) để kết hợp một cách hiệu quả nhiều

đặc trưng.

+ toán máy học AdaBoost của Freund và Schapire

Phương pháp này đã được tích hợp trong thư viện mã nguồn mở OpenCV.

Để phát hiện khuôn mặt, ảnh được quét bằng một cửa sổ con chứa một đặc trưng Haar-like. Ứng với một đặc trưng Haar-like, một bộ phân lớp yếu được xây dựng. Nếu giá trị của đặc trưng Haar-like tại cửa sổ con vượt quá một ngưỡng nào đó thì đó là khuôn mặt; ngược lại thì không phải là khuôn mặt. Đặc trưng Haar-like được Papageorgiou đề xuất lần đầu tiên. Viola và Jones đã sử dụng thành công đặc trưng Haar-like này nhận dạng đối tượng một cách nhanh chóng. Lợi ích của các đặc trưng Haar-like là nó diễn đạt được tri thức về các đối tượng trong ảnh, bởi vì nó biểu diễn mối liên hệ giữa các bộ phận của đối tượng, điều mà bản thân từng điểm ảnh không diễn đạt được.

Từ ảnh gốc ban đầu, hệ thống sẽ chia ảnh thành vô số vùng nhỏ để tính các đặc trưng. Tiếp theo, chương trình sẽ xác định các vùng khả quan có thể là khuôn mặt, cuối cùng các vùng này sẽ được đưa vào một bộ phân loại để tiến hành xác định vùng nào là mặt người.

Ta thấy quá trình huấn luyện, bộ phân lớp phải duyệt qua tất cả các đặc trưng của các mẫu trong tập huấn luyện, việc này tốn rất nhiều thời gian. Tuy nhiên, trong các mẫu đưa vào, không phải mẫu nào cũng thuộc loại khó nhận dạng, có những mẫu nền (background) rất dễ nhận ra (ta gọi đây là những mẫu nền đơn giản). Đối với những mẫu này, ta chỉ cần xét một hay vài đặc trưng đơn giản là có thể nhận diện được chứ không cần xét tất cả các đặc trưng. Nhưng đối với các bộ phân lớp thông thường thì cho dù mẫu cần nhận dạng là dễ hay khó thì nó vẫn sẽ xét tất cả các đặc trưng mà nó rút ra được trong quá trình học. Do đó, chúng tốn thời gian xử lý.

Mô hình phân lớp nhiều tầng được xây dựng chính là nhằm rút ngắn thời gian xử lý, giảm tỉ lệ phát hiện sai cho bộ phân lớp. Mô hình phân tầng gồm nhiều bộ phân lớp. Một mẫu để được phân loại là phù hợp thì nó cần phải đi qua hết tất cả các bộ phân lớp. Các bộ phân lớp sau được huấn luyện bằng những mẫu mà bộ phân lớp trước nó nhận dạng sai, tức là nó sẽ tập trung học từ các mẫu khó hơn, do đó sự kết hợp các bộ phân lớp này lại sẽ làm giảm tỉ lệ phát hiện sai. Với cấu trúc này, những mẫu không phù hợp để nhận diện sẽ bị loại ngay từ những phân lớp đầu tiên, giúp đáp ứng tốt nhất đối với độ phức tạp gia tăng của các mẫu đưa vào, đồng thời giúp rút ngắn thời gian xử lý. Vì vậy cấu trúc phân tầng giải quyết được bài toán thời gian thực thi.

- Khớp ảnh:

Với một ảnh đầu vào, hệ thống sẽ phát hiện khuôn mặt sau đó khởi tạo hình dạng trung bình và đặt hình dạng trung bình đó lên ảnh đầu vào. Áp dụng mô hình AAM luận văn đã trình bày như trên, hệ thống sẽ áp dụng thuật toán AASA (Active Appearance Search Algorithm) khớp mô hình vào ảnh đầu vào sao cho sai số giữa ảnh đầu vào và ảnh khớp được là nhỏ nhất.

Vector biểu thị mức độ khác biệt giữa ảnh đầu vào và ảnh được tổng hợp bằng mô hình diện mạo (*appearance*) được mô tả bằng công thức:

$$\delta I = I_i - I_m.$$

Trong đó,

$I_i$  là vector chỉ mức độ sáng (vector vân) của ảnh đầu vào.



$I_m$  là vector chỉ mức độ sáng (vector vân) của ảnh hiện tại được tổng hợp bằng mô hình diện mạo.

Để xác định vị trí phù hợp cho mô hình và hình ảnh, ta phải giảm thiểu độ lớn của vector riêng  $\Delta = |\delta I|^2$  bằng cách thay đổi các thông số của mô hình c. Vì vậy, ta sẽ tìm mối liên hệ giữa  $\delta I$  và tham số  $\delta c$ . Mối liên hệ tuyến tính giữa  $\delta I$  và lỗi trong mô hình tham số được thể hiện qua công thức:

$$\delta c = A \delta I$$

Để tìm A, ta thực hiện phép hồi quy tuyến tính đa biến trên một mẫu đã biết trước để tạo ra tập huấn luyện bằng cách thay đổi các thông số mô hình cho ảnh huấn luyện như vị trí, tỉ lệ và hướng ( $\delta I$  sẽ được thay bằng  $\delta g$ ).

Các thông số được xét là thông số mô hình c và thông số hình dạng t :

$$t = (s_x, s_y, t_x, t_y)^T$$

Trong đó:

$s_x$  là tham số tỷ lệ kết hợp với tham số góc quay  $s_x = s \cos(\theta) - 1$

$s_y$  là tham số tỷ lệ kết hợp với tham số góc quay  $s_y = s \sin(\theta)$

$t_x$  là tham số dịch chuyển theo chiều x

$t_y$  là tham số dịch chuyển theo chiều y

Với  $c_0$  đã biết trước ( $c_0$  là tham số vân của ảnh đầu vào)

$$c = \delta c + c_0.$$

Sử dụng công thức (\*) ta tính được x (tham số shape) và  $g_m$  (mức độ xám).

Sử dụng phép biến đổi (*warped*) x chuẩn hoá  $g_s$ .

Tính độ lỗi ảnh:  $\delta g = g_s - g_m$

Thực hiện thay thế ngẫu nhiên tham số mô hình trong mỗi ảnh huấn luyện ta sẽ nhận được  $\delta c$  và  $\delta g$ .

Thực hiện phép hồi quy tuyến tính ta được  $\delta c = A \delta g$ .

Giá trị  $\delta c$  được sử dụng trong suốt thuật toán.

Thuật toán AASA như sau:

(1) Tính vector lỗi  $\delta g_0 = g_s - g_m$

(2) Tính độ lỗi  $E_0 = |\delta g_0|^2$

(3) Tính giá trị hiệu chỉnh  $\delta c = A \delta g_0$

(4) Cập nhật  $c_1 = c_0 - k \delta c$  với  $k = 1$

(5) Tính lại vector lỗi  $\delta g_1$

(6) Nếu  $|\delta g_1|^2 < E_0$  thì chấp nhận ước lượng mới  $c_1$

(7) Ngược lại quay lại (4) và chọn  $k = 1.5, 0.5, 0.25, \dots$

Quá trình này được thực hiện cho đến khi giá trị lỗi hội tụ ( $\delta g$  không được thay thế).

- Truy vấn ảnh:

Bên cạnh đó, với sự phát triển của phần cứng và phần mềm, khối lượng ảnh lưu trữ ngày càng lớn. Trong thực tế, tìm kiếm hình ảnh là nhu cầu tất yếu của con người. Ngoài tìm kiếm bằng từ khóa, tìm kiếm hình ảnh tương đồng trong cơ sở dữ liệu hình ảnh đang được sự quan tâm, nghiên cứu, ứng dụng rộng rãi và mang lại nhiều lợi ích cho con người. Nhiều phần mềm ứng dụng truy vấn ảnh đã được sử dụng trong việc tìm kiếm hình ảnh tương đồng. Người sử dụng có thể thay thế các từ khóa ấy bằng các hình ảnh, tìm những hình ảnh liên quan được người dùng cung cấp. Truy vấn ảnh giúp lọc bỏ những hình ảnh sai, hoặc không giống, tiết kiệm thời gian tìm kiếm và độ chính xác cao. Chính vì vậy, truy vấn ảnh được ứng dụng trong rất nhiều lĩnh vực. Trong lĩnh vực an ninh, truy vấn ảnh có thể nhận dạng khuôn mặt hoặc vân tay từ CSDL có sẵn. Bên cạnh đó, truy vấn ảnh còn được ứng dụng rộng rãi trong các lĩnh vực khác: truy tìm logo, truy tìm hình ảnh, y tế, sinh học...

Trước khi tìm kiếm hình ảnh tương đồng, cần thiết phải xây dựng sẵn một CSDL cho một tập hợp các hình ảnh. Việc xây dựng CSDL rất lâu, đây được gọi là quá trình "offline" để phục vụ cho quá trình tìm kiếm hình ảnh.

Để so sánh hai ảnh có tương đồng với nhau hay không, một hàm tương đồng/không tương đồng được xây dựng. Một hệ truy vấn ảnh bao gồm [1]

- + Rút trích đặc trưng ảnh.
- + Xác định mức độ tương đồng giữa hai ảnh.
- + Sắp xếp chỉ mục cho CSDL ảnh.

Một số mô hình truy vấn ảnh:

- + Mô hình rút trích đặc trưng thời gian thực: Mô hình này có ưu điểm đơn giản, dễ cài đặt. Tuy nhiên mô hình có nhược điểm là: Khi đưa ảnh vào truy vấn, hệ thống sẽ rút trích đặc trưng của ảnh truy vấn, tiếp đến là rút trích đặc trưng ảnh của CSDL ảnh, cuối cùng tính độ đo sự tương đồng và xuất kết quả. Do đó, kết quả truy vấn sẽ rất chậm.
- + Mô hình tách riêng CSDL đặc trưng: là cải tiến của mô hình rút trích đặc trưng thời gian thực. Trong mô hình tách riêng CSDL đặc trưng, các đặc trưng của CSDL ảnh được rút trích và lưu vào CSDL các đặc trưng của ảnh. Khi người dùng cần truy vấn ảnh, hệ thống sẽ rút trích đặc trưng của ảnh truy vấn và so sánh với CSDL các đặc trưng của ảnh. Sau đó, tính độ tương đồng và xuất ra kết quả. Mô hình này hoạt động nhanh hơn mô hình thời gian thực. Tuy nhiên, việc cài đặt thuật toán phức tạp hơn mô hình thời gian thực; nếu CSDL ảnh bị thay đổi, CSDL đặc trưng các ảnh sẽ không khớp.
- + Mô hình truy vấn ảnh với cơ sở tri thức: hệ thống truy vấn ảnh với cơ sở tri thức giải quyết được bài toán tối ưu truy vấn, đó là: rút trích đặc trưng ảnh, xác định mức độ tương đồng giữa hai ảnh và sắp xếp chỉ mục cho CSDL ảnh.

GHI CHÚ

## KẾT QUẢ VÀ ĐÓNG GÓP CỦA ĐỀ TÀI

Trên cơ sở nghiên cứu về bài toán nhận dạng mặt người, đề tài đã tìm hiểu mô hình AAM và kết hợp với một hệ thống truy vấn ảnh để nhận dạng khuôn mặt. Khi nghiên cứu về vấn đề này, đề tài không những chỉ nghiên cứu về mô hình AAM và các thuật toán để xử lý mà còn tìm hiểu thêm một số phương pháp nhận dạng khác, thuật toán AASA so khớp khuôn mặt, hệ thống truy vấn ảnh, các độ đo tương tự thông dụng và các cơ sở lý thuyết liên quan đến đề tài.

Bên cạnh đó, trong quá trình xây dựng chương trình ứng dụng, đề tài đã tìm hiểu về thư viện mã nguồn mở OpenCV, qua đó biết cách sử dụng các hàm trong thư viện OpenCV để thực hiện bước phát hiện khuôn mặt đối với một ảnh đầu vào trước khi tiến hành nhận dạng.

Hệ thống thử nghiệm được xây dựng sử dụng bộ thư viện mã nguồn mở OpenCV 2.4.0 trên nền tảng Visual Studio 2010. Ứng dụng nhận dạng được xây dựng trên Matlab 7.8 R2009a. Quá trình phát hiện khuôn mặt trong ảnh được thực hiện nhờ hàm *cvHaarDetectObject* trong bộ thư viện OpenCV. Do tập dữ liệu là ảnh được chụp trực diện nên đạt kết quả tốt.

Về chương trình ứng dụng, khi thử nghiệm trên tập dữ liệu mẫu tốt (tức các ảnh mặt người được chụp bằng máy ảnh kĩ thuật số, cùng điều kiện ánh sáng, ảnh chụp thẳng, trực diện, ảnh chỉ gồm chân dung của một khuôn mặt, khoảng cách từ máy ảnh đến khuôn mặt trong các ảnh bằng nhau, chất lượng ảnh tốt) thì đạt kết quả tốt, thời gian phát hiện nhanh, tuy nhiên đối với các ảnh chụp nghiêng với một góc lớn, ảnh có khuôn mặt bị che khuất thì kết quả nhận dạng chưa tốt hoặc không hội tụ.

Nhận dạng mặt người có nhiều ứng dụng trong cuộc sống, các ứng dụng trong nhận dạng mặt người có thể kể đến như: xác minh tội phạm, ứng dụng trong ATM để lưu trữ khuôn mặt của người rút tiền, quản lý nhân viên, đăng nhập máy tính cá nhân sử dụng khuôn mặt,... Vì vậy, việc xây dựng một ứng dụng nhận dạng là một nhu cầu thiết yếu. Tuy nhiên các phương pháp nhận dạng như: Eigenface, Snakes, ASM, mạng Nơron, SVM (Support Vector Machine) được áp dụng để xây dựng các ứng dụng đều tốn nhiều thời gian, không tận dụng hết những đặc trưng trên khuôn mặt. Do đó các ứng dụng này chưa đáp ứng được yêu cầu. Mô hình AAM đã khắc phục được những nhược điểm trên. So với các phương pháp trước đây thì AAM mô tả được đầy đủ các thông tin khuôn mặt. Còn so với các phương pháp gần đây thì AAM thực hiện nhận dạng nhanh hơn và có khả năng khử nhiễu.

Hầu hết các ứng dụng nhận dạng đều cho kết quả là một ảnh trong CSDL nhưng ảnh này có thể không chính xác (có thể có một vài người có khuôn mặt gần giống nhau). Chính vì vậy đề tài đã kết hợp mô hình AAM cùng với một hệ thống truy vấn ảnh để cho kết quả đầu ra là một tập ảnh, nâng cao độ chính xác cho quá trình nhận dạng. So với các công trình nghiên cứu trước đây, chúng em đã tìm hiểu và xây dựng thành công mô hình truy vấn đơn giản ảnh từ vector đặc trưng rút trích được từ mô hình AAM.

### GHI CHÚ

## KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Tìm hiểu mô hình AAM và ứng dụng mô hình AAM trong nhận dạng và truy vấn ảnh hai chiều là một đề tài rộng tương đối phức tạp. Mặc dù chúng em đã cố gắng xây dựng một hệ thống có khả năng nhận dạng mặt người dựa vào thông tin khuôn mặt xuất hiện trên ảnh tốt nhất, nhưng với những hạn chế về mặt trình độ cũng như thời gian thực hiện, chúng em kết thúc đề tài với nghiên cứu về một mảng nhỏ trong lĩnh vực thị giác máy tính và những định hướng dưới đây sẽ hy vọng ngày càng hoàn thiện khả năng tương tác giữa người với máy tính qua hệ thống nhận dạng và truy vấn khuôn mặt trong tương lai.

Để nâng cao khả năng nhận dạng, hệ thống cần tăng số lượng mẫu trong tập huấn luyện. Tập mẫu tốt nên bao gồm ảnh của nhiều người ở những lứa tuổi khác nhau, số lượng ảnh nam và nữ trong tập huấn luyện nên bằng nhau. Tập huấn luyện nên bổ sung thêm các ảnh của những khuôn mặt của con người ở các châu lục khác nhau. Bên cạnh đó, để kết quả truy vấn tốt nhất, nên áp dụng giải thuật phân cụm để truy vấn ảnh. Với một ảnh đưa vào, hệ thống sẽ phát hiện khuôn mặt, tiến hành nhận dạng, sau đó kiểm tra xem ảnh thuộc phân cụm nào và tiến hành truy vấn ảnh, đưa ra những ảnh gần giống với ảnh đầu vào nhất trong cụm đó. Trong tương lai, hướng phát triển sẽ nghiên cứu đối tượng là ảnh có khuôn mặt với nhiều hướng nhìn khác nhau, khuôn mặt chuyển động trên video hoặc hệ thống được ứng dụng để nhận dạng một sinh viên có phải là sinh viên của khoa CNTT hay không.

### GHI CHÚ

## TÀI LIỆU THAM KHẢO

- [1] Nguyễn Bá Công, Nguyễn Hữu Đức (2005), *Thiết kế hệ thống truy vấn tìm ảnh dựa vào đặc trưng cấp thấp và đặc trưng cấp cao*, Luận văn cử nhân tin học, Đại học Khoa Học Tự Nhiên, Tp HCM.
- [2] Ion Marqués (2010), *Face Recognition Algorithms*, Universidad del País Vasco
- [3] Lindsay I Smith, *A tutorial on Principal Components Analysis*, February 26, 2002.
- [4] T. F. Cootes, G. J. Edwards, and C. J. Taylor (1998), *Active appearance models*, In H. Burkhardt and B. Neumann, editors, 5<sup>th</sup> European Conference on Computer Vision, volume 2, pages 484–498. Springer, Berlin.
- [5] T. F. Cootes, G. J. Edwards, and C. J. Taylor, *A comparative evaluation of active appearance model algorithms*, In P. Lewis and M. Nixon, editors, 9<sup>th</sup> British Machine Vision Conference, volume 2, pages 680–689, Southampton, UK, Sept. 1998. BMVA Press.
- [6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graha (1995), *Active Shape Models - Their Training and Application*, Vol. 61, N O. 1, January, pp. 38-59, Department of Medical Biophysics, University of Manchester, Oxford Road, Manchester M13 9PT, England.
- [7] Laurens van der Maaten, Emile Hendriks, *Capturing Appearance Variation in Active Appearance Models*, Delft University of Technology / UC San Diego.

<b>GHI CHÚ</b>

<b>NGƯỜI THỰC HIỆN ĐỀ CƯƠNG</b>		
<b>SINH VIÊN THỰC HIỆN N 1</b>	Nguyễn Thị Diệu	K35.103.007
<b>SINH VIÊN THỰC HIỆN N 2</b>	Huỳnh Thị Yến Phương	K35.103.059

<b>ĐÁNH GIÁ CỦA GIÁO VIÊN HƯỚNG DẪN</b>		
<b>NGƯỜI HƯỚNG DẪN</b>	Âu Bửu Long	
<b>Ý KIẾN</b>		
	<input type="checkbox"/> đồng ý cho bảo vệ	<input type="checkbox"/> không đồng ý cho bảo vệ