

BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG

DƯƠNG ANH HÙNG

**NGHIÊN CỨU ỨNG DỤNG EIGENFACES
VÀ ĐẶC TRƯNG CỤC BỘ LBP
CHO BÀI TOÁN NHẬN DẠNG MẶT NGƯỜI**

Chuyên ngành: Khoa học máy tính

Mã số: 60.48.01

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

Đà Nẵng - Năm 2013

Công trình được hoàn thành tại
ĐẠI HỌC ĐÀ NẴNG

Người hướng dẫn khoa học: PGS.TS. VÕ TRUNG HÙNG

MSc. T.Hoang Ngan Le (PhD Candidate)

Phản biện 1: TS. NGUYỄN VĂN HIỆU

Phản biện 2: PGS.TS. LÊ MẠNH THẠNH

Luận văn được bảo vệ trước Hội đồng chấm Luận văn tốt nghiệp thạc sĩ Kỹ thuật họp tại Đại học Đà Nẵng vào ngày 18 tháng 5 năm 2013.

Có thể tìm hiểu luận văn tại:

- Trung tâm Thông tin - Học liệu, Đại Học Đà Nẵng

MỞ ĐẦU

1. Lý do chọn đề tài

Trong những năm gần đây, các nghiên cứu và ứng dụng của trí tuệ nhân tạo (*Artificial Intelligence*) và học máy (*Machine Learning*) thu hút được sự quan tâm của rất nhiều nhà khoa học. Một trong những lĩnh vực liên quan tới công nghệ trí thức mà hiện nay được ứng dụng rất nhiều vào trong cuộc sống là nhận dạng mẫu (*Pattern Recognition*). Các hệ thống nhận dạng phổ biến hiện nay như: nhận dạng chữ viết (đánh máy hoặc viết tay), nhận dạng chữ ký, nhận dạng vân tay, nhận dạng tròng mắt (*iris*), nhận dạng mặt người,...

Một trong các bài toán được khá nhiều người quan tâm cho đến thời điểm này là nhận dạng khuôn mặt (*Face Recognition*). Khuôn mặt đóng vai trò quan trọng trong quá trình giao tiếp giữa người với người và cũng mang một lượng thông tin giàu có, chẳng hạn như có thể xác định giới tính, tuổi tác, chủng tộc hay trạng thái cảm xúc,... Nhận dạng khuôn mặt không phải là bài toán mới nhưng nó vẫn là một thách thức lớn vì một bài toán nhận dạng mặt người chứa nhiều các bài toán khác như: phát hiện mặt người (*face detection*), đánh dấu (*facial landmarking*), rút trích đặc trưng (*feature extraction*), gán nhãn, phân lớp (*classification*). Ngoài ra, ảnh khuôn mặt trong thực tế chứa đựng nhiều vấn đề như: độ sáng, độ nhòe/mờ, độ nhiễu, độ phân giải, góc ảnh,...

Trong thực tế, nhận dạng khuôn mặt người (*Face Recognition*) là một hướng nghiên cứu được rất nhiều nhà khoa học lớn quan tâm. Ở các trường đại học hàng đầu về Công Nghệ Thông Tin như Massachusetts Institute of Technology (*MIT*), Carnegie Mellon University (*CMU*), Stanford, Berkeley và các công ty lớn như

Microsoft, Apple, Google, Facebook đều có các trung tâm về sinh trắc học (*Biometrics Center*) và nghiên cứu về nhận dạng khuôn mặt người là một trong những lĩnh vực nghiên cứu chính cho đến nay. Sự kiện gần đây nhất là vụ nổ bom ở Boston, USA, hệ thống nhận dạng mặt người của FBI đã hỗ trợ nhiều trong việc tìm kiếm hung thủ.

Cùng với sự phát triển của lĩnh vực thị giá máy tính (*computer vision*) và học máy (*machine learning*), có rất nhiều các hệ thống nhận dạng khuôn mặt khác nhau đã được phát triển. Kết quả nhận dạng của từng hệ thống cũng rất khác nhau vì tùy thuộc vào dữ liệu thử. Trong khuôn khổ của luận văn này, một phương pháp rất nổi tiếng và kinh điển là EigenFace được nghiên cứu và tìm hiểu sâu. Mặc dù EigenFace không phải là phương pháp mới nhất, tốt nhất nhưng nhờ vào tính năng hiệu quả và khả năng bền vững, độc lập dữ liệu, cho đến thời điểm này thì EigenFace được chọn làm baseline để so sánh, đánh giá một hệ thống nhận dạng. Ngoài ra, giải pháp dùng eigen (bao gồm eigenvalue và eigenvector) để tìm subspace Principal Component Analysis (*PCA*) đóng vai trò rất quan trọng trong lĩnh vực học máy. Principal Component Analysis (*PCA*) là nền tảng của một loại các bộ gán nhãn (*classifier*) sau này như Kernel Principal Component Analysis (*KPCA*), Linear Discriminant Analysis (*LDA*), Locality Preserving Projections (*LPP*), Kernel Discriminant Analysis (*KDA*), Suport Vector Machine (*SVM*),... Ngoài ra, luận văn chọn đặc trưng Local Binary Pattern (*LBP*) là một đặc trưng rất phổ biến trong các bài toán liên quan đến ảnh khuôn mặt nói chung và nhận dạng khuôn mặt nói riêng. Nhờ vào khả năng bất biến với độ sáng (đây là một trong những thách thức lớn nhất của các bài toán liên

quan đến ảnh mặt người) nên Local Binary Pattern (*LBP*) được chọn trong hệ thống nhận dạng khuôn mặt.

2. Mục đích của đề tài

Nghiên cứu cách thức xây dựng một hệ thống nhận dạng nói chung và nhận dạng khuôn mặt nói riêng.

Nghiên cứu về đặc trưng rất thông dụng trong các bài toán liên quan đến ảnh khuôn mặt nói chung và nhận dạng khuôn mặt nói riêng: Local Binary Pattern (*LBP*).

Tìm hiểu kỹ về lý thuyết toán và lập luận của bài toán rất nổi tiếng, kinh điển EigenFace: là hệ thống nhận dạng mặt người dựa trên subspace Principal Component Analysis (*PCA*).

Bên cạnh nghiên cứu sâu về đặc trưng *LBP* và subspace *PCA*, luận văn còn mở rộng tìm hiểu và giới thiệu một số các đặc trưng cũng như các bộ phân lớp phổ biến khác trong bài toán nhận dạng.

3. Ý nghĩa khoa học đạt được

Đề tài góp phần giới thiệu các bước tiến hành để xây dựng hệ thống nhận dạng nói chung và nhận dạng khuôn mặt nói riêng.

Nghiên cứu và kiểm chứng lý thuyết toán về bài toán eigen (*eigenvalues*, *eigenvector*) và các tính chất toán học của *PCA* trong khuôn khổ ứng dụng nhận dạng khuôn mặt.

Cài đặt thử nghiệm và đánh giá bằng thực nghiệm các kỹ thuật trong rút trích đặc trưng cục bộ Local Binary Pattern (*LBP*) và phân loại đối tượng dựa trên subspace Principle Component Analysis (*PCA*).

Những nghiên cứu này sẽ hỗ trợ cho việc áp dụng *PCA* để giải quyết nhiều bài toán khác đặt ra trong thực tế như: giảm số chiều

(*dimensionality reduction*), hồi phục ảnh (*reconstruction*), phân đoạn ảnh (*segmentation*),...

4. Ý nghĩa thực tiễn đạt được

Luận văn nghiên cứu một kỹ thuật rất kinh điển, nổi tiếng và quan trọng trong lĩnh vực nhận dạng khuôn mặt là EigenFace.

Luận văn đã xây dựng thành công hệ thống nhận dạng khuôn mặt trên đặc trưng cục bộ Local Binary Pattern (*LBP*) và subspace Principle Component Analysis (*PCA*).

Bên cạnh đó, đề tài này góp phần xây dựng chương trình nhận dạng khuôn mặt người để áp dụng cho nhiều ứng dụng thực tiễn, nhằm đáp ứng cho các yêu cầu như nhận dạng, bảo mật ngày càng cao.

Ngoài ra, đề tài đã xây dựng cơ sở dữ liệu trong thực tế trên khuôn mặt người Việt Nam, đóng góp thêm vào bộ cơ sở dữ liệu chung của cả nước.

5. Mục tiêu và nhiệm vụ

Với mục đích như trên, mục tiêu và nhiệm vụ của luận văn được xác định như sau:

Tìm hiểu mô hình bài toán và các bước thực hiện để xây dựng một hệ thống nhận dạng khuôn mặt người.

Tìm hiểu, nắm các công nghệ đã và đang được áp dụng cho từng bước có trong hệ thống.

Hiểu rõ lý thuyết toán học và tính chất đặc trưng của những kỹ thuật sẽ được sử dụng để giải quyết bài toán trong khuôn khổ luận văn.

Tìm hiểu và sử dụng các công cụ hữu ích nhằm phục vụ cho bài toán nhận dạng mặt người cũng như trong quá trình xây dựng hệ thống: MASM, OpenCV, Matlab.

Ngoài cơ sở dữ liệu chuẩn MBGC công bố tại địa chỉ trang web <http://www.nist.gov/itl/iad/ig/mbgc.cfm>, tôi sẽ tiến hành xây dựng dữ liệu thực tế trên khuôn mặt người Việt Nam nhằm phục vụ cho chương trình thực nghiệm.

6. Đối tượng và phạm vi nghiên cứu

Các phương pháp, giải thuật phục vụ cho việc phát hiện và nhận dạng khuôn mặt người trên ảnh.

Phần mềm Matlab, đặc biệt là những thư viện phục vụ cho việc xử lý ảnh, phát hiện và nhận dạng mặt người.

Bộ cơ sở dữ liệu chuẩn Multiple Biometric Grand Challenge (MBGC) và bộ cơ sở dữ liệu do sinh viên tự thu thập.

Phạm vi nghiên cứu:

Việc xử lý và nhận dạng ảnh khuôn mặt thỏa mãn các điều kiện sau:

Ánh sáng đều, không có chiều sáng (*no illumination*), không có ánh sáng mạnh (*no strong lighting*);

Góc ảnh: trực diện (*frontal*) hoặc gần như trực diện;

Không bị che khuất (*no occlusion*);

Ảnh chất lượng cao (*high quality images*).

7. Kết quả dự kiến

Chương trình (ở mức thử nghiệm) nhận dạng mặt người trên dữ liệu chuẩn Multiple Biometric Grand Challenge Database (MBGC) đạt trên 90%. Thử nghiệm chương trình trên dữ liệu thực tế (*Vietnamese database*) và khả năng nhận dạng đạt được 70% - 80%.

8. **Bố cục luận văn**

Bố cục luận văn gồm có 3 chương chính:

Chương 1: Cơ sở lý thuyết

Chương này giới thiệu tổng quát các bước cần thực hiện trong bài toán nhận dạng. Bên cạnh việc giới thiệu các kỹ thuật liên quan phổ biến, đối tượng và phạm vi nghiên cứu cũng được định nghĩa, trình bày chi tiết trong chương này.

Chương 2: Lựa chọn giải pháp và kỹ thuật

Tập trung trình bày về đặc trưng cục bộ Local Binary Pattern (*LBP*) và Principal Component Analysis (*PCA*). Các lý thuyết toán, lập luận được phân tích và trình bày rõ trong từng bước thực hiện. Đặc điểm và tính chất của hai kỹ thuật cũng được phân tích kỹ trong chương này.

Chương 3: Hệ thống nhận dạng khuôn mặt dựa trên LBP và PCA

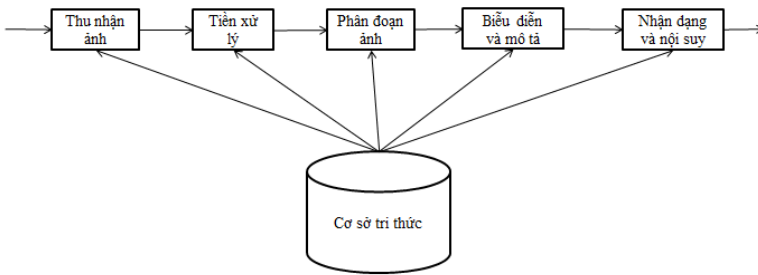
Xây dựng được demo chương trình nhận dạng khuôn mặt, trong phạm vi của luận văn thì bộ dữ liệu chuẩn Multiple Biometric Grand Challenge Database (*MBGC*) và bộ dữ liệu do chính sinh viên tự thu thập được sử dụng. Ngoài ra thì khả năng và hiệu quả của hệ thống cũng được so sánh, đánh giá với các phương pháp khác.

CHƯƠNG 1. CƠ SỞ LÝ THUYẾT

1.1. XỬ LÝ ẢNH

1.1.1 Giới thiệu

Các bước cơ bản trong xử lý ảnh



Hình 1.1. Các bước cơ bản trong xử lý ảnh

Phản thu nhận ảnh: Ảnh có thể nhận qua camera. Có thể là màu hoặc đen trắng với độ phân giải khác nhau.

Tiền xử lý: lọc nhiễu, nâng cao chất lượng ảnh.

Phân vùng ảnh hay phân đoạn: tách ảnh thành các vùng hoặc đối tượng quan tâm.

Biểu diễn ảnh: Việc chọn các tính chất để thể hiện ảnh gọi là trích chọn đặc trưng (*feature extraction*).

Nhận dạng và nội suy ảnh: Nhận dạng ảnh là quá trình xác định ảnh. Quá trình này thường thu được bằng cách so sánh với mẫu chuẩn đã được học (hoặc lưu) từ trước. Nội suy là phán đoán theo ý nghĩa trên cơ sở nhận dạng.

Cơ sở tri thức: Nhằm giúp quá trình xử lý và phân tích ảnh theo cách làm của con người.

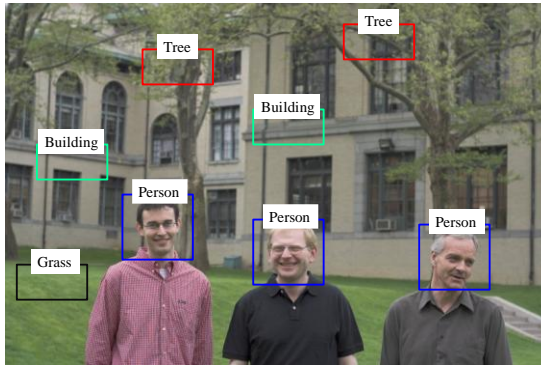
1.1.2. Một số kỹ thuật xử lý ảnh liên quan đến bài toán nhận dạng khuôn mặt

1.2. BÀI TOÁN NHẬN DẠNG MẪU

1.2.1. Giới thiệu và định nghĩa bài toán

1.2.2. Các vấn đề liên quan đến nhận dạng

1.2.3. Một số lĩnh vực ứng dụng



Hình 1.1. Ảnh nhận dạng đối tượng trong tự nhiên

1.3. DÒ TÌM KHUÔN MẶT

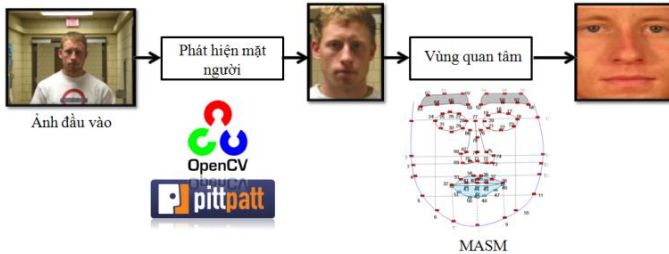
1.3.1. Giới thiệu và định nghĩa bài toán

1.3.2. Một số hướng tiếp cận trong dò tìm khuôn mặt

1.3.3. Phương pháp dò tìm dùng trong khuôn khổ luận văn

Trong khuôn khổ của luận văn, quá trình dò tìm khuôn mặt và các điểm mốc khuôn mặt được thực hiện bằng công cụ *Modified Active Shape Model (MASM)*. Công cụ này do giáo viên hướng dẫn, hiện đang công tác ở Biometrics Lab, Carnegie Mellon University (CMU), Pittsburgh, PA, USA cung cấp. Bộ công cụ MASM là một

phần trong dự án nhận dạng khuôn mặt được cục FBI, USA đầu tư phát triển từ năm 2010.



Hình 1.2. Tiến trình dò tìm, định vị khuôn mặt

1.4. NHẬN DẠNG KHUÔN MẶT

1.4.1. Giới thiệu và định nghĩa bài toán

1.4.2. Một số khó khăn, thử thách trong nhận dạng khuôn mặt

Một số khó khăn, thử thách trong nhận dạng khuôn mặt: Hướng (*pose*), sự có mặt của các chi tiết không phải là đặc trưng riêng của khuôn mặt người như: râu quai nón, mắt kính,các nét mặt (*facial expression*), mặt người bị che khuất, điều kiện ảnh, đặc biệt là về độ sáng và chất lượng ảnh.

Để hạn chế các khó khăn trên, bài toán được định nghĩa và tiến hành trong các điều kiện sau: Ánh sáng đều, không có chiếu sáng (*no illumination*), không có ánh sáng mạnh (*no strong lighting*), góc ảnh trực diện (*frontal*) hoặc gần như trực diện, không bị che khuất (*no occlusion*), ảnh chất lượng cao

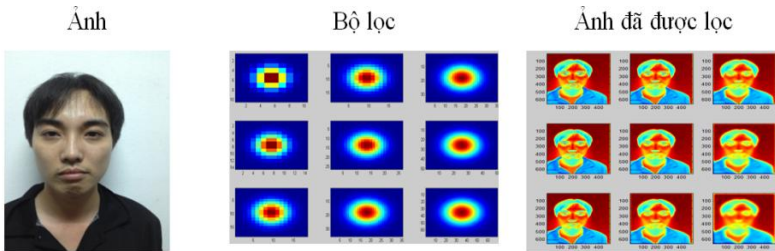


**Hình 1.3. Một số hình ảnh trong database MBGC (dòng 1)
và ảnh do sinh viên tự thu thập (dòng 2)**

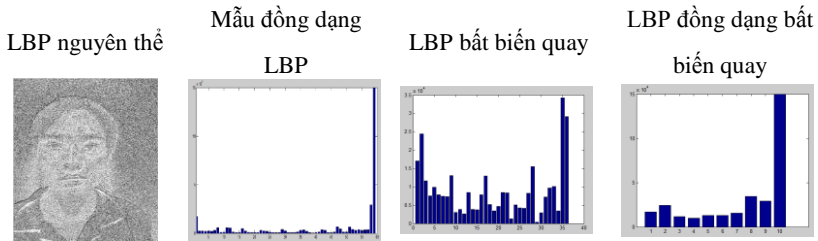
1.4.3. Đặc trưng thông dụng trong nhận dạng khuôn mặt

Một số đặc trưng thông dụng trong nhận dạng (tổng quát) được nghiên cứu và trình bày như sau:

Đặc trưng Gaussian với các trị sigma (σ) và tỷ lệ khác nhau.



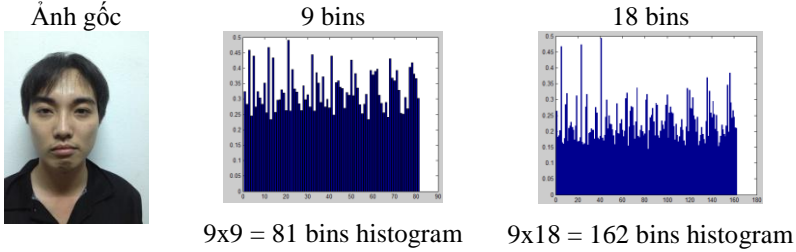
**Hình 1.4. Ảnh minh họa sau khi lọc với Gaussian
Đặc trưng Local Binary Pattern (LBP) và các biến thể**



Hình 1.5. Ví dụ về biến thể của LBP

Histogram of Oriented Gradients (HOG)

Thay vì tính toán gradient (tích vô hướng) trên 2 chiều là x và y , HOG được sử dụng để tính gradient cho tần số cao (*high frequency*) trên các chiều khác nhau.



Hình 1.6. Đặc trưng HOG tương ứng với kích thước khối ảnh 3x3 với 9 hướng (cột 2) hoặc 18 hướng (cột 3)

1.4.4. Bộ phân lớp

Bộ phân lớp dùng trong bài toán nhận dạng cũng rất đa dạng, tùy thuộc vào dữ liệu. Một số bộ phân lớp phổ biến trong nhóm phân lớp tuyến tính được trình bày như sau:

Principal Component Analysis (PCA): Mục tiêu của PCA là tìm ra một không gian đặc trưng mới với số chiều giảm hơn nhiều so với số chiều ban đầu nhưng vẫn đảm bảo đặc tính của tập dữ liệu.

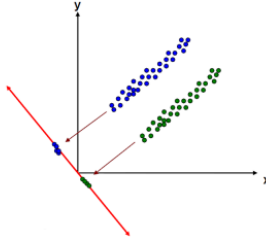


Biểu đồ mô tả các điểm khác nhau trên trục tọa độ.

Góc này mô tả giá trị lớn nhất của các điểm trên trục mới được vẽ lại.

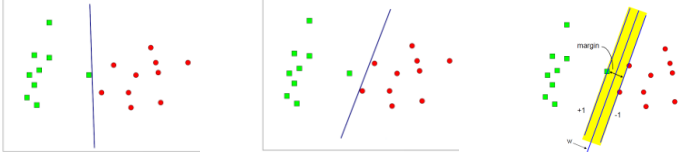
Hình 1.7. Minh họa cho phương pháp tìm trục chính của PCA

Linear Discriminant Analysis (LDA): có thể đưa 2 điểm dữ liệu trên cùng 1 lớp nằm gần về với nhau. Các điểm dữ liệu không cùng nằm về 1 lớp thì chúng sẽ cách xa nhau hơn.



Hình 1.8. Phương pháp biểu diễn dữ liệu LDA

Support Vector Machine (SVM): tìm và lựa chọn ranh giới giữa hai thể loại sao cho khoảng cách từ các ví dụ huấn luyện (*support samples*) tới ranh giới là xa nhất có thể.



Hình 1.9. Ví dụ về phân lớp dữ liệu bằng phương pháp thông thường NN (2 cột đầu) và bằng SVM (cột cuối)

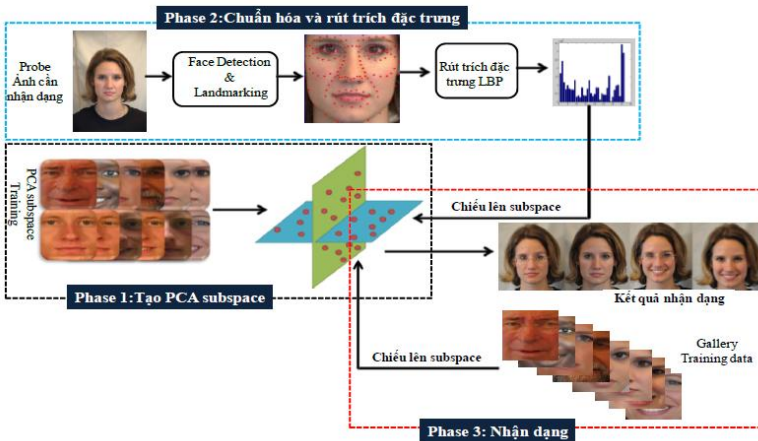
CHƯƠNG 2. LỰA CHỌN GIẢI PHÁP VÀ KỸ THUẬT

2.1. GIỚI THIỆU HƯỚNG TIẾP CẬN

2.1.1. Giới thiệu

2.1.2. Hướng tiếp cận trong luận văn

Hệ thống nhận dạng đề xuất trong khuôn khổ luận văn bao gồm 3 bước và được tóm tắt như trong hình vẽ sau:



Hình 2.1. Hướng tiếp cận nhận dạng của luận văn

Bước 1: Tạo PCA subspace dựa trên bài toán eigen

Bước 2: Đầu vào của bước này là ảnh có chứa mặt người cần nhận dạng. Đầu ra của bước này là đặc trưng dùng trong quá trình nhận dạng. Bước này được thực hiện thông qua nhiều giai đoạn như sau: Tiền xử lý, phát hiện khuôn mặt, landmarking, cắt vùng khuôn mặt, rút trích đặc trưng.

Bước 3: Trong bước này, đặc trưng của ảnh cần nhận dạng và đặc trưng của dữ liệu học mẫu được chiếu lên PCA subspace. Dựa vào hệ số chiếu, ta sẽ có được kết quả nhận dạng

2.2 ĐẶC TRƯNG LBP

2.2.1 Giới thiệu

2.2.2. LBP cơ bản

Thông tin LBP của pixel tại trung tâm của mỗi khối ảnh sẽ được tính dựa trên thông tin của các pixel lân cận. Có thể tóm tắt các bước tiến hành như sau:

Bước 1: Xác định bán kính làm việc.

Bước 2: Tính giá trị LBP cho pixel ở trung tâm (x_c, y_c) khối ảnh dựa trên thông tin của các pixel lân

$$LBP_{P,R}(x_c, y_c) = \sum_{p=1}^P s(g_p - g_c) 2^p$$

cận:

Trong đó, (g_p) là giá trị grayscale của các pixel lân cận, (g_c) là giá trị grayscale của các trung tâm và (s) là hàm nhị phân được xác định như sau: $s(z) = 1$ nếu giá trị $z \geq 0$.

Ví dụ:

47	51	65
62	70	70
80	83	78

-23	-19	-5
-8	*	0
10	13	8

0	0	0
0	*	1
1	1	1

$$\begin{aligned}
 & 1 * 2^0 + 1 * 2^1 + 1 * 2^2 \\
 & + 1 * 2^3 + 0 * 2^4 + \\
 & 0 * 2^5 + 0 * 2^6 + 0 * 2^7 \\
 & = 15
 \end{aligned}$$

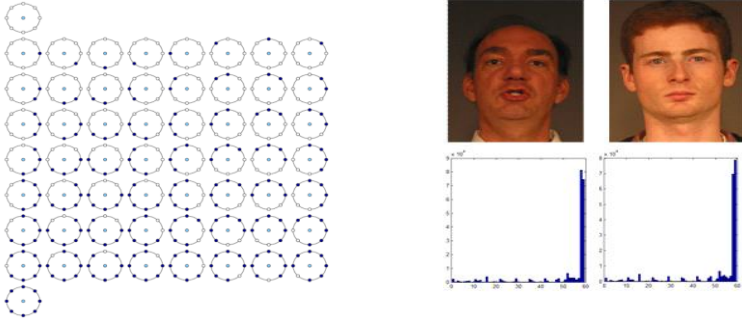
2.2.3. Các biến thể của LBP

LBP đồng dạng

Một mẫu nhị phân được gọi là đồng dạng khi xét chuỗi bit xoay vòng thì có nhiều nhất là 2 lần thay đổi (*transitions*) từ giá trị bit 0 sang 1 hoặc từ giá trị bit 1 sang 0. Ví dụ: 00000000 có 0 transitions, 01110000 có 2 transitions, 11001111 có 2 transitions nên đây là uniform LBP. 11001001 có 4 transitions, 01010011 có 6 transitions nên không phải là uniform LBP.

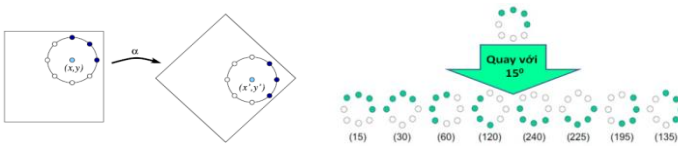
Dựa trên định nghĩa này, bảng ánh xạ cho bán kính làm việc P-neighbours sẽ có $P(P-1) + 3$ nhãn. Có nghĩa là có 59 nhãn trong trường hợp làm việc với 8-neighbour. Hình vẽ sau đây thể hiện 59 nhãn (mẫu) và minh họa về histogram của đặc trưng LBP đồng dạng.

Bảng 1.1. Bảng thống kê các mẫu của uniform LBP



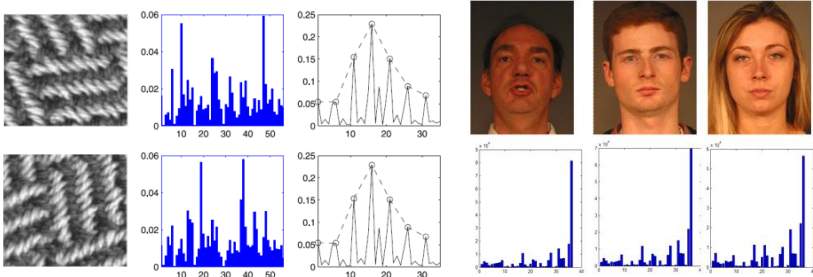
LBP bất biến với phép quay

Giả sử $I_\alpha(x, y)$ là ảnh quay góc (α) của ảnh $I(x, y)$. Với phép quay này điểm ảnh (x, y) sẽ nằm tại vị trí (x', y') như hình vẽ sau đây (hình trái). Trong ví dụ này (hình phải): tất cả 8 mẫu LBP bên dưới được ánh xạ về mẫu LBP đầu tiên vì mẫu đầu tiên cho giá trị nhỏ nhất.



Hình 2.8. Minh họa về các trường hợp của LBP sau khi quay với góc 15 độ

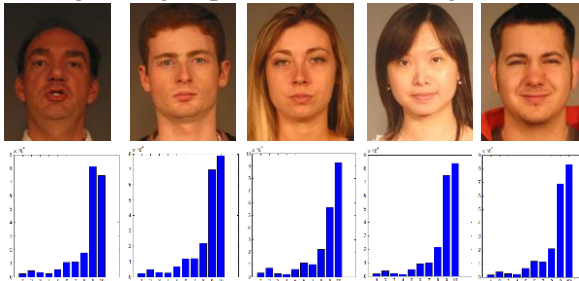
Khả năng bất biến với phép quay của đặc trưng này được minh họa qua ví dụ sau (hình 2.3 trái). Rút trích đặc trưng này trên ảnh khuôn mặt được thể hiện trong hình 2.3 phải



Hình 2.2. Minh họa đặc trưng LBP bất biến với phép quay trên khuôn mặt

LBP đồng dạng có khả năng bất biến với phép quay

Kết hợp của mẫu LBP đồng dạng và LBP bất biến với phép quay có thể tạo nên một dạng biến thể khác của LBP (*uniform rotation – invariant LBP*). Dựa trên định nghĩa này, bảng ánh xạ cho bán kính làm việc P-neighbors sẽ có P + 2 nhãn (*label*). Có nghĩa là có 10 nhãn trong trường hợp làm việc với 8-neighbour



Hình 2.3. Minh họa đặc trưng LBP đồng dạng và bất biến với phép quay

2.2.4. Tính chất của LBP trong nhận dạng

2.3. PHƯƠNG PHÁP PCA

2.3.1. Giới thiệu về PCA

Phương pháp phân tích thành phần chính (PCA) là một công cụ rất phổ biến trong việc nhận dạng, được biết đến như là Eigenfaces trong ứng dụng nhận dạng mặt người sử dụng PCA. Cụm từ Eigenfaces được hình thành dựa trên ứng dụng nhận dạng mặt người và bài toán PCA giải bài phương pháp phân tích eigenvalue. PCA được xem như là phương pháp chuẩn dùng để đánh giá mức độ hiệu quả của các phương pháp khác (baseline benchmark).

Có hai phương pháp cơ bản dùng để giải quyết bài toán PCA là: phân tích eigenvalue và Singular Value Decomposition (SVD). Trong khi SVD giải quyết bài toán PCA dựa trên ma trận dữ liệu (*data matrix*) thì phương pháp phân tích eigenvalue làm việc trên hiệp phương sai của dữ liệu. Trong khuôn khổ luận văn, bài toán PCA được giải quyết dựa trên phương pháp phân tích eigenvalue.

2.3.2. Các khái niệm trong PCA

Trị trung bình (mean) của 1 vector là một giá trị đơn (scalar) và được tính theo công thức: $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$.

Độ lệch chuẩn (standard deviation): cho biết độ khác biệt hay độ lệch của các phần tử trong dữ liệu so với trị trung bình. Cho trước dữ liệu X và trị trung bình của X là μ , độ lệch chuẩn của X được tính như sau: $\sigma = \sqrt{E[X - \mu]^2}$.

Ma trận hiệp phương sai (covariance matrix): Giả sử cho trước một ma trận $X = [X_1 \ X_2 \ \dots X_N]$, trong đó X_i là vector. Ma trận hiệp phương sai, ký hiệu là Σ , được định nghĩa như sau: $\Sigma = \text{cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)]$.

2.3.3. Bài toán Eigen

Trong ngữ cảnh của bài toán eigen, nếu tồn tại vector x thỏa điều kiện song song với Ax thì x được gọi là eigenvectors của ma trận A . Tính song song được biểu diễn như sau: $Ax = \lambda x$, trong đó λ là 1 trị đơn (scalar) và được xem là eigenvalue của ma trận A . Các tính eigenvectors và eigenvalues được tổng quát như sau:

Tính định thức (determinant) của $A - \lambda I$.

Tìm eigenvalues bằng cách giải quyết $\det(A - \lambda I) = 0$, trong đó λ là nghiệm.

Tương ứng với mỗi giá trị eigenvalue λ_i , ta tính được eigenvector bằng cách giải phương trình $Ax = \lambda x$.

2.3.4. Đặt vấn đề và giải pháp của PCA

Đặt vấn đề

Một trong những câu hỏi cơ bản trong bài toán PCA là: cho trước 1 tập dữ liệu, tìm trục chiều mà có thể phân biệt được dữ liệu tốt nhất, tức là biến thể (variation) lớn nhất. Xem xét ví dụ sau đây: Mục đích của bài toán PCA là đi tìm trục chiều tốt nhất ở hình vẽ bên phải.



Hình 2.4. Ví dụ về phép chiếu từ không gian 2D xuống 1D

Giải quyết vấn đề: Gọi ω là trục chiều cần tìm. Phép chiếu của dữ liệu x trên trục chiều này là $\omega^T x$. Mục đích của PCA là tìm ω sao cho biến thiên của $\omega^T x$ lớn nhất. Bài toán được biểu diễn như sau:

$\arg \max_{(\omega)} \{ \text{var}(\omega^T x) \}.$

Giá trị biến thiên của phép chiếu dữ liệu x lên trục ω được thể hiện như sau:

$$\begin{aligned}\text{Var}(w^T x) &= E(w^T x - E(w^T x))^2 \\ &= w^T E(x - E(x))(x - E(x))^T w \quad (**).\end{aligned}$$

Trong đó $E[x]$ là trị trung bình của x . Đặt $\mathring{a} = E(x - E(x))(x - E(x))^T$ là ma trận hiệp phương sai (covariance matrix). Công thức (**) tương đương $\text{Var}(w^T x) = w^T \mathring{a} w$, với $\mathring{a} = E(x - E(x))(x - E(x))^T$ và $w^T w = 1$. Lưu ý là trong ngữ cảnh của bài toán, w được chuẩn hóa về vector đơn vị (*unit norm*), có nghĩa là $w^T w = 1$.

Mục tiêu của bài toán PCA là đi tìm trục w sao cho giá trị biến thiên của phép chiếu $\omega^T x$ lớn nhất (tức là đi tìm $\max(\text{Var}(w^T x))$). Như vậy ta có thể phát biểu bài toán như sau:

$$\arg \max_{(w)} \{ \text{var}(w^T x) \}, \text{ s.t. } w^T w = 1 \quad (***)$$

Giải pháp đơn giản nhất để giải quyết (***) là sử dụng phương pháp tối ưu Lagrangian (*Lagrangian optimization*).

Từ (***) ta đưa về dạng hàm số Lagrangian: $L(w, l) = w^T \mathring{a} w - l(w^T w - 1)$, trong đó l là hệ số Lagrangian. Lấy đạo hàm theo w , đặt nó có giá trị bằng không (0), ta có công thức sau: $\nabla_w L(w, l) = 0$

$$\Leftrightarrow 2 \sum \omega - 2 \lambda \omega = 0 \Leftrightarrow \sum \omega - \lambda \omega = 0 \Leftrightarrow \sum \omega = \lambda \omega$$

Đây chính là bài toán eigen, trong đó l là eigenvalues, PCA subspace w là eigenvector của ma trận hiệp phương sai $\mathring{a} = E(x - E(x))(x - E(x))^T$.

2.3.5. Dùng PCA để xây dựng subspace

Cho trước tập huấn luyện (training data) gồm N ảnh (x_1, x_2, \dots, x_N) mỗi ảnh được biểu diễn dưới dạng vector cột có M chiều $(x_i \in \mathbb{R}^M)$. Sắp xếp các ảnh lại ta được 1 ma trận X. PCA subspace được xây dựng qua các bước như sau:

Bước 1: Chuẩn hóa dữ liệu: $\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$ và

$$X = (X - \bar{X}) = \begin{bmatrix} x_1 - \bar{X} & x_2 - \bar{X} & \dots & x_N - \bar{X} \end{bmatrix}$$

Bước 2: Tính ma trận hiệp phương sai của X: $\Sigma = X^T X$.

Bước 3: Tìm nghiệm của phương trình $\Sigma w = \lambda w$ và được giải bằng phương pháp eigen. Ma trận hiệp phương sai Σ kích thước $N \times N$ nên sẽ có N nghiệm khi tìm lời giải cho bài toán eigen. Tập hợp của các eigenvector sẽ hình thành 1 PCA subspace.

2.3.6. Ứng dụng PCA trong bài toán nhận dạng mặt người

Bước 1: Thực hiện xây dựng không gian con PCA (PCA subspace): $w = [w_1 w_2 \dots w_m]^T \in \mathbb{R}^{M \times m}$ từ tập ảnh huấn luyện X.

Bước 2: Thực hiện nhận dạng.

Đầu vào: Dữ liệu huấn luyện Y (khác với dữ liệu huấn luyện X) và ảnh cần kiểm chứng y.

Tiến hành: Chiếu tập huấn luyện Y lên PCA subspace ω , ta sẽ có S hệ số $\alpha_1, \alpha_2, \dots, \alpha_s$, trong đó $a_1 = w^T Y_1, a_2 = w^T Y_2, \dots, a_{n_1} = w^T Y_{n_1}$ và S là tổng số các mẫu có trong tập huấn luyện Y.

Chiếu dữ liệu được kiểm tra y lên PCA subspace ω : $a_y = w^T y$.

Thực hiện so sánh hệ số (coefficients) giữa a_y và $a_1 a_2 \dots a_s$.

CHƯƠNG 3. HỆ THỐNG NHẬN DẠNG MẶT NGƯỜI DỰA TRÊN LBP VÀ PCA

3.1. MÔ TẢ ỨNG DỤNG

3.1.1. Công cụ và ngôn ngữ phát triển

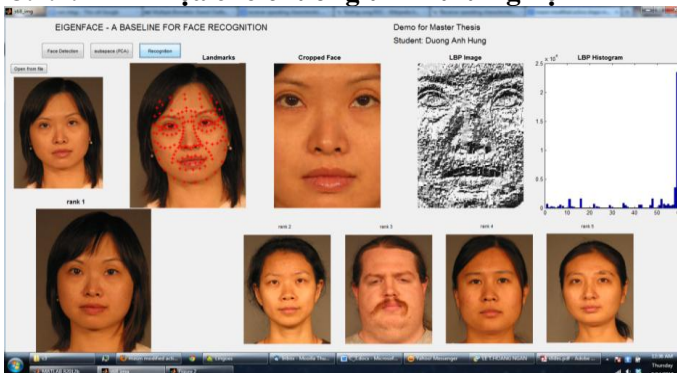
Công cụ: MASM để dò tìm khuôn mặt và xác định điểm khuôn mặt. Là một công cụ hỗ trợ đắc lực cho thể thống nhận dạng được FBI đầu tư (do giáo viên hướng dẫn cung cấp).

Ngôn ngữ: Matlab 2012b (32 bit) chạy trên hệ điều hành Windows 7.



Hình 3.1. MASM cho việc phát hiện các đặc điểm trên khuôn mặt

3.1.2. Minh họa cho chương trình thử nghiệm



Hình 3.2. Giao diện của chương trình thử nghiệm (rút trích đặc trưng và nhận dạng)



Hình 3.3. Minh họa cho PCA subspace dùng trong chương trình

3.2. DỮ LIỆU VÀ ĐỘ ĐO

3.2.1. Dữ liệu chuẩn MGBC

Dữ liệu:

Dữ liệu tạo PCA Subspace: 10.000 ảnh từ MBGC

Dữ liệu học mẫu (train) và thử nghiệm (test) được tiến hành trên hai bộ:

MBGC: 860 ảnh học mẫu và 213 ảnh thử nghiệm của 43 đối tượng khác nhau.

3.2.2. Dữ liệu thực tế

Do sinh viên thu thập: 150 ảnh học mẫu và 53 ảnh thử nghiệm của 13 đối tượng khác nhau.

3.2.3 Độ đo:

Độ chính xác (*Accuracy*): được xác định dựa trên số lượng ảnh nhận dạng đúng so với tổng số lượng ảnh có trong database. Thông thường khi nói đến độ chính xác, ta thường nói kết quả nhận dạng đầu tiên RANK-1.

ROC curve - Receiver Operating Characteristic: là một bảng tương quan (*mapping*) giữa tỷ lệ dương tính giả (*false positive rate - FPR*) và tỷ lệ dương tính thật (*true positive rate - TPR*).

3.3. KẾT QUẢ THỰC HIỆN

3.3.1. Kết quả trên dữ liệu chuẩn

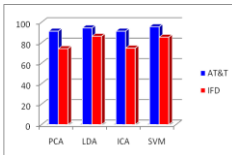
3.3.2. Kết quả trên dữ liệu thực tế

3.3.3. Một số sai sót trong quá trình nhận dạng

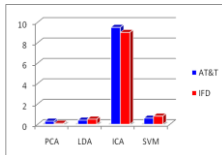
3.3.4. So sánh đánh giá

Như đã phân tích và trình bày ngay từ Chương 1, PCA là thật toán tốt, chạy ổn định trên mọi kiểu dữ liệu, chính là vì PCA được chọn làm baseline.

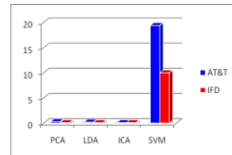
Độ chính xác (%)



Thời gian học mẫu



Thời gian nhận dạng



KẾT LUẬN

Mục đích của đề tài là tìm hiểu và đánh giá một số thuật toán về xử lý ảnh, nhận dạng để từ đó xây dựng một hệ thống nhận dạng khuôn mặt hoàn chỉnh nhằm phục vụ cho ứng dụng trong thực tế. Nội dung đề tài đã thực hiện đầy đủ việc tìm hiểu và nghiên cứu một số thuật toán tiêu biểu tại từng giai đoạn xử lý của quá trình nhận dạng khuôn mặt. Trọng tâm nghiên cứu là các thuật toán có khả năng thực hiện một cách chính xác nhất cho quá trình nhận dạng. Cụ thể là tìm hiểu và trình bày hai thuật toán quan trọng nhất là nhị phân cục bộ LBP nhằm rút trích đặc trưng cho khuôn mặt và bộ phân lớp PCA cho quá trình nhận dạng.

Về thực tế thì luận văn đã xây dựng được một công cụ nhận dạng khuôn mặt hoàn chỉnh xác, có thể đem ứng dụng trực tiếp vào thực tế. Hệ thống nhận dạng đã hoạt động khá chính, cho kết quả đạt được là hơn 90% với bộ thư viện chuẩn MBGC và đạt yêu cầu đề ra là từ 70-80% cho bộ dữ liệu mà sinh viên tự mình thu thập được.

Mặc dù đã có nhiều cố gắng, nhưng do thời gian nghiên cứu chưa đủ, kinh nghiệm nghiên cứu chưa có nhiều, vì vậy luận văn này không tránh khỏi một số sai sót, hạn chế nhất định như: việc thu thập dữ liệu còn ít, ảnh chụp chưa được tốt, demo chương trình còn mang tính minh họa.

Tuy nhiên, xử lý ảnh hay nhận dạng khuôn mặt vẫn còn là một hướng nghiên cứu khá mới mẻ. Vì vậy, trong tương lai, đề tài này hứa hẹn có thể được phát triển thêm và triển khai rộng rãi trong nhiều ứng dụng nhận dạng các cá nhân, gương mặt khác nhau, phục vụ cho các nhu cầu mới phát sinh, đáp ứng được yêu cầu ngày càng tăng về tính bảo mật.