**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**🙠🕮🙢**

**ĐỀ ÁN MÔN CÁC HỆ CƠ SỞ TRI THỨC**

**TÊN ĐỀ TÀI**

**NHẬN DẠNG MẶT NGƯỜI SỬ DỤNG BIẾN ĐỔI GABOR WAVELET KẾT HỢP VỚI MẠNG NEURAL**

GVHD: TS. Lê Hoàng Thái

Nhóm thực hiện:

Nhóm 08 – Ngành Khoa Học Máy Tính – Khóa 22

1. Nguyễn Khắc Văn
2. Hoàng Văn Hiếu
3. Nguyễn Quốc Trung
4. Trần Đăng Khoa

*Tp. Hồ Chí Minh, tháng 11 năm 2012*

Mục lục

[Lời mở đầu 4](#_Toc339575625)

[Phần I. Tổng Quan Về Nhận Dạng Khuôn Mặt 5](#_Toc339575626)

[1.1 Các khái niệm liên quan đến nhận dạng khuôn mặt 5](#_Toc339575627)

[1.2 Bài toán nhận dạng khuôn mặt 5](#_Toc339575628)

[1.3 Mô hình bài toán nhận dạng mặt người 6](#_Toc339575629)

[1.4 Quy trình nhận dạng khuôn mặt 7](#_Toc339575630)

[1.5 Các hướng tiếp cận chính trong lĩnh vực nhận dạng khuôn mặt 8](#_Toc339575631)

[Phần II. Tổng quan về Gabor Wavelet và MLP Neural Network 11](#_Toc339575632)

[1.1. Tổng quan về Gabor Wavelet 11](#_Toc339575633)

[1.1.1. Giới thiệu 11](#_Toc339575634)

[1.1.2. Biểu diễn hình ảnh bằng Gabor Wavelet 2D 12](#_Toc339575635)

[1.1.3. Trích chọn đặc trưng khuôn mặt bằng biến đổi sóng Gabor 2D 12](#_Toc339575636)

[1.2. Tổng quan về mô hình mạng Neural MLP 14](#_Toc339575637)

[1.2.1. Giới thiệu chung về mô hình mạng Neural MLP 14](#_Toc339575638)

[1.2.2. Huấn luyện mạng neuron MLP 15](#_Toc339575639)

[o Thuật toán lan truyền ngược 16](#_Toc339575640)

[Phần III. Cài đặt chương trình và thử nghiệm 17](#_Toc339575641)

[1.1. Mô tả chung về chương trình 17](#_Toc339575642)

[1.2. Cài đặt chương trình 19](#_Toc339575643)

[1.3. Kết quả của hệ thống đối với các trường hợp ảnh bị thay đổi biểu cảm, độ chiếu sáng 20](#_Toc339575644)

[Tài liệu tham khảo 22](#_Toc339575645)

# Lời mở đầu

Cùng với sự phát triển mạnh mẽ của truyền thông và khoa học máy tính thì sự quan tâm đến vấn đề bảo mật ngày càng tăng. Hiện nay, công nghệ hiện đại đã cho phép việc xác thực dựa vào “bản chất” của từng cá nhân. Công nghệ này dựa trên lĩnh vực được gọi là *sinh trắc học*. Kiểm soát bằng sinh trắc học là những phương pháp tự động cho phép xác thực hay nhận dạng một cá nhân dựa vào các *đặc trưng sinh lý học* của người đó như đặc điểm vân tay, gương mặt, gen,…hoặc dựa trên những đặc điểm liên quan đến *đặc trưng hành vi* như dạng chữ viết, giọng nói… Vì những hệ thống nhận dạng bằng sinh trắc học sử dụng thông tin sinh trắc học của con người nên kết quả chính xác và đặc biệt là rất khó bị giả mạo.

Nhận dạng khuôn mặt là phương pháp nhận dạng dựa vào đặc trưng sinh lý học cho kết quả chính xác cao đồng thời rất thuận tiện khi sử dụng. Do đó, việc nghiên cứu các đặc tính của khuôn mặt người đã thu hút được sự quan tâm của các nhà nghiên cứu. Hiện nay có nhiều phương pháp khác nhau được đưa ra để giải quyết bài toán nhận dạng khuôn mặt, các phương pháp thường được sử dụng như phương pháp dựa vào mẫu và các đặc trưng hình học, phương pháp dựa trên việc xem xét tổng thể hay từng phần trên khuôn mặt, phương pháp dùng mạng neural và máy học. Tất cả những phương pháp tiếp cận này đều có những thế mạnh riêng và chúng thường được kết hợp vào trong một hệ thống nhận diện khuôn mặt nhằm tăng độ chính xác cho việc giải quyết bài toán. Phương pháp dựa và mẫu và các đặc trưng hình học kết hợp với việc sử dụng mạng neural, máy học là phương pháp tổng hợp được đánh giá là đem lại những hiệu quả khả quan nhất, trong phương pháp dựa vào mẫu, các phương pháp thống kê được sử dụng để biểu diễn toàn bộ ảnh như phương pháp phân tích thành phần chính (PCA), phương pháp phân tích thành phần độc lập (ICA), phương pháp dùng biến đổi bước sóng Gabor (Gabor Wavelet), …

Các phương pháp nhận dạng mặt người từ những giai đoạn đầu đã tổng kết được những điểm quan trọng trên khuôn mặt người đó là : Mắt, đỉnh mũi, miệng, cạnh cằm và mối quan hệ về khoảng cách giữa các điểm trên để diễn tả cho mỗi khuôn mặt. Tuy nhiên, những phương pháp đã đưa ra đã không đạt được kết quả tốt trong trường hợp khuôn mặt bị xoay đi một góc và trường hợp sự chiếu sáng có sự thay đổi lớn. Sự thay đổi lớn về độ co giãn và mức chiếu độ chiếu sáng là những vấn đề quan trọng được đặt ra đối với việc nhận dạng mặt người cần phải giải quyết. Phương pháp so khớp bằng đồ thị co giãn và các biểu đồ topo sử dụng biến đổi sóng nhằm miêu tả ảnh về mặt hình dáng sinh học của khuôn mặt. Trong phương pháp này thì các vector sẽ được tạo ra bằng cách lấy mẫu qua kết quả biến đổi sóng của ảnh mặt người tại các nút của đồ thị so khớp.

Trong đề tài này nhóm em đã tìm hiểu và đề cập đến một phương pháp hữu hiệu đối với trường hợp bị thay đổi cảm xúc của ảnh và sự thay đổi trong điều kiện chiếu sáng. Phương pháp này sử dụng biến đổi sóng Gabor để xây dựng vector đặc trưng của khuôn mặt dựa trên việc chọn các “năng lượng cao – điểm chóp” (những điểm có sự đột biến mạnh) của ảnh kết quả sau phép biến đổi sóng Gabor. Những điểm này sẽ được chọn làm các điểm đặc trưng cho khuôn mặt thay vì sử dụng các nút đồ thị đã được xác định trước trong phép so khớp đồ thị co giãn. Những điểm đặc trưng đó thường tập trung ở: Mắt, lông mày, mũi, miệng, cạnh cằm, má lúm đồng tiền, nốt ruồi, vết sẹo,… Đây là những điểm có ý nghĩa rất cao trong việc nhận dạng.

# Phần I. Tổng Quan Về Nhận Dạng Khuôn Mặt

## Các khái niệm liên quan đến nhận dạng khuôn mặt

* Hệ thống sinh trắc học

Hệ thống sinh trắc học là một hệ thống được thiết kế để xác minh và nhận dạng một người dựa vào những đặc trưng sinh học duy nhất của người đó.

* Hệ thống nhận dạng khuôn mặt

Hệ thống nhận dạng khuôn mặt là một hệthống được thiết kếđểtìm thông tin của một người. Kĩ thuật nhận dạng là kiểm tra sự phù hợp dựa trên phép so sánh một - nhiều cụ thể là tìm ra một người là ai trong số những người đã được lưu trữ trong hệ thống dựa vào thông tin khuôn mặt.

* Hệ thống xác minh hay xác thực khuôn mặt là gì?

Hệ thống xác minh/xác thực khuôn mặt là một hệ thống được thiết kế để xác minh thông tin của một người. Xác minh là kiểm tra sự phù hợp trên phép so sánh một - một cụ thể là đối chiếu thông tin mới nhận về một người với thông tin đã lưu trữ về người này có khớp hay không dựa trên thông tin khuôn mặt. Xác thực quá trình chưa biết thông tin về người cần xác thực và kết quả đặt ra là xác thực xem người này là ai.

## Bài toán nhận dạng khuôn mặt

Với sự phát triển mạnh mẽ của khoa học kỹ thuật số và mạng toàn cầu, vấn đề đảm bảo an toàn về thông tin cũng như vật chất trở nên ngày càng quan trọng. Phần lớn các hệ thống thông tin trước đây và cho đến nay với các thông tin xác thực người truy cập không đặc trưng cho chính bản thân họ và chỉ đơn thuần là những gì họ đang sở hữu: Số chứng minh thư nhân dân, chìa khoá, mật mã, thẻ tín dụng hoặc họ tên. Rõ ràng các thông tin hay vật dụng này không mang tính đặc trưng mà chỉ mang tính xác thực với người sử dụng, nếu chúng bị đánh cắp hoặc sao chép thì kẻ trộm hoàn toàn có quyền truy cập và sử dụng hệ thống. Hiện nay các công nghệ hiện đại đã cho phép việc xác thực dựa vào bản chất của từng cá nhân.

Các lĩnh vực nhân trắc học được dùng trong xác thực

Nhận dạng gương mặt là phương pháp nhận dạng dựa vào đặc tính sinh lý và ngày nay nhận dạng mặt người đã được quan tâm và được coi như là một trong các lĩnh vực quan trọng nhất trong khoa học nhận dạng là do nhu cầu ứng dụng nhận dạng mặt người trong các lĩnh vực (kinh tế, thi hành luật pháp, điều tra tội phạm, bảo mật,…) ngày càng nhiều.

Vậy nhận dạng mặt người là gì? Đó là sự so sánh các ảnh mặt người với nhau, chỉ ra mức độ tương tự giữa các mặt được so sánh và đưa ra quyết định về tính đồng nhất giữa chúng.

## Mô hình bài toán nhận dạng mặt người

Nhận dạng mặt người là một trong những bài toán phức tạp trong lĩnh vực nhận dạng ảnh. Một gương mặt người không chỉ là đối tượng ba chiều mà còn là một thực thể mang tính động rất cao. Ngoài ra, nhận dạng mặt người còn phụ thuộc vào nhiều yếu tố: Chất lượng ảnh của thiết bị thu, nền phức tạp, sự dịch chuyển, xoay theo ba chiều, cảm xúc thể hiện trên khuôn mặt,... Hiện nay chất lượng của các hệ thống dò tìm và nhận dạng ảnh khuôn mặt người chưa cao. Người ta vẫn chưa tìm ra được phương pháp trích chọn các đặc trưng của khuôn mặt mà vẫn chỉ sử dụng các đặc trưng của ảnh. Một số phương pháp được triển khai ứng dụng nhưng tốc độ xử lý, nhận dạng ảnh còn chậm, không đáp ứng thời gian thực.

Nhận dạng mặt người có thể được ứng dụng rộng rãi trong rất nhiều lĩnh vực khác nhau. Ở mỗi lĩnh vực chúng ta sẽ có mô hình cụ thể để giải quyết bài toán nhận dạng mặt người riêng. Chúng ta có thể chia các loại hệ thống nhận dạng mặt người thành các mô hình chính sau đây:

* Nhận dạng Online:

Mô hình này sử dụng thiết bị Camera để theo rõi, giám sát. Đây là những ứng dụng thường được triển khai ở các nhà ga, cửa khẩu, sân bay... nơi có đông người qua lại. Hệ thống liên tục tách các ảnh khuôn mặt từ những ảnh chụp được, đối sánh với Cơ sở dữ liệu để xác định những đối tượng cần quan tâm. Về yêu cầu :

* Đảm bảo độ chính xác.
* Đảm bảo thời gian thực.
* Nhận dạng trong tập cơ sở dữ liệu hữu hạn:

Xác thực một người có đúng là người mà anh ta tự nhận hay không (hệ thống cửa bảo mật tự động, kiểm tra giấy chứng minh nhân dân, thẻ căn cước, giấy phép lái xe, giấy ra vào cơ quan...).Về yêu cầu:

* Đảm bảo về độ chính xác.
* Thời gian cho phép nhiều hơn so với mô hình nhận dạng Online. Ví dụ: Hệ thống cửa bảo mật tự động thì thời gian có thể chấp nhận được khoảng từ 15s đến 30s.
* Nhận dạng trong tập cơ sở dữ liệu lớn:

Xác nhận một người là ai trong Cơ sở dữ liệu của hệ thống như Cơ sở dữ liệu về đối tượng hình sự, Cơ sở dữ liệu về chứng minh thư nhân dân...Về yêu cầu:

* Đảm bảo độ chính xác.
* Thời gian cần để xác thực cho phép lâu hơn các mô hình khác.

Việc định vị chính xác vị trí khuôn mặt và tìm một số điểm đặc trưng trên khuôn mặt có ý nghĩa quyết định đến hiệu quả của hệ thống nhận dạng mặt người. Từ đó áp dụng vào các ứng dụng thực tế: hệ thống chống trộm, quản lý nhân công bằng ảnh thẻ,...

Cấu trúc điển hình của hệ thống nhận dạng mặt người gồm có 2 bộ phận chính:

* Bộ dò tìm (Face Image Detector): Có nhiệm vụ xác định vị trí của gương mặt trong một ảnh bình thường.
* Bộ phận nhận dạng hay phân lớp gương mặt (Face Recognizer): Để xác định người có gương mặt tương ứng là ai trong cơ sở dữ liệu.

## Quy trình nhận dạng khuôn mặt

Tiền xử lý

Theo vết - Dò tìm khuôn mặt

Định vị trí - kích thước và hình dạng

Rút trích đặc trưng

Dữ liệu ảnh của người dùng

So khớp đặc trưng

Chứng thực ảnh

Ảnh/Video

Ảnh đã qua xử lý

Vector đặc trưng

Sơ đồ luồng xử lý của quá trình nhận dạng khuôn mặt

* Giai đoạn theo vết – dò tìm khuôn mặt:

Dò tìm khuôn mặt trong ảnh tĩnh hay theo vết khuôn mặt nếu quan sát qua video. Đây là giai đoạn quan trọng ảnh hưởng đến kết quả nhận dạng, việc phát hiện khuôn mặt gắn với điều kiện không ổn định của dữ liệu ảnh: Độ sáng của ảnh, kích thước và hướng của khuôn mặt, góc quay,… Vì vậy, với hệ thống nhận dạng mặt người cần phải có một số tiêu chuẩn ràng buộc:

* Ảnh với màu tự nhiên chiếu sáng và độ đậm đều (không quá tối và không quá sáng).
* Kích thước của khuôn mặt không quá lớn hoặc quá nhỏ.
* Khuôn mặt phải được chụp trực diện, mắt mở tự nhiên.
* Không có vật gì che khuôn mặt: mũ, khăn, tóc.
* Giai đoạn tiền xử lý:

Giai đoạn này để nâng cao chất lượng của ảnh số giúp cho việc phát hiện vị trí mặt người và quá trình nhận dạng đạt hiệu quả hơn. Ảnh sau tiền xử lý chỉ được nâng cao chất lượng (chứ không thay đổi tư thế và biểu cảm trên khuôn mặt) và đưa sang khối trích đặc trưng. Các công việc trong giai đoạn này có thể là: Chuẩn hoá kích cỡ ảnh (*Image size normalization*), hiệu chỉnh độ sáng, tối của ảnh (*Histogram equalization, illumination normalization*), lọc nhiễu (*median filtering*).

* Giai đoạn rút trích đặc trưng:

Giai đoạn này tìm ra các đặc trưng chính của khuôn mặt, từ các đặc trưng này hình thành các véc tơ đặc trưng, các vector này được sử dụng để đối sánh sự giống nhau giữa ảnh khuôn mặt cần tìm và ảnh trong cơ sở dữ liệu.

* Giai đoạn so khớp đặc trưng:

Thực hiện việc so sánh giữa các vector đặc trưng để chọn ra khoảng cách giữa ảnh cần tìm và ảnh trong cơ sở dữ liệu.

## Các hướng tiếp cận chính trong lĩnh vực nhận dạng khuôn mặt

Hiện nay có nhiều phương pháp khác nhau được đưa ra để giải quyết bài toán nhận dạng khuôn mặt, các phương pháp thường được sử dụng như phương pháp dựa vào mẫu và các đặc trưng hình học, phương pháp dựa trên việc xem xét tổng thể hay từng phần trên khuôn mặt, phương pháp dùng mạng neural và máy học. Tất cả những phương pháp tiếp cận này đều có những thế mạnh riêng và chúng thường được kết hợp vào trong một hệ thống nhận diện khuôn mặt nhằm tăng độ chính xác cho việc giải quyết bài toán.

Một bài toán nhận dạng khuôn mặt cần xác định hai vấn đề chính: Dùng thông tin, đặc trưng nào để nhận dạng ví dụ như chân mày, cặp mắt, mũi, môi, tai, hay kết hợp các thông tin trên. Và dùng phương pháp nào để huấn luyện cho máy nhận dạng dùng nguồn thông tin, đặc trưng đó. Nhận dạng khuôn mặt trên máy tính đã trải qua nhiều bước thăng trầm với các kết quả như sau:

* Wenyi Zhao, Arvindh Krishnaswamy, Rama Chellappa, Danie L.Swets, John Weng (1998)[1] sử dụng phương pháp PCA (phân tích thành phần chính) kết hợp LDA (phân tích độc lập tuyến tính). Bước 1, chiếu ảnh khuôn mặt từ không gian ảnh thô sang không gian các không gian khuôn mặt (Mỗi lớp khuôn mặt được nhận dạng sẽ được mô hình hóa bằng một không gian khuôn mặt) dùng PCA. Bước 2, sửdụng phương pháp LDA để tạo bộphân loại tuyến tính có khảnăng phân lớp các lớp khuôn mặt.
* Emmanuel Viennet và Francoise Fogelman Soulie (1998), sử dụng phương pháp mạng neural nhân tạo để xử lý và nhận dạng khuôn mặt.
* Antonio J.Colmenarez và Thomas S.Huang (1998), sử dụng kỹ thuật học thị giác và phù hợp mẫu 2-D. Ông quan niệm bài toán dò tìm khuôn mặt là thao tác phân loại khuôn mặt trong đó khuôn mặt thuộc về một lớp và các đối tượng khác thuộc về lớp còn lại bằng cách ước lượng mô hình xác suất cho mỗi lớp, và việc dò tìm sử dụng luật quyết định Maximum-likeli hood.
* Kazunori Okada, Johannes Steffens, Thomas Maurer, Hai Hong, Egor Elagin, Hartmut Neven, and Christoph (1998), nhận dạng khuôn mặt dựa vào sóng Gabor và phương pháp phù hợp đồ thị bó. Với ý tưởng dùng đồ thị để biểu diễn khuôn mặt, ảnh khuôn mặt được đánh dấu tại các vịtrí đã được xác định trước trên khuôn mặt, gọi các vị trí này chính là các vị trí chuẩn. Khi thực hiện thao tác so khớp đồthị với một ảnh, các điểm chuẩn (Jets) sẽ trích ra từ ảnh và so sánh các điểm chuẩn này với tất cả các điểm chuẩn tương ứng trong các đồ thị khác nhau, và đồ thị nào phù hợp nhất với ảnh sẽ được chọn.
* Baback Moghaddam và Alex Pentland (1998), đưa ra phương pháp phù hợp thị giác trực tiếp từ cá cảnh cần sử dụng cho mục đích nhận dạng khuôn mặt và dùng độ đo xác suất để tính độ tương tự.
* Massimo Tistaelli và Enrico Grosso (1998), đưa ra kỹ thuật thị giác động. Vì khả năng quan sát các chuyển động của khuôn mặt và xử lý các tính huống theo dự định là thông tin rất quan trọng, từ đó nhận được mô tả đầy đủ hơn về khuôn mặt cho mục đích thu thập mẫu và nhận dạng.
* Jeffrey Huang, Chengjun Liu, và Harry Wechsler (1998)[8], đề xuất thuật toán căn cứ trên tính tiến hóa (Evolutionary computation) và di truyền (Genetic) cho các tác vụ nhận dạng khuôn mặt. Đối với cách tiếp cận này, hai mắt sẽ được dò tìm trước tiên và thông tin này được xem là vết để quan sát khuôn mặt, trình xử lý dò tiếp mắt bằng cách sử dụng một thuật toán lai để kết hợp thao tác học và tiến hóa trong quá trình học.
* Daniel Bgraham và Nigel M Allinson (1998), sử dụng phương pháp được gọi là tạo bản sao không gian đặc trưng để biểu diễn và nhận dạng hướng di chuyển của khuôn mặt.
* Oi Bin Sun, Chian Prong Lam và Jian Kang Wu (1998), sửdụng phương pháp tìm vùng hai chân mày, hai mắt, mũi, miệng và cằm. Ảnh khuôn mặt thẳng ban đầu được chiếu theo chiều ngang để tìm các giá trị điểm ảnh thỏa ngưỡng cho trước, đồ thị biểu diễn theo trục ngang sẽ định vị vị trí biên trên và biên dưới của hình chữ nhật bao các đặc trưng cục bộ khuôn mặt. Tương tự với chiều đứng để tìm ra đường biên bên trái và phải cho các vùng đặc trưng.
* Ara V.Nefian và Monson H.Hayes III (1998), trình bày hướng tiếp cận theo mô hình mô hình Markovẩn (HMM) trong đó ảnh mẫu khuôn mặt được lượng hóa thành chuỗi quan sát trên khuôn mặt theo quan niệm dựa trên thứtự xuất hiện các đặc trưng khuôn mặt {hai chân mày, hai lông mi, mũi, miệng, cằm}. Trong chuỗi quan sát đó, mỗi quan sát lại là một vector nhiều chiều và mỗi vector quan sát này được sửdụng để đặc trưng cho mỗi trạng thái trong chuỗi trạng trạng thái của HMM. Mỗi người được ước lượng bằng một mô hình của HMM.
* Guodong Guo, Stan Z.Li, Kap Luk Chan (17 January2001), dùng phương pháp SVM đểnhận dạng khuôn mặt. Sử dụng chiến lược kết hợp nhiều bộphân loại nhịphân để xây dựng bộ phân loại SVM đa lớp.

Nhận xét:

* Nhận dạng mặt người là một trong những bài toán khó khăn nhất trong lĩnh vực nhận dạng ảnh. Một khuôn mặt người không chỉ là đối tuợng ba chiều mà còn là một thực thể mang tính động rất cao. Ngoài ra, do ảnh mặt người thường được chụp trong điều kiện môi trường tự nhiên nên thông thường nền ảnh rất phức tạp và độ chiếu sáng có thể rất kém là một ví dụ về một bức ảnh với nền phức tạp có chứa mặt người.
* Các yếu tố xuất hiện trên ảnh tạo nên khó khăn cho hệ thống nhận dạng có thể được phân thành các loại sau đây:
* Máy ảnh không rõ và nhiễu.
* Nền phức tạp.
* Độ sáng.
* Sự dịch chuyển, xoay, biến đổi tỉ lệ giữa các thành phần.
* Cảm xúc thể hiện trên gương mặt .
* Hoá trang, kiểu tóc.
* Do đó, có thể nói những thách thức đối với bài toán nhận dạng khuôn mặt là:
* Những biến đổi quá lớn giữa các ảnh khuôn mặt khác nhau từ một người cần nhận dạng gồm trạng thái cảm xúc trên khuôn mặt, ánh sáng, và các thay đổi vị trí của khuôn mặt..vv.
* Giới hạn về số ảnh cần thiết cho việc nhận dạng, tập học 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.
* Hiện nay có rất nhiều nghiên cứu về lĩnh vực nhận dạng ảnh mặt người. Nhưng nhiều phương pháp đưa ra mới chỉ dựa vào các đặc trưng của ảnh chứ chưa tiếp cận được với các đặc trưng của khuôn mặt trong ảnh (tức những đặc trưng *sinh trắc học* của khuôn mặt).

Trong đề tài này, nhóm em tìm hiểu và đề cập đến một phương pháp hữu hiệu đối với trường hợp thay đổi cảm xúc và sự thay đổi trong điều kiện chiếu sáng. Đó là phương pháp sử dụng sóng Gabor. Phương pháp này sử dụng những điểm có sự đột biến mạnh trong biến đổi sóng Gabor. Những điểm này thường tập trung ở: Mắt, lông mày, mũi, miệng, cạnh cằm, má lúm đồng tiền, nốt ruồi, vết sẹo, v.v… (những điểm có ý nghĩa rất cao trong việc nhận dạng). Vì vậy, đề tài sẽ cố gắng tiếp cận với phương pháp nhận dạng mặt người dựa vào các đặc trưng *sinh trắc học* của mặt người và thực hiện cài đặt chương trình nhận dạng mặt người theo mô hình “nhận dạng với tập Cơ sở dữ liệu hữu hạn” làm cơ sở để triển khai ứng dụng trong thực tế như hệ thống cửa bảo mật tự động... Đồng thời đưa ra các cải tiến để giải quyết các bài toán nhận dạng mặt người trong thực tế như nhận dạng đối với các đối tượng bịt mặt,… bên cạnh đó quá trình huấn luyện và nhận dạng khuôn mặt sẽ được thực hiện bởi mạng MLP Neural Network.

# Phần II. Tổng quan về Gabor Wavelet và MLP Neural Network

## Tổng quan về Gabor Wavelet

### Giới thiệu

Hàm Gabor được đề xuất đầu tiên bởi Dennis Gabor như là 1 công cụ dùng để phát hiện tín hiệu từ tạp nhiễu.



Sau đó, bộ lọc Gabor được Daugman tổng quát hóa thành dạng 2 chiều. Bộ lọc Gabor 2 chiều hiện nay được sử dụng rộng rãi trong nhiều ứng dụng xử lý ảnh trên máy tính.



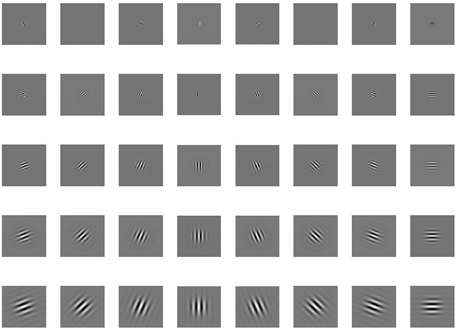
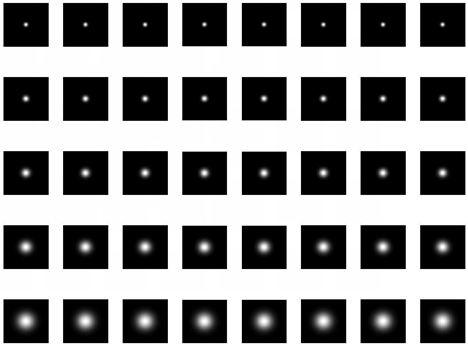
Mỗi  là một sóng phẳng đặc trưng bởi vector hàm sóng Gauss , với  là độ lệch tiêu chuẩn của hàm sóng Gauss (thông thường  = 2), j là giá trị ảo trong trường số phức. Tần số trung tâm của bộ lọc Gabor thứ i được cho bởi vector sóng đặc trưng:



Một số thông số quan trọng đối với biến đổi sóng Gabor:

*  mô tả hướng của hàm sóng Gabor. Thông số này khiến cho sóng Gabor quay quanh tâm của nó. Trong đa số các trường hợp  có giá trị trong khoảng 0 đến . Ở đây ta sẽ chọn  từ  với bước nhảy :  với u = 1,....,8.
*  = fv mô tả bước sóng của sóng sin hoặc tần số trung tâm của sóng Gabor. Ở đây ta sẽ chọn v có giá trị trong khoảng (0…4), và f được chọn là . Sóng Gabor có bước sóng lớn sẽ đáp ứng với những thay đổi cường độ của ảnh. Sóng Gabor có bước sóng nhỏ sẽ đáp ứng với những cạnh.

Biểu diễn dạng ảnh của bộ lọc Gabor Wavelet với 5 tỉ lệ và 8 hướng, chúng ta sẽ có 40 mặt phẳng sóng như sau (phần biên độ và phần thực của bộ lọc Gabor).



### Biểu diễn hình ảnh bằng Gabor Wavelet 2D

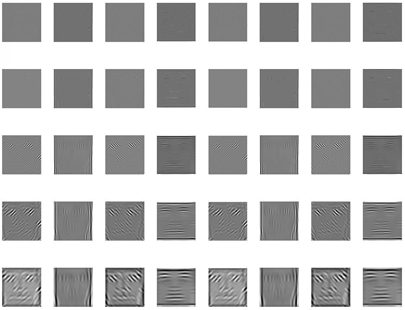
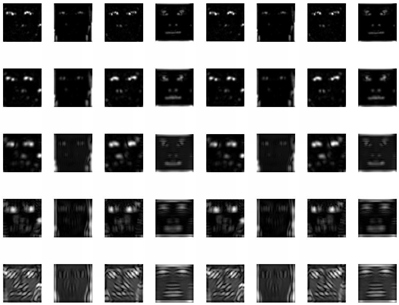
Biến đổi sóng Gabor của một ảnh I cho trước được cho bởi công thức sau:

 ( là cường độ ảnh tại vị trí )

Một hình ảnh có thể được biểu diễn bởi biến đổi sóng Gabor cho phép mô tả cả cấu trúc của không gian tần số và mối liên hệ không gian. Chập hình ảnh với tổ hợp bộ lọc Gabor cho 5 không gian tần số (v = 0,.., 4) và 8 hướng (u = 1,...,8) thu được toàn bộ mật độ phổ tần số, cả biên độ và pha.

Cho ảnh gốc:

Biểu diễn ảnh gốc bằng sóng Gabor (phần thực và phần biên độ của ảnh sau khi biến đổi sóng Gabor):



### Trích chọn đặc trưng khuôn mặt bằng biến đổi sóng Gabor 2D

Việc nhận dạng khuôn mặt dùng biến đổi sóng Gabor 2D là việc sử dụng các vector đặc trưng để mã hóa khuôn mặt thay vì dùng biểu đồ các điểm đặc trưng trên khuôn mặt. Các điểm có năng lượng cao sẽ được sử dụng để so sánh các khuôn mặt. Điều này không chỉ giúp giảm khối lượng tính toán mà còn tăng độ chính xác của thuật toán do không cần phải xác định các điểm đặc trưng trên khuôn mặt bằng tay

Chiết xuất đặc trưng:

* Xác định các điểm đặc trưng
* Tính vector đặc trưng

#### Xác định các điểm đặc trưng

Ở bước này, vector đặc trưng được tạo ra từ các điểm có nội dung thông tin cao trên hình ảnh khuôn mặt. Trông hầu hết các phương pháp dựa vào đặc trưng, các đặc trưng trên khuôn mặt là mắt, mũi, miệng. Tuy nhiên, ở đây chúng ta sẽ không áp đặt vị trí và số lượng các điểm đặc trưng. Số lượng điểm đặc trưng và vị trí của chúng có thể thay đổi nhằm mục đích biểu diễn các đặc trưng nhân diện đa dạng của nhiều khuôn mặt khác nhau như nếp nhăn, nốt ruồi,...cũng là những đặc điểm có thể sử dụng để nhận dạng khuôn mặt.



Từ đáp ứng của khuôn mặt qua bộ lọc Gabor, ta sẽ tìm các điểm đặc trưng theo phương pháp sau:

Một điểm là điểm đặc trưng nếu thỏa mãn các điều kiện sau:

 (3)

 (4)



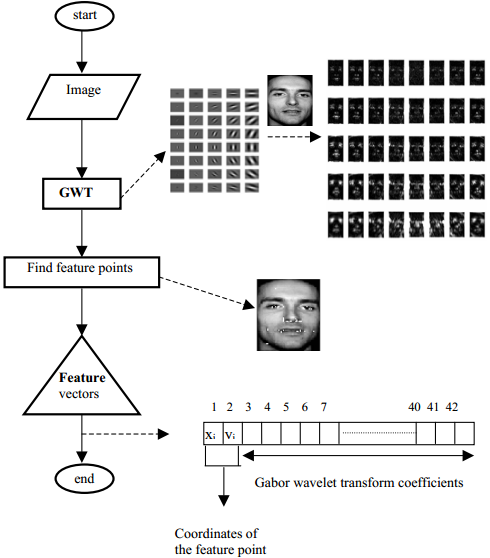
Với là đáp ứng của ảnh khuôn mặt đối với bộ lọc Gabor thứ j; là kích thước của ảnh khuôn mặt; là một hình vuông có cạnh WxW pixel có tâm tại . Kích thước W là một thông số quan trọng trong thuật toán này, nó phải được chọn đủ nhỏ để thu được những đặc điểm quan trọng, và đủ lớn để tránh phần dư thừa. Phương trình (4) được dùng để tránh bị mắc trong 1 khu vực địa phương mà không tìm ra được đinh của đáp ứng. Ở đây chúng ta sẽ chọn W = 9 để tìm các điểm đặc trưng của khuôn mặt qua đáp ứng đối với bộ lọc Gabor. Một sơ đồ đặc trưng cho khuôn mặt được tạo nên bằng cách lần lượt thực hiện quá trình trên đối với 40 bộ lọc Gabor.

#### Xuất vector đặc trưng

Vector đặc trưng được tạo ra tại các điểm đặc trưng ở dạng tổ hợp các hệ số của biến đổi sóng Gabor. Vector đặc trưng thứ k của hình ảnh khuôn mặt mẫu thứ i được định nghĩa như sau:



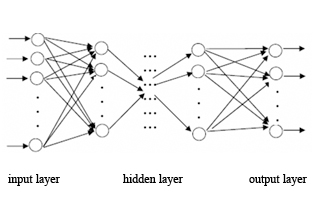
Với 40 bộ lọc Gabor, vector đặc trưng có 42 thành phần. 2 thành phần đầu tiên trong vector đặc trưng biểu diễn vị trí của điểm đặc trưng theo tọa độ (x,y), 40 thành phần tiếp theo của vector đặc trưng là các giá trị của đáp ứng qua bộ lọc Gabor của hình ảnh khuôn mặt tại vị trí tương ứng. Việc xây dựng vector đặc trưng bằng các giá trị của biến đổi sóng Gabor là rất quan trọng. Vector đặc trưng cho phép thể hiện cấu trúc không gian tần số và cả mối quan hệ không gian của vùng ảnh địa phương xung quanh điểm đặc trưng tương ứng.



Lưu đồ của các bước trích chọn đặc trưng ảnh khuôn mặt dùng biến đổi sóng Gabor

## Tổng quan về mô hình mạng Neural MLP

### Giới thiệu chung về mô hình mạng Neural MLP

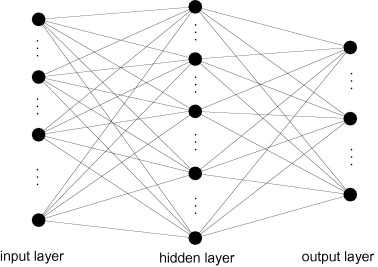
Mô hình mạng neural MLP (multilayer feedforward perceptron) là một dạng cơ bản và được sử dụng rộng rãi nhất của mô hình mạng neural. Mô hình này gồm một số hữu hạn các lớp. Mỗi lớp gồm một số hữu hạn neural. Mỗi neural của mỗi lớp được kết nối với mỗi neuron của mỗi lớp kế tiếp một cách tuần tự. Những kết nối này gọi là *synapses*. Dòng dữ liệu từ một lớp sẽ đi tới lớp kế tiếp. Một mạng neural MLP tổng quát là mạng có n (n≥3) lớp, lớp đầu tiên gọi là lớp nhập, lớp cuối cùng gọi là lớp kết quả, các lớp ở giữa gọi là các lớp ẩn (ví dụ minh họa hình bên dưới)

#### Kiến trúc của một mạng MLP tổng quát có thể mô tả như sau:

* + Đầu vào là các vector (x1, x2, ..., xp) trong không gian p chiều, đầu ra là các vector (y1, y2, ..., yq) trong không gian q chiều. Đối với các bài toán phân loại, p chính là kích thước của mẫu đầu vào, q chính là số lớp cần phân loại.
  + Mỗi nơron thuộc tầng sau liên kết với tất cả các nơron thuộc tầng liền trước nó.
  + Đầu ra của neural tầng trước là đầu vào của neural thuộc tầng liền sau nó.

Hoạt động của mạng MLP như sau: Tại tầng đầu vào các neural nhận tín hiệu vào xử lý (tính tổng trọng số, gửi tới hàm truyền) rồi cho ra kết quả (là kết quả của hàm truyền), kết quả này sẽ được truyền tới các nơron thuộc tầng ẩn thứ nhất, các neural tại đây tiếp nhận như là tín hiệu đầu vào, xử lý và gửi kết quả đến tầng ẩn thứ 2, … quá trình tiếp tục cho đến khi các neural thuộc tầng ra cho kết quả.

Thông thường người ta chỉ sử dụng một lớp ẩn. Vì vậy đôi khi người ta hay đồng nhất mô hình mạng neural MLP với mô hình mạng neural MLP 3 lớp (ví dụ minh họa hình bên dưới)



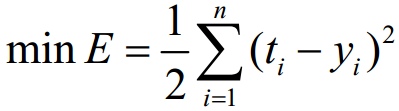
### Huấn luyện mạng neuron MLP

#### Các luật học cơ bản của mạng neural nhân tạo

* + Có hai kỹ thuật học trong mạng neural nhân tạo đó là học tham số (parameter learning) và học cấu trúc (structure learning). Học tham số là việc thay đổi trọng số của các liên kết giữa các neural trong một mạng, còn học cấu trúc là việc điều chỉnh cấu trúc của mạng bao gồm thay đổi số lớp neural, số neural của mỗi lớp và cách liên kết giữa chúng. Hai vấn đề này có thể được thực hiện đồng thời hoặc tách biệt.
  + Về mặt phương pháp học, có thể chia ra làm ba loại: Học có giám sát hay còn gọi là học có thầy (supervised learning), học tăng cường (reinforcement learning) và học không có giám sát hay còn gọi là học không có thầy ( unsupperviced learning).

#### Quá trình huấn luyện của mô hình mạng neural MLP

* + Mạng MLP được huấn luyện theo thuật toán lan truyền ngược lỗi (back propagation), sao cho sai lệch giữa tín hiệu ra của mạng và giá trị đích (target) là nhỏ nhất.



Trong đó:

* + - * yi : Tín hiệu ra của mạng
      * ti : Tín hiệu đích mong muốn
      * n : Kích thước không gian dữ liệu (thông số chẩn đoán).

### Thuật toán lan truyền ngược

Ta sử dụng một số kí hiệu sau:

* j: Neural thứ j (hay nút thứ j)
* *X*j: Vector đầu vào của nút thứ j
* *W*j: Vector trọng số của nút thứ j
* xji: Đầu vào của nút thứ j từ nút thứ i
* wji: Trọng số trên xji
* bj: Ngưỡng tại nút thứ j
* oj: Đầu ra của nút thứ j
* tj: Đầu ra mong muốn của nút thứ j
* Downstream(j): Tập tất cả các nút nhận đầu ra của nút thứ j làm một giá trị đầu vào.
* η: Tốc độ học
* f: Hàm truyền với f(x) = 1 / (1 + e-x)

Thuật toán lan truyền ngược được mô tả như sau:

**Input**:

* Mạng MLP với ni đầu vào, nh nút ẩn và no đầu ra.
* Hệ số học η
* Tập dữ liệu huấn luyện D = {là vector đầu vào, là vector đầu ra mong muốn}.

**Output**: Các vector trọng số

**Thuật toán**:

**Bước 1**: Khởi tạo trọng số bởi các giá trị ngẫu nhiên nhỏ.

**Bước 2**: Lặp lại cho tới khi thỏa mãn điều kiện kết thúc.

Với mỗi mẫu, thực hiện các bước sau:

2.1. Tính đầu ra oj cho mỗi nút j:

**oj = f(d – bj)** với **d = Σxjiwji**

2.2. Với mỗi nút k thuộc tầng ra, tính δk theo công thức:

**δk = (tk – ok)(1 – ok)ok**

2.3. Với mỗi nút h thuộc tầng ẩn, tính δh theo công thức:

**δh = oh(1 – oh) Σδkwkh** với **k** ∈ **Downstream(j)**

2.4 Cập nhật: **wji = wji + Δwji**

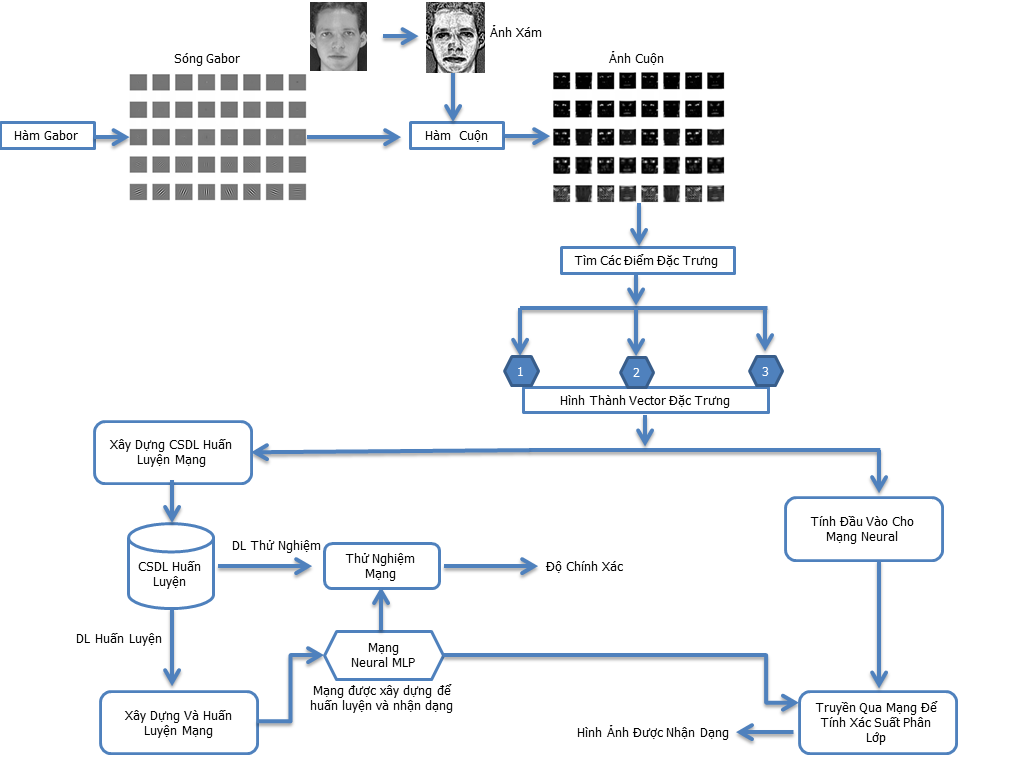
Trong đó **Δwji *= η*δkxji**

Việc đưa mẫu huấn luyện qua mạng, tính toán và cập nhật trọng số được tiến hành với tất cả phần tử trong bộ mẫu (có thể chọn ngẫu nhiên hoặc tuần tự). Quá trình sẽ dừng lại khi sai số trung bình (hoặc tổng sai số) nhỏ hơn một giá trị cho trước hoặc thay đổi không đáng kể (tức là quá trình huấn luyện hội tụ).

# Phần III. Cài đặt chương trình và thử nghiệm

## Mô tả chung về chương trình

Chương trình nhận dạng ảnh khuôn mặt sử dụng biến đổi sóng Gabor kết hợp sử dụng mạng Neural MLP.



Sơ đồ logic của chương trình nhận dạng mặt người dựa trên biến đổi sóng Gabor

**Mô tả các bước trong sơ đồ logic:**

Từ hàm Gabor, chúng ta sẽ tạo ra 40 mặt nạ sóng Gabor, các mặt nạ này khi ta chuyển về biểu diễn trên không gian ảnh thì có hình dạng như trong sơ đồ (các mặt nạ sóng Gabor).

Ảnh đầu vào đã được tách vùng khuôn mặt và được thực hiện các bước tiền sử lý trước đó, sau đó ảnh được chuyển về dạng ảnh xám trước khi ảnh được thực hiện biến đổi sóng Gabor. Với 40 mặt nạ sóng Gabor chúng ta có 40 ảnh cuộn của khuôn mặt.

Từ 40 ảnh cuộn chúng ta sẽ thực hiện tìm kiếm các điểm đặc trưng và từ các điểm đặc trưng này sẽ hình thành vector đặc trưng của khuôn mặt làm đầu vào cho quá trình huấn luyện mạng và quá trình nhận dạng khuôn mặt.

Quá trình hình thành vector đặc trưng từ những điểm đặc trưng có thể được thực hiện thông qua 1 trong 2 quy trình như sau:

* + Trên mỗi ảnh cuộn ta sẽ tìm những điểm đặc trưng nhất trong một ô có kích thước 9x9 pixel. Hình thành các vector đặc trưng cho mỗi điểm đặc trưng, sau đó ta cộng các điểm đặc trưng này để tạo thành một vector đặc trưng của khuôn mặt.
  + Chúng ta có 40 ảnh cuộn, mỗi ảnh cuộn ta chọn một điểm đặc trưng nhất của khuôn mặt. Từ đó ta có được 40 điểm đặc trưng để hình thành vector đặc trưng của khuôn mặt.

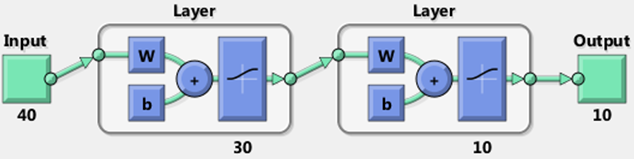
Từ vector đặc trưng của từng khuôn mặt, với các khuôn mặt khác nhau chúng ta sẽ có một cơ sở dữ liệu phục vụ cho quá trình huấn luyện, kiểm thử và nhận dạng khuôn mặt. Quá trình huấn luyện và nhận dạng khuôn mặt ở đây sẽ được thực hiện bởi việc sử dụng mạng neural truyền thẳng 3 lớp.

## Cài đặt chương trình

* **Dữ liệu ảnh huấn luyện và nhận dạng**

Dữ liệu ảnh của 10 người khác nhau từ thư viện ảnh của Yale và Olivetti (92 x 112 pixel – xám hóa). Mỗi người chọn 7 ảnh, một ảnh dùng huấn luyện và 6 ảnh dùng để kiểm tra kết quả nhận dạng. Ảnh để huấn luyện chọn là ảnh có độ sáng đồng nhất, khuôn mặt bình thường không biểu cảm, tư thế đầu thẳng ngay ngắn. Các ảnh chọn để nhận dạng kiểm nghiệm có các biểu cảm khác nhau trên khuôn mặt, độ sáng trên ảnh không đồng nhất.

* Việc hình thành vector đặc cho làm dữ liệu đầu vào cho mạng neural truyền thẳng 3 lớp được thực hiện theo 3 phương pháp:
  + Trên mỗi ảnh cuộn ta sẽ tìm những điểm đặc trưng nhất trong một ô có kích thước 9x9 pixel. Hình thành các vector đặc trưng cho mỗi điểm đặc trưng, sau đó ta cộng các điểm đặc trưng này để tạo thành một vector đặc trưng của khuôn mặt.
  + Chúng ta có 40 ảnh cuộn, mỗi ảnh cuộn ta chọn một điểm đặc trưng nhất của khuôn mặt. Từ đó ta có được 40 điểm đặc trưng để hình thành vector đặc trưng của khuôn mặt.
* Mạng neural được xây dựng là mạng neural truyền thẳng 3 lớp, với các thông số khởi tạo ban đầu như sau:
  + Kiểu học: là học có giám sát.
  + Lớp nhập có 40 neuron
  + Lớp ẩn có 30 neuron
  + Lớp xuất có 10 neuron
  + Mỗi neuron dùng phép hợp tuyến tính và hàm truyền phi tuyến sigmoid.
  + Ngưỡng của mỗi lớp là 1
  + Hệ số học là: 0.4
  + Sai số so với kết quả mong muốn là 0.001

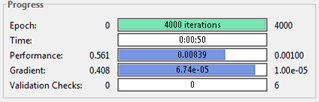


Chương trình nhận dạng thử nghiệm được xây dựng trên môi trường Matlab để tận dụng khả năng hỗ trợ các hàm xử lý Neural. Hệ thống được kết hợp xây dựng ở chế độ giao diện để dễ dàng cho người sử dụng.

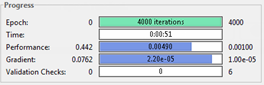
## Kết quả của hệ thống đối với các trường hợp ảnh bị thay đổi biểu cảm, độ chiếu sáng

#### Quá trình huấn luyện

* Phương pháp 1:
* Thời gian huấn luyện: khoảng 50 giây.
* Số lần học đã đạt giới hạn: 4000 lần.
* Sai số so với kết quả mong muốn: khoảng 0.008



* Phương pháp 2:
* Thời gian huấn luyện tới khi hội tụ : khoảng 50 giây.
* Số lần họ đã đạt giới hạn: 4000 lần
* Sai số so với kết quả mong muốn: khoảng 0.004



#### Quá trình nhận dạng

* Đây là vector đâu ra sau khi nhận dạng bằng mạng neuron đã huấn luyện theo 2 phương pháp. Dưới đây là kết quả tượng trưng cho 2 phương pháp với 1 mẫu học làm đại diện
* Phương pháp 1 (đề xuất)

|  |  |
| --- | --- |
| Mẫu test | Kết quả |
| Mẫu học 1 (01.png) | (0.9129,0.0502,0.0403,0.0054,0.0148,0.0536,0.0511,0.0022,0.0162,0.0345) |
| Ảnh tối của mẫu học 01(01\_dark.png) | (0.8592,0.0576,0.0689,0.0064,0.0072,0.1358,0.0567,0.0059,0.0162,0.0612) |
| Ảnh sáng của mẫu học 01(01\_bright.png) | (0.8975,0.0537,0.0320,0.0084,0.0207,0.0427,0.0620,0.0018,0.0168,0.0327) |
| Ảnh happy 01(subject01\_happy.png) | (0.7399,0.1433,0.0315,0.0086,0.0100,0.2570,0.0456,0.0048,0.0230,0.0318) |
| Ảnh sad 01(subject01\_sad.png) | (0.3226,0.1906,0.0958,0.0059,0.0076,0.1899,0.0825,0.0175,0.0226,0.0106) |
| Ảnh supride 01(subject01\_suprise.png) | (0.7149,0.2632,0.0629,0.0068,0.0176,0.0349,0.0814,0.0152,0.0107,0.0088) |
| Ảnh wink  01(subject01\_wink.png) | (0.8154,0.0614,0.0348,0.0061,0.0111,0.3152,0.0875,0.0055,0.0145,0.0221) |

* Phương pháp 2 (gốc)

|  |  |
| --- | --- |
| Mẫu test | Kết quả |
| Mẫu học 1 (01.png) | (0.9213,0.0191,0.0305,0.0232,0.0317,0.0314,0.0337,0.0150,0.0339,0.0252) |
| Ảnh tối của mẫu học 01(01\_dark.png) | (0.6215,0.0226,0.1150,0.0089,0.0097,0.2132,0.0632,0.0298,0.0245,0.0348) |
| Ảnh sáng của mẫu học 01(01\_bright.png) | (0.9170,0.0181,0.0296,0.0262,0.0357,0.0301,0.0387,0.0148,0.0378,0.0244) |
| Ảnh happy 01(subject01\_happy.png) | (0.8760,0.0254,0.0326,0.0201,0.0198,0.0731,0.0373,0.0144,0.0721,0.0183) |
| Ảnh sad 01(subject01\_sad.png) | (0.5066,0.1983,0.1159,0.0014,0.0101,0.0752,0.1262,0.0641,0.2136,0.0121) |
| Ảnh supride 01(subject01\_suprise.png) | (0.8355,0.0578,0.0215,0.0103,0.0171,0.0388,0.1085,0.0227,0.0381,0.0181) |
| Ảnh wink  01(subject01\_wink.png) | (0.7644,0.0329,0.0271,0.0234,0.0087,0.1836,0.1503,0.0250,0.0907,0.0080) |

* Tỉ lệ nhận dạng của từng phương pháp trong bộ ảnh test:
  + Với ngưỡng chấp nhận 0.2
* Tỉ lệ nhận dạng của phương pháp 1: 5/7
* Tỉ lệ nhận dạng của phương pháp 2: 6/7
  + Với ngưỡng chấp nhận 0.3
* Tỉ lệ nhận dạng của phương pháp 1: 6/7
* Tỉ lệ nhận dạng của phương pháp 2: 7 /7
* Nhận xét: Qua kết quả của 2 phương pháp nhận thấy phương pháp 1 cho kết quả tương đối tốt với thời gian khá nhanh. Phương pháp 2 cho kết quả tốt nhất nhưng thời gian tạo ra vector đặc trưng khá lâu.

#### Đánh giá chương trình thử nghiệm

Các kết quả thử nghiệm cho thấy hệ thống đặc biệt hiệu quả khi nhận dạng các ảnh mang kính, biểu cảm khuôn mặt thay đổi hay cường độ chiếu sáng bị thay đổi (nhưng không thay đổi góc chụp của camera). Tỉ lệ chính xác của phương pháp 1 đạt trên 70% còn tỉ lệ chính xác của phương pháp 2 đạt trên 90%.

#### Kết luận

Trong đề tài này, nhóm em đã trình bày về bài toán nhận dạng ảnh mặt người dùng biến đổi Gabor wavelet và mạng neural truyền thẳng 3 lớp. Kết quả của chương trình cho thấy được sự hiệu quả của phương pháp sử dụng biến đổi Gabor wavelet kết hợp với việc sử dụng mạng neural.

# Tài liệu tham khảo

1. Vinay Kumar and B. Shreyas (2007) “Face Recognition Using Gabor Wavelets.
2. Stan Z. Li Anil K. Jain “Handbook of Face Recognition”.
3. Anil K. Jain - Patrick Flynn - Arun A. Ross “Handbook of Biometrics”.