

TỔNG QUAN CÁC PHƯƠNG PHÁP NHẬN DẠNG KHUÔN MẶT DỰA TRÊN ĐẶC TRƯNG CẠNH

OVERVIEW OF FACE RECOGNITION BASED ON EDGE FEATURES

Đặng Nguyên Châu, Đỗ Hồng Tuấn

Trường Đại học Bách khoa Tp. Hồ Chí Minh; chaudn@hcmut.edu.vn, do-hong@hcmut.edu.vn

Tóm tắt - Nhận dạng khuôn mặt là một trong những vấn đề quan trọng trong hướng nghiên cứu về nhận dạng của ngành thị giác máy tính. Do tính giống nhau của khuôn mặt nên việc trích ra các đặc trưng của khuôn mặt dùng cho nhận dạng là rất khó. Trong các đặc trưng của khuôn mặt dùng để nhận dạng thì đặc trưng về cạnh là một đặc trưng chỉ mới được nghiên cứu và phát triển trong những năm gần đây. Bài báo này sẽ giới thiệu tổng quát về một hướng nghiên cứu nhận dạng khuôn mặt dựa trên bản đồ cạnh (edge map) của khuôn mặt. Việc tính toán sự trùng khớp sẽ dựa trên khoảng cách Hausdorff. Các mô phỏng sẽ so sánh sự chính xác của việc nhận dạng khuôn mặt dựa vào bản đồ cạnh, với phương pháp rất phổ biến của nhận dạng khuôn mặt là Eigenface. Các kết quả cũng chỉ ra rằng việc nhận dạng khuôn mặt dựa vào bản đồ cạnh cho kết quả nhận dạng chính xác cao hơn phương pháp Eigenface trong hầu hết các so sánh.

Từ khóa - nhận dạng khuôn mặt; bản đồ cạnh; khoảng cách Hausdorff; đặc trưng khuôn mặt; Eigenface.

1. Giới thiệu

Tự động nhận dạng khuôn mặt là một hướng nghiên cứu thú vị đã thu hút được rất đông các nhà nghiên cứu trong khoảng hơn 20 năm qua. Từ khi bắt đầu ra đời đến nay, hướng nghiên cứu về nhận dạng khuôn mặt đã thu hút được rất nhiều nhà nghiên cứu trên toàn thế giới. Chính vì là một hướng nghiên cứu thu hút nên đã có rất nhiều phương pháp khác nhau về nhận dạng khuôn mặt đã được đề xuất. Các nghiên cứu về nhận dạng khuôn mặt có thể được chia thành các nhóm chính sau: Eigenface, mạng thần kinh nhân tạo, mô hình Markov ẩn, nhận dạng dựa vào các đặc trưng hình học (geometrical feature matching) và nhận dạng mẫu (template matching). Trong khi các hướng nghiên cứu về mạng thần kinh nhân tạo và template matching cho tỷ lệ nhận dạng chính xác cao, nhưng yêu cầu phải có nhiều ảnh làm cơ sở dữ liệu cho cùng một đối tượng, thì sẽ không thích hợp với các ứng dụng mà chỉ có một ảnh cho một đối tượng để nhận dạng. Trong khi đó, các nghiên cứu dựa vào mô hình Markov ẩn và đặc trưng hình học thì tỷ lệ chính xác lại phụ thuộc rất nhiều vào trực giác chọn các thông số huấn luyện cũng như thời gian nhận dạng khá lớn. Eigenface tuy là một hướng nghiên cứu đã lâu nhưng lại đơn giản, cho kết quả nhận dạng chính xác tương đối và dễ dàng áp dụng cho các ứng dụng đòi hỏi chỉ có một ảnh làm cơ sở dữ liệu cho một đối tượng.

Trong vấn đề nhận dạng, cạnh cũng là một đặc trưng rất hay được sử dụng. Tuy nhiên trong nhận dạng khuôn mặt, đặc trưng cạnh của khuôn mặt vẫn không được sử dụng khi nghiên cứu. Takács [1] là người đầu tiên sử dụng đặc trưng cạnh của khuôn mặt trong việc nhận dạng khuôn mặt. Tuy nhiên các nghiên cứu của Takács và sau này chỉ dựa vào các điểm trên cạnh của khuôn mặt nên không cho thông tin chính xác cao về khuôn mặt. Y. Gao và K. H. Leung [2] đã đưa ra một phương pháp nhận dạng khuôn mặt dựa trên các

Abstract - Face recognition is an important issue in research of pattern recognition in computer vision. Due to the similarity of human faces, face feature extraction for recognition presents a significant challenge for researchers. Edge map, one of human face's feature, however, has not been used for face recognition until recent years. This paper presents a brief review of face recognition based on edge map. Matching of human faces is carried out using Hausdorff distance. Experiment results of face recognition based on edge map will be compared with the Eigenface method, which is a common method in face recognition. It is very encouraging that in most comparison experiments, face recognition based on edge map gives higher accuracy than that based on Eigenface.

Key words - face recognition; edge map; Hausdorff distance; face feature; Eigenface.

đường trong bản đồ cạnh của khuôn mặt (Line Edge Map – LEM). Phương pháp này cho tỷ lệ nhận dạng chính xác rất cao, và cũng cho thấy sự bền vững của việc nhận dạng khuôn mặt trong các điều kiện khác nhau về ánh sáng, cũng như việc thay đổi cảm xúc trên khuôn mặt. Bài báo này sẽ trình bày về hướng nghiên cứu nhận dạng khuôn mặt dựa trên bản đồ cạnh, từ việc sử dụng các điểm trong bản đồ cạnh của Takács cho đến việc sử dụng các đường trong bản đồ cạnh của Y. Gao. Chúng tôi sẽ sử dụng phương pháp Eigenface như là một phương pháp so sánh với các phương pháp được trình bày trong bài báo này.

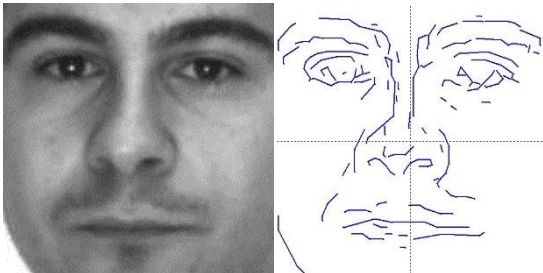
Phần còn lại của bài báo sẽ được trình bày như sau: Phần 2 giới thiệu về các phương pháp khác nhau sử dụng bản đồ cạnh để nhận dạng khuôn mặt. Phần 3 mô phỏng và so sánh kết quả của các thuật toán được nêu trong phần trên. Bài báo sẽ kết thúc tại Phần 4 với một số bình luận.

2. Nhận dạng khuôn mặt dựa vào bản đồ cạnh

2.1. Bản đồ cạnh của một khuôn mặt

Trong vấn đề xử lý ảnh, cạnh (edge) được định nghĩa là sự thay đổi độ sáng đột ngột giữa các pixel, là kết quả của sự thay đổi về cấu trúc hình học của các thành phần trong vật thể. Vì thế, cạnh cũng là một đặc trưng quan trọng và được dùng để nhận dạng. Có rất nhiều phương pháp tìm cạnh trong một bức ảnh [5-7] đã được đề xuất. Mỗi phương pháp tìm cạnh sẽ có ưu và nhược điểm khác nhau. Heath [8] đã chỉ ra rằng không có một phương pháp tìm cạnh nào là ưu điểm tuyệt đối và rất khó để có thể tìm được một phương pháp tìm cạnh tốt nhất cho một bức ảnh bất kỳ. Trong các mô phỏng của bài báo này, chúng tôi sẽ sử dụng chương trình LEM Generation của tác giả Y. Gao để tạo ra các bản đồ cạnh cho một bức ảnh khuôn mặt. Trong chương trình này, tác giả sử dụng phương pháp tìm cạnh của Babu [9] kết hợp với phương pháp làm mỏng cạnh để tạo ra các

cạnh có độ dày 1 pixel của bức ảnh. Tuy nhiên số pixel của cạnh vẫn còn rất lớn, dẫn tới việc tăng khối lượng tính toán không cần thiết. Áp dụng thuật toán Dynamic-two-Strip Dyn2S [10] trên bản đồ cạnh này ta sẽ tìm được các điểm trội. Các điểm trội này có độ uốn cong lớn nhất. Trong giới hạn bài báo này, chúng tôi chỉ tóm tắt các ý chính của thuật toán Dyn2S. Chi tiết của thuật toán Dyn2S được trình bày tại [10]. Tại mỗi điểm trên đường cong, chúng ta sử dụng hai dải chữ nhật ở bên trái và bên phải của điểm đó. Các điểm còn lại trên đường cong nếu nằm trong hai dải đó được xem xấp xỉ như trên cùng một đường thẳng với điểm đang xét. Nếu ta đặt tỉ số của chiều dài và chiều rộng của mỗi dải là E , góc tạo bởi giữa hai dải là θ thì tại mỗi điểm, ta sẽ có một chỉ số merit được tính bởi $W = E^{left} \cdot S \cdot E^{right}$, trong đó $S = |180^\circ - \theta|$, E^{left} và E^{right} là tỉ số chiều dài và rộng của hai dải bên trái và phải của điểm chúng ta đang xét. Chiều dài của các dải là như nhau. Các dải sẽ được thay đổi chiều rộng trong một khoảng giới hạn và thay đổi góc quay để làm sao có được nhiều điểm nằm trong dải nhất. Tại tất cả các điểm trên cạnh, ta đều làm tương tự. Sau đó, các điểm trội trên cạnh sẽ được chọn như sau: i) những điểm có chỉ số W nhỏ hơn hai điểm bên cạnh sẽ được bỏ đi; ii) với những điểm còn lại, những điểm nào có thể được xấp xỉ là một đường thẳng với điểm đang xét thì ta bỏ đi, chỉ giữ lại điểm cuối của hai dải bên trái và bên phải. Như vậy, sau khi kết hợp giữa thuật toán tìm cạnh và Dyn2S lên một bức ảnh khuôn mặt, ta sẽ được bản đồ cạnh như Hình 1.



Hình 1. Bản đồ cạnh (edge map) của khuôn mặt

2.2. Khoảng cách Hausdorff

Khoảng cách Hausdorff là khoảng cách được dùng để tính cho khoảng cách giữa hai tập hợp điểm, nhưng không cần xét đến sự tương ứng điểm – điểm giữa hai tập hợp như các khoảng cách khác. Huttenlocher [11] đã ứng dụng khoảng cách Hausdorff để so sánh sự giống nhau giữa các bức ảnh.

Cho tập hợp các điểm trên bản đồ cạnh của một bức ảnh trong cơ sở dữ liệu là $C = \{c_1, c_2, \dots, c_p\}$ và của một bức ảnh cần nhận dạng là $N = \{n_1, n_2, \dots, n_m\}$. Khoảng cách Hausdorff của hai tập điểm được định nghĩa:

$$h(C, N) = \max(h(C, N), h(N, C)) \quad (1)$$

trong đó:

$$h(C, N) = \max_{c_i \in C} \min_{n_j \in N} \|c_i - n_j\| \quad (2)$$

và $\|c_i - n_j\|$ là khoảng cách Euclid giữa hai điểm c_i và n_j . Khoảng cách $h(C, N)$ được gọi là khoảng cách Hausdorff từ ảnh C đến ảnh N . Như vậy, với cách định nghĩa trên thì

khoảng cách Hausdorff giữa hai bức ảnh $H(C, N)$ chính là chỉ số dùng để đo sự khác nhau giữa hai bức ảnh.

Tuy nhiên với cách định nghĩa trên, khoảng cách Hausdorff rất nhạy với các điểm đặc biệt trong bức ảnh. Ví dụ, nếu vì một lý do nào đó mà trong bản đồ cạnh có vài điểm, hoặc thậm chí một điểm nằm cách biệt ra ngoài, khác với những điểm khác; thì với cách định nghĩa của khoảng cách Hausdorff như phương trình (1) và (2) thì khoảng cách này sẽ do điểm đặc biệt đó quyết định. Như vậy, như vật thể hoặc khuôn mặt có hình dáng rất khác nhau, nhưng nếu có một vài hoặc một điểm đặc biệt thì sẽ trở nên rất giống nhau, nếu chỉ xét khoảng cách Hausdorff như trên. Dubuisson và Jain [12] đã chỉ ra một phương pháp cải tiến của khoảng cách Hausdorff (Modified Hausdorff Distance – MHD) để loại bỏ nhược điểm của phương pháp tính khoảng cách Hausdorff trực tiếp như phương trình (2). Phương pháp MHD này đã được Takács [1] ứng dụng vào việc tính khoảng cách giữa các bản đồ cạnh của khuôn mặt trong việc nhận dạng khuôn mặt. Tuy nhiên, bản đồ cạnh của khuôn mặt được Takács sử dụng để tính toán chỉ là cạnh của bức ảnh khi được áp dụng phương pháp tìm cạnh kết hợp với phương pháp làm mỏng cạnh, chứ chưa sử dụng thuật toán Dyn2S để tìm các điểm trội như đã nói ở trên. Khoảng cách MHD được định nghĩa như sau:

$$h_{MHD}(C, N) = \frac{1}{P} \sum_{c_i \in C} \min_{n_j \in N} \|c_i - n_j\| \quad (3)$$

trong đó P là số điểm trong C . Với cách định nghĩa này sẽ làm giảm sự tác động của các điểm đột biến đến khoảng cách giữa hai bức ảnh.

Gao [13] đã đề xuất một phương pháp khác cũng dựa trên việc tính toán khoảng cách Hausdorff để nhận dạng ảnh. Trong phương pháp này, Gao đã sử dụng thuật toán Dyn2S để tìm các điểm trội trên cạnh khuôn mặt. Chính việc dùng thuật toán Dyn2S đã làm giảm số lượng điểm trên cạnh rất nhiều, dẫn đến việc làm giảm khối lượng tính toán. Trong phương pháp của mình, Gao đã định nghĩa một khoảng cách Hausdorff từ ảnh C đến ảnh N như sau:

$$h_{MMHD}(C, N) = \frac{1}{\sum_{c_i, n_j} W_{c_i, n_j}} \sum_{c_i \in C} W_{c_i, n_j} \min_{n_j \in N} \|c_i - n_j\| \quad (4)$$

trong đó $W_{c_i, n_j} = 1/2 (W_{c_i} + W_{n_j})$ là trung bình của chỉ số merit tại hai điểm c_i và n_j trong thuật toán Dyn2S. Như vậy, khoảng cách Hausdorff của hai bức ảnh sẽ là:

$$H_{MMHD}(C, N) = \max(h_{MMHD}(C, N), h_{MMHD}(N, C)) \quad (5)$$

Dựa trên kết quả các mô phỏng trong [13] của Gao được thực hiện trên cơ sở dữ liệu AR và Bern, các kết quả cho thấy rằng tỷ lệ chính xác của phương pháp MHD của Takács và MMHD của Gao là tương đương nhau. Như vậy, với đề xuất phương pháp tính khoảng cách Hausdorff như tại phương trình (4) và (5) thì không làm tăng thêm tính chính xác của thuật toán so với phương trình (3). Vậy, đóng góp của Gao trong bài báo này thực chất là việc áp dụng thuật toán Dyn2S vào bản đồ cạnh để làm giảm bớt đi số điểm cần tính toán, dẫn đến làm giảm khối lượng tính toán mà không làm giảm đi độ chính xác. Việc áp dụng thuật toán Dyn2S làm giảm đến 80%

số điểm cần tính toán trên bản đồ cạnh. Vì tỷ lệ chính xác của MHD và MMHD như nhau nên trong các mô phỏng của bài báo này, chúng tôi sẽ sử dụng thuật toán MHD, nhưng trên bản đồ cạnh khuôn mặt đã được áp dụng thuật toán Dyn2S như Hình 1.

Một nhược điểm của thuật toán MHD và MMHD là việc coi tất cả các điểm trên bản đồ cạnh là các điểm độc lập, vì vậy, hai điểm cạnh nhau cũng giống như hai điểm nằm xa nhau và không có thông tin về sự liên hệ giữa các điểm với nhau. Gao và Leung [2] cho rằng các điểm trên cùng một cạnh được nối với nhau bằng một đường thẳng. Như vậy, thay vì dùng khoảng cách Hausdorff giữa hai tập hợp điểm, thì Gao đã đưa ra cách tính khoảng cách Hausdorff giữa hai tập hợp đường, được gọi là Line Segment Hausdorff Distance (LHD). Khác với các phương pháp khác sử dụng cách tính khoảng cách giữa tập hợp các đường, giống như MHD và MMHD, phương pháp LHD không phụ thuộc vào sự tương ứng giữa các đường – đường trong ảnh cơ sở dữ liệu và ảnh cần nhận dạng. Chính vì ưu điểm này mà phương pháp LHD vẫn có thể được dùng ngay cả khi một bức ảnh bị mất một số cạnh do lỗi trong quá trình phân đoạn ảnh.

Cho tập hợp các đường trên bản đồ cạnh của một bức ảnh trong cơ sở dữ liệu là $C^l = \{c_1^l, c_2^l, \dots, c_p^l\}$ và t của một bức ảnh cần nhận dạng là $N^l = \{n_1^l, n_2^l, \dots, n_m^l\}$. Khoảng cách giữa 2 đường bất kỳ được định nghĩa như sau:

$$d(c_i^l, n_j^l) = \sqrt{d_\theta^2(c_i^l, n_j^l) + d_{ss}^2(c_i^l, n_j^l) + d_{vg}^2(c_i^l, n_j^l)} \quad (6)$$

trong đó: $d_\theta(c_i^l, n_j^l)$, $d_{ss}(c_i^l, n_j^l)$, $d_{vg}(c_i^l, n_j^l)$ lần lượt là khoảng cách góc, khoảng cách song song và khoảng cách vuông góc của hai đường c_i^l và n_j^l . Trong đó, khoảng cách góc được định nghĩa là:

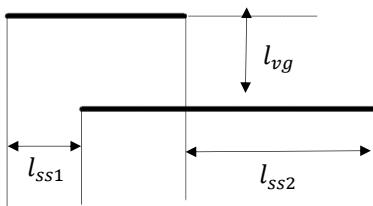
$$d_\theta(c_i^l, n_j^l) = \theta^2(c_i^l, n_j^l) / W \quad (7)$$

với $\theta(c_i^l, n_j^l)$ là góc giao nhau nhỏ nhất giữa hai đường c_i^l và n_j^l . Hệ số W là một trọng số tùy chọn và được xác định trong quá trình huấn luyện tập ảnh.

Để tính khoảng cách song song và vuông góc giữa hai đường thẳng, chúng ta sẽ xoay đường thẳng có độ dài ngắn hơn để song song với đường thẳng dài hơn. Khi hai đường thẳng song song như Hình 2, khoảng cách song song và vuông góc sẽ được tính như sau:

$$d_{ss}(c_i^l, n_j^l) = \min(I_{ss1}, I_{ss2}) \quad (8)$$

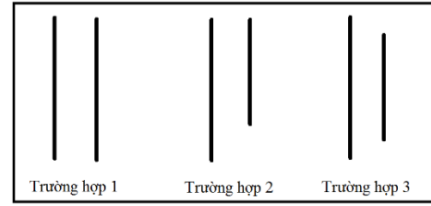
$$d_{vg}(c_i^l, n_j^l) = I_{vg} \quad (9)$$



Hình 2. Khoảng cách giữa hai đường thẳng song song

Khoảng cách song song sẽ được tính là khoảng cách nhỏ nhất giữa điểm ngoài cùng bên trái và ngoài cùng bên phải của hai đường thẳng. Khoảng cách song song sẽ bằng

0 khi hai đường thẳng thuộc một trong ba trường hợp được mô tả trong Hình 3.



Hình 3. Các trường hợp $d_{ss} = 0$

Như vậy khoảng cách LHD gốc (Primary LHD – pLHD) giữa hai bản đồ cạnh LEM được định nghĩa như sau:

$$H_{pLHD}(C^l, N^l) = \max(h(C^l, N^l), h(N^l, C^l)) \quad (10)$$

với:

$$h(C^l, N^l) = \frac{1}{\sum_{c_i^l \in C^l} I_{c_i^l}} \sum_{c_i^l \in C^l} \min_{n_j^l \in N^l} d(c_i^l, n_j^l) \quad (11)$$

và $I_{c_i^l}$ là chiều dài của đoạn thẳng c_i^l . Trong phương trình (11), ta thấy rằng khoảng cách của mỗi đường thẳng sẽ được nhân với một trọng số là chiều dài của chính nó.

Tuy nhiên, với khoảng cách pLHD được tính như phương trình (11) thì có một điểm yếu. Giả sử chúng ta có N^l là LEM của khuôn mặt cần nhận dạng, n_j^l là một đường trong LEM đó; C_g^l và C_k^l lần lượt là LEM của khuôn mặt nhận dạng đúng của N^l và khuôn mặt khác với N^l trong cơ sở dữ liệu. Nếu vì một lý do nào đó mà đường tương ứng với n_j^l trong C_g^l là c_{gj}^l bị mất đi, khi đó, đường n_j^l trong N^l sẽ có khoảng cách gần nhất tới một đường khác trong C_g^l , giả sử là đường c_{gi}^l . Khi đó khoảng cách $d(c_{gi}^l, n_j^l)$ có thể sẽ lớn hơn rất nhiều so với $d(c_{kj}^l, n_j^l)$, với c_{kj}^l là đường trong C_k^l có khoảng cách gần nhất tới n_j^l . Điều này dẫn tới việc N^l sẽ có khoảng cách tới C_k^l gần hơn so với C_g^l và dẫn tới nhận dạng sai.

Để khắc phục điều này, Gao đưa thêm vào một thông số nữa vào khoảng cách Hausdorff, đó là tỉ số tin cậy. Nếu một đường thẳng c_i^l trong LEM C^l có khoảng cách gần nhất tới đường thẳng n_j^l trong LEM N^l , và hai đường nào có khoảng cách góc nhỏ hơn một lượng K_g , và khoảng cách giữa hai trung điểm của hai đường thẳng nhỏ hơn một lượng K_{vt} , thì khi đó đường n_j^l được xem là tin cậy với đường c_i^l . Khi đó, tỉ số tin cậy của một bức ảnh được định nghĩa là tỉ số giữa tổng số đường tin cậy D_{tc} và tổng số đường trong LEM D_{total} của bức ảnh.

$$R = \frac{D_{tc}}{D_{total}} \quad (12)$$

Như vậy, tổng số đường nằm ngoài vùng tin cậy của hai bức ảnh sẽ là:

$$D_n = 1 - \frac{R_C + R_N}{2} \quad (13)$$

Vậy khoảng cách Hausdorff giữa hai bức ảnh sẽ là:

$$H_{LHD}(C', N') = \sqrt{H_{pLHD}^2(C', N') + (W_n D_n)^2} \quad (14)$$

với W_n là một trọng số cho tổng số các đường nằm ngoài vùng tin cậy.

Chúng tôi sử dụng phương pháp mô phỏng annealing [14], là một phương pháp dựa trên kỹ thuật tối ưu ngẫu nhiên, trên việc nhận dạng các khuôn mặt trong cơ sở dữ liệu Bern và tìm ra được các chỉ số như sau: $W = 30$, $K_g = 20$, $K_{vt} = 5$ và $W_n = 6$. Chúng tôi sử dụng các giá trị này cho các mô phỏng còn lại.

Mihir et al. [15] đã đề xuất phương pháp trích xuất các vùng là mắt, mũi, miệng từ bản đồ cạnh để nhận dạng. Mục đích của việc này là làm giảm bớt số đường cạnh dư thừa trong bản đồ cạnh, mà theo tác giả là không cần thiết trong việc tính toán vào trong khoảng cách Hausdorff. Để có thể trích xuất các vùng này, tác giả đã tạo ra bản đồ cạnh nhân tạo để thể hiện các vùng mắt, mũi, miệng của một khuôn mặt người, và sử dụng sự tương quan của mô hình này với các bản đồ cạnh để trích xuất ra các vùng trên bản đồ cạnh. Như vậy, khoảng cách giữa 2 bức ảnh sẽ là tổng khoảng cách LHD như [2] cho từng vùng và được cộng lại theo hệ số mắt:mũi:miệng là 2:1:1. Các kết quả tại [15] cho thấy phương pháp này cho kết quả nhận dạng chính xác cao hơn phương pháp LHD của Gao [2] khoảng 5%.

3. Kết quả thí nghiệm

Trong bài báo này, chúng tôi sẽ mô phỏng và tính toán độ chính xác của các phương pháp nhận dạng với các điều kiện khác nhau của một bức ảnh khuôn mặt: i) điều kiện lý tưởng; ii) điều kiện ánh sáng khác nhau và iii) điều kiện các góc chụp khác nhau của khuôn mặt. Để so sánh giữa các phương pháp nhận diện khác nhau, chúng tôi sẽ sử dụng phương pháp Eigenface như là phương pháp chuẩn để so sánh tỷ lệ chính xác, với các phương pháp nhận diện khuôn mặt dựa trên bản đồ cạnh là MHD và LHD. Trong bài báo này, chúng tôi không chọn phương pháp của Mihir [15] để mô phỏng và so sánh, vì lý do các thông số của thuật toán không được công bố; ngoài ra việc huấn luyện cũng không được đề cập nên việc nghiên cứu này sử dụng chung một cơ sở dữ liệu cho cả quá trình huấn luyện, và nhận dạng cũng là một vấn đề chưa sáng tỏ. Vì thế, trong bài báo này chúng tôi chỉ giới hạn ở việc giới thiệu ý tưởng của thuật toán chứ không đi sâu vào việc mô phỏng và phân tích.

Trong các mô phỏng, chúng tôi sẽ sử dụng hai cơ sở dữ liệu chuẩn về khuôn mặt, thường được dùng trong các nghiên cứu về nhận diện khuôn mặt là cơ sở dữ liệu Bern [3] và AR [4]. Trong cơ sở dữ liệu của Bern có ảnh của 30 người khác nhau với các điều kiện ánh sáng lý tưởng như Hình 4. Mỗi người sẽ có 10 bức ảnh xám với các góc chụp khác nhau của khuôn mặt (2 ảnh chụp chính diện, 2 đầu xoay sang trái, 2 ảnh đầu xoay sang phải, 2 ảnh ngược nhìn lên và 2 ảnh ngược nhìn xuống). Trong cơ sở dữ liệu của AR Database, có tất cả ảnh của 100 người, gồm 50 nam và 50 nữ. Mỗi người sẽ có tất cả 26 ảnh khác nhau, trong đó có 13 ảnh gồm ảnh chụp trong điều kiện lý tưởng, các ảnh với các điều kiện ánh sáng khác nhau (gồm chiếu sáng bên phải, bên trái và chiếu sáng từ 2 bên), với các thay đổi khác nhau khi mang các vật dụng như kính, khăn choàng; 13 ảnh còn lại cũng

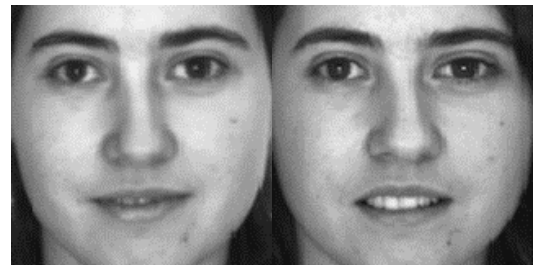
giống như 13 ảnh đầu tiên nhưng được chụp sau đó 2 tuần, để tạo ra sự thay đổi trên khuôn mặt như Hình 5. Tất cả các bức ảnh sẽ được tiền xử lý để tạo thành các bức ảnh xám 160x160 pixel, và sau đó tạo ra các bản đồ cạnh để phục vụ cho việc mô phỏng phương pháp MHD và LHD.

3.1. Nhận dạng khuôn mặt dưới các điều kiện lý tưởng

Các bức ảnh chụp chính diện dưới các điều kiện ánh sáng lý tưởng của cơ sở dữ liệu Bern và AR được dùng để mô phỏng các thuật toán nhận dạng. Đối với Bern, ta có 30 ảnh dùng làm cơ sở dữ liệu và 30 ảnh khác để nhận dạng. Đối với AR database, chúng ta có 100 ảnh để làm cơ sở dữ liệu và 100 ảnh để nhận dạng. Tỷ lệ nhận dạng chính xác của các thuật toán với 2 cơ sở dữ liệu được liệt kê trong Bảng 1.



Hình 4. Cặp ảnh chính diện của một người trong cơ sở dữ liệu Bern



Hình 5. Cặp ảnh chính diện của một người trong cơ sở dữ liệu AR chụp cách nhau 2 tuần

Bảng 1. Kết quả nhận dạng với điều kiện lý tưởng

Thuật toán	Bern	AR
Eigenface – 20 eigenvectors	100%	53%
MHD	100%	69%
LHD	100%	93%

Đối với cơ sở dữ liệu Bern, tất cả các thuật toán đều đạt nhận dạng chính xác đến 100%, là do rất ít sự khác nhau giữa các bức ảnh dùng làm cơ sở dữ liệu và dùng để nhận dạng. Tuy nhiên, đối với cơ sở dữ liệu AR, lại có sự khác nhau rất lớn giữa cặp ảnh vì các tấm ảnh được chụp cách nhau 2 tuần. Đối với cơ sở dữ liệu AR, các thuật toán có tỷ lệ chính xác khác nhau rất lớn. Bảng 2 so sánh độ chính xác của các thuật toán đối với cơ sở dữ liệu AR.

Bảng 2. Tỷ lệ chính xác trên cơ sở dữ liệu AR

Thuật toán	Tỷ lệ chính xác
LHD	93%
MHD	69%
Eigenface – 20 eigenvectors	53%
Eigenface – 60 eigenvectors	60%
Eigenface – 100 eigenvectors	63%

Tỷ lệ chính xác của phương pháp Eigenface phụ thuộc rất nhiều vào số vector riêng được chọn. Càng tăng số lượng vector riêng, sẽ càng làm tăng tỷ lệ chính xác. Tuy nhiên, vì số lượng ảnh trong tập cơ sở dữ liệu là 100, nên số lượng vector riêng lớn nhất là 100. Vậy, tỷ lệ chính xác lớn nhất của Eigenface đối với cơ sở dữ liệu AR là 63%, thấp hơn nhiều so với phương pháp MHD và LHD.

3.2. Nhận dạng khuôn mặt với các điều kiện khác nhau của ánh sáng

Trong mô phỏng này, chúng tôi sử dụng 100 ảnh chính diện của 100 người trong cơ sở dữ liệu AR để làm cơ sở dữ liệu. Với mỗi người sẽ có 2 tấm ảnh với đèn chiếu từ bên trái, 2 tấm với đèn chiếu từ bên phải và 2 tấm với đèn chiếu từ cả 2 bên, vậy ta sẽ có 300 ảnh để nhận dạng. Tỷ lệ nhận dạng chính xác của các thuật toán được liệt kê trong Bảng 3.

Trong tất cả các kết quả của các điều kiện ánh sáng khác nhau tác động đến khuôn mặt, phương pháp LHD đều cho kết quả cao hơn hẳn so với Eigenface, cho dù ta chọn số vector riêng cực đại – đạt tỷ lệ chính xác cao nhất. Trong phương pháp Eigenface, ba thành phần cơ bản đầu tiên, tương ứng là ba vector riêng có trị riêng lớn nhất, là các thành phần nhạy cảm với sự thay đổi của điều kiện ánh sáng của bức ảnh. Nếu ta bỏ đi ba thành phần đó, tỷ lệ chính xác sẽ tăng lên nhưng vẫn không bằng với phương pháp MHD, và nhất là phương pháp LHD. Kết quả mô phỏng này chứng tỏ phương pháp LHD rất bền vững với các điều kiện ánh sáng khác nhau, phù hợp với thực tế ứng dụng của việc nhận dạng khuôn mặt.

Bảng 3. Kết quả nhận dạng với các điều kiện chiếu sáng khác nhau

Điều kiện chiếu sáng	Eigenface		MHD	LHD
Trái	100 eigenvectors	5%	74%	93%
	100 eigenvector w/o 1 st 3	48%		
Phải	100 eigenvectors	6%	66%	87%
	100 eigenvector w/o 1 st 3	40%		
Trái + Phải	100 eigenvectors	2%	50%	71%
	100 eigenvector w/o 1 st 3	43%		

3.3. Nhận dạng khuôn mặt với các góc chụp khác nhau của khuôn mặt

Trong mô phỏng này, chúng tôi dùng 30 ảnh chính diện của 30 người trong cơ sở dữ liệu Bern để làm cơ sở dữ liệu. Mỗi người sẽ có 8 ảnh khác nhau về góc chụp khuôn mặt, vậy chúng ta có 240 ảnh cần nhận diện. Kết quả của việc ảnh hưởng của góc chụp đến tỷ lệ chính xác được nêu trong Bảng 4. Trong mô phỏng này, chúng tôi sử dụng 30 vector riêng cho phương pháp Eigenface, là số lượng vector riêng lớn nhất có thể chọn.

Bảng 4. Tỷ lệ nhận dạng chính xác với các góc chụp khác nhau của khuôn mặt

Góc chụp	Eigenface – 30 eigenvectors	MHD	LHD
Nhìn trái	70%	56,6%	46,67%
Nhìn phải	70%	50%	53,33%

Nhìn lên	50%	60%	70%
Nhìn xuống	43,33%	53,3%	63,33%
Trung bình	58,33%	55%	58,33%

Trong mô phỏng này, chúng ta thấy rằng góc chụp của khuôn mặt ảnh hưởng lớn đến kết quả chính xác của nhận dạng. Phương pháp Eigenface cho kết quả cao hơn so với LHD khoảng 1% là không quá nhiều.

3.4. Nhận dạng khuôn mặt với các cảm xúc khác nhau của khuôn mặt

Trong mô phỏng này, chúng tôi sử dụng 100 ảnh chính diện của 100 người trong cơ sở dữ liệu AR để làm cơ sở dữ liệu. Với mỗi người, sẽ có 3 tấm ảnh về cảm xúc là cười, giận dữ và hốt hoảng. Vậy ta sẽ có 100 tấm ảnh cho mỗi cảm xúc để nhận dạng. Tỷ lệ nhận dạng chính xác của các thuật toán được liệt kê trong Bảng 5.

Bảng 5. Tỷ lệ nhận dạng chính xác với các cảm xúc khác nhau của khuôn mặt

Cảm xúc	Eigenface		MHD	LHD
Cười	100 eigenvectors	91%	42%	81%
	100 eigenvector w/o 1 st 3	74%		
Giận dữ	100 eigenvectors	88%	85%	90%
	100 eigenvector w/o 1 st 3	71%		
Hốt hoảng	100 eigenvectors	38%	16%	32%
	100 eigenvector w/o 1 st 3	20%		

Trong mô phỏng này, chúng ta thấy phương pháp LHD cho kết quả chính xác cao hơn so với MHD nhưng thấp hơn so với Eigenface. Tuy nhiên trong điều kiện ứng dụng thực tế với ánh sáng không lý tưởng, ta cần bỏ đi 3 vector riêng tương ứng 3 trị riêng lớn nhất cho phương pháp Eigenface. Để nâng cao tỷ lệ nhận dạng chính xác thì phương pháp LHD cho kết quả tốt hơn Eigenface.

4. Kết luận

Trong bài báo này, chúng tôi đã trình bày khái quát về một hướng nghiên cứu mới trong nhận diện khuôn mặt, là kết hợp giữa khoảng cách Hausdorff và bản đồ cạnh của khuôn mặt, và phương pháp tốt nhất là LHD. Đây là phương pháp còn khá mới trong việc nhận diện khuôn mặt.

Các kết quả mô phỏng cho thấy phương pháp MHD và LHD cho kết quả vượt trội so với phương pháp Eigenface, vốn rất hay được dùng trong các ứng dụng về nhận dạng khuôn mặt. Phương pháp LHD còn cho thấy sự bền vững với các điều kiện khác nhau của bức ảnh khuôn mặt như điều kiện ánh sáng thay đổi, góc chụp thay đổi và cảm xúc khuôn mặt thay đổi. Với các ưu điểm đó, phương pháp nhận diện khuôn mặt dựa trên đặc trưng cạnh, nhất là phương pháp LHD, là phương pháp rất tốt để ứng dụng cho việc nhận dạng khuôn mặt.

TÀI LIỆU THAM KHẢO

- [1] B. Takács, "Comparing face images using the modified Hausdorff distance", *Pattern Recognition*, Vol. 31, 1998, pp. 1873-1881.
- [2] Y. Gao and M. K. Leung, "Face recognition using line edge map",

- IEEE Trans. on Pattern and Analysis and Machine Intelligence*, Vol. 24, No. 6, Jun 2002, pp. 764-779.
- [3] Bern University Database, <http://www.fki.inf.unibe.ch/databases/iam-faces-database/FullFaces.tgz>
- [4] A.M. Martinez and R. Benavente, *The AR Face Database*, CVC Technical Report, No. 24, June 1998.
- [5] L. A. Iverson and S. W. Zucker, "Logical/linear operators for image curves", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, 1995, pp. 982-996.
- [6] I. E. Sobel, *Cameral model and machine perception*, PhD thesis, Stanford University, USA, 1970.
- [7] J. Canny, "A computational approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, 1986, pp. 679-698.
- [8] Heath, M.D., Sarkar, S., Sanocki, T., and Bowyer, K.W., "Comparison of edge detectors: A methodology and initial study", *Comput. Vis. Image Underst.*, Vol. 69, 1998, pp. 38-54.
- [9] Nevatia, R., and Babu, K.R., "Linear feature extraction and description", *Comput. Graph. Image Process.*, Vol. 13, 1980, pp. 257-269.
- [10] Leung, M.K.H., and Yang, Y.H., "Dynamic two-strip algorithm in curve fitting", *Pattern Recognition*, Vol. 23, 1990, pp. 69-79.
- [11] Huttenlocher, D.P., Klanderman, G.A., and Rucklidge, W.J., "Comparing images using the Hausdorff distance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, 1993, pp. 850-863.
- [12] Dubuisson, M.P., and Jain, A.K., "A modified Hausdorff distance for object matching", *Proc. 12th Int. Conf. on Pattern recognition*, Jerusalem, Israel, 1994, pp. 566-568.
- [13] Y. Gao, "Efficiently comparing face images using a modified Hausdorff distance", *IEE Proc. Vision, Image and Signal Processing*, Vol. 150, No. 6, Dec 2003, pp. 346-350.
- [14] V. Granville, M. Krivanek, J. Rasson, "Simulated annealing: A proof of convergence", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 6, 1994, pp. 652-656.
- [15] J. Mihir, M. Suman and J. Naresh, "Eye detection and face recognition using line edge map", *Proc. National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, Gandhinagar, India, 2008.

(BBT nhận bài: 14/02/2017, hoàn tất thủ tục phản biện: 10/05/2017)