

CHƯƠNG 4

PHÁT HIỆN KHUÔN MẶT DỰA TRÊN ĐẶC TRƯNG LỖI LỒM

4.1 Tập mẫu học

Trong luận văn này chúng tôi thử nghiệm trên tập dữ liệu ảnh thu thập bởi Markus Weber [32]. Đây là tập dữ liệu chụp các khuôn mặt thẳng, chụp dưới nền phức tạp và đa dạng gồm cả trong nhà và ngoài trời. Điều kiện chiếu sáng khá đa dạng gồm cả các ảnh chụp cùng chiều chiếu sáng và ngược sáng. Ngoài ra còn có các ảnh chụp phức tạp do một phần của khuôn mặt bị che khuất. Ngoài ra, tập ảnh này còn chứa cả ảnh chụp lại từ tranh họa mặt người.



(a) Ảnh chụp trong nhà



(b) Ảnh chụp ngoài trời



(c) Ảnh chụp ngược sáng



(d) Ảnh chụp phức tạp do che khuất



(e) Ảnh chụp phức tạp do biểu cảm



(f) Ảnh chụp từ tranh họa

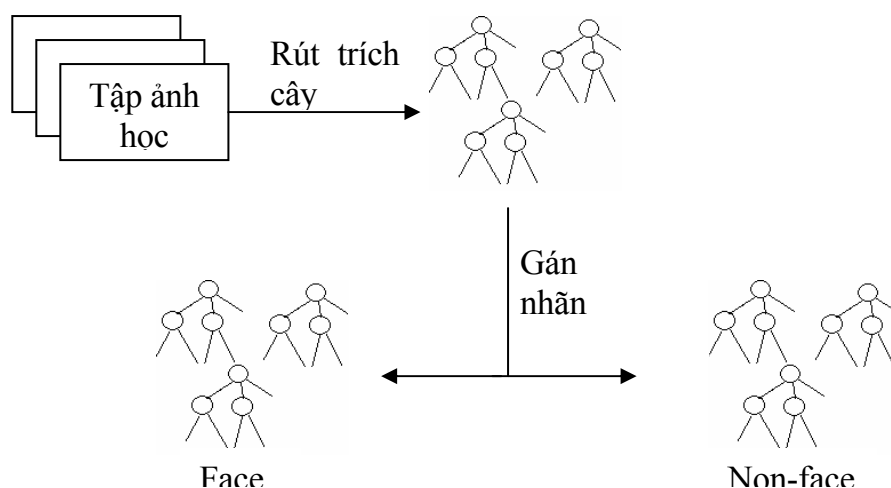
Hình 4.1 Một số ảnh trong tập ảnh Weber

Trong tập ảnh này có tất cả 450 ảnh. Chúng tôi sử dụng 200 ảnh để học và 250 ảnh còn lại để kiểm thử.

4.2 Mô hình thống kê

4.2.1 Gán nhãn

Từ tập ảnh học, chúng tôi dùng phương pháp trình bày trong chương 3 để rút trích các cây lỗi lờm trên ảnh. Các cây này được gán nhãn thủ công. Mỗi cây bất kỳ được gán nhãn là mặt (*face*) hay không phải khuôn mặt (*non-face*). Như vậy, từ một tập ảnh học ta rút trích được hai tập cây. Tập cây biểu diễn khuôn mặt và tập cây biểu diễn đối tượng không phải là khuôn mặt.



Hình 4.2 Mô hình gán nhãn cho tập học

4.2.2 Thống kê

Mục đích của bước này là xấp xỉ được một cấu trúc cây khuôn mặt. Trong bước gán nhãn, chúng tôi tiến hành rút trích cây và gán nhãn. Kết quả thu được 2036 cây thuộc lớp “*khuôn mặt*”. Các cây này được dùng làm dữ liệu học thống kê. Mô hình thống kê chúng tôi sử dụng như sau:

Đặt $S(T_x)$ là tổng khoảng cách của một cây T_x bất kỳ đến tất cả các cây còn lại trong tập học.

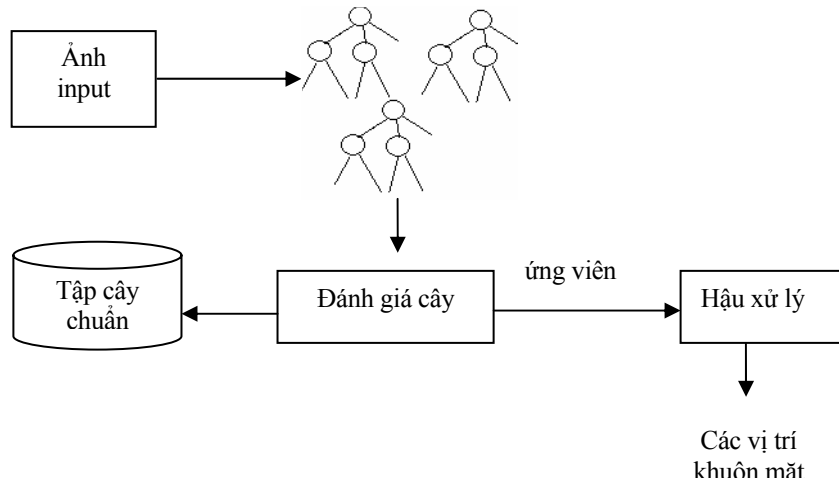
$$S(T_x) = \sum_{i=1}^{i=n \wedge i \neq x} TS(T_x, T_i) \quad (4.1)$$

Ta có tập học sẽ trở thành:

$$Training = \{(T_i, S(T_i)), \forall T_i \in FaceSet\} \quad (4.2)$$

Trong đó $FaceSet$ là tập các cây được gán nhãn là “*khuôn mặt*”. Việc học thống kê theo mô hình của chúng tôi khá đơn giản. Trước tiên, chọn trong tập $Training$ ra k cây có giá trị $S(T)$ bé nhất. k cây này sẽ là đại diện tiêu biểu cho một lớp khuôn mặt.

Từ kết quả k cây rút trích được, tạm gọi là tập cây chuẩn, chúng ta có thể dùng chúng để phát hiện khuôn mặt theo mô hình như sau:



Hình 4.3 Mô hình phát hiện khuôn mặt

Trong mô hình này, trước tiên từ ảnh đầu vào ta dùng phương pháp rút trích lỗi lờm ta sẽ nhận được một danh sách các cây lỗi lờm. Ứng với mỗi cây ta dựa vào tập cây chuẩn để kiểm tra xem liệu đây có phải là cây biểu diễn “khuôn mặt” hay không. Nếu thỏa điều kiện đánh giá nó sẽ trở thành ứng viên khuôn mặt. Trong bước cuối cùng, hậu xử lý, các ứng viên sẽ được kiểm tra ràng buộc. Và sau cùng là khử trùng lặp khuôn mặt do có một số cây có một phần của nút gốc chồng lên nhau.

4.2.3 Đánh giá dùng cho phát hiện khuôn mặt

Dựa trên tập cây chuẩn, một cây bất kỳ được đánh giá có khả năng là khuôn mặt hay không dựa trên trung bình khoảng cách của cây đang xét đến tất cả các cây trong tập cây chuẩn.

$$\overline{D}(t) = \frac{\sum_{i=1}^k \|t, T_i\|}{k} \quad (4.3)$$

Trong đó:

$\overline{D}(t)$: là trung bình khoảng cách của cây t đến các cây trong tập chuẩn

t : là cây đang xét

T_i : là các cây trong tập k cây chuẩn

Căn cứ vào giá trị \overline{D} , cây t được đánh giá là ứng viên mặt hay không dựa theo công thức sau:

$$t \in FaceCandidate \Leftrightarrow \overline{D}(t) \leq \delta_1 \quad (4.4)$$

Ngưỡng δ_1 là ngưỡng thực nghiệm, ngưỡng này sẽ được lấy khoảng 0.2 đến 0.4. Ngoài cách tính trung bình khoảng cách, ta có thể đánh giá ứng viên thông qua việc tính trung bình độ tương đồng của cây ứng viên với tập cây chuẩn.

$$\overline{S}(t) = \frac{\sum_{i=1}^k TS(t, T_i)}{k} \quad (4.5)$$

Trong đó:

$\overline{S}(t)$: là trung bình độ tương đồng của cây t so với các cây trong tập chuẩn

t : là cây đang xét

T_i : là các cây trong tập k cây chuẩn

Tương tự như trường hợp trên, giá trị $\overline{S}(t)$ sẽ được dùng để đánh xem cây t có khả năng là khuôn mặt hay không.

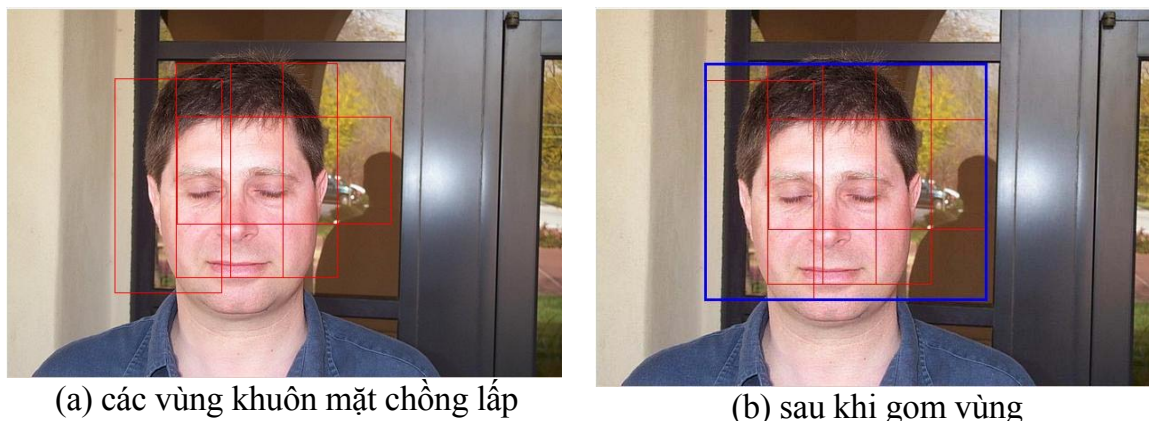
$$t \in FaceCandidate \Leftrightarrow \overline{S}(t) \geq \delta_2 \quad (4.6)$$

Dễ thấy ngưỡng δ_2 là số gần với một, có thể nói $\delta_2 \cong 1 - \delta_1$. Ngưỡng này sẽ được lấy trong khoảng từ 0.6 đến 0.8. Tuy nhiên, về bản chất hai cách này hoàn toàn giống nhau.

4.2.4 Hậu xử lý

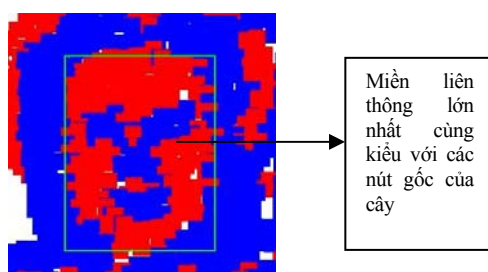
Các kết quả thu được cho thấy rằng có một vùng khuôn mặt được phát hiện có một phần chồng lấp lên nhau. Có thể phát hiện ra nhiều cây ứng viên được đánh giá là “*khuôn mặt*”, nhưng thật sự chúng là sự chồng lấp của nhiều cây lên một vị trí

khuôn mặt. Vì vậy, chúng ta cần một thuật toán tốt để giảm những phát hiện các ứng viên trùng này. Trước hết chúng tôi sử dụng một đánh giá đơn giản để liên kết các vùng trùng lặp. Nếu hai ứng viên khuôn mặt có vùng diện tích bị chồng lấp lên nhau quá một phần ba diện tích của nó, chúng được xem là các mặt trùng và được liên kết lại thành một vùng lớn hơn. Vùng diện tích xem xét đó chính là vùng diện tích của nút gốc của cây ứng viên.



Hình 4.4 Một ví dụ về giảm trùng lặp

Sau khi gom nhóm được vùng, chúng tôi thu được một hay nhiều vùng chứa khuôn mặt. Tuy nhiên, do sai số của kích thước của cây và kích thước khuôn mặt, nên chúng tôi luôn có được vùng khuôn mặt phát hiện được lớn hơn khuôn mặt thực tế. Chúng tôi áp dụng tiếp kỹ thuật xác định vùng lỗi lõm với mức nhỏ hơn nhằm để phát hiện chính xác vùng. Chúng tôi chỉ áp dụng thuật toán dò tìm đặc trưng lỗi lõm trên vùng bao khuôn mặt, sau đó liên kết các vùng lỗi lõm cùng loại lại với nhau. Vùng khuôn mặt thật sự sẽ là vùng liên thông lớn nhất trong đó. Với cách như vậy, chúng ta sẽ thu nhỏ, và xác định chính xác hơn vị trí khuôn mặt.

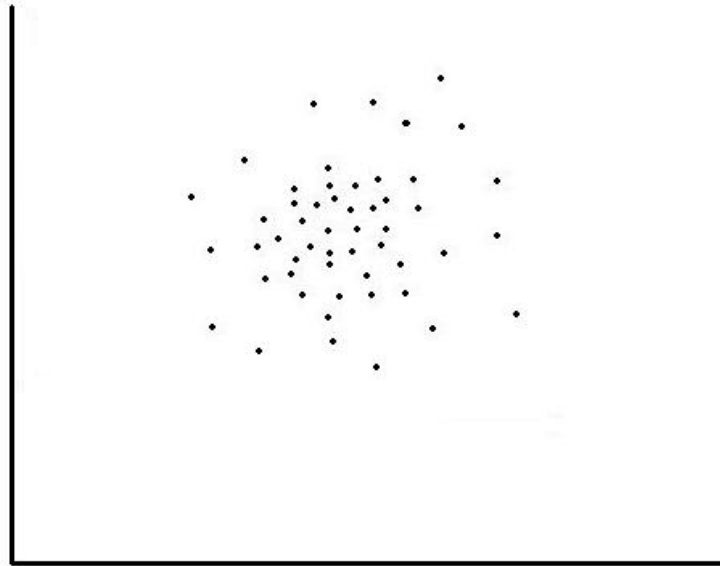


Hình 4.5 hậu xử lý

4.3 Cơ sở lý thuyết của mô hình

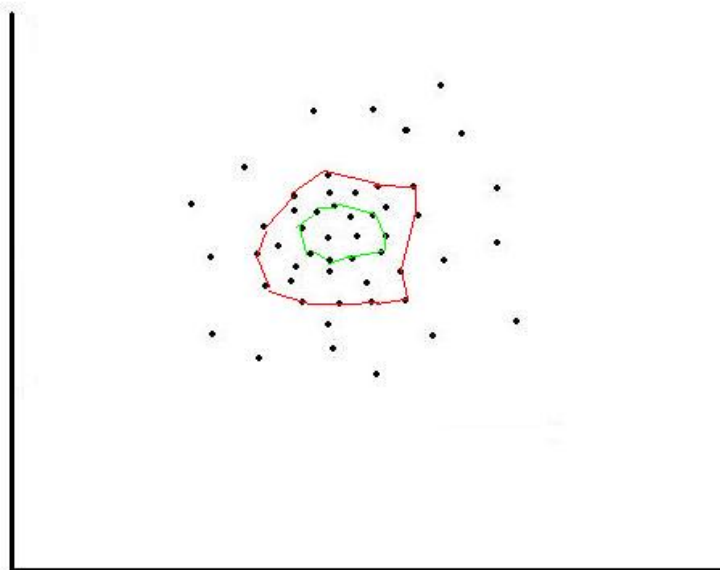
Như đã trình bày trong phần 3.4.3 ở chương 3, mỗi cây có thể được xem là một điểm trong *không gian cây*, trong đó mỗi cây chứa cấu trúc lỗi lõm của các đối

tượng quan tâm. Trong bài toán này ta quan tâm đến các cây biểu diễn cấu trúc khuôn mặt. Từ tập học, gồm các cây được đánh nhãn “*khuôn mặt*”, ta có thể biểu diễn nó trong một không gian hữu hạn chiều, là các điểm (xem hình 4.6).



Hình 4.6 Các cây khuôn mặt trong không gian cây

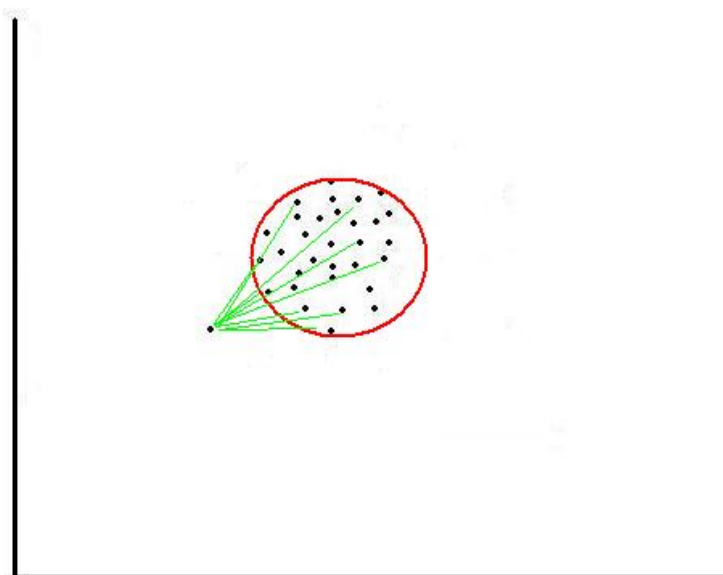
Trong hình trên, các điểm tương ứng với các cây được gán nhãn “*khuôn mặt*” trong tập học. Việc chọn ra k cây cho tập chuẩn có thể hiểu theo nghĩa là ta chọn k điểm đại diện. k điểm này nằm ngay trung tâm của lớp khuôn mặt (xem hình 4.7). Phương pháp gần giống với phương pháp k người láng giềng gần nhất nhưng có thay đổi cho phù hợp với bài toán phân lớp này nhằm tối ưu tốc độ.



Hình 4.7 Chọn k cây chuẩn theo tiêu chí trọng tâm

Vấn đề ở đây là xác định số k hợp lý để sao cho k cây trung tâm đó đủ tổng quát để đại diện cho cấu trúc khuôn mặt.

Sau đó, trên cơ sở tập cây mẫu này, một cây mới được so sánh và tính trung bình khoảng cách đến tất cả các cây trong tập mẫu.



Hình 4.7 Đánh giá điểm thuộc lớp “*khuôn mặt*” hay không

Như vậy, xét về mặt toán học đây việc tính khoảng cách trung bình từ điểm đang xét so với điểm chuẩn tương ứng với việc xét vị trí tương đối của nó so với cả nhóm điểm chuẩn. Nếu điểm nằm hoàn toàn bên trong cụm, chắc chắn khoảng cách trung bình sẽ rất nhỏ và sẽ được xếp vào lớp “*khuôn mặt*”. Ngược lại, nếu điểm đang xét nằm bên ngoài cụm, khoảng cách trung bình sẽ tăng lên và khả năng được xếp vào lớp “*khuôn mặt*” là rất thấp.

Mô hình thống kê này cũng phù hợp cho bài toán phân loại đối tượng cho thị giác robot. Phù hợp với cách học thích ứng, cho các đối tượng được học trước và cho cả đối tượng chưa biết được học tăng cường trong quá trình sống của robot.