**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**🙠🕮🙢**

**ĐỀ ÁN MÔN CÁC HỆ CƠ SỞ TRI THỨC**

**ĐỀ TÀI**

**MIÊU TẢ CỤC BỘ ĐẶC TRƯNG KHUÔN MẶT  
 SỬ DỤNG LOCAL BINARY PATTERN**

GVHD: TS. Lê Hoàng Thái

Nhóm thực hiện:

Nhóm 07 – Ngành Khoa Học Máy Tính – Cao học khóa 23

1. Đỗ Đặng Minh
2. Huỳnh Công Toàn
3. Dương Xuân Long
4. Hồ Văn Tấn

*Tp. Hồ Chí Minh, tháng 03 năm 2014*

Mục lục

[Thông tin nhóm 2](#_Toc381771894)

[Danh mục các kí hiệu, chữ viết tắt và ý nghĩa 2](#_Toc381771895)

[Danh mục hình vẽ, đồ thị 3](#_Toc381771896)

[Chương 1. Tổng quan về nhận dạng khuôn mặt 4](#_Toc381771897)

[1.1 Nhận dạng khuôn mặt 4](#_Toc381771898)

[1.2 Miêu tả khuôn mặt 6](#_Toc381771899)

[1.3 Miêu tả đặc trưng khuôn mặt 7](#_Toc381771900)

[Chương 2. Tổng quan về LBP 9](#_Toc381771901)

[2.1 Tổng quan về LBP 9](#_Toc381771902)

[2.2 LBP trong miền không gian (LBP-2D) 10](#_Toc381771903)

[2.3 LBP trong miền không gian và thời gian (LBP-3D) 13](#_Toc381771904)

[2.4 LBP đa mức 14](#_Toc381771905)

[2.5 Miêu tả khuôn mặt sử dụng LBP 15](#_Toc381771906)

[2.5.1 Biểu diễn khuôn mặt với hình ảnh tĩnh 15](#_Toc381771907)

[2.5.2 Miêu tả với khuôn mặt tuần tự 17](#_Toc381771908)

[2.6 Nhận dạng khuôn mặt sử dụng LBP 18](#_Toc381771909)

[2.7 LBP trong các lĩnh vực khác 21](#_Toc381771910)

[Chương 3. Bài toán ứng dụng 22](#_Toc381771911)

[3.1 Mô tả bài toán 22](#_Toc381771912)

[3.2 Phương pháp thực hiện 23](#_Toc381771913)

[Chương 4. Thực nghiệm 25](#_Toc381771914)

[4.1 Mô tả chương trình 25](#_Toc381771915)

[4.2 Môi trường thực nghiệm 27](#_Toc381771916)

[4.3 Dữ liệu ảnh huấn luyện và nhận dạng 28](#_Toc381771917)

[4.4 Kết quả thực nghiệm 28](#_Toc381771918)

[4.5 Nhận xét 32](#_Toc381771919)

[Tài liệu tham khảo 33](#_Toc381771920)

# Thông tin nhóm

|  |  |  |  |
| --- | --- | --- | --- |
| **MSHV** | **Họ tên** | **Số điện thoại** | **E-mail** |
| 13 11 015 | Đỗ Đặng Minh | 0168-993-5242 | [masterminh219@gmail.com](mailto:masterminh219@gmail.com) |
| 13 11 026 | Huỳnh Công Toàn | 0121-516-1090 | [alex7huynh@gmail.com](mailto:alex7huynh@gmail.com) |
| 13 11 048 | Dương Xuân Long | 097-357-0042 | [kht\_vvkt@yahoo.com.vn](mailto:kht_vvkt@yahoo.com.vn) |
| 13 11 058 | Hồ Văn Tấn | 090-290-9334 | [tanhv90@gmail.com](mailto:tanhv90@gmail.com) |

# Danh mục các kí hiệu, chữ viết tắt và ý nghĩa

|  |  |  |
| --- | --- | --- |
| **Từ viết tắt** | **Nghĩa tiếng Anh** | **Nghĩa tiếng Việt** |
| LBP | Local Binary Pattern | Mẫu nhị phân cục bộ |
| LBP-TOP | LBP from Three Orthogonal Planes | Mẫu nhị phân cục bộ từ ba miền trực giao |
| CS-LBP | Center-symmetric local binary pattern | Mẫu nhị phân cục bộ đối xứng trung tâm |
| MB-LBP | Multiscale block LBP | Mẫu nhị phân cục bộ khối đa mức |
| LDA | Linear Discriminant Analysis | Phân tích phân biệt tuyến tính |
| PCA | Principal Component Analysis | Phân tích thành phần chính |

# Danh mục hình vẽ, đồ thị

[**Hình 1:** tình huống sử dụng hệ thống MRTD ở sân bay (hình trái) và ứng dụng dựa trên MRTD (hình phải) 3](#_Toc381077825)

[**Hình 2:** các lĩnh vực trong sinh trắc học 4](#_Toc381077826)

[**Hình 3:** Khuôn mặt được chia thành những thành phần đặc trưng 6](#_Toc381077827)

[**Hình 4:** Đặc trưng nhận dạng khuôn mặt 7](#_Toc381077828)

[**Hình 5:** Mối quan hệ của LBP với phương pháp kết cấu bề mặt trước đó 9](#_Toc381077829)

[**Hình 6:** Toán tử LBP cơ bản 10](#_Toc381077830)

[**Hình 7:** Tập lân cận của các (P, R) khác nhau. Giá trị điểm ảnh song tuyến nội suy nếu điểm lấy mẫu không nằm ở trung tâm của điểm ảnh. 10](#_Toc381077831)

[**Hình 8:** Ví dụ kết cấu bề mặt gốc phát hiện bởi LBP (vòng tròn trắng đại diện 1, đen đại diện cho 0) 11](#_Toc381077832)

[**Hình 9:** **a.** Ba miền của kết cấu bề mặt động; **b.** Biểu đồ tần số cho mỗi miền; **c.** Nối các biểu đổ lại với nhau 12](#_Toc381077833)

[**Hình 10:** Ba LBP4,R gần kề và kết hợp các luật bất khả thi. 13](#_Toc381077834)

[**Hình 11:** Biểu diễn khuôn mặt sử dụng LBP 15](#_Toc381077835)

[**Hình 12:** Ví dụ về nét đặc trưng theo khối. **(a)** Những khối mặt **(b)** Đặc trưng LBP bởi ba mặt phẳng trực giao **(c)** Liên kết một khối mặt theo sự xuất hiện và chuyển động. 16](#_Toc381077836)

[**Hình 13: (a)** Một ví dụ về hình ảnh khuôn mặt được chia theo ma trận 7x7 **(b)** các trọng đặt ra cho χ2 trọng biện pháp đồng dạng. các hình vuông màu đen chỉ trọng lượng 0.0, màu xám 1.0, ánh sáng màu xám và trắng 2.0 4.0 18](#_Toc381077837)

[**Hình 14:** Ví dụ về Thư viện hình ảnh và thăm dò từ cơ sở dữ liệu FRGC, và tương ứng của họ hình ảnh lọc với chuỗi tiền xử lý của Tan và Triggs 19](#_Toc381077838)

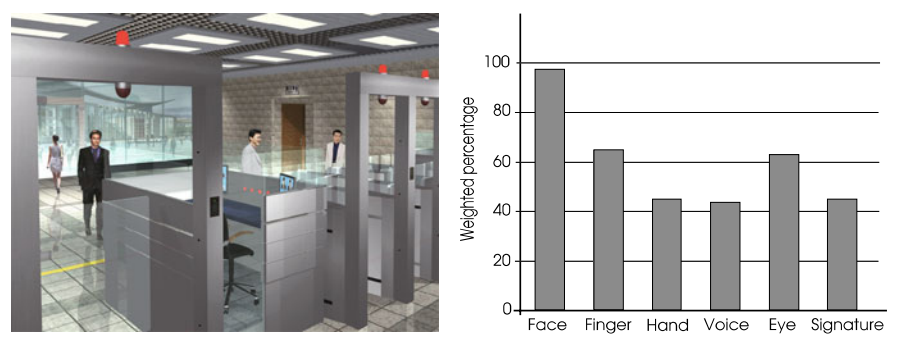
[**Hình 15:** Các cặp cảm xúc khuôn mặt khác nhau 21](#_Toc381077839)

# Chương 1. Tổng quan về nhận dạng khuôn mặt

## Nhận dạng khuôn mặt

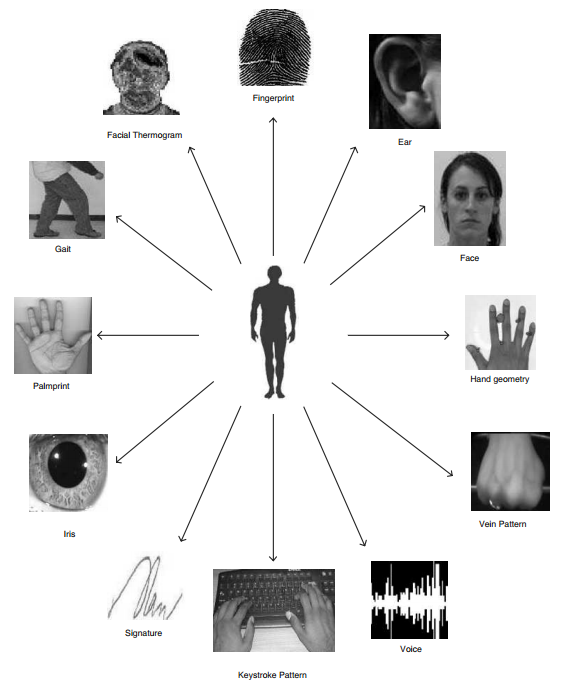
Nhận diện khuôn mặt là một điều mà con người có thể thực hiện một cách dễ dàng trong cuộc sống hiện nay. Chúng ta có sẵn những chiếc máy tính mạnh mẽ nhưng lại không hề đắt tiền và hệ thống tính toán được nhúng sẵn đã tạo ra một sức mạnh trong việc tự xử lý ảnh số trong hầu hết các chương trình, bao gồm cả xác minh sinh trắc học, giám sát, tương tác con người và máy tính, và quản lý đa phương tiện truyền thông. Nghiên cứu và phát triển trong hệ thống tự nhận dạng khuôn mặt người theo cách tự nhiên nhất.

Nhận diện khuôn mặt có nhiều lợi ích hơn là những dạng nhận diện sinh trắc học khác như: ngón tay và mắt, ngoài việc tự nhiên và không cần cung cấp tài liệu, ích lợi quan trọng nhất của nhận diện khuôn mặt là có thể chụp được trong một khoảng cách và có cách thức đảo ngược lại hình. Giữa những sinh trắc học thì nhận dạng khuôn mặt có độ tương thích cao nhất trong hệ thống Machine Readable Travel Documents (MRTD) dựa trên số lượng nhân tố đánh giá như sự kết nạp, sự hồi phục, yêu cầu máy và sự nhận biết chung. Nhận diện khuôn mặt, một trong những ngành công nghệ sinh trắc học chính, đang ngày càng phát triển để chiếm giữ vị trí quan trọng trong thiết bị chụp ảnh (đặc biệt là camera trên thiết bị di động), đang có một số lượng lớn các ảnh trên các trang web và yêu cầu bảo mật thì ngày càng cao.



**Hình 1:** tình huống sử dụng hệ thống MRTD ở sân bay (hình trái)   
và ứng dụng dựa trên MRTD (hình phải)

Hệ thống nhận dạng khuôn mặt tự động đầu tiên được phát triển bởi Takeo Kanade trong luận văn tiến sĩ của mình vào năm 1973. Sau đó là thời kỳ yên ắng của nhận dạng khuôn mặt tự động cho đến khi Sirovich và Kirby lưu trữ miêu tả khuôn mặt với dung lượng thấp, bắt nguồn từ PCA. Đây là sự tiên phong của Turk và Pentland trên Eigenface làm bùng nổ nghiên cứu nhận dạng khuôn mặt. Những cột mốc đáng nhớ khác trong quá trình phát triển nhận dạng khuôn mặt như: phương pháp Fisherface (áp dụng LDA sau bước PCA để đạt được hiệu quả chính xác hơn), sử dụng phương pháp lọc cục bộ như Gabor để cung cấp những đặc trưng khuôn mặt hiệu quả hơn, và thiết kế của AdaBoost dựa vào kiến trúc phân lớp học lan truyền cho thời gian thực tìm kiếm khuôn mặt.



**Hình 2:** các lĩnh vực trong sinh trắc học

Công nghệ nhận diện khuôn mặt ngày nay phát triển lớn mạnh hơn rất nhiều so với thời mà Eigenface được giới thiệu, Trong hoàn cảnh ràng buộc, ví dụ kèm theo ánh sáng, tư thế, biểu hiện khuôn mặt cũng sẽ được điều khiển, nhận dạng khuôn mặt tự động có thể vượt trội nhận diện cử chỉ con người, đặc biệt dữ liệu chứa nhiều ảnh. Tuy nhiên, nhận diện khuôn mặt cũng đối mặt với những thách thức khi mà khuôn mặt đòi hỏi không được thay đổi dưới điều kiện môi trường.

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.

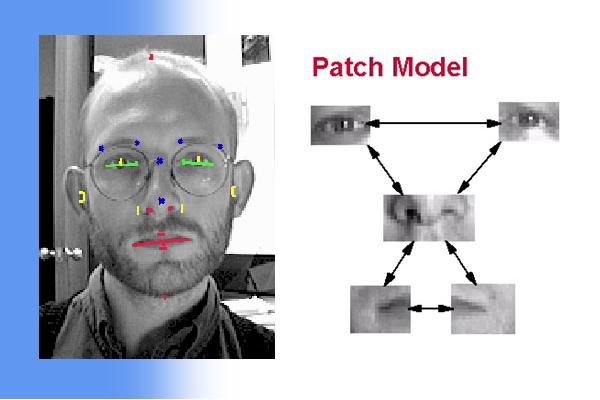
## Miêu tả khuôn mặt

Phát triển nhận dạng khuôn mặt bao gồm hai vấn đề là miêu tả khuôn mặt và bộ thiết kế phân lớp. Mục đích của miêu tả khuôn mặt là dùng để dẫn xuất được các đặc trưng từ ảnh gốc, sau đó có thể nhận diện được nhiều thể hiện khác nhau thuộc về khuôn mặt của một cá nhân với số ảnh ít nhất có thể và có thể tách biệt được khuôn mặt đó với nhiều khuôn mặt khác nhau. Vì thế, những đặc tính của miêu tả khuôn mặt cần có để thiết kế hệ thống nhận dạng khuôn mặt là:

* Phân biệt được những khuôn mặt của các cá thể khác nhau
* Trích xuất dễ dàng từ ảnh gốc để tăng tốc độ xử lý
* Lưu trữ bộ nhớ thấp

Trên thực tế thì để thỏa mãn những điều kiện trên là không dễ dàng bởi vì khuôn mặt của con người luôn có khả năng biến đổi lớn do biểu cảm, hướng, tư thế,… Vì thế, những nhân tố quan trọng này được dành cho nhiệm vụ khác của nhận dạng sinh trắc học mặt người như: tìm kiếm, cục bộ hóa, đăng ký, xác minh. Nên vấn đề của nhận dạng khuôn mặt là tìm kiếm những biểu hiện đặc trưng khuôn mặt thích hợp.

Có rất nhiều phương pháp được đề cử trong việc mô tả khuôn mặt để nhận dạng khuôn mặt. Trong những nỗ lực đầu tiên thì Kanade đã biểu diễn khuôn mặt xét trong điều kiện quan hệ hình học như cạch và góc giữa những bộ phận của khuôn mặt. Sau đó, kỹ thuật dựa vào ngoại hình được đề xuất. Phương pháp này xem khuôn mặt như là những điểm ảnh của mảng hai chiều và cố gắng dẫn xuất bộ mô tả cho ngoại hình khuôn mặt. Đi theo những cách này thì những pháp tổng thể khác nhau như: Pricipal Component Analysis (PCA), Linear Discriminant Analysis (LCA) và 2D PCA ngày càng được nghiên cứu rộng rãi. Sau đó, bộ mô tả cục bộ đạt được sự chú ý ngày càng nhiều do có độ vững trước sự thay đổi của tư thế và ánh sáng. Trong số các bộ mô tả đó là bộ lọc Gabor và mẫu cục bộ nhị phân (Local Binary Pattern) được giới thiệu rất thành công trong việc mã hóa khuôn mặt.



**Hình 3:** Khuôn mặt được chia thành những thành phần đặc trưng

## Miêu tả đặc trưng khuôn mặt

Lúc đầu, thuật toán để nhận dạng khuôn mặt được chia thành:

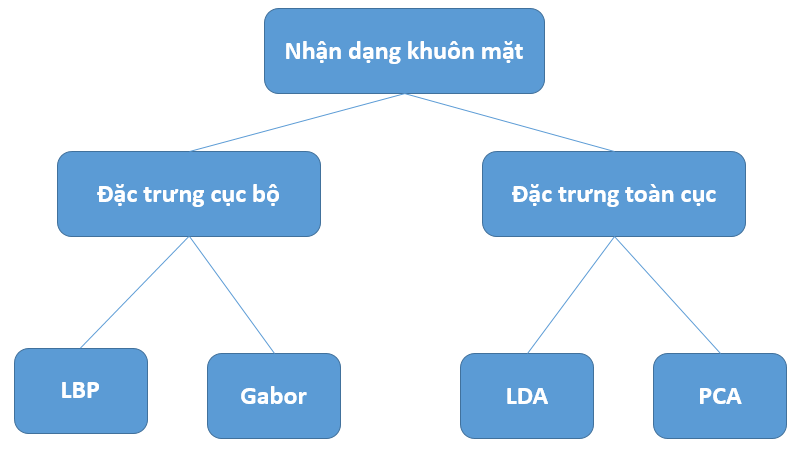
* Ngoại hình cơ bản
* Đặc trưng cơ bản
* Cách tiếp cận hybrid

Sự phân loại này được chấp nhận rộng rãi và được sử dụng trong thuật toán nhận dạng khuôn mặt, cục bộ hóa và xác minh.

Nhưng để tối ưu hóa cho việc miêu tả bộ phận khuôn mặt nên chỉ tập trung cho phần đặc trưng khuôn mặt và cách tiếp cận hybrid. Cho nên, gộp hai phần đó lại rồi chia thành:

* Phương pháp tổng quát dựa trên những hình ảnh tổng quát xử lý đặc trưng như: cạnh, đường thẳng,…
* Đặc trưng mẫu dựa vào phương pháp: dùng để phát hiện đặc trưng của khuôn mặt như: mắt, mũi,…
* Cách thức cấu trúc khớp nhau: dùng để ràng buộc các đặc trưng của khuôn mặt.

Theo quan điểm thì hướng tiếp cận tổng thể và đặc trưng mẫu dựa vào phương pháp là giống nhau vì cả hai đều là học từ mẫu quét từ cửa sổ hay là mẫu thể hiện và phát hiện khuôn mặt hoặc bộ phận khuôn mặt. Phương pháp nổi tiếng nhất là bộ nhận dạng Viola-Jones và PCA hoặc là LDA được tính toán mẫu không gian phụ và những công việc sản sinh. Những phương pháp này hiệu quả nhưng không bao gồm Haar-cascades của Viola-Jones và mẫu không gian phụ của PCA và LDA do những cách này không phải là mẫu đặc trưng chung. Những mẫu này được xem như học thống kê hoặc là bộ mô tả thuật toán cho chính mình. Và cấu trúc khớp nhau cũng không thuộc phạm vi trong phần nghiên cứu này.



**Hình 4:** Đặc trưng nhận dạng khuôn mặt

Việc chọn đặc trưng thích hợp tương đương với việc chọn đặc trưng và nhiệm vụ trích xuất được diễn ra trong hầu hết thị giác máy tính và phân tích ảnh. Nhưng cần chọn đặc trưng nào cho thích hợp? Câu trả lời tốt nhất là để cho việc kết nối và học đặc trưng của con người được tính từ những vị trí cục bộ khác nhau.

Thị giác máy tính và xử lý ảnh bao gồm nhiều đặc trưng khác nhau và trích xuất đặc trưng. Trong sinh trắc học khuôn mặt thì chỉ đặc trưng nhất định giữ lại và tiếp tục trong việc tạo ra kết quả trạng thái nghệ thuật. Được chấp nhận rộng rãi là những đặc trưng được cấu thành từ phản hồi của bộ lọc Gabor trong nhiều hướng và qui mô khác nhau. Gần đây phổ biến hơn là mẫu cục bộ nhị phân (LBP). Chúng ta cùng xem lại:

* Gabor filter và những phương pháp tiếp cận tần số hướng cục bộ dùng để cho kết quả mới nhất trong việc phát hiện và nhận diện khuôn mặt.
* Bộ mô tả SIFT phân loại hiển thị đối tượng và so khớp khuôn mặt.
* Gray-level patch thì cục kỳ đơn giản
* LBP dùng thuật toán để cấu thành những đặc trưng.

Ta có LBP và bộ lọc Gabor cho ra kết quả tốt nên được khuyến khích. Việc chia lớp đặc trưng như phức tạp và sóng mượt, bộ lọc có thể lái được và sự khác biệt của Gaussians, cho đến những cách của Gabor, bởi vì không có sự khác biệt cơ bản giữa chúng và tận dụng đúng cách thì sẽ cho kết quả tốt. Tương tự, SIFT, LBP và pha mô tả Daugman có cùng đặc tính. Độ linh hoạt của LBP rất thích hợp cho sinh trắc học khuôn mặt vì LBP có nhiều cách khởi tạo với nhiều tham số khác nhau.

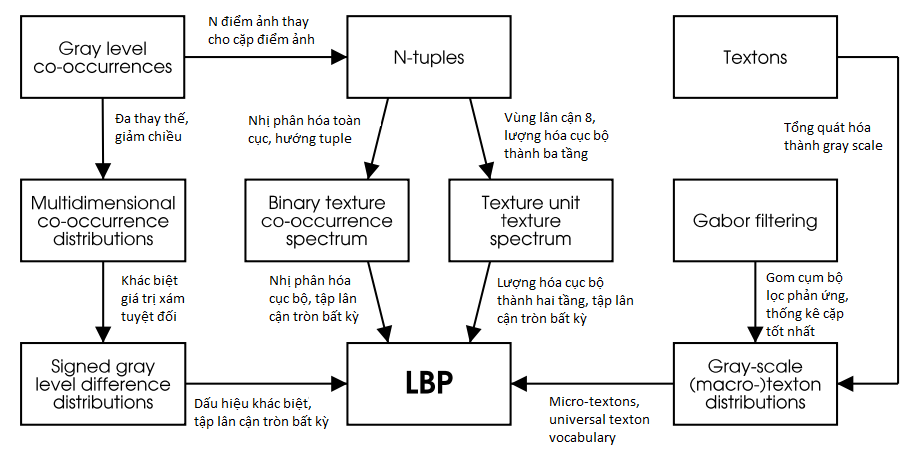
# Chương 2. Tổng quan về LBP

## Tổng quan về LBP

Việc sử dụng LBP trong phân tích khuôn mặt bắt đầu từ năm 2004 khi cách biểu diễn khuôn mặt mới cho việc nhận dạng khuôn mặt được đề xuất. Trong cách tiếp cận này, ảnh được chia thành nhiều vùng để đặc trưng LBP được trích xuất và nối lại thành một biểu đồ tần số đặc trưng cải tiến, sau đó dùng làm ký hiệu khuôn mặt. Cách tiếp cận tiến hóa dần và nhanh chóng thành công. Sau đó nhiều nhóm nghiên cứu và công ty trên khắp thế giới áp dụng.

Toán tử LBP và các biến thể của nó không chỉ dùng trong nhận diện khuôn mặt mà còn dùng trong nhiều lĩnh vực khác liên quan đến khuôn mặt như phát hiện khuôn mặt, phát hiện cảm xúc khuôn mặt, phân lớp giới tính, ước tính độ tuổi và nhận dạng tiếng nói. Sự thành công của LBP trong mô tả khuôn mặt là do khả năng phân biệt, sự đơn giản trong tính toán của toán tử, và khả năng áp dụng tốt cho biến đổi trong ảnh monotonic gray scale như thay đổi độ sáng. Việc dùng biểu đồ tần số làm đặc trưng cũng làm cho cách tiếp LBP thích hợp cho các thay đổi tư thế hay khuôn mặt.

LBP có liên quan đến nhiều toán tử phân tích kết cấu bề mặt nổi tiếng khác như hình bên dưới. Dấu mũi tên đại diện cho mối quan hệ giữa các phương pháp khác nhau, chữ bên dưới là điểm khác nhau tổng quát giữa chúng. LBP có thể được xem như sự kết hợp giữa các toán tử lọc dẫn xuất cục bộ với đầu ra được lượng tử hóa bằng ngưỡng cho trước.

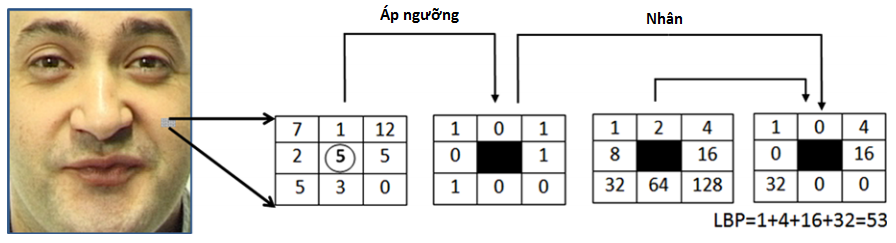


**Hình 5:** Mối quan hệ của LBP với phương pháp kết cấu bề mặt trước đó

## LBP trong miền không gian (LBP-2D)

Toán tử LBP dùng phân tích kết cấu bề mặt được xác định dựa trên độ đo sự bất biến giá trị độ xám (gray-scale) của kết cấu bề mặt ảnh, được xuất phát từ một định nghĩa tổng quan về các kết cấu bề mặt cục bộ lân cận.

Toán tử LBP thuần túy (the original LBP operator) được dùng để đánh nhãn các điểm ảnh của một ảnh bởi các bộ 3x3. Trong mỗi bộ 3x3 như vậy thì điểm ảnh được khảo sát là điểm ảnh trung tâm so với các điểm ảnh lân cận xung quanh của nó và được biểu diễn bởi một chuỗi nhị phân có trật tự. Biểu đồ tần số bao gồm 28 = 256 giá trị nhãn khác nhau được sử dụng để mô tả hình ảnh.



**Hình 6:** Toán tử LBP cơ bản

Hình 2.2 minh họa dạng toán tử LBP cơ bản. Các toán tử LBP đã được mở rộng để sử dụng các vùng lân cận có kích thước khác nhau. Việc sử dụng các vùng lân cận hình tròn và phép nội suy song tuyến tính tại các tọa độ không nguyên cho phép ta sử dụng bất kì bán kính và số lượng các điểm lấy mẫu. (Hình 2.3 biểu diễn các vùng lân cận hình tròn khác nhau). Ở đây ta có ký hiệu (P, R) được sử dụng cho các nhóm điểm ảnh lân cận; có nghĩa là P điểm lấy mẫu trên một vòng tròn có bán kính R.



**Hình 7:** Tập lân cận của các (P, R) khác nhau. Giá trị điểm ảnh song tuyến nội suy nếu điểm lấy mẫu không nằm ở trung tâm của điểm ảnh.

Một mở rộng khác toán tử thuần túy là định nghĩa mẫu đồng nhất (uniform patterns). Mở rộng này được lấy cảm hứng từ thực tế là một số mẫu nhị phân xuất hiện thường xuyên hơn là các kết cấu bề mặt ảnh. Một mẫu nhị phân cục bộ (LBP) được gọi là đồng nhất nếu mẫu nhị phân chứa tối đa hai sự chuyển đổi bit từ 0 đến 1 hoặc từ 1 đến 0. Các mẫu 00000000 (0 sự chuyển đổi), 01110000 (2 sự chuyển đổi) và 11001111 (2 sự chuyển đổi) là ví dụ cho các mẫu đồng nhất; trong khi đó mẫu 11001001 (4 sự chuyển đổi) và 01010011 (6 sự chuyển đổi) thì không phải.

Trong tính toán các nhãn LBP, mẫu đồng nhất được sử dụng để có một nhãn riêng biệt cho mỗi mẫu đồng nhất và tất cả các mẫu không đồng nhất được đánh nhãn với một nhãn duy nhất. Ví dụ, khi sử dụng vùng lân cận (8, R), có tổng cộng 256 mẫu, trong đó có 58 là đồng nhất nên ta sẽ có tất cả 59 nhãn khác nhau.

Ojala nhận thấy trong thí nghiệm của ông với những kết cấu bề mặt ảnh thì những mẫu đồng nhất tính toán cho gần 90% khi sử dụng vùng lân cận (8, 1), và gần 70% cho vùng lân cận (16, 2). Họ tìm thấy 90.6% trong số những mẫu trong vùng lân cận (8, 1) và 85.2% các mẫu ở vùng lân cận (8, 2) là đồng nhất trong trường hợp tiền xử lý hình ảnh khuôn mặt FERET.

Toán tử LBPP, R đồng nhất được ký hiệu là . Với chuỗi LBP có chiều dài là P thì mẫu có tối đa hai sự chuyển đổi (mẫu đồng nhất) là P (P – 1). Có hai mẫu không có sự chuyển đổi nào là mẫu 0 hoặc 1. Việc sử dụng mẫu LBP đồng nhất có hai lợi điểm. Thứ nhất là tiết kiệm bộ nhớ, vì trường hợp LBP tổng quát chúng ta có 2Pmẫu có thể, nhưng nếu xét riêng mẫu đồng nhất thì chúng ta có tối đa P (P + 1) + 2 mẫu nếu chúng ta sử dụng toán tử . Thứ hai là mẫu LBPu2 chỉ phát hiện những mẫu kết cấu bề mặt cục bộ quan trọng như các điểm cuối đường thẳng, cạnh biên và các góc.



**Hình 8:** Ví dụ kết cấu bề mặt gốc phát hiện bởi LBP  
(vòng tròn trắng đại diện 1, đen đại diện cho 0)

Sau khi LBP đánh nhãn hình ảnh đã thu được, biểu đồ tần số LBP có thể được định nghĩa như sau:

|  |  |
| --- | --- |
|  | (1.1) |

trong đó n là số lượng nhãn khác nhau do toán tử LBP tạo ra và

Các biểu đồ tần số phải được chuẩn hóa để có một mô tả nhất quán khi mà các miếng ảnh có biểu đồ tần số không cùng kích thước.

|  |  |
| --- | --- |
|  | (1.2) |

## LBP trong miền không gian và thời gian (LBP-3D)

Toán tử LBP thuần túy được định nghĩa chỉ để xử lý thông tin về mặt không gian, nhưng gần đây nó đã được mở rộng sang biểu diễn cả không gian và thời gian để phân tích kết cấu bề mặt động. Chính việc này đã dẫn đến tên gọi toán tử LBP khối (Volume Local Binary Pattern - VLBP). Ý tưởng của VLBP là xem kết cấu bề mặt động là một tập (X,Y,T) - trong đó X và Y biểu diễn tọa độ không gian, còn T là chỉ số khung hình (thời gian). Vùng lân cận của mỗi điểm ảnh vì thế được định nghĩa trong một không gian ba chiều. Tương tự như LBP, các volume texton có thể được định nghĩa và rút ra thành biểu đồ tần số. Chính vì vậy mà VLBP kết hợp cả cử động và diện mạo thành một mô tả kết cấu bề mặt động.



**Hình 9:** **a.** Ba miền của kết cấu bề mặt động; **b.** Biểu đồ tần số cho mỗi miền;  
 **c.** Nối các biểu đổ lại với nhau

LBP trên mặt phẳng ba miền trực giao (LBP-TOP) được giới thiệu để giúp VLBP dễ tính toán và dễ mở rộng hơn. LBP-TOP chứa ba miền trực giao: XY, XT, YT và kết nối các thống kê hội tụ LBP vào ba hướng này. Các vùng lân cận hình tròn được tổng quát hóa thành mẫu elip để phù hợp với thống kê không gian - thời gian. Các luật LBP trích từ miền XY, XT và YT biểu diễn thành XY-LBP, XT-LBP và YT-LBP cho tất cả các điểm ảnh và thống kê của ba miền khác nhau được nối lại thành một biểu đồ tần số duy nhất. Quy trình này trình bày ở Hình 2.5 ở trên. Trong cách biểu diễn này, kết cấu bề mặt động được mã hóa bởi XY-LBP, XT-LBP và YT-LBP.

Không hợp lí khi dùng cùng một bán kính cho trục không gian và thời gian nên trong miền XT và YT, chúng ta phải dùng bán kính khác nhau để gán cho các mẫu lân cận trong không gian và thời gian. Tổng quát hơn, bán kính trong trục X,Y,T và số các điểm lân cận trong miền XY, XT và YT có thể được biểu thị bằng RX, RY, RT, PXY, PXT và PYT. Đặc trưng tương ứng được biểu thị là .

Giả sử cho một kết cấu bề mặt động X x Y x T (). Một biểu đồ của kết cấu bề mặt động có thể được định nghĩa như sau:

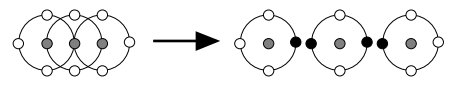
|  |  |
| --- | --- |
|  | (1.3) |

trong đó là số nhãn khác nhau do toán tử LBP tạo ra trong miền thứ *j* (j = 0: XY, 1: XT và 2: YT) và biểu thị luật LBP của điểm ảnh trung tâm (x, y, t) trong miền thứ *j*. Tương tự như LBP thuần túy, biểu đồ tần số phải được chuẩn hóa để có một mô tả chặt chẽ dùng để so sánh các kết cấu bề mặt động:

|  |  |
| --- | --- |
|  | (1.4) |

## LBP đa mức

Bởi vì LBP đặc trưng được tính toán trong một vùng lân cận 3 x 3 không thể bao quát các cấu trúc quy mô lớn nên LBP đa mức (multi-scale LBP) được đề xuất để giải quyết hạn chế này.



**Hình 10:** Ba LBP4,R gần kề và kết hợp các luật bất khả thi.

Ở hình 2.5 biểu diễn ba luật LBP 4-bit gần kề. Giả sử bit đầu tiên ở luật trái nhất là 0, bit thứ ba ở luật bên phải của nó phải là 1. Tương tự, bit đầu trong luật trung tâm và bit thứ ba ở luật phải nhất phải khác nhau hoặc cùng bằng 1. Phần bên phải hình trên cho thấy một sự kết hợp bất khả thi các luật. Mỗi luật LBP giới hạn số tập các luật gần kế nó, làm cho “vùng tác động” của một luật lớn hơn vùng 3x3 điểm ảnh.

Cách trực tiếp để phóng to vùng không gian hỗ trợ là kết hợp thông tin cung cấp bởi *N* toán tử LBP với các giá trị *P* và *R* khác nhau. Bằng cách này, mỗi điểm ảnh trong một hình sẽ có *N* luật LBP khác nhau. Thông tin chính xác nhất sẽ nhận được bằng cách phân phối có điều kiện các luật trên. Tuy nhiên, việc phân phối như vậy sẽ rất thưa thớt dù kích thước ảnh ở mức vừa phải. Vì vậy, chúng ta chỉ xem xét phân phối nhỏ các toán tử khác nhau. Mặc dù các luật LBP với các bán kính khác nhau không độc lập hoàn toàn trong các trường hợp đặc trưng nhưng việc phân tích đa quyết định thường cải thiện khả năng phân biệt đặc trưng kết quả. Với hầu hết ứng dụng, cách xây dựng toán toán tử LBP đa mức đa đem lại độ chính xác rất cao.

LBP đa mức được mở rộng thành LBP khối đa mức (multiscale block LBP) và được dùng chủ yếu khi phân tích ảnh khuôn mặt. Ý tưởng chính của MB-LBP là so sánh giá trị điểm ảnh trung bình trong khối nhỏ thay vì so sánh các giá trị điểm ảnh. Toán tử này luôn xem xét 8 vùng lân cận, tạo ra các nhãn từ 0 đến 255. Ví dụ như, mỗi khối có kích thước 3 x 3 điểm ảnh, toán tử MB-LBP tương ứng so sánh giá trị xém trung bình của khối trung tâm với giá trị trung bình của 8 khối lân cận cùng kích thước và miền phụ hợp của toán tử là 9 x 9 điểm ảnh.

## Miêu tả khuôn mặt sử dụng LBP

## Biểu diễn khuôn mặt với hình ảnh tĩnh

Trong cách tiếp cận LBP cho việc phân loại kết cấu bề mặt, sự xuất hiện của các mã LBP trong một hình ảnh được sẽ được thu thập vào một biểu đồ. Sau đó, sự phân loại này sẽ được thực hiện bằng việc tính toán những biểu đồ đơn giản tương đương nhau. Tuy nhiên, xem xét một cách tiếp cận tương tự này sẽ cho ra những kết quả đại diện hình ảnh khuôn mặt với một sự mất mát thông tin không gian. Do đó, cách làm này nên hệ thống hóa các thông tin kết cấu bề mặt với vị trí của nó. Một cách để đạt được mục tiêu này là sử dụng các mô tả kết cấu bề mặt LBP để xây dựng một số mô tả địa phương mặt và kết hợp chúng thành một mô tả toàn diện. Chẳng hạn như biểu diễn cục bộ đã đạt dược tầm quan trọng trong thời gian gần đây, điều này dễ hiểu cho những hạn chế của biểu diễn tổng thể. Những phương pháp đặc trưng địa phương này dường như thiết thực hơn với sự đa dạng trong sắp đặt bố cục hoặc sự chiếu sáng so với các phương pháp tồng thể.

Các phương pháp cơ bản cho LBP dựa trên mô tả khuôn mặt như sau: các mặt hình ảnh được chia thành các khu vực địa phương và những LBP miêu tả kết cấu bề mặt được chiết xuất từ vùng khác nhau một cách độc lập. Các miêu tả được nối với nhau thành một mô tả toàn thể tổng quát như thể hiện trong hình. 2.6.



**Hình 11:** Biểu diễn khuôn mặt sử dụng LBP

Biểu đồ cơ bản được sử dụng để thu thập thông tin về các mã LBP trong một hình ảnh có thể được mở rộng thành một biểu đồ không gian cải tiến mà nó mã hóa mã hóa cả sự xuất hiện và quan hệ không gian của các vùng trên khuôn mặt. Như các vùng trên khuôn mặt R0, R1,..., Rm-1 đã được xác định, các biểu đồ về không gian nâng cao được xác định như sau:

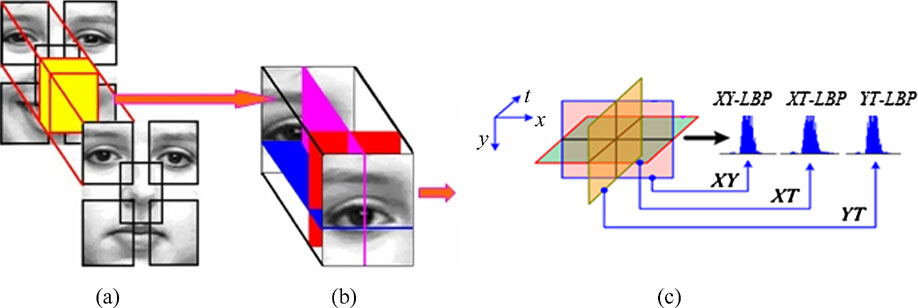
|  |  |
| --- | --- |
|  | (1.4) |

Biểu đồ này mang tính hiệu quả có sự mô tả của khuôn mặt trên ba cấp độ khác nhau trong cục bộ: các nhãn LBP cho biểu đồ chứa thông tin về các khuôn mẫu trên một cấp độ điểm ảnh, các nhãn được tóm tắt trong một khu vực nhỏ để sản sinh ra thông tin trên cấp độ khu vực và các biểu đồ khu vực được kết nối với nhau để xây dựng một mô tả khuôn mặt toàn thể. Cần lưu ý rằng khi sử dụng biểu đồ dựa phương pháp các vùng R0, R1,..., Rm-1 không cần phải là hình chữ nhật. Các vùng này cũng không nhất thiết phải có cùng kích thước hoặc hình dạng, và chúng cũng không nhất thiết phải bao phủ toàn bộ hình ảnh. Nó cũng có thể là một vùng không gian chồng chéo nhau.

Từ đây, những điều trên đã phác thảo ra các LBP nguyên thủy ban đầu để biểu diễn khuôn mặt, sau đó thông qua các nhiệm vụ khác nhau trên khuôn mặt phân tích hình ảnh. Hình 2.6 cho thấy một ví dụ về một biểu diễn khuôn mặt dựa trên LBP.

## Miêu tả với khuôn mặt tuần tự

Phát hiện tâm vật lý chỉ ra rằng sự vận đông khuôn mặt có thể cung cấp những thông tin có giá trị cho việc phân tích khuôn mặt. Vì vậy, sự biểu diễn khuôn mặt một cách hiệu quả nên mã hóa bao gồm cả xuất hiện và chuyển động của khuôn mặt. Do đó, một cách là mô tả sự biểu diễn khuôn mặt dựa trên LBP không gian - thời gian dựa trên phân tích khuôn mặt trong video sử dụng các mô tả vùng ghép nối lại với nhau.



**Hình 12:** Ví dụ về nét đặc trưng theo khối. **(a)** Những khối mặt **(b)** Đặc trưng LBP bởi ba mặt phẳng trực giao **(c)** Liên kết một khối mặt theo sự xuất hiện và chuyển động.

Một mô tả LBP tính toán thông qua một chuỗi khuôn mặt hoàn toàn mã hóa chỉ có sự xuất hiện của các vi mô hình mà không bất kỳ dấu hiệu về vị trí của chúng. Để khắc phục tình trạng này, một sự biểu diễn mà trong đó hình ảnh khuôn mặt được chia thành nhiều khối chồng lên nhau được sử dụng. Các biểu đồ LBP-TOP trong mỗi khối được tính toán và ghép nối với nhau thành một biểu đồ đơn nhất, như minh họa trong hình 2.7. Tất cả các đặc trưng được rút trích từ ​​mỗi vùng sẽ được kết nối để biểu diễn cho sự xuất hiện và chuyển động của khuôn mặt trong chuỗi. Các đặc trưng VLBP cơ bản cũng có thể được xem xét và rút trích dựa trên cơ sở các vùng chuyển động theo cùng một cách như các tính năng LBP-TOP.

Các biểu đồ LBP-TOP trong từng tập khối được tính toán và nối vào một biểu đồ duy nhất. Tất cả các đặc trưng rút trích từ mỗi tập khối sẽ được kết nối để biểu diễn cho sự xuất hiện và chuyển động của khuôn mặt. Đây là cách hiệu quả có một mô tả của khuôn mặt trên ba cấp độ khác nhau của cục bộ. Các nhãn trong biểu đồ chứa thông tin từ ba mặt phẳng trực giao, mô tả sự xuất hiện và thông tin thời gian ở cấp độ điểm ảnh. Các nhãn được tóm tắt thông qua những khối nhỏ để sản sinh thông tin về cấp độ khu vực thể hiện những đặc tính cho sự xuất hiện và chuyển động của khuôn mặt tại các địa điểm cụ thể, và tất cả các thông tin từ cấp độ khu vực này sẽ được kết nối để xây dựng một mô tả toàn diện cho khuôn mặt.

## Nhận dạng khuôn mặt sử dụng LBP

Phần này sẽ mô tả các ứng dụng của mô tả khuôn mặt đi đến nhận diện khuôn mặt. Thường là một nguyên tắc phân lớp vùng gần nhất được sử dụng trong việc nhận dạng khuôn mặt. Điều này là do thực tế, số lượng hình ảnh huấn luyện (bộ sưu tập) cho mỗi chủ đề là thấp, thường chỉ có một. Tuy nhiên, ý tưởng về một biểu đồ không gian nâng cao có thể khai thác thêm khi xác định các biện pháp khoảng cách cho sự phân lớp. Một thuộc tính cục bộ của một phương pháp biểu diễn khuôn mặt đề xuất là mỗi phần tử trong biểu đồ tăng cường tương ứng với một khu vực nhỏ nào đó của khuôn mặt. Dựa trên những phát hiện tâm vật lý, trong đó cho thấy một số đặc điểm khuôn mặt (như mắt) đóng một vai trò quan trọng hơn trong nhận dạng khuôn mặt con người hơn các tính năng khác, nó có thể được kỳ vọng ​​rằng một số khu vực trên khuôn mặt đóng góp nhiều hơn những khu vưc khác. Bằng cách sử dụng giả định này các khu vực có thể được cân nhặc dựa trên tầm quan trọng của thông tin mà nó có. Hình 2.8 là một ví dụ trọng lượng của khu vực trên khuôn mặt khác nhau.



**Hình 13: (a)** Một ví dụ về hình ảnh khuôn mặt được chia theo ma trận 7x7 **(b)** các trọng đặt ra cho χ2 trọng biện pháp đồng dạng. các hình vuông màu đen chỉ trọng lượng 0.0, màu xám 1.0, ánh sáng màu xám và trắng 2.0 4.0

Độ cân nhắc Chi về khoảng cách vuông được định nghĩa là:

|  |  |
| --- | --- |
|  | (1.4) |

trong đó x và ξ là những biểu đồ tăng cường bình thường hóa để được so sánh, các chỉ số i và j tham khảo thứ i bin tương ứng với thứ j khu vực địa phương và wj là độ cân nhắc cho khu vực j.

Thực tế, nhóm của Ahonen đã biểu diễn một tập hợp các thí nghiệm trên hình ảnh khuôn mặt FERET. Kết quả cho thấy phương pháp LBP đạt được nhận dạng khuôn mặt cao hơn so với thuật toán điều. Để đạt được tốt hơn sự hiểu biết về việc công nhận kết quả thu được là dựa trên ý tưởng chung về việc tính toán các đặc trưng kết cấu bề mặt từ các vùng trên khuôn mặt cục bộ hoặc do sự phân biệt LBP năng lực của các toán tử mô hình nhị phân cục bộ, các nghiên cứu cũng so sánh LBP với ba kết cấu bề mặt mô tả khác, cụ thể là sự khác biệt biểu đồ màu xám cấp, kết cấu bề mặt đồng nhất mô tả và một phiên bản cải tiến của biểu đồ texton. Kết quả xác nhận tính hợp lệ của cách tiếp cận LBP và cho thấy hiệu suất của LBP trong mô tả khuôn mặt vượt quá mà các mô hình kết cấu bề mặt khác như thể hiện trong Bảng 4.2. Nguyên nhân chính cho hiệu suất tốt hơn của phương pháp LBP so với mô tả kết cấu bề mặt khác là độ dung sai đối với sự thay đổi thang xám đơn điệu. Hơn nữa, phương pháp này có them lợi thế là sự tính toán có hiệu quả và sự tránh xám bình thường trước các toán tử LBP.



**Hình 14:** Ví dụ về Thư viện hình ảnh và thăm dò từ cơ sở dữ liệu FRGC, và tương ứng của họ hình ảnh lọc với chuỗi tiền xử lý của Tan và Triggs

Gần đây, Tan và Triggs phát triển một chuỗi tiền xử lý rất hiệu quả cho những hình ảnh khuôn mặt và thu nhận được những kết quả xuất sắc bằng việc sử dụng nhận dạng khuôn mặt dựa trên LBP cho cơ sở dữ liệu FRGC. Kể từ đó, nhiều người khác đã áp dụng tiền xử lý của họ dây chuyền cho các ứng dụng xử lý với các biến thể chiếu sáng trầm trọng. Hình 2.9 cho thấy một ví dụ về bộ sưu tập và thăm dò hình ảnh từ cơ sở dữ liệu và các FRGC hình ảnh được lọc bằng phương pháp tiền xử lý tương ứng.

Nhóm của Chan coi LBP đa mức và nguồn gốc mô tả khuôn mặt mới từ LDA của những biểu đồ mẫu nhị phân cục bộ đa quy mô. Hình ảnh khuôn mặt trước tiên được phân chia thành nhiều khu vực không chồng chéo. Trong mỗi khu vực, mỗi biểu đồ LBP đa quy mô đồng nhất sẽ được rút trích và kết nối với nhau thành một đặc trưng khu vực. Các đặc trưng này sau đó được chiếu lên không gian LDA được sử dụng như một mô tả khuôn mặt phân biệt. Phương pháp này đã được thử nghiệm trong nhận dạng khuôn mặt trên tiêu chuẩn cơ sở dữ liệu FERET và xác thực khuôn mặt trên cơ sở dữ liệu XM2VTS với kết quả rất hứa hẹn.

Nhóm của Zhang sử dụng phương pháp LBP cho nhận dạng khuôn mặt và thuật toán được sử dụng thuật giải học AdaBoost để lựa chọn một thiết lập tối ưu của các khu vực cục bộ và độ quan trọng của chúng. Điều này mang lại cho một vectơ đặc trưng nhỏ hơn so với các vector đặc trưng được sử dụng trong phương pháp LBP nguyên thủy. Tuy nhiên, phương pháp này không nâng cao hiệu suất đáng kể. Sau đó, nhóm của Huang đã đề xuất một biến thể của AdaBoost gọi JSBoost lựa chọn các thiết lập tối ưu của LBP đặc trưng cho nhận dạng khuôn mặt.

Để xử lý các vấn đề với các biến thể chiếu sáng mạnh, nhóm của Li phát triển rất thành công hệ thống kết hợp hồng ngoại gần (NIR – Near Infrared System) chụp ảnh với mô hình nhị phân cục bộ đặc trưng và thuận giải học AdaBoost. Các hằng số của LBP liên quan đến những thay đổi cấp độ màu xám đơn điệu với làm cho các đặc trưng rút trích từ các hằng số chiếu sáng của ​​NIR.

Bên cạnh đó, Rodriguez và Marcel đã đề xuất một cách tiếp cận dựa trên biểu đồ LBP đặc tả của người dung một cách thích hợp cho nhiệm vụ xác thực khuôn mặt. Phương pháp xem xét biểu đồ cục bộ như là phân bố xác suất và tính tỷ lệ log-likeligood thay vì χ2 tương tự. Một mô hình khuôn mặt tổng quát được coi là một bộ dữ liệu ban đầu của LBP biểu đồ. Sau đó, một mô hình người dùng cụ thể thu được bằng kỹ thuật chuyển đổi từ mô hình tổng quát trong một khuôn mẫu xác suất. Các báo cáo kết quả thực nghiệm cho thấy những phương pháp đề xuất mang lại hiệu suất tốt trên hai tiêu chuẩn cơ sở dữ liệu (XM2VTS và BANCA). Sau đó, Ahonen và Pietikäinen có thêm nâng cao hiệu quả hoạt động kiểm tra mặt trên cơ sở dữ liệu BANCA bằng cách phát triển một phương pháp mới cho việc ước lượng các bản phân bố cục bộ của nhãn LBP. Phương pháp này là dựa trên đánh giá mật độ hạt nhân trong không gian xy, và nó cung cấp không gian chính xác tốt hơn nhiều so với phương pháp dựa trên khối của Rodriguez và Marcel.

## LBP trong các lĩnh vực khác

Cách tiếp cận LBP được áp dụng vào nhiều việc phân tích khuôn mặt khác như phát hiện biểu cảm khuôn mặt, phát hiện giới tính, phân lớp độ tuổi, phát hiện khuôn mặt, nhận dạng tròng mắt, ước lượng tư thế đầu và nhận dạng khuôn mặt 3D. Ví dụ như LBP dùng với Active Shape Model (ASM) để cục bộ hóa và biểu diễn điểm trọng yếu của khuôn mặt vì việc cục bộ hóa chính xác rất quan trọng đối với việc phân tích khuôn mặt và tổng hợp vấn đề. Diện mạo cục bộ của các điểm trọng yếu trong ảnh khuôn mặt được mô hình hóa bằng bản mở rộng của LBP (ELBP). ELBP được đề xuất không chỉ để mã hóa thông tin dẫn xuất đầu tiên của ảnh khuôn mặt mà còn vận tốc biến đổi cục bộ. Các thí nghiệm đã cho thấy sự kết hợp ASM-ELBP đã làm tăng độ chính xác liên kết khuôn mặt so với phương pháp ASM ban đầu.



**Hình 15:** Các cặp cảm xúc khuôn mặt khác nhau

Ký hiệu mô tả LBP không gian - thời gian, đặc biệt là LBP-TOP được ứng dụng thành công trong các ứng dụng video như nhận diện biểu cảm khuôn mặt động, nhận diện tiếng nói và nhận diện giới tính. Chúng có thể mô tả hình dáng, cử động ngang dọc từ chuỗi video khá hiệu quả. Cách tiếp cận dựa trên LBP-TOP có thể mở rộng để thêm đặc trưng đa quyết định tính toán từ khác khối có kích thước khác nhau, mẫu lân cận khác nhau và dùng AdaBoost để chọn lớp đặc trưng cho tất cả lớp biểu cảm hoặc mỗi cặp lớp biểu cảm, từ đó cải thiện hiệu năng với vectơ đặc trưng ngắn. Sau đó, trên nền tảng của lớp được chọn, xem xét vị trí và loại đặc trưng của đặc trưng phân biệt nhất cho mỗi lớp. Hình 2.7 cho thấy các đặc trưng được chọn cho mỗi cặp biểu cảm. Chúng khác nhau và phụ thuộc vào từng biểu cảm.

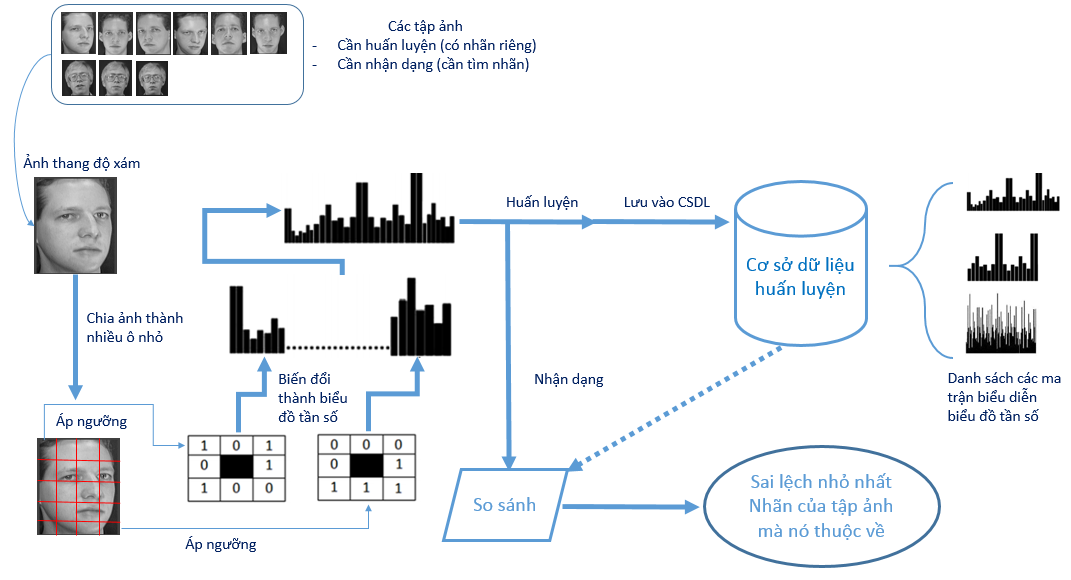
# Chương 3. Bài toán ứng dụng

## Mô tả bài toán

Chương trình có các bộ dữ liệu mẫu gồm nhiều tập ảnh. Mỗi tập ảnh gồm nhiều ảnh khuôn mặt của một người ở nhiều góc cạnh khác nhau như nhìn thẳng, nhìn nghiêng,… Mỗi tập ảnh này có một nhãn riêng để phân biệt với các tập ảnh khác. Các ảnh này đều đã được chuẩn hóa thành ảnh thang độ xám định dạng pgm để phù hợp cho việc sử dụng thuật toán LBP.

Yêu cầu đặt ra khi cung cấp một ảnh khuôn mặt thang độ xám và xác định xem ảnh này liệu có xuất hiện trong cơ sở dữ liệu đã có hay không. Nếu ảnh này có xuất hiện thì thuộc vào tập ảnh nào. Ảnh khuôn mặt này có thể ở góc cạnh nhìn thẳng, nhìn nghiêng,… có thể có hoặc không có trong cơ sở dữ liệu mẫu.

## Phương pháp thực hiện



Bước 1

Chương trình tiến hành đọc các dữ liệu mẫu để huấn luyện. Với mỗi ảnh đọc được, tiến hành gọi phương thức *createLBPHFaceRecognizer* với tham số thiết lập trước để thiết lập bộ nhận dạng dùng biểu đồ tần số của LBP. Các tham số này bao gồm bán kính *R*, số điểm lấy mẫu *P*, kích thước lưới *grid\_x* *grid\_y* của biểu đồ tần số và ngưỡng trên *threshold*.

Bước 2

Mỗi ảnh sẽ được chia nhỏ thành thành các ô. Với mỗi ô sẽ tiến hành xây dựng biểu đồ tần số. Chương trình sẽ gọi đến phương thức *elbp\_* và *spatial\_histogram* để xây dựng biểu đồ tần số và toán tử LBP.

|  |  |
| --- | --- |
|  | (3.1) |

với là điểm ảnh trung tâm với cường độ . là cường độ các điểm ảnh lân cận. Hàm s được định nghĩa như sau:

Với điểm cho trước, có thể tính được vị trí của điểm lân cận như sau:

|  |  |
| --- | --- |
|  | (3.2) |

Sau khi tính lần lượt cho tất cả các điểm lân cận, chúng ta có được biểu đồ tần số biểu diễn bằng ma trận với giá trị 0 và 1. Sau đó, nối các biểu đồ tần số lại để có biểu đồ tần số cho duy nhất một ảnh. Như vậy mỗi ảnh sẽ được biểu diễn bằng một biểu đồ tần số dưới dạng ma trận và một nhãn để biết nó thuộc về tập nào. Tiến hành tương tự cho tất cả các ảnh còn lại muốn huấn luyện, chúng ta được một vector lưu lại danh sách các ma trận này.

Bước 3

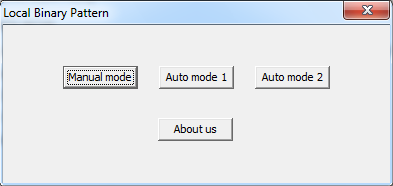
Khi tiến hành truy vấn một ảnh X xem có thuộc tập nào trong cơ sở dữ liệu mẫu, chương trình gọi phương thức *predict* tiến hành chuyển ảnh X thành biểu đồ tần số tương tự như các bước trên.

Sau đó tiến hành so sánh biểu đồ tần số của ảnh X lần lượt với các ma trận trong danh sách các ma trận đã có từ bước huấn luyện, thu được một giá trị tạm gọi là *dist*. Nếu giá trị *dist* thu được khi so sánh với một ma trận của ảnh Y nằm dưới ngưỡng trên *threshold* và *dist* là nhỏ nhất thì nhãn của Y cũng chính là nhãn của X.

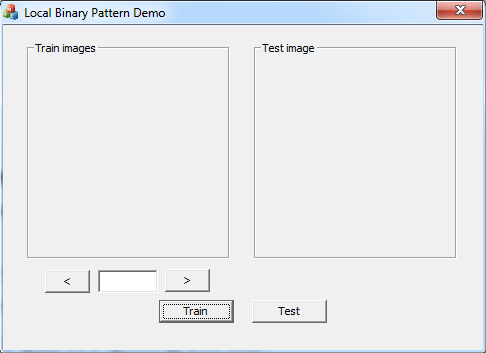
Nếu nhãn ban đầu (mặc định là -1) sau khi kết thúc quá trình so sánh vẫn không thay đổi thì có kết luận ảnh X không nằm trong bộ dữ liệu mẫu đã được huấn luyện.

# Chương 4. Thực nghiệm

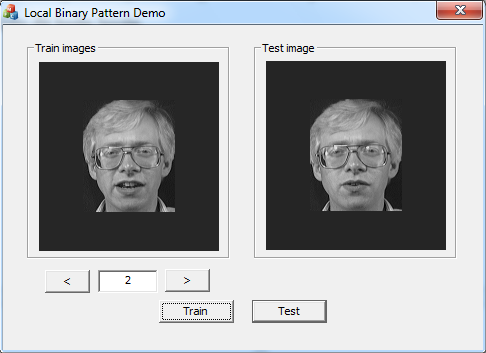
## Mô tả chương trình



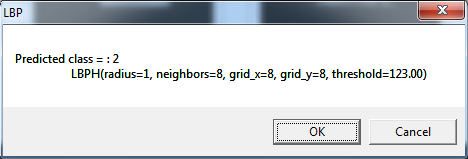
* Manual mode:



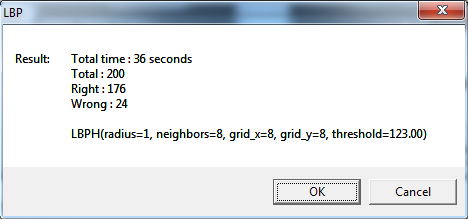
* Chọn “Train”, nhập nhãn của tập ảnh cần huấn luyện, sau đó chọn ảnh cần huấn luyện. Lặp lại bước này cho đến khi chọn hết các ảnh cần huấn luyện.
* Chọn “Test”, chọn một ảnh cần phân tích.



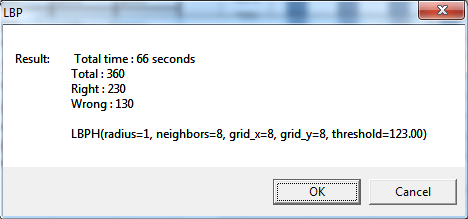
* Hệ thống sẽ xử lý và thông báo ảnh cần phân tích thuộc tập nào trong cơ sở dữ liệu đã huấn luyện.



* Auto mode 1:
* Chọn một tập tin huấn luyện trong thư mục Database.
* Chọn một tập tin để kiểm tra trong thư mục Database.
* Ví dụ như chọn DataTN1-Train1.csv để huấn luyện và DataTN1-Test1.csv để kiểm tra. Chương trình sẽ huấn luyện toàn bộ các ảnh trong Train1 thành một bộ dữ liệu mẫu, sau đó sẽ lần lượt truy vấn tất cả các ảnh trong Test1.



* Auto mode 2: cách thực hiện tương tự auto mode 1.
* Chọn một tập tin huấn luyện trong thư mục Database\TN2.
* Chọn một tập tin để kiểm tra trong thư mục Database\TN2.
* Ví dụ như chọn DataTN2-Train1.csv để huấn luyện và DataTN2-Test1.csv để kiểm tra. Chương trình sẽ huấn luyện toàn bộ các ảnh trong Train1 thành một bộ dữ liệu mẫu, sau đó sẽ lần lượt truy vấn tất cả các ảnh trong Test1. Sau mỗi lần truy vấn xong, hệ thống sẽ bổ sung ảnh vừa truy vấn vào bộ dữ liệu mẫu. Chính vì vậy phương pháp này có độ chính xác cao hơn nhưng quá trình chạy tốn nhiều thời gian hơn.



## Môi trường thực nghiệm

Môi trường lập trình: Visual Studio 2010

Ngôn ngữ lập trình: C++

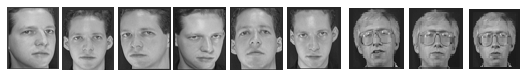
Thư viện sử dụng: OpenCV 2.4.8

Hệ điều hành: Windows 7 Professional 32-bit

Processor: Intel® Core™ 2 Quad CPU Q8400 @ 2.66 GHz

## Dữ liệu ảnh huấn luyện và nhận dạng

Dữ liệu thực nghiệm là bộ dữ liệu của AT&T Laboratories, Cambridge. Bộ dữ liệu bao gồm 400 ảnh chia thành 40 tập, mỗi tập có 10 ảnh. Mỗi tập ảnh là ảnh khuôn mặt thang độ xám ở nhiều góc cạnh khác nhau của một người.



Dữ liệu thực nghiệm được chia thành 2 phần là thực nghiệm 1 (TN1) và thực nghiệm 2 (TN2). Mỗi phần thực nghiệm bao gồm 10 tập tin để huấn luyện và 10 tập tin để kiểm tra.

Mỗi tập tin trong TN1 đều chứa 200 ảnh (50% tổng số ảnh).

Trong TN2, Train1 sẽ chứa 10% tổng số ảnh, Test1 chứa 90% tổng số ảnh. Các bộ train kế tiếp sẽ tăng 5% tổng số ảnh, test sẽ giảm 5% tổng số ảnh. Vì vậy Train10 sẽ chứa 55% tổng số ảnh, Test10 chứa 45% tổng số ảnh.

## Kết quả thực nghiệm

Thực hiện thực nghiệm với dữ liệu đã được chuẩn bị như ở mục 4.3. Trong đó, dữ liệu thực nghiệm được thể hiện thông qua 10 cặp file csv (train và test) tương ứng cho 10 lần chạy thực nghiệm 1 và 10 cặp file csv cho lần chạy thực nghiệm 2. Trong quá trình tìm hiểu thuật toán LBP khi được cài đặt thực tế, các thông số được thiết lập bao gồm:

* Radius = 1
* Grid = 8x8
* Threshold = 123

Ở đây, thông số threshold được xem là một trong những yếu tố quyết định đến việc xác định ảnh được đưa vào test có thuộc cùng tập với các ảnh đã được train và đánh dấu. Do đó, phương pháp thực nghiệm sẽ được tiến hành trên các thông số threshold là 40, 123 và 160 để đánh giá sự hiệu quả nhận dạng của chương trình.

Thực nghiệm 1: 50% dành cho huấn luyện, 50% để kiểm tra. Lặp lại 10 lần.

* Với thông số threshold = 40, kết quả chương trình:
* Với thông số threshold = 123, kết quả chương trình:
* Với thông số threshold = 160, kết quả chương trình:

Thực nghiệm 2: K-folds, dữ hiệu huấn luyện tăng dần, huấn luyện đến khi hội tụ. Lặp lại 10 lần. Tuy nhiên ở đây, thực hiện 10 lần và huấn luyện khi ảnh train đạt mức 55%, ảnh test 45%.

* Với thông số threshold = 40, kết quả chương trình:
* Với thông số threshold = 123, kết quả chương trình:
* Với thông số threshold = 160, kết quả chương trình:

Như vậy, với các thông số đầu vào khác nhau, mà ở đây, sự thay đổi của threshold đầu vào, đã ảnh hưởng khá lớn tới thuật toán nhận dạng. Với thông số threshold 40, chương trình chỉ có thể nhận dạng gần 50% ảnh test. Còn với các thông số threshold 123, 160 hoặc các thông số khác mà do điều kiện không cho phép, chương trình cho kết quả nhận dạng là gần 90%. Đồng thời, với hai phương pháp thực nghiệm cũng chỉ ra rằng, với ảnh được chia 50% cho train và test như ở phương pháp thực nghiệm một, chương trình cho ra tỷ lệ nhận dạng không có sự thay đổi nhiều (88% - 91%). Ở phương pháp thực nghiệm hai, chương trình cho ra tỷ lệ nhận dạng có sự tăng lên tỷ lệ thuận với số ảnh được train. Do đó, thuật toán này cũng chỉ ra rằng tỷ lệ nhận dạng sẽ ngày càng tăng nếu số lượng ảnh được train tăng lên. Tuy nhiên, có một điểm đáng lưu ý là thời gian thực hiện chương trình cho các phần thực nghiệm lại không có sự biến đổi nhiều (cho dù sự thay đổi của ảnh train hay ảnh test đầu vào). Như vậy, với các thông số tối ưu cho thuật toán LBP cùng với tập dữ liệu chuẩn, thuật toán có thể nhận dạng tới gần 90% ảnh đưa vào test.

## Nhận xét

Một khuyết điểm dễ nhận thấy ở LBP và các phương pháp mô tả cục bộ bằng lượng hóa vector là việc thay đổi thông số đầu vào. Theo nghĩa thông thường ảnh đầu vào thay đổi ít thì thay đổi ở kết quả đầu ra cũng không đáng kể, nhưng LBP lại không làm được.

LBP cũng không hiệu quả trên ảnh bị nhiễu hạt hay bề mặt phẳng. Nhiều biến thể của LBP đã được nghiên cứu để giải quyết các vấn đề này.

Dưới đây là bảng so sánh một số phương pháp nhận dạng trên tập dữ liệu AT&T.

Như vậy chúng ta có thể thấy được độ chính xác của LBP cũng khá cao (trên 80%) nhưng vẫn còn khá thấp hơn so với nhiều phương pháp khác. Cũng vì vậy mà LBP thường không được dùng riêng mà hay kết hợp với một số phương pháp khác như AdaBoost để có được hiệu quả tốt hơn.

# Tài liệu tham khảo

**Tiếng Anh**

1. Joni-Kristian Kämäräinen, Abdenour Hadid, and Matti Pietikäinen, *“Handbook of Face Recognition”*, 2nd Edition, pp. 79-108.
2. Matti Pietikäinen, Abdenour Hadid, Guoying Zhao, Timo Ahonen, “*Computer Vision Using Local Binary Patterns*”, Springer, pp. 13-65
3. Ajeet Singh, BK Singh, Manish Verma, “*Comparison of HGPP, PCA, LDA, ICA and SVM*”

**Trang web**

1. <http://docs.opencv.org/trunk/modules/contrib/doc/facerec/facerec_tutorial.html>
2. <http://www.bytefish.de/blog/local_binary_patterns>