

ĐỀ CƯƠNG NGHIÊN CỨU

THÔNG TIN NHÓM

Thành viên 1: Nguyễn Huỳnh Xuân Mai – 1712091

Thành viên 2: Nguyễn Gia Thuận – 1712174

Thành viên 3: Mai Công Trình – 1712840

THÔNG TIN ĐỀ TÀI

Tên đề tài: Face Recognition Using Near Infrared Images

Nguồn: Chương 15 – Handbook of Face Recognition

Từ khóa: Near Infrared Face Recognition

Nội dung trình bày:

1. Giới thiệu

1.1. Thông tin thành viên nhóm

1.2. Giới thiệu tổng quan chủ đề:

- Lí do tại sao cần đến nhận diện mặt cận hồng ngoại.
- Phương pháp nhận diện cận hồng ngoại hiệu quả khắc phục được những gì?
- Liệt kê các đề mục.

2. Thiết kế phần cứng:

- Giới thiệu phần cứng hoạt động: LED
- Mô hình hóa hình ảnh NIR hoạt động:
 - Nêu công thức hình ảnh $I(x, y)$ dưới nguồn sáng điểm được hình thành: $I(x, y) = \rho(x, y)\mathbf{n}(x, y)\mathbf{s}$, giải thích rõ các thành phần trong công thức ấy
 - Khi ánh sáng của mặt theo hướng phía trước là kết quả của thiết kế phần cứng, hình ảnh được xấp xỉ là:

$$I(x, y) \propto \rho(x, y)n_z(x, y)$$

3. Thuật toán

3.1. Dựa trên tính năng: LBP

- LBP code 3x3 window (có minh họa hình ảnh: local window, thresholded, weights).
- Áp dụng toán tử LBP cho hình ảnh NIR hoạt động sẽ tạo ra các tính năng bất biến chiếu sáng cho khuôn mặt.
- Mô tả cụ thể Phương pháp khớp khuôn mặt dựa trên LBP.
 - Gán ID.
 - Sử dụng mô hình cửa sổ trượt.
 - Biểu diễn lại ảnh bằng Histogram.
 - Nhận dạng: phương pháp sử dụng để xác định khoảng cách giữa 2 histogram là CHI-SQUARE.
- Mô tả cụ thể quy trình trích xuất các tính năng biểu đồ LBP.

3.2. AdaBoost

- Nêu lí do tại sao sử dụng phương pháp học dựa trên AdaBoost để chọn các tính năng LBP tốt nhất và xây dựng trình phân loại khuôn mặt.
- Vì AdaBoost về cơ bản học một trình phân loại hai lớp, nên chuyển đổi vấn đề đa lớp thành vấn đề hai lớp bằng cách sử dụng ý tưởng về sự khác biệt trong nội bộ và ngoại khóa (giải thích và đưa ví dụ cụ thể).
- Giải thích cụ thể biến đổi hình thành công thức kết hợp tuyến tính các phân loại yếu theo cách tối ưu thành phân loại mạnh hơn.

4. Xây dựng hệ thống

- 4.1. Phát hiện mặt và mắt: Nêu ví dụ, giải thích, hình ảnh minh họa.
- 4.2. Phát hiện mặt: Nêu ví dụ, giải thích, hình ảnh minh họa.

5. Thực nghiệm

- 5.1. Đánh giá cơ bản: Mô tả tập dữ liệu train và test, đưa hình ảnh kết quả
- 5.2. Chiếu sáng yếu: Mô tả tập dữ liệu train và test, đưa hình ảnh kết quả
- 5.3. Mắt kính: Mô tả tập dữ liệu train và test, đưa hình ảnh kết quả
- 5.4. Thời gian trôi: Mô tả tập dữ liệu train và test, đưa hình ảnh kết quả
- 5.5. Môi trường bên ngoài: Mô tả tập dữ liệu train và test, đưa hình ảnh kết quả

Nội dung báo cáo:

1. Giới thiệu

- 1.1. Thông tin thành viên nhóm + Phân công + Tiến độ/mức độ hoàn thành

1.2. Giới thiệu tổng quan chủ đề:

- Lí do tại sao cần đến nhận diện mặt cận hồng ngoại.
- Phương pháp nhận diện cận hồng ngoại hiệu quả khắc phục được những gì?
- Nêu ngắn gọn/liệt kê những đề mục.

2. Hệ thống NIR hoạt động

- Giới thiệu phần cứng hoạt động: LED
- Nêu các cầu chính cho hình ảnh hệ thống NIR
- Sự chọn lọc ánh sáng NIR từ hệ thống hình ảnh đạt được bằng phương pháp nào
- Minh họa hình ảnh thiết bị NIR và khuôn mặt (có giải thích ngắn gọn bên dưới)
- Minh họa hình ảnh so sánh ảnh được chụp bởi camera màu và hệ thống NIR (có giải thích ngắn gọn bên dưới)
- Tóm tắt các điểm chính của hệ thống hình ảnh được nâng cao lên

3. Chiếu sáng đại diện khuôn mặt bất biến

- Phân tích sử dụng mô hình hình ảnh bề mặt Lambertian để chỉ ra rằng hình ảnh NIR chứa thông tin nội tại, phù hợp nhất về khuôn mặt, chỉ chịu một hằng số nhân hoặc biến đổi đơn điệu do thay đổi cường độ ánh sáng. Sau đó, trình bày mô hình nhị phân cục bộ (LBP) để sửa đổi mức độ tự do của biến đổi đơn điệu để đạt được một đại diện bất biến chiếu sáng cho nhận dạng khuôn mặt.

3.1. Mô hình hóa hình ảnh NIR hoạt động:

- Nêu công thức hình ảnh $I(x, y)$ dưới nguồn sáng điểm được hình thành:

$$I(x, y) = \rho(x, y) \mathbf{n}(x, y) \mathbf{s}$$
 giải thích rõ các thành phần trong công thức ấy.
- Yếu tố quan trọng nhất ảnh hưởng đến hiệu suất nhận diện khuôn mặt là hướng chiếu sáng sự cố so với bề mặt bình thường. Sản phẩm của $\rho(x, y)$ và $\mathbf{n}(x, y)$ là thuộc tính nội tại của khuôn mặt ở tư thế cố định và là yếu tố duy nhất cần thiết để phát hiện và nhận diện khuôn mặt. Do đó, \mathbf{s} là thuộc tính bên ngoài cần được loại bỏ.
- Nêu lên sự thay đổi công thức thành: $I(x, y) = \kappa \rho(x, y) \cos \theta(x, y) \rightarrow$ hình ảnh khuôn mặt $\rho(x, y) \cos \theta(x, y)$ thay đổi khi hướng ánh sáng thay đổi. Thiết kế phần cứng hiện tại là nhằm bảo tồn tài sản nội tại trong khi giảm thiểu sự biến đổi do yếu tố bên ngoài của đèn môi trường.

- Lúc này, ảnh sẽ ở dạng công thức: $I(x,y) = \kappa \rho(x,y) n z(x,y)$

3.2. Đền bù cho hệ thống đơn điệu

- LBP code 3x3 (có minh họa hình ảnh: local window, thresholded, weights).
- Hình ảnh NIR và các tính năng LBP cùng nhau dẫn đến một đại diện bất biến chiếu sáng của khuôn mặt. Nói cách khác, áp dụng toán tử LBP cho hình ảnh NIR hoạt động sẽ tạo ra các tính năng bất biến chiếu sáng cho khuôn mặt.
- Mô tả cụ thể Phương pháp khớp khuôn mặt dựa trên LBP.
 - Gán nhãn ID.
 - Sử dụng mô hình cửa sổ trượt.
 - Biểu diễn lại ảnh bằng Histogram.
 - Nhận dạng: phương pháp sử dụng để xác định khoảng cách giữa 2 histogram là CHI-SQUARE.
- Mô tả cụ thể quy trình trích xuất các tính năng biểu đồ LBP.

4. Nhận dạng khuôn mặt NIR

4.1. Dựa trên tính năng: AdaBoost

- Nêu lí do tại sao sử dụng phương pháp học dựa trên AdaBoost để chọn các tính năng LBP tốt nhất và xây dựng trình phân loại khuôn mặt.
- Vì AdaBoost về cơ bản học một trình phân loại hai lớp, nên chuyển đổi vấn đề đa lớp thành vấn đề hai lớp bằng cách sử dụng ý tưởng về sự khác biệt trong nội bộ và ngoại khóa (giải thích và đưa ví dụ cụ thể).
- Giải thích cụ thể biến đổi hình thành công thức kết hợp tuyến tính các phân loại yếu theo cách tối ưu thành phân loại mạnh hơn.

4.2. Phân lớp LDA

- Nêu lí do tại sao sử dụng phân lớp LDA
- Giải thích cụ thể biến đổi hình thành công thức “khoảng cách cosin” dùng để matching.

5. Thực nghiệm:

- Minh họa những ưu điểm của phương pháp nhận dạng khuôn mặt NIR. Hiệu suất của động cơ khớp mặt LBP + AdaBoost và LBP + LDA NIR được so sánh với một số động cơ khớp mặt và đường cơ sở hiện có.

5.1. Đánh giá cơ bản: mô tả tập dữ liệu train và test, đưa hình ảnh kết quả

5.2. Chiều sáng yếu: mô tả tập dữ liệu train và test, đưa hình ảnh kết quả

5.3. Mắt kính: mô tả tập dữ liệu train và test, đưa hình ảnh kết quả

5.4. Thời gian trôi: mô tả tập dữ liệu train và test, đưa hình ảnh kết quả

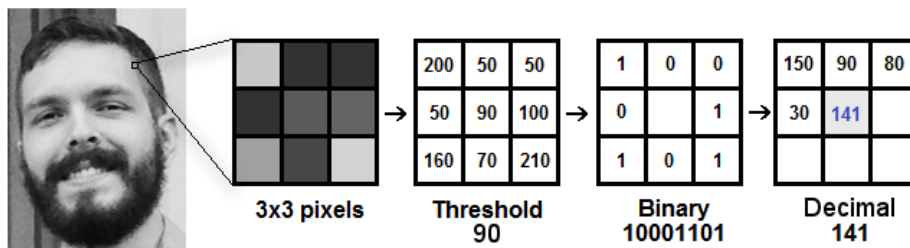
5.5. Môi trường bên ngoài: mô tả tập dữ liệu train và test, đưa hình ảnh kết quả

6. Kết luận

Xây dựng demo cho chủ đề nghiên cứu

- Phương pháp giải quyết vấn đề:

- Demo tự cài đặt
- Phương pháp sử dụng: Local Binary Pattern
- Sau đây là các bước tiếp cận của demo:
 - *Model LBP sử dụng:* radius = 1, neighbors = 8
 - *Chuẩn bị dữ liệu:* Gán nhãn (ID) cho mỗi ảnh mặt người trong tập dữ liệu có sẵn, thuật toán sẽ sử dụng những thông tin này để nhận dạng 1 ảnh input được đưa vào --> trả về ID phù hợp nhất.
 - *Áp dụng thuật toán:* Sử dụng mô hình cửa sổ trượt.



- Biểu diễn lại ảnh sau khi áp dụng LBP bằng Histogram
- Nhận dạng: đến bước này, mỗi ảnh trong tập dataset đều được biểu diễn bằng 1 histogram tương ứng. Vì vậy, khi nhận 1 ảnh đầu vào, ta thực hiện lại các bước trên để biểu diễn nó bằng 1 histogram. Output đầu ra được xác định bằng ID của ảnh có giá trị của histogram gần với histogram của ảnh đầu vào nhất.
- * Có nhiều phương pháp để xác định 'khoảng cách' giữa 2 histogram. Phương pháp được sử dụng trong demo là CHI-SQUARE.

- **Dữ liệu thực nghiệm:**

- Gồm dữ liệu của 1000 khuôn mặt của 50 người khác nhau (20 hình mỗi người)
- Nguồn: <http://www.yongxu.org/databases.html>
- Dữ liệu được phân loại lại để tiện cho quá trình huấn luyện:
https://drive.google.com/drive/folders/1sbK6RiQxdQtoUSoixOLj0Hteaan8LgiT?fbclid=IwAR3tmx8uC2Lrc8amPH0LH8Asxx2GSe1dAeUAu_eT-OjD3Oitt0giWXIFdxo

- **Thực nghiệm và đánh giá:**

<i>Ngữ cảnh</i>	<i>Mô tả</i>	<i>Kết quả dự kiến</i>
Chiếu sáng yếu	Hình ảnh camera được thu ở điều kiện ánh sáng yếu	Giá trị trung bình và phương sai là 0,6675 và 0,0377 cho các cặp trong lớp và 0,3492 và 0,0403 cho các cặp ngoài lớp Đạt độ chính xác nhận dạng 91.8%
Ảnh hưởng của khuôn mặt có đeo kính và không đeo kính	1500 hình ảnh của 30 đối tượng, 50 hình ảnh cho mỗi đối tượng trong đó 25 hình ảnh có kính và 25 hình ảnh không có. Các hình ảnh không có kính được sử dụng làm bộ train và các hình ảnh có kính làm bộ test	Với FAR = 0.1%, VR là 91.8%
Ảnh hưởng của tuổi tác	Sử dụng 750 hình ảnh của 30 người, 25 hình ảnh mỗi người; trong số 25 hình ảnh, 10 bức ảnh được chụp một năm trước và được sử dụng làm bộ train và 15 hình ảnh hiện tại được sử dụng làm bộ test	Với FAR = 0.1%, VR là 91.8%
Môi trường ánh sáng ngoài trời nắng gắt	Dùng đèn halogen 1000W chiếu sáng khi chụp ảnh	FAR = 0.01, VR = 85%

THÔNG TIN TỰ ĐÁNH GIÁ TIẾN ĐỘ

- Tài liệu nhóm kiểm được chưa đủ nhiều nên phần nội dung (trình bày + báo cáo) sẽ dần được bổ sung.
- Chưa cài đặt được Adaboost cho demo.
- Tập huấn luyện còn nhiều hạn chế, chưa phân loại tốt những hình ảnh với nhiều góc mặt khác nhau.
- Việc đánh giá và thực nghiệm là đánh giá mô hình dựa vào các điều kiện ngoại cảnh khác nhau thì nhóm vẫn đang phân vân rằng sẽ test mô hình và so sánh với các dữ liệu với trong sách.