LUẬN VĂN CAO HỌC

LỜI CẢM ƠN

Em xin chân thành cảm ơn Thầy Trương Quốc Bảo đã tận tình giúp đỡ và hướng dẫn để em có thể hoàn thành luận văn một cách tốt nhất.

Xin gửi lời cảm ơn trân quí đến nhà Trường Đại học Cần Thơ, khoa Sau đại học và Khoa Công nghệ thông tin và Truyền thông đã tổ chức khóa đào tạo cao học ngành Hệ thống thông tin để em có cơ hội được học tập và nâng cao, phát triển bản thân.

Con xin gửi lời cảm ơn đến gia đình, người thân và bạn bè, những người đã luôn ở bên cạnh là nguồn động lực để con có thể vượt qua những khó khăn và thực hiện tốt luận văn.

MỤC LỤC

DANH MỤC CÁC CHỮ VIẾT TẮT VÀ KÝ HIỆU

|  |  |  |
| --- | --- | --- |
| **Từ viết tắt** | **Cụm từ đầy đủ** | **Ý nghĩa** |
| AAM | Active Appearance Model | Thuật toán thị giác máy tính so khớp mô hình thống kê của hình dạng đối tượng |
| ASM | Active Shape Model | Mô hình thống kê về hình dạng đối tượng |
| PCA | Principal Component Analysis | Phân tích thành phần chính |
| AUs | Action Units | Các đơn vị vận động |
| Open CV | Open Source Computer Vision Library | Thư viện về thị giác máy tính mã nguồn mở |
| SVM | Support Vector Machine | Máy học véc tơ hỗ trợ |
| ANNs | Artificial Neutral Networks | Mạng nơ ron nhân tạo |
| HOG | Histogram of oriented gradients | Bộ mô tả tính năng dùng để phát hiện đối tượng |
| BSD | Berkeley Software Distribution | Bản quyền phân phối phần mềm nguồn mở |
|  |  |  |
|  |  |  |

DANH MỤC BẢNG

DANH MỤC HÌNH

LỜI MỞ ĐẦU

1. Tình hình nghiên cứu hiện tại

Thế giới đang bước vào cuộc cách mạng công nghiệp 4.0 với những bước chuyển biến tích cực và mạnh mẽ của tất cả các ngành nghề trong mọi lĩnh vực của đời sống hiện đại. Cùng với đó là sự phát triển vượt bậc của khoa học và công nghệ, sự hỗ trợ thiết yếu của công nghệ thông tin trong việc tin học hóa và đơn giản hóa các quy trình làm việc của mọi lĩnh vực khác nhau. Để thực hiện được điều đó, cần có sự tương tác giữa người và máy để máy tính có thể hiểu và thực hiện những công việc tự động theo mong muốn của con người.

Thật vậy, để máy tính có thể giao tiếp với con người thì chúng ta cần có những phương pháp và kỹ thuật cụ thể, một trong số chúng là khả năng nhận dạng được cảm xúc của con người. Một trong những phương pháp vật lý và có hiệu quả để nhận dạng được cảm xúc của con người là thông qua cả biểu cảm trên gương mặt. Đây là chủ đề chính trong các nghiên cứu của các nhà khoa học ở nhiều năm trước đây, và hiện tại nó vẫn được xem là một đề tài hấp dẫn vì tính ứng dụng và độ phổ biến.

Các ứng dụng thực tế của việc nhận dạng cảm xúc là rất đa dạng và hữu ích. Trong đời sống, các ứng dụng di động nhận dạng được cảm xúc của người dùng để gán các biểu tượng cảm xúc tương ứng như snow, magic, polygram đang được giới trẻ rất ưa chuộng. Trong y học, các bác sĩ có thể theo dõi các thay đổi cảm xúc của bệnh nhân để đưa ra những chẩn đoán bệnh chính xác và điều trị bệnh hiệu quả với các ca thần kinh hay rối loạn cảm xúc. Trong thương mại, các tập đoàn, doanh nghiệp, nhà sản xuất có thể thu thập, phân tích và thống kê cảm xúc của khách hàng để đưa ra những quyết định kịp thời và đúng đắn nhằm tối ưu hóa lợi nhuận.

2. Tính cấp thiết của đề tài

Từ những ứng dụng thiết thực trên, việc nhận dạng cảm xúc của con người dựa vào biểu cảm trên gương mặt là một chủ đề rất hay, tuy không mới nhưng tính khoa học và thực tiễn cao. Các nhà nghiên cứu đi trước đã có những công trình nghiên cứu hay với các phương pháp nghiên cứu khác nhau đã cho kết quả và độ chính xác nhất định, đây cũng làm một động lực để các đề tài sau có cơ sở khoa học để so sánh và đánh giá, cải tiến hơn nữa hiệu quả mang lại.

Với sự phát triển và phổ biến của mạng xã hội và công nghệ thông tin như hiện nay, việc tạo ra các ứng dụng để hiểu được người dùng hơn cả người dùng hiểu chính họ là một điều tuyệt vời. Và việc nhận dạng cảm xúc của con người dựa trên gương mặt là một nền tảng cốt lõi cho các ứng dụng này.

3. Mục tiêu của đề tài

Mục tiêu của đề tài là tìm hiểu về các loại cảm xúc của con người, các đặc trung của gương mặt người và sự tương quan giữa cảm xúc và các đặc trưng đó. Nghiên cứu các kỹ thuật, phương pháp, thuật toán để thực hiện các công việc cụ thể trong toàn bộ quá trình hoạt động của hệ thống. Thông qua đó, thực hiện huấn luyện được một tập dữ liệu có khả năng nhận dạng được cảm xúc của con người thông qua gương mặt. Cuối cùng, xây dựng được một hệ thống nhận dạng cảm xúc con người dựa trên các hình ảnh đầu vào.

4. Đối tượng, phạm vi nghiên cứu của đề tài

Đối tượng nghiên cứu: phương pháp phân khúc ảnh và tiền xử lý hình ảnh dựa vào đặc trung HOG; các kỹ thuật nhận dạng và xác định vị trí của các thành phần trên gương mặt như AAM/ASM, PCA, AUs; mô hình máy học SVM và mạng nơ ron nhân tạo, thư viện OpenCV.

Phạm vi nghiên cứu: cơ sở lý thuyết và ứng dụng thực tế của các phương pháp, kỹ thuật hỗ trợ nhận dạng để áp dụng vào hệ thống; dữ liệu đầu vào là các ảnh tĩnh, hoặc video.

5. Phương pháp nghiên cứu

Nghiên cứu cơ sở lý thuyết của các kỹ thuật hỗ trợ nhận dạng cảm xúc dựa trên mặt người. Nghiên cứu và so sánh kết quả thực nghiệm với những đề tài cùng chủ đề đã được thực hiện trước đó. Xây dựng một chương trình thực tế để kiểm tra độ chính xác và khả năng mở rộng của dự án. Ứng dụng kết quả nghiên cứu vào một ngữ cảnh thực tế của đời sống xã hội.

6. Kết quả dự kiến

Tìm hiểu được các loại cảm xúc cơ bản của con người, kiến thức nền tảng về các phương pháp hỗ trợ nhận dạng cảm xúc dựa trên mặt người, xây dựng chương trình kiểm tra với kết quả chấp nhận được và ứng dụng vào một lĩnh vực cụ thể.

CHƯƠNG 1. CÁC LOẠI CẢM XÚC CỦA CON NGƯỜI VÀ ĐẶC TRƯNG TRÊN GƯƠNG MẶT NGƯỜI

* 1. Các trạng thái cảm xúc của con người

Cảm xúc có tác động rất lớn đến cuộc sống của con người. Cảm xúc là một thứ rất phức tạp, có thể thay đổi nhanh chóng, một người có thể có hơn một cảm xúc tại một thời điểm. Vì vậy, cảm xúc là gì và con người có tất cả bao nhiêu cảm xúc là các câu hỏi chưa có một đáp án thuyết phục.

Vào thế kỷ IV trước công nguyên, Aristole đã nhận định có 14 loại cảm xúc cơ bản nhất, bao gồm: hài lòng, tử tế, tranh đua, ganh tị, đáng thương, câm phẫn, sợ hãi, bình tĩnh, tự tin, thù địch, giận dữ, xấu hổ, bằng hữu và vô liêm sĩ. Những năm gần đây, sáu là con số đại diện cho số lượng cảm xúc cơ bản nhất của con người được các nhà tâm lý học công bố, gồm có: vui vẻ, buồn bã, ngạc nhiên, sợ hãi, ghê tởm, giận dữ.

Ở một nhận định khác, TS. Rachael Jack của đại học Glasgow, Vương quốc Anh cho rằng con người chỉ có bốn loại cảm xúc cơ bản khi mà những biểu hiện của các cơ trên gương mặt trong hai cảm xúc sợ hãi và ngạc nhiên là như nhau, tương tự với giận dữ và ghê tởm. Thật sự là hai loại cảm xúc trong từng cặp ở trên có sự giống nhau trong quá trình vận động các nhóm cơ mặt để hình thành cảm xúc, chúng chỉ khác nhau khi được thể hiện đầy đủ và xong xuôi, vì thực chất chúng là những cảm xúc khác nhau.

Đề tài được thực hiện nhận dạng sáu loại cảm xúc cơ bản nhất của con người (vui vẻ, buồn bã, ngạc nhiên, sợ hãi, ghê tởm, giận dữ).

* 1. Đặc trưng của gương mặt người

Gương mặt là một phần của cơ thể người, bộ phận trung tâm để bộ lộ cảm xúc và là phương tiện truyền tải cảm xúc giữa người với người. Gương mặt người có các thành phần đặc trưng giống nhau như chân mày, mắt, mũi, miệng, v.v. nhưng mỗi người lại có một gương mặt khác nhau và là duy nhất. Do đó, gương mặt là đặc trưng tốt nhất để phân biệt một người với những người khác.

Gương mặt là nơi biểu hiện của những cảm xúc. Một nụ cười thể hiện cho một niềm vui, một cái chau mày đồng nghĩa với việc không tán thành hay khó chịu. Chính vì thế, nhận dạng cảm xúc của người đối diện qua gương mặt có một vai trò quan trọng trong giao tiếp. Con người có thể đọc được cảm xúc của người khác nhờ biểu cảm trên gương mặt của người đó, từ đó dự đoán được khả năng xảy ra của các hành vi tiếp theo.

Cơ mặt đóng một vai trò nổi bật trong việc mô tả cảm xúc con người, cùng với các đặc trưng khác trên gương mặt mang đến sự đa dạng trong sự biểu lộ nhiều cảm xúc khác nhau.

* 1. Cảm xúc dựa trên mặt người

Gương mặt là nơi cảm xúc được bộc lộ rõ nhất và dễ dàng nhận thấy nhất. Chính vì thế các biểu hiện trên gương mặt người được sử dụng như là một nền tảng cốt lõi để nhận dạng cảm xúc con người.

Tương ứng với từng cảm xúc riêng biệt là những biểu hiện khác nhau trên gương mặt. Như vậy, với một loại cảm xúc nhất định sẽ có một tập các biểu hiện nhất định của các nhóm cơ và các đặc trưng khác trên gương mặt.

Đối với một người, việc quan sát gương mặt và nhận dạng cảm xúc của một người khác là một điều rất dễ dàng. Vậy đối với máy tính thì như thế nào. Chúng cần được học, được huấn luyện để có được khả năng này như con người.

Chương 2. CÁC PHƯƠNG PHÁP GIÚP NHẬN DẠNG CẢM XÚC DỰA TRÊN GƯƠNG MẶT

* 1. Phương pháp dựa trên đặc trưng của gương mặt

Sử dụng phương pháp phân tích thành phần chính PCA. Phương pháp này sẽ lấy ra được các thành phần chính của gương mặt (thể hiện bằng các eigenvectors) trong tập ảnh huấn luyện và tạo ra không gian mặt. Tiếp tục sử dụng các thuật toán máy học để huấn luyện tạo thành các tập dữ liệu huấn luyện là các lớp tướng ứng với các loại cảm xúc cơ bản.

(Hình ảnh)

PCA là một công cụ mạnh mẽ cho việc xác định hình dạng của gương mặt, phân tích các thành phần trên gương mặt.

* 1. Phương pháp sử dụng các đơn vị vận động trên gương mặt

Cảm xúc được xác định dựa trên sự chuyển động của các đơn vị trên khuôn mặt, các action units. Có tất cả 64 action units tương ứng với 64 biểu hiện khác nhau của các nhóm cơ trên gương mặt [Ekman và Friesen 1978]. Việc xác định cảm xúc chỉ đơn giản là việc xác định có bao nhiêu action unit cùng xuất hiện trên gương mặt tại một thời điểm, và sự kết hợp của chúng cho ra một cảm xúc duy nhất. Ví dụ

Ưu điểm

Nhược điểm

* 1. Sử dụng mô hình AAM kết hợp tương quan điểm.

Đầu tiên, mô hình AAM được sử dụng để phát hiện vùng mặt, sau đó dựa vào sự phân bố hiện tại của các landmark và tỉ lệ giữa chúng để biết được tương ứng với loại cảm xúc nào.

Phương pháp này có ưu điểm rất lớn khi được sử dụng trong nhận dạng cảm xúc thời gian thực vì có tốc độ xử lý nhanh, thường được dùng trong nhận dạng cảm xúc người dùng qua video hay người dùng thực tế.

Một hạn chế của phương pháp này là việc xác định được ngưỡng tỉ lệ để nhận định đó là cảm xúc nào.

Chương 3. CƠ SỞ LÝ THUYẾT

3.1 Principal Component Analysis

PCA - Phân tích thành phần chính – là một trong những phương pháp phân tích dữ liệu nhiều biến đơn giản nhất, là một kỹ thuật dùng để làm nổi bật sự thay đổi và đưa ra các mô hình chính yếu trong tập dữ liệu, mô hình hóa tập dữ liệu, giúp cho việc phân tích và khám phá tập dữ liệu dễ dàng và chính xác.

PCA là một kỹ thuật có tính ứng dụng cao trong việc nhận dạng gương mặt, phổ biến để phát hiện mẫu của dữ liệu nhiều chiều.

Các tính năng chính của PCA:

Giảm số chiều của dữ liệu quan sát. Được sử dụng khi các mẫu và thông tin trong dữ liệu khó nhận ra trong không gian đa chiều.

PCA xây dựng một không gian mới với trục tọa độ mới có số chiều ít hơn, nhưng có khả năng thể hiện dữ liệu tương đương hoặc tốt hơn không gian cũ, nghĩa là đảm bảo độ biến thiên của dữ liệu trên mỗi chiều mới.

Các trục tọa độ trong không gian mới là tổ hợp tuyến tính của các trục tọa độ ở không gian cũ.

Trong không gian mới, các thông tin của dữ liệu có thể được phát hiện ở một khía cạnh khác mà ở không gian của không tìm thấy được, vì các liên kết tìm ẩn của dữ liệu có thể được khai phá.

Một số khái niệm toán học quan trọng được sử dụng trong PCA: độ lệch chuẩn (standard deviation), phương sai (variance), hiệp phương sai (covariation), giá trị riêng (aigenvalue) và vector riêng (eigenvector).

3.1.1 Độ lệch chuẩn

Độ lệch chuẩn là một đại lượng dùng để đo khoảng cách giữa các phần tử trong tập dữ liệu. Độ lệch chuẩn còn được hiểu là khoảng cách trung bình từ trung bình mẫu (X ngang) đến các điểm của dữ liệu.

Giả sử ta có tập dữ liệu X = [x1, x2, ..., x­­n]. Vậy trung bình mẫu X ngang là một giá trị nằm giữa của tập X:

X ngang =

Công thức tính độ lệch chuẩn (s):

s =

3.1.2 Phương sai

Phương sai là bình phương của độ lệch chuẩn. Phương sai là một đại lượng khác dùng để biểu diễn dữ liệu: đo khoảng cách giữa các phần tử trong tập dữ liệu.

Công thức tính phương sai (s2):

s2 =

3.1.3 Hiệp phương sai

Hai khái niệm độ lệch chuẩn và phương sai được sử dụng để biểu diễn dữ liệu một chiều, nhưng dữ liệu trong thực tế có nhiều hơn một chiều và có sự liên hệ với nhau mật thiết. Do đó, đại lượng hiệp phương sai ra đời để có thể tính toán và biểu diễn được dữ liệu đa chiều.

Hiệp phương sai thực chất chỉ tính toán được sự biến thiên của hai chiều dữ liệu. Nên ta có thể tính từng cặp chiều dữ liệu cho toàn bộ chiều của tập dữ liệu.

Công thức tính hiệp phương sai:

3.1.4 Vector riêng

3.1.5 Giá trị riêng

3.1.6 Các bước thực hiện cơ bản của PCA

Bước 1: Lấy dữ liệu đầu vào. Tính trung bình mẫu, hay còn gọi là vector kỳ vọng của toàn bộ dữ liệu.

Công thức:

Bước 2: Trừ mỗi điểm dữ liệu đi một lượng vector kỳ vọng của toàn bộ dữ liệu. Xét chiều dữ liệu ở chiều x đều có một giá trị trung bình mẫu. Thực hiện trừ lần lượt các giá trị chiều x cho trung bình mẫu.

Công thức:

Bước 3: Tính ma trận hiệp phương sai.

Công thức:

Bước 4: Tính các vector riêng và giá trị riêng của ma trận hiệp phương sai. Sắp xếp kết quả theo thứ tự giảm dần của giá trị riêng.

Bước 5: Chọn ra các thành phần chính. Chọn ra K vector riêng ứng với K giá trị riêng lớn nhất để xây dựng ma trận UK có các cột tạo thành một hệ trực giao. K vector này còn được gọi là các thành phần chính, tạo thành không gian con gần với phân bố của dữ liệu ban đầu đã chuẩn hóa. Tùy vào số lượng thành phần chính yêu cầu, lấy lần lượt các thành phần (các vector riêng) tương ứng có các giá trị riêng cao nhất.

Bước 6: Chiếu dữ liệu ban đầu đã chuẩn hóa xuống không gian con tìm được. Dữ liệu mới chính là tọa độ của các điểm trong không gian mới.

3.1.7 Một số hạn chế của PCA

Chỉ làm việc với dữ liệu numeric.

Nhạy cảm với các điểm outlier/extreme.

Do PCA hoàn toàn dựa trên các biến đổi tuyến tính, nên không phù hợp với các mô hình phi tuyến.

3.2 Mô hình ASM/AAM

3.3 Đặc trưng HOG, SHIF, Haar-like

HOG (Histogram of Oriented Gradients) là một bộ mô tả tính năng được sử dụng để phát hiện đối tượng trong thị giác máy tính và xử lý ảnh. HOG được tính toán trên một lưới dày đặc các ô và chuẩn hóa sự tương phản giữa các khối để nâng cao độ chính xác. HOG được dùng chủ yếu để phát hiện và mô tả hình dạng của một đối tượng trong ảnh. Bài toán tính toán HOG thông thường gồm năm bước chính:

Bước 1: Chuẩn hóa hình ảnh.

Bước 2: Tính gradient theo x và y.

Bước 3: Thống kê thành phần cùng trọng số trong mỗi ô.

Bước 4: Chuẩn hóa các khối.

Bước 5: Thu thập tất cả các biểu đồ cường độ gradient định hướng để tạo ra vector tính năng cuối cùng.

3.4 Open CV

OpenCV (Open Source Computer Vision Library) là một thư viện nguồn mở cho thị giác máy tính, xử lý ảnh và máy học, gồm các hàm chức năng tính toán được tích hợp sẵn, được phân phối dưới giấy phép BSD. OpenCV được thiết kế để tính toán hiệu quả và tập trung chủ yếu vào các ứng dụng thời gian thực.

OpenCV được sử dụng rộng rãi trong các ứng dụng như: kiểm tra và giám sát tự động, robot và xe hơi tự hành, phân tích hình ảnh y tế, tìm kiếm và phục hồi hình ảnh hay video, thực tế ảo, ….

3.5 Máy học SVM

(Ghi trong sách)

3.6 Mạng nơ ron nhân tạo

(Ghi trong sách)

CHƯƠNG 4. QUY TRÌNH THỰC HIỆN

4.1 Sơ đồ tổng quan

(Vẽ sơ đồ tổng quan của hệ thống)

4.2

CHƯƠNG 5. CÀI ĐẶT VÀ THỰC NGHIỆM

5.1 Yêu cầu phần cứng và phần mềm

5.2 Giao diện chương trình

5.3 Kiểm thử và kết quả

CHƯỚNG 6. ĐÁNH GIÁ VÀ KẾT LUẬN

6.1 Đánh giá kết quả đạt được

(Kèm theo các so sánh các phương pháp)

6.2 Kết luận

6.3 Thách thức trong nhận dạng cảm xúc dựa trên mặt người

Nhận dạng cảm xúc con người dựa trên các biểu cảm trên gương mặt là một bài toán khó và mang tính tương đối, vì cảm xúc của con người là vô cùng đa dạng. Ngoài sáu cảm xúc cơ bản, con người còn có rất nhiều cảm xúc khác với sự sai khác không rõ ràng và rất khó phân biệt.

Bên cạnh đó, phương pháp nhận dạng cảm xúc dựa trên gương mặt mang tính tương đối là do chúng ta chỉ dựa vào các nhóm cơ mặt và vị trí các thành phần trên gương mặt để xác định cảm xúc, và cảm xúc này có thể chưa thật sự đúng đắn.

Một khó khăn khác về môi trường thực hiện, hệ thống sẽ cho các kết quả không tốt với độ chính xác thấp nếu ảnh đầu vào có chất lượng kém, điều kiện ánh sáng không tốt, ảnh mờ hay vùng gương mặt khó nhận dạng, kích thước quá nhỏ.

6.4 Hướng phát triển

Cải tiến để có thể nhận dạng thêm được nhiều loại cảm xúc phức tạp hơn của con người.

Nghiên cứu và áp dụng hệ thống vào một ứng dụng thực tế và hữu ích.

Cải tiến độ chính xác của đề tài với nhiều dữ liệu huấn luyện hơn.

TÀI LIỆU THAM KHẢO

(Học cách lập tài liệu tham khảo tự động)