**ĐỀ CƯƠNG THỊ GIÁC MÁY TÍNH**

*VC GROUP – 2020*

***(Nguồn: Vở ghi trên lớp + Giáo trình Thị Giác Máy Tính)***

1. **Thế nào là lọc ảnh (định nghĩa, cách sử dụng và áp dụng cho các bài toán này)?**

* Lọc ảnh là thao tác biến đối dữ liệu đầu vào để sinh ra 1 ảnh đầu ra nhằm mục đích lọc tín hiệu không mong muốn (lọc nhiễu) hoặc tang cường tính chất nào đó của ảnh (độ sắc nét)
* Các phương pháp lọc ảnh:

+ Lọc ảnh trong miền không gian: Sử dụng hàm biến đổi nhận giá trị đầu vào là một bộ lọc (ma trận con) và lân cận của vị trí điểm ảnh đang xét để sinh ra 1 giá trị output tương ứng ở vị trí ảnh output

+ Lọc ảnh trong miền tần số:

/Lọc thông thấp: Chỉ cho đi qua các tần số thấp, loại bỏ các tần số cao, có tác dụng làm mờ ảnh, loại bỏ nhiễu

/Lọc thông cao: Chỉ cho đi qua các tần số cao, loại bỏ các tần số thấp,làm ảnh sắc nét hơn

/Lọc dải thông: Chỉ giữ lại tần số trong một dải nhất định

/Lọc chặn dải thông: Loại bỏ tần số trong một dải nhất định

1. **Trình bày về phép lọc trung vị và ứng dụng?**

* Lọc trung vị sử dụng 1 cửa số kích thước k\*k(k lẻ) di chuyển qua tất cả các điểm ảnh của ảnh input, tại một vị trí, lấy giá trị trung vị của lân cận tương ứng có điểm đang xét là trung tâm và gán cho điểm ảnh output
* Đặc điểm:

+ Rất hiệu quả với nhiễu muối tiêu vì các nhiễu đó tương ứng với các giá trị cực trị

+ Cho các giá trị điểm ảnh là thật

+ Cũng làm mịn nhưng không nhiều

* Ứng dụng:

+ Lọc nhiễu muối tiêu ảnh (rất hiệu quả)

1. **Trình bày về phép lọc Gaussian và ứng dụng?**

* Phép lọc Gaussian sử dụng một phép nhân chập một ảnh với một nhân Gauss có các trọng số được sinh theo công thức phân bố Gauss

Trong đó: (x,y) là vị trí trong nhân (với điểm trung tâm là (0,0)), là một hằng số dương điều chỉnh độ cao của nhân Gaussian

* Đặc điểm:

+ Có tác dụng lọc nhiễu và làm mịn tốt nhất

+ Có tốc độ nhanh

1. **Trình bày về thao tác nhân chập và ứng dụng?**

* Nhân chập: Sử dụng một nhân là một ma trận bao gồm các trọng số, di chuyển qua các điểm ảnh của ảnh Input. Tại mỗi vị trí, lấy tổng các tích của các điểm ảnh nằm trong nhân với trọng số của nhân đó và gán cho điểm ảnh Output
* Nhân của phép nhân chập là một ma trận các số thực được thiết kế theo quy tắc để phục vụ cho mục đích của phép lọc ảnh sử dụng nhân đó
* Ứng dụng: Phục vụ cho việc lọc ảnh

1. **Thế nào là phân lớp, một bài toán phân lớp? Trình bày các hiểu biết về phương pháp phân lớp kNN và SVM?**

***5.1. Phân lớp*** là thao tác dự đoán và gán nhãn (danh tính cho một bức ảnh hoặc là đối tượng trong một bức ảnh) hay còn gọi là nhận dạng. Phân lớp là một bài toán ở mức chung, còn nhận dạng là bài toán ở mức chi tiết

* Để thực hiện phân lớp, máy tính sẽ tìm một mô hình khớp nhất với dữ liệu có trước đó (tập dữ liệu huấn luyện/học) để đưa ra dự đoán và dữ liệu mới
  1. ***Phương pháp phân lớp KNN:***

+ Trong phương pháp này, có N lớp dữ liệu, mỗi lớp sẽ có ni phần tử, mỗi phần tử được biểu diễn bằng một vector đa chiều đặc trưng (có hang tram đến hang ngàn giá trị) được tính toán bằng cách sử dụng một thuật toán trích chọn đặc trưng và một nhân. Tập dữ liệu được chia thành 2 tập con là tập huấn luyện và tập kiểm thử

+ Một số k lẻ và một hàm khoảng cách sẽ được lựa chọn

+ Khi cần phân lớp/dự đoán/nhận dạng 1 mẫu (chưa có nhãn), kNN sẽ đi tìm k mẫu trong tập huấn luyện (tham chiếu) có độ tương đồng lớn nhất với mẫu cần nhận dạng bằng cách sử dụng hàm khoảng cách (càng lớn thì độ tương đồng càng bé)

+ Nếu có ít nhất (k/2+1) mẫu trong số k mẫu tương đồng nhất cùng thuộc về 1 lớp thì gán nhãn của các mẫu đó cho mẫu cần nhận dạng

+ kNN là một phương pháp phân lớp đơn giản, không đòi hỏi phải có bước học và hiệu quả trong đa số các bài toán, tuy nhiên độ chính xác phụ thuộc nhiều vào hàm khoảng cách

* 1. ***Phương pháp phân lớp SVM***

+ Là phương pháp xuất phát từ lĩnh vực học máy dựa trên ý tưởng: Một bộ phân lớp tốt không cần hiểu biết nhiều về dữ liệu

***+ Phân lớp SVM nhị phân:***

Input: Hai tập dữ liệu là huấn luyện và kiểm thử, mỗi phần từ là một vector N chiều có nhãn là +1 và -1

Output: Một mô hình cho phép dự đoán bất cứ mô hình dữ liệu nào theo format tương tự của tập huấn luyện

Thuật toán:

B1: Khởi tạo hệ số của một mặt phẳng trong không gian N chiều

B2: Giải một hệ phương trình tuyến tính bậc 1 để tìm các hệ số của mặt phẳng sao cho:

+ Mặt phẳng cách đều điểm gần nhất của 2 tập hợp điểm tương ứng với nhãn +1 và -1

+ Khoảng cách ngắn nhất từ mỗi tập hợp đến mặt phẳng là bằng nhau và là lớn nhất

+ Kết quả là một siêu phẳng tối ưu

B3: Sử dụng siêu phẳng dể phân lớp bằng cách tính khoảng cách từ điểm cần phân lớp đến siêu phẳng, nếu dấu của khoảng cách là âm thì có nhãn là -1, ngược lại là +1

SVM nhị phân nhanh với số chiều nhỏ (vài nghìn tới vài chục nghìn) và có độ chính xác cao

***+ Phân lớp SVM đa nhãn:***

Input: Tập dữ liệu huấn luyện và kiểm thử gồm các phần tử dưới dạng vector N chiều với các nhãn (với số nhãn lớn hơn 2)

Output: Siêu phẳng – mô hình phân lớp được nhãn

Thuật toán: Nếu có K lớp, ta sẽ đi tìm K siêu phẳng nhị phân:

Với mỗi lớp thứ i:

/Các lớp thứ i có nhãn là +1

/Các lớp còn lại có nhãn là -1

/Tìm siêu phẳng nhị phân giúp phân lớp một phần tử dữ liệu thuộc lớp i hoặc không

Trên thực tế, bộ phân lớp SVM không mấy hiệu quả vì:

+ Độ chính xác hạn chế

+ Chậm vì phải tìm K siêu phẳng

1. **Trình bày về mạng nơ ron sử dụng cho bài toán phân lớp (các lớp Input, Hidden, Ouput, hàm kích hoạt) và cách tính số tham số của 1 mạng nơ ron?**

- Là cấu trúc mô phỏng cấu trúc và cách hoạt động của bộ não động vật cao cấp

- Mạng noron sẽ bao gồm nhiều lớp, mỗi lớp có nhiều nơ ron/nút mô phỏng các nơ ron sinh học

***- Cấu trúc một mạng nơ non:***

+ Có một lớp Input để nhận dữ liệu mà bài toán cần phải xử lý, có số nút bằng với số thuộc tính mô tả mỗi phần tử dữ liệu

+ Có một lớp Output có số nút bằng với số giả thiết khác nhau có thể có của bài toán

+ Có 2-3 lớp nằm giữa Input và Output gọi là các lớp ẩn có số các noron khác nhau tùy thuộc vào bài toán và thực hiện các thao tác tính toán chính của mạng

+ Mạng nơ ron truyền thống sử dụng cac kết nối đầy đủ, mỗi nơ ron của một lớp sẽ kết nối với tất cả các nút trong lớp ngay sau đó

+ Ở mỗi lớp, thao tác tính toán sẽ thực hiện qua các bước:

/Sử dụng một hàm tính tổng tuyến tính của các tham số nhận được từ các nơ ron ở lớp trước và các trọng số tương ứng với các kết nối đó – hàm f

/Sử dụng một hàm kích hoạt để điều chỉnh giá trị của hàm f về khoảng quy ước và mô phỏng quan hệ phức tạp giữa cac noron trong mạng

+ Một mạng nơ ron sẽ được khởi tạo với các tham số (các trọng số và các biến độ nghiêng) hoàn toàn ngẫu nhiên

+ Mạng nơ ron phải trải qua quá trình huấn luyện sử dụng tập dữ liệu huấn luyện, quá trình này sẽ tiến hành điều chỉnh hệ số của mạng

1. **Thế nào là trích chọn đặc trưng và thuật toán trích chọn đặc trưng? Trình bày phương pháp trích chọn đặc trưng LBP và ứng dụng? Trình bày phương pháp trích chọn đặc trưng SIFT và ứng dụng?**
   1. ***Trích chọn đặc trưng***

- Trích chọn đặc trưng: Đặc trưng của một bức ảnh là các đặc điểm có tính phân biệt, cho phép phân biệt ảnh đó với các ảnh khác

- Một phương pháp trích chọn đặc trưng là một thuật toán/hàm biến đổi/hàm tính toán nhận dữ liệu input là một bức ảnh, thực hiện biến đổi và tính ra các giá trị đại diện cho đặc trưng của bức ảnh. Các giá trị đại diện này tạo thành một vector, gọi là vector đặc trưng. Vector đặc trưng có hang tram tới hang ngàn giá trị

* 1. ***LBP***

+ Toán tử LBP là một phương pháp tính toán dùng để trích chọn đặc trưng cục bộ bằng cách so sánh mỗi điểm ảnh Input với P điểm lân cận nằm cách đều nhau trên một đường tròn bán kính R và điểm đang xét là tâm của một đường tròn. Nếu điểm lân cận có giá trị lớn hơn hoặc bằng điểm ở tâm thì kết quả là một bit 1, ngược lại là một bit 0 tại mỗi vị trí điểm lân cận. Sau đó, dãy các bit được ghép lại và chuyển sang hệ só 10, giá trị nhận được gọi là nhãn LBP của điểm đang xét

+ Thuật toán:

Input: Một ảnh xám, một toán tử LBP, tham số chia MxN

Output: Một vector LBP có kích thước M\*N\*59 giá trị

Thuật toán:

B1: Áp dụng toán tử LBP để sinh ảnh LBP

B2: Chia ảnh LBP bằng các vùng con hình chữ nhật bằng nhau, không giao nhau, tính histogram của các vùng con đó. Mỗi dãy histogram có 256 giá trị

B3: Ghép các dãy histogram của các vùng con (theo thứ tự) lại thành một vector histogram có độ dài M\*N\*256

B4: Sử dụng các màu đồng nhất (các giá trị LBP có số lần chuyển đổi từ 0 sang 1 không vượt quá 2) để đánh số lại các giá trị nhị phân trong các chuỗi histogram -> Giảm độ dài vector LBP từ N\*N\*256 xuống M\*N\*59

+ Ứng dụng:

/ Giải quyết các bài toán phân lớp vân ảnh

/ Sau đó LBP được mở rộng và được áp dụng rất hiệu quả cho bài toàn trích chọn mặt người

+ Tính chất LBP:

/Nhanh, đơn giản, hiệu quả

/Bất biến đối với thay đổi về cường độ sáng đơn điệu

/Rất nhiều phương pháp khác dựa vào LBP

* 1. ***SIFT***

- Dựa trên các điểm mốc của ảnh, có tính bất biến, không thay đổi ở các tỉ lệ/kích thước khác nhau của ảnh. Các đặc trưng là các giá trị tính được, là kết quả của việc tính toán trong các điểm mốc

- SIFT hiệu quả, có bản quyền

- Các bước của thuật toán SIFT

B1: Xác định các điểm keypoint

B2: Sàng lọc các điểm keypoint tiềm năng ở B1

B3: Gán hướng

B4: Biểu diễn đặc trưng SIFT cho mỗi điểm keypoint

B5: Đối sánh hai ảnh dựa trên vector SIFT

- Tính chất: Chậm, hiệu quả, sử dụng trong nhiều ứng dụng thị giác máy tính

1. **Trình bày phương pháp giảm số chiều PCA và ứng dụng**

= PCA là một thuật toán thống kê sử dụng biến đổi trực giao để biến đổi một tập các biến quan sát (có thể có tương quan với nhau) thành một tập các biến không có tương quan tuyến tính với nhau

- Thuật toán PCA (Giáo trình Thị Giác Máy Tính trang 51 – Rất phức tạp và đặc sệt công thức tính toán)

- Ứng dụng:

+ Giảm số chiều cho các dữ liệu lớn

+ Được dùng như một thuật toán trích xuất đặc điểm cho hệ thống eigenfaces để nhận dạng mặt

+ Nén dữ liệu ảnh

+ Biểu diễn trực quan các dữ liệu đa chiều trong thống kê

1. **Mô tả các thành phần và các bước xây dựng một hệ thống nhận dạng đối tượng (trong ảnh)? Cho ví dụ cụ thể về hệ thống nhận dạng mặt/vân tay?**

*Các bước xây dựng một hệ thống nhận dạng đối tượng (trong ảnh): Một hệ thống nhận dạng đối tượng thông thường xử lý 4 bước sau:*

*• Thu nhận tín hiệu (hình ảnh) và tiền xử lý.*

*• Trích chọn đặc trưng.*

*• Phát hiện đối tượng (detection).*

*• Phân lớp đối tượng (classification)*

*Ảnh đầu vào được thu nhận và tiền xử lý sau đó các ảnh được qua công đoạn tiền xử lý nhằm tăng độ chính xác cho hệ thống. Các ảnh sau đó được trích chọn đặc trưng để tạo ra các vector đặc trưng trong bước rút trích đặc trưng. Những vector đặc trưng này sẽ là dữ liệu đầu vào cho 1 mô hình đã được huấn luyện trước. Phát hiện đối tượng: dò tìm và định vị những vị trí đối tượng xuất hiện trong ảnh. Những đối tượng thu được qua bước phát hiện sẽ tiếp tục được phân lớp thành từng lớp riêng biệt để nhận dạng.*

*Ví dụ cụ thể về hệ thống nhận dạng mặt: Một hệ thống nhận dạng mặt (cục bộ) gồm có nhiều modules khác nhau có chức năng riêng:*

*• Module phát hiện khuôn mặt (face detection): nhằm mục đích phát hiện vùng ảnh chứa mặt cần nhận dạng, module này còn thực hiện cắt lấy phần ảnh chứa đúng khuôn mặt để nhận dạng.*

*• Module tiền xử lý (preprocessing): thực hiện việc chuẩn hóa ánh sáng nhằm làm giảm ảnh hưởng của yếu tố ánh sáng (ví dụ bằng cách thực hiện cân bằng histogram).*

*• Tiếp đến là module quan trọng nhất thực hiện việc trích chọn các đặc điểm của ảnh (feature extraction): ở bước này LBP hay SIFT có thể được sử dụng*

*• Tiếp đến là bước giảm số chiều của vector đặc điểm (feature vectors) nhận được từ các bước trước nhằm loại bỏ các thông tin thừa, tăng tính phân biệt giữa các bức ảnh của các đối tượng khác nhau.*

*• Cuối cùng là bước phân lớp (classification) nhằm tìm ra định danh hoặc đưa ra quyết định về bức ảnh cần nhận dạng, ở đây có thể sử dụng kNN hoặc SVM. Ở bước phân lớp, các ảnh sẽ được biểu iễn dưới dạng các vector nên việc đo khoảng cách hay độ tương đồng có thể được tính dựa trên các hàm khoảng cách trong không gian Ơclit.*

*Ví dụ về hệ thống nhận dạng vân tay: Sử dụng biến đổi sóng nhỏ Gabor và Minutiae matching:*

*• Bước đầu tiên của thuật toán (chuẩn hóa, tìm hướng của các ridge, sử dụng phép biến đổi Gabor) nhằm mục đích tăng cường chất lượng ảnh trong khi thuật toán làm mảnh (thinning) nhằm mục đích làm nổi bật các đặc điểm của vân tay cho các bước trích chọn đặc điểm (các đặc điểm minutiae), bước cuối cùng là thực hiện phân lớp bằng thuật toán đối sánh mẫu với các mẫu là các đặc điểm minutiae.*

*• Để thực hiện chuẩn hóa (normalization), thuật toán cân bằng histogram được thực hiện trên ảnh input, kết quả nhận được ảnh có độ tương phản tốt hơn.*

*• Tiếp đến sử dụng toán tử Sobel được sử dụng để tính gradient của ảnh, sau đó hướng của các đặc điểm ridge sẽ được tính toán dựa trên gradient đối với các vùng ảnh có kích thước WxW (W = 16).*

*• Bộ lọc Gabor được sử dụng để loại bỏ các nhiễu và giữ lại các cấu trúc của vân tay (ridge/valley). Sau đó ảnh được nhị phân hóa với một ngưỡng được chọn từ các thử nghiệm thực tế sao cho kết quả nhận được là tốt nhất.*

*• Thuật toán làm mảnh được chọn để làm nổi bật các đặc điểm của vân tay.*

*• Ở bước trích chọn đặc điểm, các đặc điểm minutiae (ridge pixel, ridge – ending and ridge – bifurcation) được trích xuất dựa trên các luật.*

*• Cuối cùng ở bước đối sánh mẫu, ta có một bảng điểm minutiae. Các điểm minutiae được lưu ở dạng tập 3 điểm một và việc đối sánh hai vân tay được thực hiện bằng cách khớp giữa các tập 3 điểm này (thuật toán brute force).*

10. Thế nào là histogram của một ảnh ? Mô tả các bước tính histogram của 1 ảnh và đoạn giả mã lệnh (thuật toán) thực hiện việc đó ?

* Histogram của một ảnh là một biểu diễn trực quan (dạng hình ảnh) về tấn suất xuất hiện của các giá trị cường độ sáng (màu sắc) trong ảnh
* Để tính histogram của ảnh ta phải thực hiện phương pháp đếm phân phối (đếm số lần xuất hiện của các giá trị cường độ sáng trong ảnh)

Giả mã lệnh :

* histogram[256];
* memset(histogram, 0, 256\*sizeof(int));
* for(int i=0;i<img.rows;i++)
* for(int j=0;j<img.cols;j++)
* histogram[img.at<int>(i,j)] ++;