



Đề cương ôn tập Học phần

(2 trang)

Phần A: Lý thuyết

1. Thế nào là lọc ảnh (định nghĩa, cách sử dụng và áp dụng cho các bài toán này)?
2. Trình bày về phép lọc trung vị và ứng dụng?
3. Trình bày về phép lọc Gaussian và ứng dụng?
4. Trình bày về thao tác nhân chập và ứng dụng?
5. Thế nào là phân lớp, một bài toán phân lớp? Trình bày các hiểu biết về phương pháp phân lớp kNN và SVM?
6. Trình bày về mạng nơ ron sử dụng cho bài toán phân lớp (các lớp Input, Hidden, Output, hàm kích hoạt) và cách tính số tham số của 1 mạng nơ ron?
7. 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? phương pháp trích chọn đặc trưng Gabor Wavelet?
8. Trình bày phương pháp giảm số chiều PCA và ứng dụng.
9. 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?

Phần B: Bài tập Yêu cầu:

1. Hiểu được cách thực hiện phép nhân chập
2. Hiểu được cách thực hiện thuật toán lọc ảnh (trung vị, trung bình, Gaussian)
3. Hiểu được cách thực hiện phân lớp kNN.
4. Hiểu được cách thực hiện thuật toán trích chọn đặc trưng LBP

Bài tập:

5. Xem lại các bài tập trong bài giảng Tham khảo:
 1. Bài giảng trên lớp.
 2. Medium.com



Đề cương ôn tập Thị giác máy tính



Nhân chập (convolution): Cho ảnh 4x4 sau:

1	2	10	6
2	8	3	1
6	9	4	5
1	0	2	3

Hãy thực hiện nhân chập ảnh trên với nhân 3x3 sau:

1	0	-1
1	0	-1
1	0	-1

	-8	7	
	0	8	

Ví dụ về kNN (k=1):

Tập tham chiếu/huấn luyện (training/reference) gồm 4 vector đặc trưng tương ứng với 3 lớp:

$V1=(1, 5)$, $v2 = (4, 6)$, $v3 = (7, 4)$, $v4 = (8, 6)$

$V1, v2$ thuộc lớp thứ nhất (label = 1).

$V3$ thuộc lớp thứ hai (label = 2).

$V4$ thuộc lớp thứ ba (label = 3).

Hỏi mẫu $T = (3, 4)$ thuộc lớp nào nếu dùng hàm khoảng cách Ơclit, ChiSquare?

Hàm khoảng cách Ơclit:

$$D(V1, T) = \sqrt{(1-3)^2 + (5-4)^2} = \sqrt{5}$$

$$D(V2, T) = \sqrt{(4-3)^2 + (6-4)^2} = \sqrt{5}$$

$$D(V3, T) = \sqrt{(7-3)^2 + (4-4)^2} = 4$$

$$D(V4, T) = \sqrt{(8-3)^2 + (6-4)^2} = \sqrt{29}$$

Do khoảng cách từ T đến $V1, V2$ là nhỏ nhất nên nhãn của T sẽ là 1.

Dùng hàm khoảng cách Chi square: (χ - bình phương)



Đề cương ôn tập Thị giác máy tính



$$\sum_{i=1}^n \frac{(x_i - y_i)^2}{(x_i + y_i)}$$

$$\sum_{i=1}^n n(x_i - y_i)^2(x_i + y_i)$$

$$D(V1, T) = (1-3)^2/(1+3) + (5-4)^2/(5+4) = 4/4 + 1/9 = 10/9$$

$$D(V2, T) = (4-3)^2/(4+3) + (6-4)^2/(6+4) = 1/7 + 4/10 = 1/7 + 2/5 = 19/35.$$

$$D(V3, T) = (7-3)^2/(7+3) + (4-4)^2/(4+4) = 16/10.$$

$D(V4, T) = (8-3)^2/(8+3) + (6-4)^2/(6+4) = 25/11 + 4/10$. Vậy khoảng cách ngắn nhất là từ V2 tới T nên nhãn của T là 1.

Cho ảnh xám như sau:

123	27	9	231	20
89	189	20	9	15
0	255	28	145	180
235	132	90	65	176
98	45	8	120	31

Hãy tính ảnh LBP^{8,1} tương ứng.

Xét tại ô (1, 1) có mức xám bằng 189: xâu nhị phân là 00000100=>4

Xét tại ô (2, 1) có mức xám bằng 20: xâu nhị phân là 10101111=>175

...

Vậy ảnh LBP^{8,1} nhận được là:

	4	175	...	
	
	



Đề cương ôn tập Thị giác máy tính



Xem lại phần phân lớp bằng mạng nơ ron, cấu trúc mạng nơ ron, cách tính số tham số của mạng nơ ron.

Cho biết cấu trúc (các thành phần) của 1 một mạng nơ ron? Giả sử cần xây dựng 1 mạng nơ ron để nhận dạng cảm xúc (7 cảm xúc) từ ảnh có kích thước 48x48 với 2 lớp ẩn (có số nơ ron tùy chọn), hãy vẽ mô hình mạng tương ứng và cho biết số tham số của mạng đó. Batch_size, epoch, training, validation, testing

Loss/error

Cho ảnh sau:

123	27	9	231	20
89	189	20	9	15
0	255	28	145	180
235	132	90	65	176
98	45	8	120	31

Hãy thực hiện nhân chập ảnh với nhân:

0	-1	0
-1	5	-1
0	-1	0

Dùng hàm khoảng cách $d(S1, S2) = \text{abs}(x1-x2) + \text{abs}(y1-y2)$, $S1 = (x1, y1)$, $S2 = (x2, y2)$.

Với $V1=(1, 5)$, $v2 = (4, 6)$, $v3 = (7, 4)$, $v4 = (8, 6)$

Hỏi mẫu $u = (4, 1)$ thuộc về lớp nào?

$D(v1, u) = \text{abs}(1-4) + \text{abs}(5-1) = 7$.

$D(v2, u) = \text{abs}(4-4) + \text{abs}(6-1) = 5$.

$D(v3, u) = \text{abs}(7-4) + \text{abs}(4-1) = 6$.

$D(v4, u) = \text{abs}(8-4) + \text{abs}(6-1) = 9$.

Vậy khoảng cách từ $v2$ tới u là nhỏ nhất nên nhãn của u là 1, hay u thuộc lớp thứ nhất.



Đề cương ôn tập Thị giác máy tính



$$G(x,y, \sigma) = e^{-(x^2+y^2)/2\sigma^2}/2\pi\sigma^2$$

0.059	0.091	0.059
0.091	0.16	0.091
0.059	0.091	0.059

Lấy $\sigma = 1$:

$$G(0,0) = 1/2\pi = 0.16$$
$$G(0,1) = e^{-1/2}/2\pi = 0.091$$

$$G(1,0) = 0.091$$

$$G(1,1) = e^{-2/2}/2\pi = 0.059$$

Tính histogram của vùng ảnh:

Phép chia nguyên, phép mod (phép chia lấy dư)