Question 1: Morphological operators and Filtering(20 pts.)

a) 1. Given the original image shown in Figure 1, and the structuring element

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

describe which morphological operator (erode, dilate, open, close) has been used to create the four images labeled A, B, C and D (Figure 1). **2 pts.**

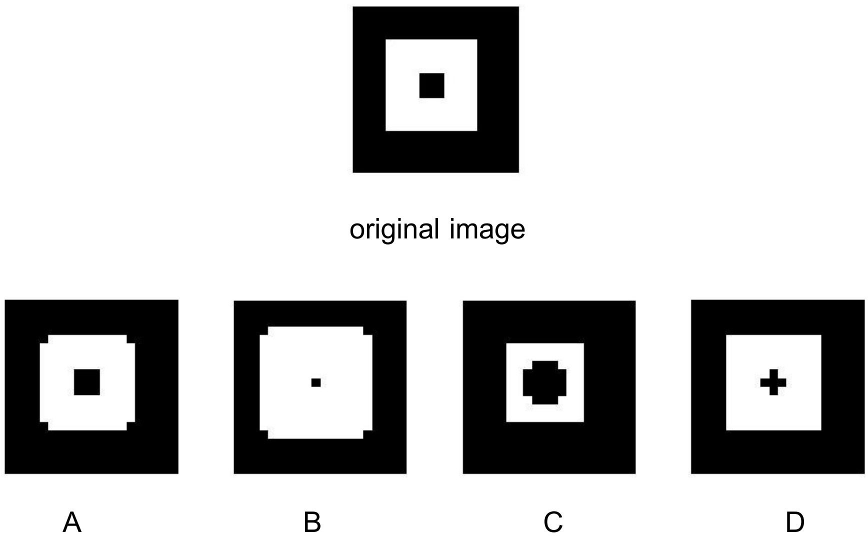


Figure 1: (Top) Input image, (Bottom) Output images

ANSWER

A: open
B: dilate
C: erode
D: close

ANSWER

2. Given the image in Figure 2(a) and the structuring element

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

perform an erosion operation. Assume the origin of the structuring element is in the center. Ignore cases where the structuring element extends beyond the image. Show your answer by writing 0's and 1's in Figure 2(b). **6 pts.**

1	1	1	1	1	1	1
1	0	0	1	0	0	1
1	1	1	1	1	1	1
1	0	0	0	0	0	1
1	1	1	1	1	1	1
1	1	1	1	1	1	1
1	1	0	0	1	1	1

(a)

(b)

Figure 2: Erosion

ANSWER

ANSWER

b) 1. The image in Figure 3(a) was filtered using one of these filters, 1. ideal lowpass filter; 2. Gaussian lowpass filter; 3. median filter, to result in the image shown in Figure 3(b). The small black square on the lower right hand corner of the original image shows the size of the mask that was used. (That small square is not part of the image.) For each of these filters listed above, explain why you think it was, or was not, the filter actually used. **3 pts.**

ANSWER

Can't be the ideal low pass filter, since most boundaries are not blurred, and there is no ringing effect either. (1 pt.) Can't be the Gaussian filter because the edges are not blurred. (1 pt.) Must be the median filter, since it reduces noise, eliminates small objects and does not blur edges. (1 pt.)

1	1	1	1	1	1	1
1	0	0	1	0	0	1
1	1	1	1	1	1	1
1	0	0	0	0	0	1
1	1	1	1	1	1	1
1	1	1	1	1	1	1
1	1	0	0	1	1	1

1	1	1	1	1	1	1
1	0	0	**0**	0	0	1
1	**0**	**0**	**0**	**0**	**0**	1
1	0	0	0	0	0	1
1	1	**0**	**0**	**0**	**0**	1
1	**0**	**0**	**0**	**0**	1	1
1	1	0	0	1	1	1

ANSWER

2. What is the minimum number of multiplications needed to compute the convolution between an image of size $N \times N$ pixels and the 3×3 filter (K) shown below? (Neglect the border issues and justify your answer.) **3 pts.**

$$K = \begin{bmatrix} 0.0749 & 0.1236 & 0.0749 \\ 0.1236 & 0.2060 & 0.1236 \\ 0.0749 & 0.1236 & 0.0749 \end{bmatrix}$$

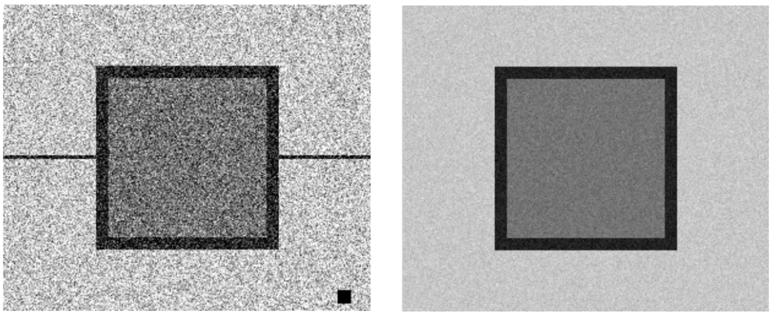
ANSWER

Since the filter is separable (1 pt), it is $6N^2$ *(2 pts).* $(9N^2$ *or* $9(N-2)^2$ *1 pt)* $(6N^2$ *or* $6(N-2)^2$ *3 pts)* $(3N^2$ *or* $3(N-2)^2$ *(buffer solution) 3 pts)*

ANSWER

3. Given a 3x3 filter shown below:

$$\begin{bmatrix} -k/8 & -k/8 & -k/8 \\ -k/8 & k+1 & -k/8 \\ -k/8 & -k/8 & -k/8 \end{bmatrix}$$



(a) Original Image (b) Filtered Image

Figure 3:

What is the effect of convolving an image with this filter for $k = 0$ ad $k = 2$. **4 pts.**

ANSWER

*No change for $k = 0$. (1 pt) For $k = 2$..the result is sharpening or edge enhancement in the original image (1 pt). because this means convolving the image with delta+2*sharpening filter (2 pts). (http://www.dspguide.com/ch24/2.htm)*

ANSWER

4. Is the bilateral filter spatially invariant? Justify your answer. **2 pts.** ANSWER

No (1 pt). Because it is a product of gaussian on distance and gaussian on intensity difference. (1 pt)

ANSWER

Question 2: Segmentation and Fourier Transform (20 pts.)

a) Imagine a classifier which decides whether a patient is infected with flu. Assume that the cost of having a false negative is five times higher than the cost of a false positive and the cost of true negatives and true positives is zero. One out of six patients that you check is infected. The ROC curve of your classifier can be parameterized as

$$y = f(x) = \frac{\log(1+x)}{\log(2)}.$$

Derive the cost function of the classifier in terms of the x and y coordinates of the ROC curve. Find the point on the ROC curve, where this cost function reaches its minimum value. **7 pts.**

ANSWER

Assuming that P and N are the numbers of positive and negative patients respectively, C_{FN} is the cost of having a false negative and C_{FP} is the cost of having a false positive: $C_{TP} = 0$ and $C_{TN} = 0$.

$$C_{FN} = (5 * C_{FP}) \tag{1}$$

Therefore, the cost function,
$C = k(5 * FN + FP)$

If $X = \frac{FP}{N}$ *and* $Y = \frac{TP}{P}$

$$\begin{aligned} C(X,Y) &= k(5(P - TP) + FP) * \frac{PN}{PN} \\ &= kPN(\frac{5}{N} - \frac{5Y}{N} + \frac{X}{P}) \end{aligned}$$

(2 pts)
Solution 1:

if $(C^{'}(X,Y) = 0)$ *then (1 pt),*

$Y^{'} = \frac{N}{5P}$ *(2 pts)*

$$f'(x) = \frac{1}{(1+x)\log(2)} = \frac{N * C_{FP}}{P * C_{FN}} = 1 \tag{2}$$

(1 pt for calculating $f'(x)$*) OR*

$$f'(x) = \frac{1}{(1+x)\log(2)} = \frac{N}{P * 5} = 1 \tag{3}$$

which leads to $(x, f(x))$ *with* $x = \frac{1}{\log(2)} - 1$*. (1 pt)*
OR

Solution 2:

$$\begin{aligned} C(X,Y) &= k(5(P - TP) + FP) * \frac{PN}{PN} \\ &= kPN(\frac{5}{N} - \frac{5Y}{N} + \frac{X}{P}) \end{aligned}$$

Minimize $C(x,y)$*, using the constraint* $f(x) - y = 0$*. Using lagranges multipliers, we need to minimize,*

$$C(x,y) + \lambda(f(x) - y)$$

(1 pt)

Therefore,

$$\frac{\partial C}{\partial x} + \lambda \frac{\partial f(x)}{\partial x} = 0$$

and

$$\frac{\partial C}{\partial y} - \lambda = 0$$

Computing these quantities, we get, $\frac{\partial C}{\partial x} = kN$ *and* $\frac{\partial C}{\partial y} = -5kP$. *Therefore,*

$$5kP + \lambda = 0$$

Therefore, $\lambda = -5kP$.

$$kN + \frac{\lambda}{(1+x)log2} = 0$$

$$-\frac{5kP}{(1+\lambda)log2} = -kN$$

which leads to $(x, f(x))$ *with* $x = \frac{1}{\log(2)} - 1$.

ANSWER

b) Draw the ROC curve of an optimal binary classifier (indicating what each axis represents) for the following:

1. Perfect Classifier, i.e. a classifier that always makes the right decision
2. A classifier that makes decisions based on the flipping of a fair coin (i.e. positive for a head and negative for a tail)
3. A classifier that makes decisions based on the flipping of an unfair coin (assuming that the probability of heads is p)

3 pts. ANSWERANSWERANSWERANSWERANSWERANSWERANSWERANSWERANSWER

1. *The ROC curve is a single point in (1,0) (TP=P and FP=0). Accepted answer: a curve starting form (0,0) passing through (1,0) and ending in (1,1).*
2. *The ROC is only one point at (0.5,0.5).*
3. *The ROC is only one point at (p,p) or diagonal curve connecting (0,0) to (1,1).*

ANSWER

c) Compute the fourier transform of the function g(x) for the following three cases:

1. $$g(x) = f_1(x) - f_2(x)$$

   where

$$f_1(x) = \begin{matrix} 1 & x \geq -a \\ 0 & x < -a \end{matrix}$$

$$f_2(x) = \begin{matrix} 1 & x \geq a \\ 0 & x < a \end{matrix}$$

2 pts.

2.

$$g(x) = \begin{matrix} 1 - |x| & |x| < 1 \\ 0 & otherwise \end{matrix}$$

3 pts.

3. g(x) is shown in Figure 4, i.e. it is a sine function in the middle and 0 elsewhere.

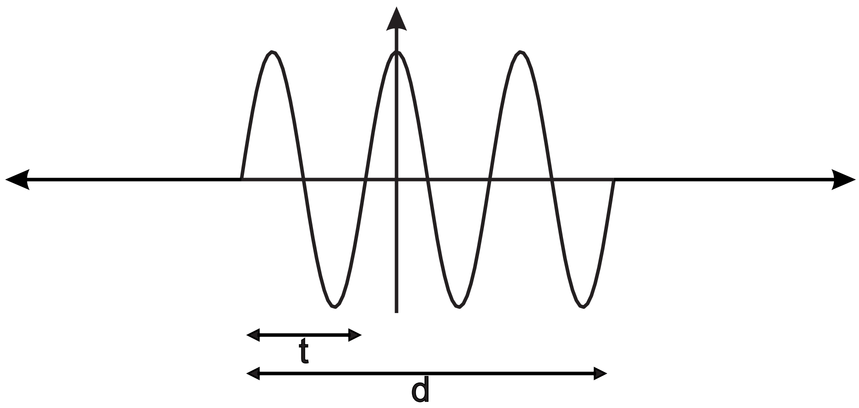


Figure 4: g(x)

5 pts.

ANSWER

1. *g(x)=Box(x) (1pt)*

   Therefore, $F(g(x)) = sinc(x)$ *(1pt)*

2. *g(x)is the triangle function (1 pt)*
   $g(x) = Box(x) * Box(x)$ *(1 pt)*
   Therefore,

   $F(g(x) = F(Box(x))XF(Box(x)) = sinc^2(x)$ *(1 pt)*

3. $g(x) = sin(x)Xbox(x)$ *(1 pt)*

$$\begin{aligned} F(g(x)) &= F(sin(x)) * F(box(x)) \\ &= \delta(\frac{1}{t}) * sinc(x) \end{aligned}$$

   (1 pt)
   Therefore the output is, 2 sinc functions centered at $\pm\frac{1}{t}$*, of width* $\frac{1}{d}$*.(3 pts)*

ANSWER

Question 3: Miscellaneous (20 pts)

a) What does sampling in the spatial domain correspond to in the frequency domain? **1 pt.** ANSWER

Multiples copies of the spectrum of the original signal -or- Convolution with a series of delta spaced each other by the sampling frequency.

ANSWER

b) 1. In the Hough Transform, a point $(x0, y0)$ in the xy-plane is mapped into a curve in the (ρ, θ)-parameter space. Write down the equation of the curve. **1 pt.**

2. If we apply the Hough transform on the image below, what would be the maximum value for the accumulator cell in the (ρ, θ) space? What is the corresponding (ρ, θ) value? **3 pts.**

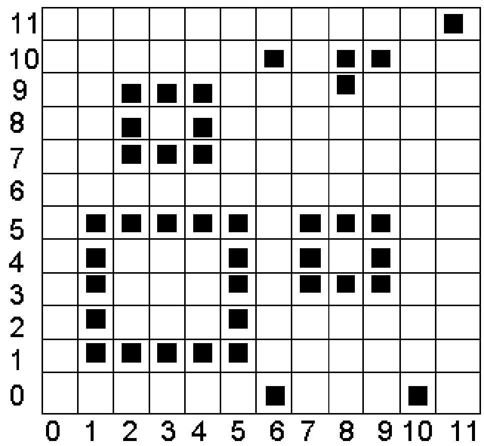


Figure 5: Image

ANSWER

1. $x\cos(\theta) + y\sin(\theta) = \rho$
2. *The maximum value is 8, corresponding to the horizontal line with vertical coordinate 5, because there are 8 points lying on this line (1 pt). This line has the parameter* $\rho = 5, \theta = \pi/2$ *(2 pts).*

ANSWER

c) What are the three properties a feature detector should have in order to be robust? Is the Harris Corner detector robust to all three of them? **3 pts.**

ANSWER

1pt per item.

- *Invariant to brightness change*
- *Invariant to shift and rotation*
- *Invariant to scale (this is where Harris Corner detector fails)*

ANSWER

d) How would you exploit temporal redundancy for video compression? Describe the three types of frames used in this method. **4 pts.** ANSWER ANSWER ANSWER ANSWER ANSWER

1pt: The idea is to predict the current frame based on previously coded frames: Temporal redundancy reduction 3pts: 3 types of frames:

- *I-frame: intra-coded frame, coded independently of all other frames.*
- *P-frame: coded frame, coded based on previously coded frame (based on previous I and P frames)*
- *B-frame: Bi-directionally predicted frame, coded based on both previous and future coded frames.*

ANSWER

e) When may temporal redundancy reduction be ineffective? **2 pts.**

ANSWER

When there are many scene changes (1pt) or/and high motion (1pt)

ANSWER

f) Which technique can be used when temporal redundancy reduction fails? Briefly describe the two main steps of the practical approach for this technique. **4 pts.**

ANSWER

The technique is motion-compensated prediction (1pt.) Practical approach: Block-matching Motion Estimation. Partition each frame into blocks and describe motion of each block by finding the best matching block in the reference frame.(3pts)

ANSWER

g) Name the two disadvantages of the block-matching algorithm. **2 pts.** ANSWER ANSWER

1pt per disadvantage. 1) It assumes translational motion model which breaks down for more complex motion. 2) often produces blocking artifacts.

ANSWER