UNIVERSITY OF COLOMBO, SRI LANKA

UNIVERSITY OF COLOMBO SCHOOL OF COMPUTING

DEGREE OF BACHELOR OF INFORMATION TECHNOLOGY (EXTERNAL)
Academic Year 2006/2007 – 3rd Year Examination – Semester 5

IT5502: Image Processing and Computer Vision

Structured Question Paper
31st March, 2007
THREE HOURS

To be completed by the candidate

BIT Examination Index No: .....

Important Instructions:

- The duration of the paper is **3 (Three) hours**.
- The medium of instruction and questions is English.
- This paper has **4 questions** and **15 pages**.
- **Answer all 4 questions: Each question carries 25 marks.**
- **Write your answers** in English using the space provided **in this question paper**.
- Do not tear off any part of this answer book.
- Under no circumstances may this book, used or unused, be removed from the Examination Hall by a candidate.
- Note that questions appear on both sides of the paper.
  If a page is not printed, please inform the supervisor immediately.
- **Non-programmable Calculators may be used.**

Questions Answered

Indicate by a cross (×), (e.g. [X]) the numbers of the questions answered.

	Question numbers				
To be completed by the candidate by marking a cross (×).	1	2	3	4	
To be completed by the examiners:					

Index No: ……...

1) (a) The following paragraph is concerned with the representation of an analogue image in a digital computer. Fill in the blanks with appropriate words.

(08 marks)

An analogue image is represented by a two dimensional light intensity function, denoted by f(x,y), where the value or amplitude of f at spatial coordinates (x,y) gives the intensity (brightness) of the image at that point. As light is a form of energy, f(x,y) must be non-zero, continuous and finite.

$0 < f(x,y) < \infty$

To be suitable for computer processing, an image function f(x,y) must be digitized both (1) ..…**spatially**……and in (2) ……**amplitude**……. . Digitization of the spatial coordinates (x,y) is called (3)……**image sampling**…., and the amplitude digitization is called (4)…..**grey-level quantization** . Image sampling creates a grid of squares which are called (5)……..…**pixels**…… . Each (6)……**pixel**……..….will be assigned an integer value depending on its (7)….. .**brightness**…... . Hence, a digital image is represented by a two dimensional array of (8)…..…**grey-levels**……... .

(b) f(x,y) is a continuous function with $0 \leq x \leq X$ and $0 \leq y \leq Y$ and g(m,n) is a continuous function with $0 \leq m \leq M$ and $0 \leq n \leq N$. Write the formula for the convolution of the two functions f*g.

(02 Marks)

ANSWER IN THIS BOX

$$f(x,y)*g(x,y) = \int_0^M \int_0^N g(m,n)\, f(x-m, y-n)\, dm\, dn$$

$0 \leq x \leq X \qquad 0 \leq y \leq Y$

(c) If f(x,y) and g(m,n) given in (b) are discrete functions and x and y are also discrete where x=0,…,X , y=0,…,Y , m=0,…,M and n=0,…,N , write a formula to calculate the convolution of the two functions f*g.

(02 marks)

ANSWER IN THIS BOX

$$f(x,y)*g(x,y) = (1/MN) \sum_{m=0}^{M} \sum_{n=0}^{N} g(m,n)\, f(x-m, y-n)$$

x=0,…,X , y=0,…,Y

(d) Write a pseudo code program for the computation of the convolution function given in (c).

(04 marks)

ANSWER IN THIS BOX

(e) What operation is achieved by convoluting an image with the following mask?

1	0	-1
2	0	-2
1	0	-1

(02 marks)

ANSWER IN THIS BOX

Detection of horizontal edge in the image

(f) Give another 3x3 mask which can be used to detect edges in an image together with the mask given in (e).

(02 marks)

ANSWER IN THIS BOX

1	**2**	**1**
0	**0**	**0**
-1	**-2**	**-1**

(g) Giving appropriate processing steps, explain how edges of an image can be detected using the above two masks.

(05 masks)

ANSWER IN THIS BOX

(1). Convolute the two masks given in (e) and (f) to obtain horizontal

($G_x(x,y)$) and

vertical ($G_y(x,y)$) grey-level variation at a pixel (x,y) in the image.

(2). Calculate the overall grey-level variation (gradient) at (x,y) as

$$G(x,y) = |Gx(x,y)| + |Gy(x,y)|$$

(3). If G(x,y) > T then (x,y) is an edge point,

where T is a user selected threshold.

(2) (a) Segmentation of an image into useful regions is an important operation in image analysis. Name two image segmentation techniques.

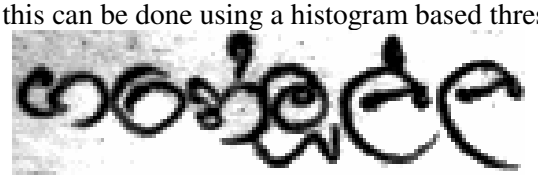
(02 marks)

ANSWER IN THIS BOX

(i) Pixel aggregation or Region growing based Segmentation

(ii)Threshold based Segmentation

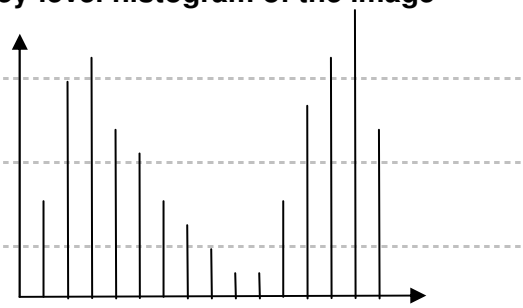
(b) The following is a handwritten image scanned from a postal envelope. In a handwritten recognition system, it is necessary to binarize the image into two levels; the characters and the background. Explain how this can be done using a histogram based thresholding technique.



(05 marks)

ANSWER IN THIS BOX

Draw the grey-level histogram of the image



Select the threshold (T) in the valley region of the histogram.

In the image f(x,y), classify the pixels (x,y) into character and background as:

If f(x,y) ≤ T, then (x,y) is a character pixel. Assign g(x,y)=1.

Else , (x,y) is a background pixel,. Assign g(x,y)=0.

g(x,y) is the binary image.

(c) Give an algorithm to obtain the threshold value automatically for the operation given in (b).

(05 marks)

ANSWER IN THIS BOX

1. ***Select an initial estimate for T.***
2. ***Segment the image using T. Let G1 be the group of pixels with grey level***

   values>T and G2 be the group of pixels with grey level values <=T.

3. ***Compare the mean (average) grey level values m1 and m2 for the pixels in***

Continued…

5. Compare a new threshold value

T= ½ (m1+m2)

6. Repeat steps 2 through 4 until the difference in successive T values is smaller than a predefined parameter.

(d) Unfortunately, in some of the character images given in (b), the illumination within the image is not uniform. Hence global thresholding techniques may not produce good results. Explain how you would modify the technique proposed above to suit such images.

(05 marks)

ANSWER IN THIS BOX

In such situations, the image is split into smaller squares and the variances of grey levels within squares examined.

If the variance within a square exceeds some value (say 100,) that image segment (square) should be thresholded separately using the iterative process.

All other image segments (squares) should be considered as a single region and segmentation performed using a single threshold.

(e) Explain briefly the following two noise removal techniques:

(i) Neighbourhood averaging
(ii) Median filtering

(04 marks)

ANSWER IN THIS BOX

(i)

Select a predefined neighbourhood (eg. 3x3, 5x5).

Replace the pixel (x,y) in the image f as

$$g(x,y) = (1/n) \sum_{(i,j) \in S} f(i,j)$$

where S is the neighbourhood and n is the number of pixels in S.

(ii)

Replace the pixel (x,y) in the image f as

Continued…

g(x,y)= median of the pixels in the neighbourhood S of (x.y) in f.

(f) Calculate the new pixel values of the shaded pixel of the following image when the above two techniques are applied separately using a 3x3 neighbourhood. Give steps of your calculations.

3	4	6	5	4
4	5	5	6	5
7	6	0	4	6
8	7	6	5	7
9	7	8	9	8

(04 marks)

ANSWER IN THIS BOX

i) New value at the shaded pixel = round ((3+4+6+4+5+5+7+6+0)/9) when (i) is used

= round(4.44)

= 4

(ii) New value at the shaded pixel = median of (3,4,6,4,5,5,7,6,0) when (ii) is used

= mid value of (0,3,4,4,5,5,6,6,7)

= 5

(3) (a) n_1 and n_2 denote the number of information-carrying units in two data sets respectively, which represent the same information. Express the data redundancy R_D of the first data set (the one characterised by n_1).

(04 marks)

ANSWER IN THIS BOX

$R_D = 1 - 1/C_R$

where $C_R = n_1/n_2$

(b) Outline the difference between Lossy and Lossless data compression techniques.

(04 marks)

ANSWER IN THIS BOX

(c) Name one Lossless image compression technique and one Lossy image compression technique.

(02 marks)

ANSWER IN THIS BOX

(i)

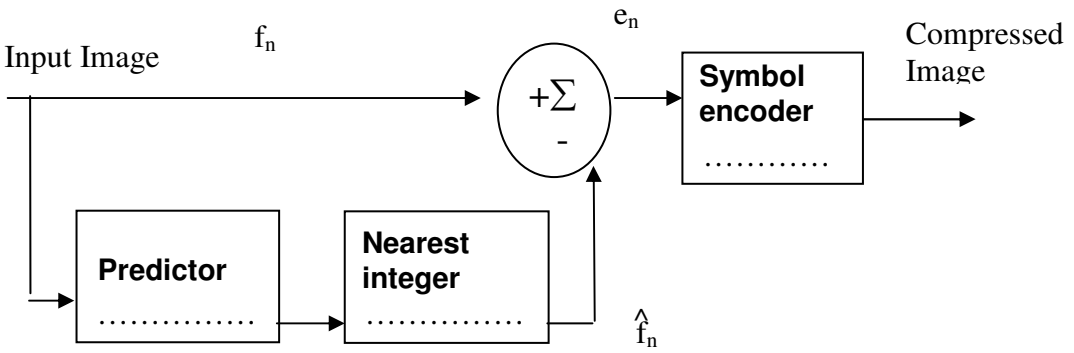
(ii)

(d) The following diagram shows the (a) encoder and (b) decoder of a Lossless predictive coding model. Fill the blank boxes with the appropriate words.

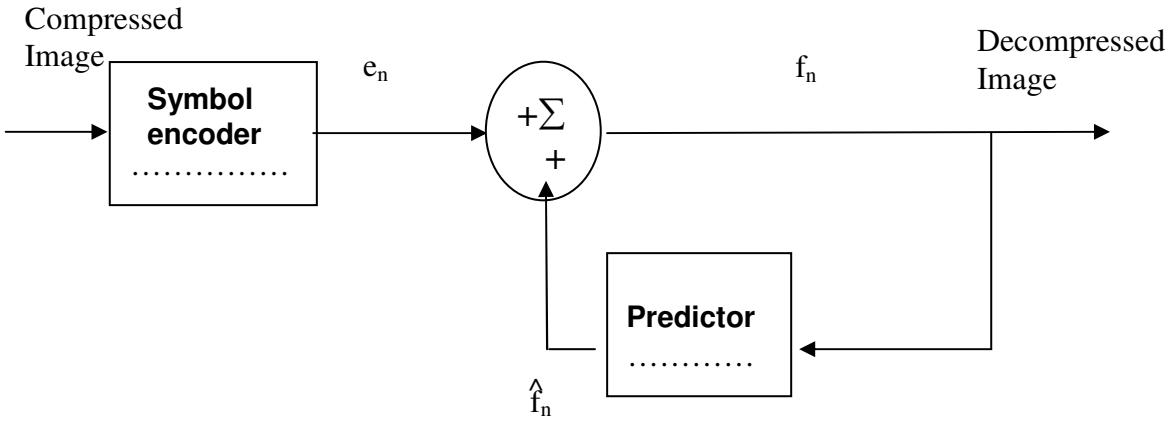
(05 marks)

ANSWER IN THIS BOX

(a)



(b)



(e) Represent the following image using

(i) Run-length coding
(ii) Chain code

0000000000000000
0000000000000000
0000000000000000
0000000100000000
0000000110000000
0000001110000000
0000011111000000
0000111111100000
0000001111100000
0000000111100000
0000000011000000
0000000001000000
0000000000000000
0000000000000000
0000000000000000
0000000000000000

(10 marks)

ANSWER IN THIS BOX

(i)

Run-length coding

a. **Possible run-lengths vary from one to sixteen=>4 bits per run are needed.**

b. **Gray level (black/white) can be coded with one bit. In addition, we have to know only the gray level of the first pixel in the row (when we have only two values and the value changes we always know what the next value will be)**

Row	value	run-lengths	binary code
0	0	16	0 1111
1	0	16	0 1111
2	0	16	0 1111
3	0	7, 1, 8	0 0110 0000 0111
4	0	7, 2, 7	0 0110 0001 0110
…	..	…	…
7	0	16	0 1111

c. **Bits needed**

There are 34 runs and 16 rows in the image=>bits needed: 16+34*4=152 (against 256 bits for uncompressed image) or 0.59 bits/pixel (against 1 bit/pixel for uncompressed image).

Continued…

(ii) **chain code**

a. **Eight directions are represented with 3 bits.**

b. **Upper left corner of the object (7,3) is selected to be the starting point**

(upper left corner of the image is(0,0)) .

c. **The edge of the object is followed clockwise until we get back to the starting point.**

We get the following chain code for the object:

7,6,7,7,6,6,5,6,3,3,3,3,4,1,1,1,2.

Since we have a binary image, we do not need gray level information.

Binary code:

Starting point		**directions**
0111	**0011**	**111 110 111 111 110 … 010**

Bits needed: 2*4+17*3=59(against 256 for uncompressed image) => 0.23

bits/pixel (against 1 bit/pixel for uncompressed image).

(4) (a) A typical Computer Vision system can be split into two sub-systems namely 'Low-Level Vision System' and 'High-Level Vision system'. Briefly explain them.

(05 marks)

ANSWER IN THIS BOX

Low-Level Vision System performs the pre-processing of the image using image processing operations (eg. Noise removal, sharpening, enhancement etc) and feature extraction (eg. Edge detection, segmentation etc).

High-Level Vision System performs the interpretation of the contents of the image using extracted features and domain knowledge using intelligent techniques (eg. Neural networks knowledge based techniques)

Parts (b) ,(c),(d) and (e) are related to the following problem.
In a busy crossing of a main road, an automated mechanism is used to identify vehicles which travel through the restricted area (in the middle of the crossing) by violating the traffic rules while the red light is on. In each of the four directions, a set of cameras is used to capture the necessary images as explained below.

- The first camera is focused on the restricted area and the sequence of images captured by this camera is processed to recognize the availability of a vehicle. In this case, a vehicle is identified by its class such as car, van, station wagon, bus, truck, etc., as it is necessary to determine the position of the registration number plate.
- Once the availability of a vehicle is confirmed, the second camera is activated and adjusted to capture the image of the registration number of the vehicle, which is later used to identify the owner.

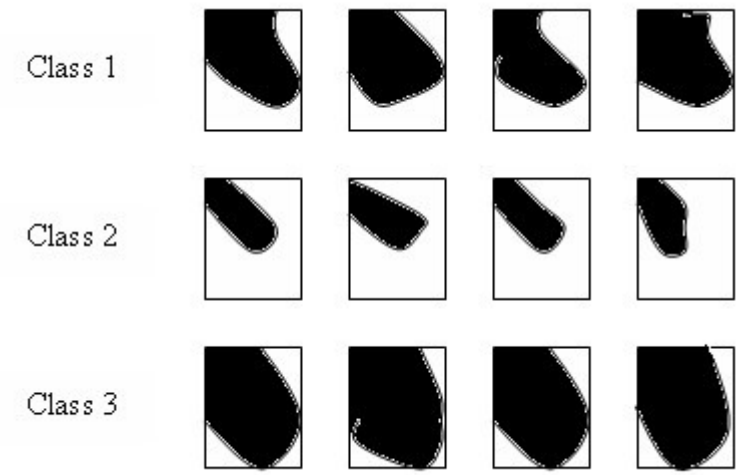
(b) After a thorough evaluation of several pattern recognition and classification methods, an Artificial Neural Network (ANN) based solution has been proposed to classify or identify the images in the above two steps.
Justify this proposal.

(05 marks)

ANSWER IN THIS BOX

Images of different vehicles (e.g. a car and a bus) have large variations, but within a class, variations are small (e.g. a Toyota car and a Nissan car). Uncertainty occurs between classes, and the objects within one class are similar.

(c) Sample images of three vehicle classes are shown in the following figure.



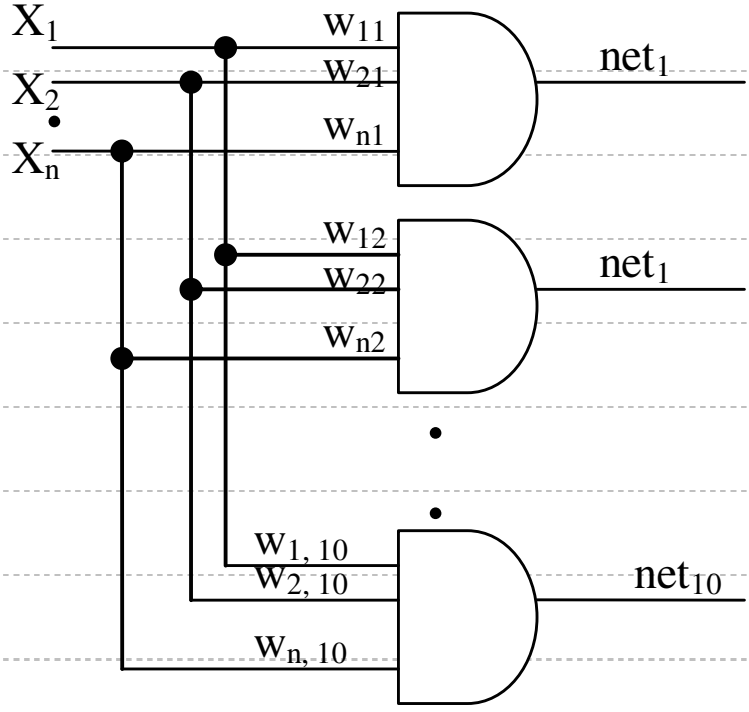
If ten such classes of vehicles have been identified, draw the ANN structure.

(05 marks)

ANSWER IN THIS BOX

Network Type – Unsupervised

Number of Output Nodes = 10



(d) What feature/s could be used for this classification effectively? (For this part of the question, consider only the given images of the three classes given in the part (c) Figure).

(05 marks)

ANSWER IN THIS BOX

EITHER a normalised version of the original image

OR one or a combination of

(a) Horizontal Projections

(b) Vertical Projections

(c) Black to White pixel ratio (i.e. Object Area)

(e) What are the pre-processing steps between an image just captured and the input to the ANN [x_1, x_2, …., x_n] ?

(05 marks)

ANSWER IN THIS BOX

(a) Noise removal

(b) Binarisation

(c) Pixel processing for horizontal, vertical projections etc.

(d) Formation of input column vector

****