The University of Nottingham

SCHOOL OF COMPUTER SCIENCE

A LEVEL 2 MODULE, SPRING SEMESTER 2017-2018

INTRODUCTION TO IMAGE PROCESSING

Time allowed ONE hour

Candidates may complete the front cover of their answer book and sign their desk card but must NOT write anything else until the start of the examination period is announced

Answer ALL THREE Questions

Dictionaries are not allowed with one exception. Those whose first language is not English may use a standard translation dictionary to translate between that language and English provided that neither language is the subject of this examination. Subject specific translation dictionaries are not permitted.

No electronic devices capable of storing and retrieving text, including electronic dictionaries, may be used.

DO NOT turn your examination paper over until instructed to do so

ADDITIONAL MATERIAL: **NONE**

INFORMATION FOR INVIGILATORS: **NONE**

1. Histogram-based Methods

(a) What is the goal of histogram equalization?

(2 marks)

To improve the contrast of an image by transforming it in such a way that the transformed image has a (nearly) uniform distribution of pixel values

(2 marks)
[knowledge]

(b) A 3-bit per pixel image has a normalized histogram as listed in the following table

Pixel value	Normalized frequency
0	0.1
1	0.5
2	0.1
3	0.1
4	0.1
5	0
6	0.05
7	0.05

Apply histogram equalization to this data and show

(i) the mapping from input pixel values to output pixel values
(ii) the normalised histogram of the output image

(10 marks)

Histogram equalization is achieved by computing the cumulative density function (CDF). The CDF provides a lookup table which specifies the mapping from input to output pixel values. To complete the mapping, the selected CDF value must be scaled back into the range of the image representation. For a 3-bit (8 value) image this is done by multiplying the CDF value by 7 and taking the integer part. The output histogram is obtained by considering each possible image value and working back through the mapping to see which bins in the input histogram map to that output bin. The output normalised frenqucy at a given bin is the sum of the input normalized frequencies of values that map to that bin.

Input pixel value	*Normalised frequency*	*CDF*	*Output pixel value*	*Output normalized frequency*
0	0.1	*0.1*	*0*	*0.1*
1	0.5	*0.6*	*4*	*0*
2	0.1	*0.7*	*4*	*0*
3	0.1	*0.8*	*5*	*0*
4	0.1	*0.9*	*6*	*0.6*
5	0	*0.9*	*6*	*0.1*
6	0.05	*0.95*	*6*	*0.15*
7	0.05	*1.0*	*7*	*0.05*

(10 marks)

[application]

(c) Why are colour histograms considered to be good representations upon which to base image retrieval methods?

(6 marks)

Colour correlates well with class identity
Human vision works hard to preserve colour constancy: presumably because colour is useful
Histograms:

- *Are invariant to translation and rotation*
- *Change slowly as viewing direction changes*
- *Change slowly with object size*
- *Change slowly with occlusion*

(6 marks)
[knowledge]

(d) Two images are being compared using Histogram Intersection. Their histogram representations are H1: [10,20,0,0,5,5,5] and H2: [5, 25, 4, 3, 5, 10, 2]. What score does Histogram Intersection assign to this match?

(2 marks)

Histogram intersection computes the sum of the minimum entries in corresponding histogram bins. In the case the computation is 5 + 20 + 0 + 0 + 5 + 5 + 2 = 37

(2 marks)
[application]

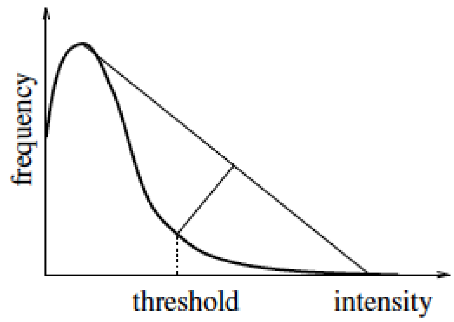
2. Thresholding and Binary Image Processing

(a) When thresholding an image to produce a binary image, manual determination of a suitable threshold can be difficult. Many automatic threshold determination techniques have therefore been developed. Briefly describe the principle behind and operation of Rosin's Unimodal method.

(8 marks)

- *Rosin's method is unusual in not assuming the image to have a bimodal histogram. Instead it assumes there is often only one, or one very dominant peak e.g. images of printed text are mainly ligtht (paper), with a small amount of dark (inked) pixels. Rosin's unimodal method*
  - *Finds the peak*
  - *Draws a line from there to the top of the furthest bin*
  - *Finds the top of the bin that is furthest from this line; that bin value is the threshold*

  The image below captures the algorithm well, and might be used in answer.



(8 marks)
[knowledge]

(b) Thresholding produces binary images. These are often noisy, in the sense that they contain incorrectly classified pixels. The mathematical morphology operations of erosion and dilation can be used to reduce such effects. The image fragment below shows a binary image in which pixels labelled 1 are considered foreground and those labelled 0 are considered background.

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

image fragment

Briefly explain the algorithm used to erode the foreground of a binary image, and show the result of applying this algorithm to this image fragment using the structuring element shown below.

1	1	1
1	1	1
1	1	1

structuring element

(8 marks)

Erosion takes two inputs

- *binary image to erode*
- *set of points to be considered; a structuring element*
- *origin is the central element*

The algorithm (erode foreground)

- *the structuring element is superimposed on each of the foreground pixels such that origin of the structuring element coincides with the input pixel position.*
- *If any of the '1' pixels in the structuring element overlap (intersect) the background then the foreground pixel is also set to background*

Result of eroding the image fragment with the structuring element given is to set all pixels to 0

(8 marks)
[knowledge]

[application]

(a) How can dilation be used to detect image edges?

(4 marks)

Dilate the foreground, subtract it from the original binary image. Edges remain

(4 marks)
[knowledge]

3. Image Filtering and Enhancement

(a) Many image processing operations rely on convolution of the image with some operator. Briefly describe the convolution process as performed in the spatial domain.

(5 marks)

The operator to be applied is represented as a 2D array, or mask, containing a set of weights. That mask is placed over each pixel (the source pixel) in the input image in turn. Mask weights are multiplied by the corresponding image pixel and the results summed over the mask area (ie it's a weighted average). This weighted average value is written to the same location (the target pixel) in the output image. Convolution cannot be applied to boundary pixels.

(5 marks)
[knowledge]

(b) Compute the result of applying

i) a 3 x 3 mean filter
ii) a 3 x 3 Sobel filter measuring gradient in the horizontal direction
iii) a 3 x 3 Laplacian filter

to the central pixel in the image fragment shown below

7	8	4
8	6	3
8	5	1

(6 marks)

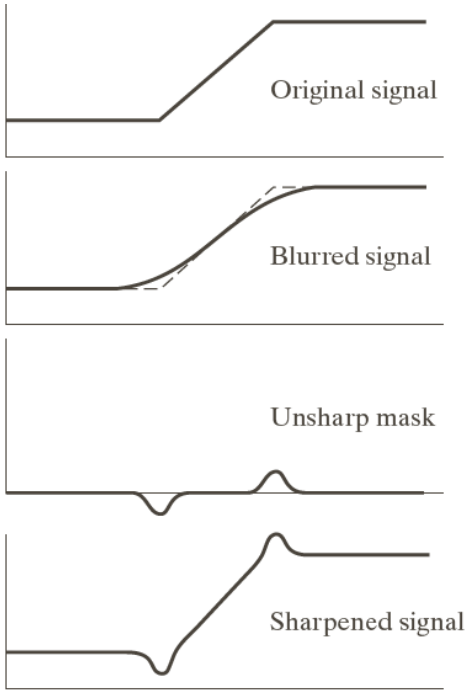
Mean filter = (7 + 8 + 4 + 8 + 6 + 3 + 8 + 5 + 1)/9 = 5.5
Sobel = (1x7 + 2x8 + 1x8 − 1x4 − 2x3 − 1x1) = 23
Laplacian = (1x8 + 1x8 + 1x 3 + 1x5 -4x6) = 0

(6 marks)
[application]

(c) Explain with the aid of a diagram how Gaussian smoothing may be used to enhance image edges via unsharp masking.

(5 marks)

Take the original image, Gaussian smooth it, Subtract the smoothed version from the original to make an unsharp mask, Add the mask to the original to make the edge appear more obvious.



(5 marks)
[knowledge]

(b) What does it mean to say that the 2D Gaussian filter is “separable”?

(4 marks)

This means you can perform a 2D Gaussian as two 1D Gaussians. First filter with a ‘horizontal’ Gaussian, then with a ‘vertical’ Gaussian. The results are then multiplied together to give the final result. The separated filter is more efficient. On an NxN image an n x n mask is O(N2n2) when applied in 2D and O(N^2 2n) when separated. (i.e. it is still order N squared in the size of the image since each pixel must be processed at least once).

(4 marks)
[knowledge]