final

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2016-2017

MEng Honours Degrees in Computing Part IV
MSc in Advanced Computing
MSc in Computing Science (Specialist)
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute

PAPER C407H

MEDICAL IMAGE COMPUTING

Tuesday 13 December 2016, 10:00
Duration: 70 minutes

Answer TWO questions

Paper contains 3 questions
Calculators required

1 **Image analysis and segmentation**

Assume a 5 x 5 image as shown below where scalar entries denote pixel intensity greylevels:

3	2	4	5
7	7	8	2
3	1	2	3
5	4	6	7

a State the purpose of histogram equalisation, and define the intensity mapping required to perform this task.

b Set up and sketch a histogram for the image shown above, clearly showing the individual bin entries and axis labels.

c Perform histogram equalisation on the image shown above, scaling the intensities between 1 – 20, and sketch the resulting image. You may round the final values down to the nearest integer value.

d Filter the original image shown above using the following filters, and discuss how you deal with image borders:

i) 3 x 3 median filter

ii) 3 x 3 mean filter

iii) 3 x 3 Gaussian filter defined as below:

1/16	2/16	1/16
2/16	4/16	2/16
1/16	2/16	1/16

e Image segmentation using Markov Random Fields use a formulation based on the following energy function:

$$E(\mathbf{Z}) = \sum_{p \in I} D_p(z_p) + \sum_{\{p,q\} \in N} V_{p,q}(z_p, z_q)$$

Explain the different terms $\mathbf{Z}, D, V$ and their interaction. How is the energy function minimized?

The five parts carry, respectively, 10%, 10%, 20%, 35%, and 25% of the marks.

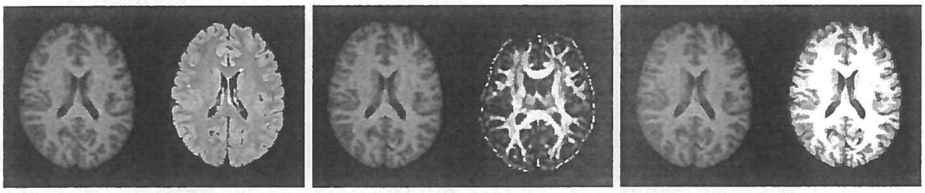
2 Image registration

a (Dis)similarity measures

i) Write down the equation for the sum of *squared* differences (SSD) measure.

ii) Explain in one or two sentences why the sum of *absolute* differences (SAD) measure can be considered to be more robust than SSD.

iii) For the following three pairs of images (a)-(c), name a suitable (dis)similarity measure for each pair that could be considered for intensity-based registration of the two images. Explain your choice in one or two sentences.



(a) (b) (c)

b Registration approaches

i) Explain in a few sentences what are the advantages and disadvantages of feature-based and intensity-based registration.

ii) Describe or sketch the iterative process of intensity-based registration.

iii) Explain in one or two sentences how convergence to local optima can be potentially avoided in intensity-based registration.

iv) Explain in one or two sentences the difference between *extrinsic* and *intrinsic* anatomical landmarks in the context of feature-based registration.

v) Explain in one or two sentences why *backward* warping is preferred over *forward* warping.

c Applications

A patient with a traumatic brain injury is undergoing MR and CT brain imaging on the same day.

Explain what type of transformation model and what (dis)similarity measure are suitable for an intensity-based registration to align the two brain images.

The three parts carry, respectively, 45%, 45%, and 10% of the marks.

3 Machine learning

a Applications

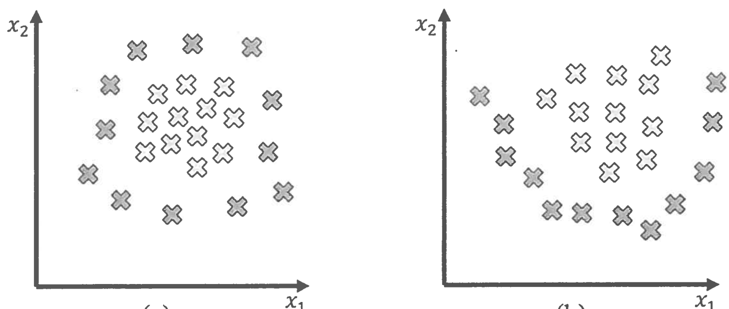
i) Describe an approach in a few sentences that utilises a database of healthy liver shapes to detect abnormal cases.

ii) Give two examples of supervised learning for solving medical image analysis tasks, one for classification and one for regression problems.

b Logistic regression

i) Write down the mathematical definition of the hypothesis function for logistic regression, and explain what each symbol stands for.

ii) Given two-class classification problems with two measurements x_1 and x_2 for each data point. Explain in one or two sentences if and how logistic regression can be used given the training examples in (a) and (b) below.



(a) (b)

c Naive Bayes Classifier

For solving classification problems, we are interested in the posterior distribution $p(c|\mathbf{x})$ where c corresponds to the class label, and $\mathbf{x}$ is the feature vector containing n individual features $\mathbf{x} = [x_1, \ldots, x_i, \ldots, x_n]^\top$.

i) Write down the equation for computing $p(c|\mathbf{x})$ with naive Bayes.

ii) Classify a subject with feature values $x_1 = 140$ and $x_2 = 90$ using a Gaussian distribution assumption and the statistics given in the table below.

c	$p(c)$	μ_1	σ_1	μ_2	σ_2
healthy	0.9	100	20	70	10
diseased	0.1	150	30	100	20

Univariate Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

The three parts carry, respectively, 20%, 40%, and 40% of the marks.