Two hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Computer Vision

Date: Thursday 19th May 2011
Time: 09:45 - 11:45

Please answer question ONE and also THREE other questions from the FIVE questions provided
Use a SEPARATE answerbook for each Question

This is a CLOSED book examination
The use of electronic calculators is NOT permitted
1. **Compulsory**

Answer any four (from 6) of the following parts:

a) How does a *median filter* reduce random noise in an image whilst tending to preserve edges? (5 marks)

b) Write down the kernel of a simple convolution filter that would give a *zero-crossing* at edges. Give an argument why a zero-crossing edge detector might be preferred to a first derivative edge detector. (5 marks)

c) How is the output of a *rank order filter* calculated? Explain the trade-off involved in choosing the neighbourhood size. (5 marks)

d) What is meant by the *skeleton* of an object? Explain with the aid of a diagram how the skeleton can change dramatically for a small change in the shape of an object. (5 marks)

e) Explain how an *adaptive thresholding* method could be implemented using a large smoothing filter and image arithmetic. Under what circumstances might you expect the method to work well? (5 marks)

f) Explain the role of *non-maximal suppression* in edge detection and describe a practical algorithm for applying it to an edge strength and orientation image. (5 marks)
2.

a) Outline the advantages, and at least one disadvantage, of using a model-based approach to medical image interpretation. (4 marks)

b) Consider the Hough Transform. Do you think this can be considered as an example of a model-based approach to image interpretation? Explain the reasoning behind your answer. (1 mark)

A student is trying to develop a robot that can lay tables. As a first step, she considers the image analysis problem of finding a single fork on a table. See Figure 1 for some example images.

![Example images of single forks.](image)

Figure 1: Example images of single forks.

c) She first tries applying the Hough transform for lines to these images. Give a brief description of the results you would expect for this. Do you think this would be a useful method to pursue? Explain your answer. (2 marks)

She then decides to try a more sophisticated approach, and decides to test both the Active Contour Method (ACM) of Kass and Witkin, and the Active Shape Model (ASM) approach of Cootes et al.

d) Describe briefly the ACM and ASM methods. Discuss what might happen if she applied these methods to the images shown above. (10 marks)

e) She now proposes to extend her methods to deal with knives and spoons as well. Given the many and varied shapes of cutlery that she finds online, discuss whether you think the ACM or ASM approach would be the more efficient approach to take, in order to solve the task of locating a single piece of cutlery amongst many on the table (with possible occlusions), so that the robotic arm can pick it up. (3 marks)
3.  

a) Give a brief outline of the three main constituents of a non-rigid pairwise image-registration algorithm. Compare and contrast an algorithm that uses a non-parametric representation of image warps, with one that uses a parametric representation.  

(b) Compare and contrast elastic image registration and fluid image registration. For each method, give an example application from biomedical imaging for which that method would be the most suitable. Explain your choices.  

(c) In a journal paper, the following is given as the objective function for a registration algorithm:

$$
\mathcal{L} = \alpha \sum_{\alpha=1}^{d} \int_{\Omega} (\nabla u_{\alpha}(r))^2 \, dr + \int_{\Omega} |I_s(r) - I_t(r_u)| \, dr,
$$

where:

- \( r = \{ r_\alpha : \alpha = 1, \ldots, d \} \),
- \( u(r) = \{ u_{\alpha}(r) : \alpha = 1, \ldots, d \} \), \( r_u \doteq r + u(r) \),
- \( \nabla \doteq \frac{\partial^2}{\partial r_1^2} + \frac{\partial^2}{\partial r_2^2} + \ldots + \frac{\partial^2}{\partial r_d^2} \).

i) Explain the intended role of each term in the objective function, and how its form allows it to fulfil that role, making sure that you identify the meaning of each function or operator that appears.  

ii) Write down the general formula for a transformation that consists of a combined rotation and scaling, followed by a translation. Hence (or otherwise) deduce one important aspect of the behaviour of the proposed registration algorithm.
4. You are asked to develop a computer vision system that can detect cars (but not motorbikes) from side views, such in the images below (Figure 2).

<table>
<thead>
<tr>
<th><img src="image1.png" alt="Motorbike" /></th>
<th><img src="image2.png" alt="Car" /></th>
<th><img src="image3.png" alt="Car" /></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image4.png" alt="Car" /></td>
<td><img src="image5.png" alt="Car" /></td>
<td><img src="image6.png" alt="Motorbike" /></td>
</tr>
</tbody>
</table>

**Figure 2**

a) Describe one method for detecting interesting feature points, which occur at a range of scales in an image (6 marks)

b) Describe one method for computing the “signature” representing the image structure around a point, assuming the scale and orientation have been estimated (6 marks)

c) Suppose that we have computed clusters of features from a training set, and determined how likely features in each cluster are to be part of a car. Describe how this information could be used in a “Bag of Features” car detector. (8 marks)
5.

a) What is the difference between motion field and optical flow? (4 marks)

b) Given the brightness constancy equations:

$$\frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0 \quad (Eq. 1)$$

i. Explain intuitively what each of the five derivatives in the brightness constancy equation measures. (5 marks)

ii. Give three different concise phenomena in the physical world that will invalidate the brightness constancy equation. (4 marks)

iii. What is the aperture problem? How does it relate to Eq. 1? (4 marks)

iv. Suppose that the camera only translates, that you know the direction of translation, and that the rest of the world is stationary. Explain how you would use this knowledge to address the aperture problem. (3 marks)