Two hours

Question ONE is COMPULSORY

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Computer Vision

Date: Monday 19th May 2014
Time: 14:00 - 16:00

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. This question is 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) What is the optical flow constraint and why does it not uniquely determine the flow at each point in an image? [5 marks]

c) Explain how the Hough transform using the \((r, \theta)\) parameterisation can be used to find straight lines in edge-detected images. [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) What is the main assumption that lies behind the idea of using image smoothing for noise suppression? [5 marks]

f) Explain and contrast the region-based and edge-based approaches to extracting structure from images. [5 marks]

End of Question 1
2.

A COMP61342 student working in medical image analysis wishes to study shape and shape variation. You may assume that she has access to a database of 2D segmented shapes, and that she can request annotation by an expert, if required.

a) Describe how she would build a statistical shape model (SSM) using this shape database, and what expert annotation she might need.

[8 marks]

b) Describe briefly what problem(s) she might experience if she decided to move from working with 2D planar shapes, to 3D surfaces. Are there any methods that she could use to help avoid these problem(s)?

[3 marks]

Another COMP61342 student wants to use a deformable model in order to locate a particular structure in a set of medical images.

c) Name three different types of deformable model that might be suitable for such a task, and describe the distinguishing characteristics of each type. In particular, you should make reference to image search and the type(s) of structures that such a method might be able to locate.

[9 marks]

End of Question 2
3. Figure 1 shows a frame from a video sequence obtained by a single video camera moving through a static indoor scene. The sequence of images is to be used to generate a 3D representation of the scene. The small white squares superimposed on the image are regions that have been automatically selected as suitably “interesting” locations to be matched between frames.

a) Explain why it is useful in interpreting the 3D structure of the scene to match locations across frames. [2 marks]

b) On what basis might the “interesting” locations be selected? Outline a method that might be used to identify candidate “interesting” positions. [6 marks]

c) Describe a method for finding the positions in a given frame of the sequence that match the interesting points identified in another frame. [6 marks]

d) Explain in outline how, having obtained matching positions in different frames, the 3D structure of the scene can be determined. You should pay particular attention to the facts that the images are obtained from the same camera and that, in general, the change in position of the camera between the frames is not accurately known. [6 marks]
4.

a) Explain what calibrated camera means. [2 marks]

b) Define disparity in stereo vision. [2 marks]

c) State the formula that relates disparity, focal length, inter-ocular separation and depth. Illustrate this relationship with a diagram. [5 marks]

d) What are the main differences between the data obtained from calibrated and uncalibrated stereo systems? Illustrate these differences with practical examples of use. [4 marks]

e) Given a pair of stereo images, what do we mean by the term image rectification? Why is it important? [3 marks]

f) Consider two ideal pinhole cameras with the following top view configuration:

![Diagram of two pinhole cameras with 45° angles]

Draw the epipole and a few epipolar lines on the front view of the two 2D images. [4 marks]

*End of Question 4*
5.

In a journal paper, the following is given as the objective function for an intensity-based non-rigid pairwise registration algorithm:

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

where:

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

a) Explain **in detail** the intended role of each term in the objective function. You should include how its form allows it to fulfil that role, making sure that you identify the meaning of *every* function or operator that appears.

[5 marks]

b) Along with an image-matching term, what other components are required when constructing a general pairwise non-rigid image registration algorithm? You should discuss both the **non-parametric** and **parametric** cases. For each basic constituent that you mention, you should give at least one example.

[12 marks]

c) Why might we wish to perform registration across an entire population? Discuss briefly why fully groupwise registration might be preferred over repeated pairwise registration in such a case.

[3 marks]

End of Question 5

END OF EXAMINATION