Exam Notes

1. If doubt exists as to the interpretation of any question, the candidate is urged to submit the answer paper with a clear statement of any assumptions made.
2. This is an OPEN BOOK EXAM. Any paper notes or textbooks are permitted.
3. **No** calculator or computer or any sort is permitted.
4. This exam contains SIX (6) questions, however FIVE (5) questions constitute a complete exam paper. The first five questions as they appear in the answer book will be marked.
5. Each question is of equal value.
6. The clarity and organization of the answers are important.
7. The exam is three hours long.

1. **Basic Definitions**

   For each of the following concepts, give a brief definition (2-4 sentences) and a description of where the concept is used or applied, in practice:

   (a) The discrete cosine transform (DCT)

   (b) Image interpolation

   (c) The Canny edge detector

   (d) Robust image features, such as SIFT or SURF

   (e) Two dimensional median filtering

   (f) Image morphology

2. **Image Filtering**

   A 2D image $f(m, n)$ of size $M_1 \times N_1$ is to be convolved with a filter array or convolution kernel $h(m, n)$ of size $M_2 \times N_2$ to produce a new resulting image $g(m, n)$:

   $$g = f * h$$

   (a) Depending on how the convolution is implemented, g could be one of three different sizes. Specify these sizes.

   (b) Write a short program (in pseudo-code, Matlab, or C) to explicitly show how to compute g from f and h.

   (c) Describe a method to compute g using Fourier transforms.

   (d) In principle we can use the FFT to calculate the Fourier transform in (c), and therefore to compute a convolution. What are the issues that need to be kept in mind in using an FFT to compute a convolution?

   (e) A convolution could be applied in 3D, for example in video processing (the three dimensions of space-space-time). Why is this a poor approach, such that convolutions are almost never used in video?

3. **Video Processing**

   Give answers to each of the following:

   (a) What are some of the basic approaches to video compression? Name one or two approaches and the conceptual strategies used in each.

   (b) Given a video (a sequence of images over time), what are some of the most fundamental image processing / computer vision operations that we apply? That is, what sorts of things do we do with video?

   (c) What are common application areas for video analysis? That is, what sorts of needs commonly occur in industry that lead to video processing?

   (d) What are the issues or tradeoffs in considering causal versus non-causal processing of video data?

   (e) What are the issues or tradeoffs in considering embedded processing of video, right on the camera itself, as opposed to transmitting the video away from the camera to a separate server for processing?

4. **Medical Imaging**

   (a) Name at least two aspects of medical imaging that make it distinct in some way from the rest of the field of image processing.

   (b) What is the Radon Transform, and why is it significant to medical imaging?

   (c) Briefly describe the significance of each of the following in the context of medical imaging:

      i. Anomaly detection,

      ii. Segmentation,

      iii. Denoising.

   (d) What is *speckle* noise and in what sorts of contexts does it occur?

   (e) What does *registration* mean and when do we use it? Define the differences between rigid, affine, and deformable registration.

5. **Application — Cell Tracking**

There is a great deal of research interest in the tracking of individual cells over time in microscope images. Significant applications include embryonic development, stem-cell research, the study and cataloguing of how proteins interact with cells, and the study of exposed cells such as the cells of the cornea or the retina.

Suppose we wish to study stem cells. These cells live, move, periodically divide, and at some point suddenly change into a particular tissue type (skin cell, bone cell, muscle cell etc).

One research interest is to understand what causes a stem cell to become a particular tissue: why did it become skin as opposed to bone? To understand this we need to track the cell over time, to see how much or little it moved, how close to or far from other cells it was etc.

Your Task:

Assume that there is a microscope system returning large images I_t at time t. Furthermore, there is a cell locating function `find_cells` already available that returns a set of (x, y) locations, measured in pixel units, believed to be a cell:

$$\begin{bmatrix} x_1 & y_1 & p_1 \\ x_2 & y_2 & p_2 \\ \vdots & \vdots & \vdots \end{bmatrix} = \texttt{find\_cells}(I_t)$$

The problem, of course, is that `find_cells` does not work perfectly, so that from time to time image noise causes a cell to be missed, or mistakenly causes a cell to be identified where none is present. The `find_cells` function attempts to identify a measure of confidence p_i in the recognition of cell i, so that a larger p reflects a greater likelihood of correct cell identification.

Images are captured on a time interval of one minute, and over this one minute cells do not typically move further than a cell diameter of 20 pixel units.

Cells may be on their own, or may be clumped in groups.

What *you need to do* is to design a system that tracks cells over time, returning the trajectory for each cell which is detected in the current image, giving the x, y location of that cell in all past frames where it was present.

Be sure to describe the system carefully. You are providing a high level design: you do *not* need to produce a computer program, although you could choose to write some pseudocode.

6. **Application — Pulse Detection**

It is possible to detect the pulse of a person via image processing. With every heartbeat the capillaries in the skin dilate very slightly, making the skin appear slightly redder. The effect is not visible to the eye, but can just barely be detected by colour cameras.

You are given images I_t over time t, ten images per second. You can assume that the person is not moving significantly, so that their location is the same in each image.

Your Task:

Devise a system to report the average pulse over the last ten seconds.

Given the colour RGB images I_t, describe your strategy, either in pseudo-code form or as a block diagram. State any assumptions you make.

A few things to think about:

(a) Localization: Where will you look in the image? Is this manually calibrated or automatically detected?

(b) Calibration: What sorts of noise or other disturbances might you encounter? How will you deal with image noise?

(c) A pulse rate is like a frequency. Will you be working in the time domain, or in the frequency domain?