Robert Fisher

Question block created by wizard

Answer Question Q1 and **either** Question Q2 **or** Question Q3. If you answer both Question Q2 and Question Q3, only Question Q2 will be marked. Ignore the exam component numbering (1, 2, ... 25).

1. **Question 1.** Answer all 10 parts (Q1a-Q1j) of question 1 with 2-3 sentences. Each question is worth 2.5 points.

2. **Q1a.** Assume that you have a video sequence taken from a stationary camera that is watching a car parking lot. Describe 3 ways that you can detect a car driving in the lot from the video data.

3. **Q1b:** Suppose you are building a robot vision system for an Amazon warehouse. One subsystem has to identify, locate and grab rectangular solid boxes from shelves. What sort of a vision system would you design to estimate the box location? Why?

4. **Q1c:** You are building a mobile robot that navigates the corridors of a building. It uses a two dimensional range sensor, where all of the distance measurements lie in a plane. So, the range image is 1 dimensional and each pixel corresponds to the distance to the scene surface at a different angle. In this case, the range image is 360 pixels long, with one pixel for the distance at each degree. The plane that is scanned is parallel to the floor and at height 50 cm above the floor. Explain how the robot might use the image data to detect a person in the corridor (i.e. so the robot can avoid the person).

5. **Q1d:** Why does using normalised RGB data help compensate for illumination variations? When would using normalised RGB make the problem harder?

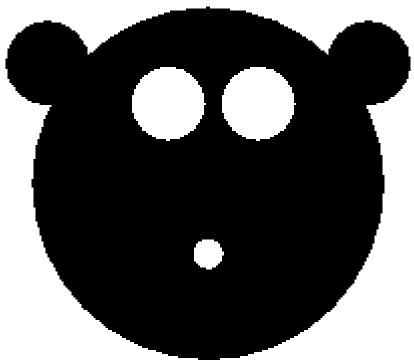
6. **Q1e:** Think of the ball tracking example in the lecture. If you were only watching the ball fall, without observing any bounce or stopping, you would only need a simple state model. Is there any advantage to using the Condensation Tracking in this scenario? Why or why not?

7. **Q1f:** Suppose you were tracking a person and they walked behind an occluding obstacle. A few seconds later a person appears from behind the obstacle. What computer vision method could you use to assess the likelihood that they were the same person as you were originally tracking? Assume that the camera is sufficiently distant that face recognition was not possible.

8. **Q1g:** Much stereo matching and object recognition work uses SIFT features. Why?

9. **Q1h:** You are building a video surveillance system with a fixed camera to watch the outside of a prison fence. You need a method for detecting changes to the scene. Would you be more likely to use an algorithm based on background subtraction or frame differencing? Why?

10. **Q1i:** Suppose that you had samples of the following flat metal part coming down a conveyor belt. And sometimes the parts are defective as well. The parts are observed with a fixed overhead camera, like in the lecture examples. What geometric model would you use to describe this part, and how would you use the model for recognition, pose estimation and inspection?



Robert Fisher

11. **Q1j:** In the second coursework, we got good 3D measurements from the ball, but poor depth values from the wall behind the ball. The ball measurements were a bit noisy but were approximately the shape of the ball. The wall measurements were large and random. What could we have done to ensure that the quality of the data from the wall was good, too?

12. **Question 2** (answer all 6 parts Q2a-Q2f)**.**

You are working for a clothing manufacturer and they wish to acquire a database of human body shapes. These would be used to improve the sizes and shapes of their clothes. This will have to be a 3D database. They would also like to analyse the database, so that they can use some parameterised body model instead of hundreds or thousands of individual samples. Assume that the volunteers who are being scanned are wearing swimsuits and they are scanned only once from the front. Also assume that they are scanned holding their arms out sideways.

13. **Q2a:** What are possible sensor systems for acquiring the 3D data? Which might be more appropriate with this application? Why? What preparations are needed to enable good data? [4 points]

14. **Q2b:** Assume that you are going to build a 3D version of the Point Distribution Model, where now the corresponding points are 3D landmarks rather than 2D landmarks. What are a good set of landmarks for capturing standard variations in human shape? What are the likely natural variations in those landmark positions? (3 points)

Robert Fisher

15. **Q2c:** Other than than the navel (belly button), the stomach area of people does not have a standard landmark. The same issue might arise with the sizes of the upper and lower arms, and similarly for the legs. How might you solve this problem? (3 points)

16. **Q2d:** For the sensor system that you proposed above in part a, what algorithms would you use to automatically locate the 3D points of interest proposed in b and c from a single set of data? Explain the steps of the algorithms. (8 points)

17. **Q2e:** How would you build the Point Distribution Model using a set of the 3D points captured from different people? How would use use the results to model the different variations that the clothing manufacturer is interested in? (4 points)

18. **Q2f:** What problems might arise with your solution and how would you overcome them? (3 points)

19. **Question 3** (answer all 6 parts Q3a-Q3c).

In this question, we are designing a system to do table tennis ball tracking, like in the first practical. In this case, we will use a stereo camera system to observe the scene. Assume that the background is plain and both cameras can see the whole scene. In this case, all of the balls will be orange. Assume that the frame rate is slow enough that it is not always obvious which balls should be paired in the next frame. Also assume that there are a lot more balls, so that collisions and occlusions are more common, and because of the slow frame rate, it is not always obvious which balls should be paired. On the other hand, assume that there are no bounces, and that the balls are not seen again once they leave the field of view.

20. **Q3a:** Where would you would place the stereo cameras? (2 points)

21. **Q3b:** How would you detect the balls? (3 points)

22. **Q3c:** What stereo constraints could you use for this problem, if any? Why? (5 points)

23. **Q3d:** Describe the state and tracking model that you would use. How would you update the tracking data given the set of detections from the new frame? How would you keep track of the many observed balls? (6 points)

24. **Q3e:** How would you use the 3D data from the stereo system? (6 points)

25. **Q3f:** What could go wrong with the process? How might you fix or prevent the problems? (3 points)