General Section (Answer 4 out of 6)

Invisible Reality

You need to hide a machine in the middle of a large room in a palace, and the machine cannot be moved. The room has intricate decoration on the walls, floor, and ceiling. A single inspector wearing a red hat will come in and walk around the room. Assume that he will not come too close to the machine. Your cameras will track his position in the room. You want to build a box made of 5 screens to cover the machine. You need to paint on the screens images that will make the machine invisible to the inspector. Explain the whole solution and provide the details for data acquisition, storage, and rendering. Discuss possible artifacts that may make the machine detectable.

Subdivision Surfaces on Fragment Processors

You have decided to implement a popular subdivision surface scheme, either Loop or Catmull-Clark (pick one). Just to make it more challenging, you want to do all of the real work in fragment shaders on the GPU. Describe how you will communicate the original mesh to the GPU, how you will store the mesh in the GPU’s memory, how you will perform the actual calculations for subdividing the mesh on the fragment processors (especially the GPU memory accesses), and how you will communicate the results from the fragment processors back to the CPU.

Lazy Trek

You are given a height field defined by an irregular planar triangulation with height samples at the vertices. The height of the terrain over any 2D point is computed by linear interpolation over a triangle containing that point. Describe how you would compute the region consisting of all points that can be reached from a point p by walking on the terrain, but never uphill. Discuss the complexity of your algorithm.

A Matter of Perspective

Initially assume that the viewer is at distance D from a screen and the near clipping plane is on the screen and the far clipping plane is at infinity. Explain how fragment depth is stored in the z-buffer and discuss the loss of accuracy due to x-buffer value quantization. How exactly does it vary with depth Z? Where is it the largest? How much accuracy improvement will you obtain by setting the near clipping plane to be at a distance 2D from the viewer? How much accuracy improvement will you obtain by setting the far clipping plane to be at distance 9D from the viewer? What do you conclude as to the relative importance of making these two clipping planes tight around the object? Explain whether using a 32-bit floating point format for the
z-buffer will improve accuracy.

The Volume of a Triangle Soup

You are given a set of triangles. Each is represented independently from the others by the coordinates of its 3 vertices. You are told that together they form a manifold triangle mesh that bounds a connected solid S and that they are properly oriented so that if you are outside that solid and see a triangle, it will appear counterclockwise to you. (You can assume that the mesh is free from self-intersections.) However, the triangles may form more than one shell (i.e. the solid may have internal cavities) and each shell may have higher genus. Provide the details of an algorithm for computing the total volume V and center of mass C of S. (We want floating point accuracy, not a solution based on sampling.) Justify your algorithm and the formulae you are using.

Anti-Aliasing Environment Maps

Environment maps are a fast way to mimic reflections. Give a description of how an environment can be created, and also describe how such an environment map is used during rendering. Pay particular attention to creating and using the pixel footprint in the environment map's parameter space in order to do proper anti-aliasing. What filtering would you suggest to get the best quality image? How would you carry out this filtering operation?

Rendering Section (Answer 2 out of 4)

Hybrid Renderer

You have been hired by Pixar to write a hybrid renderer that is half REYES (RenderMan) and half ray tracing. Here are the features that you have been assigned to include:

1) fast visibility calculation from the camera's point-of-view
2) correct reflections
3) texture mapping with high quality anti-aliasing
4) global illumination

Describe how you would combine the micropolygon-based scan conversion of REYES and a ray tracing system to give all of these features. In particular, describe how these seemingly separate approaches can be integrated together.

Cartoon Anatomy

You have decided to create a book of human anatomy in which the illustrations have been rendered directly from volumetric data from real humans. Your "look" will be cartoon-style rendering. Select a direct volume rendering technique (*not* isosurface extraction) and describe how you can modify the renderer to provide a cartoon-like style. Your renderer
should be able to create dark silhouettes and provide several shading styles including the "flat" cartoon style and hatching that follows the curvature of the surface.

Global Illumination

Select THREE different rendering algorithms that calculate indirect global illumination. Compare these methods in terms of the following:

- speed
- generality of kinds of BRDF's that can be handled
- ease of extending to effects beyond indirect illumination
- potential artifacts such as blockiness or noise
- possibility of hardware acceleration

Better Light Fields

The key to Levoy and Hanrahan's Light Field Rendering system is that they reduce the 5D radiance field to a 4D representation by noting that radiance along an unblocked ray is constant. Their main limitation, then, is that the virtual camera cannot get too close to the surface otherwise the visibility will be incorrect. Describe a modification to their light field data structure that removes this restriction and provides a true 5D radiance field. Try to create a representation that is as true to the original data structure (the "light slab") as possible. Include in your description how interpolation is handled in this new dimension. Also describe how you would compress this representation for storage on disk.

Animation Section (Answer 2 out of 4)

Evaluation of Animation

In recent times, animation has come under some criticism for how to best evaluate results of an animation system. This question is aimed at seeing how would you evaluate two different types of animations.

a) You have recently developed an animation system that lets you animate a basketball player move around with a basketball doing simple moves (like dribbling, and a jump shot, but NO slam-dunks). Assume the basketball player is mocaped and the bouncing ball is added with maya physics simulation with some timing control. How would you set up some experiments to evaluate your animation?

b) You have developed a system for animating water flowing out of a hose-pipe and you can now animate watering a lawn. The water flowing and the grass on the ground are simulated. When the water hits the grass, the grass swings on collision.

In both of the above cases, with mocap and simulation, what are the most important perceptual
cues that (a) the animator would like to show and (b) a person observing would like to see and miss if it was not there. How would you render the scenes in both instances. What would be the difference in the mocap case, vs. the simulated case? What else would you compare to and why?

Mocap Rock Climbing

Irfan, Jim, & Aaron have agreed to let you take the Vicon Mocap Equipment to CRC and capture motions of subjects climbing the wall (ie. rock climbing). Luckily Scott Robertson is there to help with the setup so you don't have to worry about setup and cleanup (yes, this is really fiction, as you will have to do cleanup yourself if you were doing this for real). After you get the data, your first reaction is to take the code you wrote for Motion Graphs (Kovar et al 2002) and use it to generate motions of rock climbing. Will it work? Why or Why not? What is different in this domain? Here are the animations you want to do:

a) Increase the height of the wall, so a subject goes twice as high.

b) Make it that the subject goes zig-zag over the wall. (assume you had a subject go diagonally left to right for this.

BONUS. Would any of the other recent effort in mocap be more appropriate? Just mention the name and one sentence as to why.

Cloth Simulation

Baraff & Witkin (SIGGRAPH 1998) and Choi & Ko (SIGGRAPH 2002) are both about cloth simulation. What is the difference between the two papers. Explain in detail and also state the pros and cons of each approach. Which one do you think is (1) more stable, why? (2) correct, why? (3) easy to implement, why? What are the limitations of each. What are the issues with integration?

OPTIONAL: Breen et al (We know, not on the reading list, except the book!) and other have expressed reservations on these approaches. Now what would these issues be?

Filling Gaps in Motion/Animation

You have been provided with a motion file and a skeleton model of a person. You are told that this motion file is of a famous US soccer player Mia Hamm, running with a few moves running side to side, but mostly running (ie, ignore that she is kicking a ball for now!). Except, you are told that for this dataset of about 1 minute at 120Hz, the middle dataset has some problems. Here are the problems:

1) from time stamp 28 sec - 32 sec, all data is missing.
2) from time stamp 25 sec - 35 sec, all lower body data is missing
3) from time stamp 25 sec - 35 sec, the root node of the skeleton (say the point on the torso) seems to be just a constant (ie does not change!).

How would you fix the dataset for all of the above 3 cases? Would you use a parametric or a non-parametric approach (feel free to specify your definition of each!?)? Would you use a model-based approach, with a dynamics or kinematics? Would data-driven methods be better then model-based ones? Why?

Perception Section (Answer 2 out of 4)

Matting and Segmentation

Matting is the task of extracting a foreground object of interest from an image by specifying an alpha value at each pixel. Suppose you are given the task of writing a program which can automatically generate mattes from a photograph given some user input. Consider the case where each pixel is assigned completely to either background or foreground (alpha values are either 0 or 1). Explain how this problem relates to the classical image segmentation problem in computer vision. One approach to matting might be to first run a standard segmentation algorithm on the image and then ask the user to select the segmented regions corresponding to the foreground object. What are the potential problems with this approach? (Choose any standard segmentation method for the purpose of this discussion). Suppose you are told that the foreground objects are all baseball players photographed during a game. How could you incorporate this knowledge to produce more useful segmentations? What is the difference between over-segmenting and under-segmenting an image? Over-segmentation has been proposed as a pre-processing step for general image analysis (the resulting segmented regions are sometime referred to as super-pixels). Given an example of a vision task where super-pixels could be a useful input representation, and an example where they could be a poor choice.

Flipbook Ordering

A flipbook is a simple form of animation consisting of a sequence of drawings or photographs that when “flipped through” by the viewer produce the illusion of motion. You have just created the world’s best flipbook, but on your way to the stapler you dropped all the pages on the floor. Reassembling the pages in their proper order by hand seems like an impossibly tedious task, but fortunately you have a digital scanner and a fast computer. Design an algorithm that would take as input bitmaps of the randomly-ordered pages from a flipbook and output a plausible ordering. First assume that the pages contain line drawings. Then consider the case where the pages contain photographs. State all of your assumptions and describe the algorithm in sufficient detail such that the major subroutines and representations are clear.

Shape Reconstruction

Suppose you have instrumented a room with multiple synchronized and calibrated cameras covering the walls and ceiling. With this camera array you can capture images of an object at
the center of the room simultaneously from different, known viewpoints. We are interested in constructing 3-D models of the object from this set of images. Compare and contrast the following two approaches to shape reconstruction: 1) Perform stereo reconstruction separately on pairs of images captured by adjacent cameras. Fit a surface model to the resulting 3-D point cloud. 2) Apply the space carving method of Kutulakos and Seitz (see “A Theory of Shape by Space Carving”) to directly construct a voxel model. Pick any standard stereo algorithm for use in method (1). Your comparison should discuss the assumptions about the object made by each method, the effect of image resolution and image noise on the result, and the computational cost. Give an example of an object for which (1) would produce better results than (2) and vice-versa.

Match Moving

In the famous match moving effect from Jurassic Park, the motion of the camera was tracked as it moved along a ridge. The estimated camera motion was used to composite running dinosaur effects with the real terrain imagery. In that case, tennis balls were placed along the path to provide reliable features for tracking the motion of the camera. Imagine repeating this effect today, and automatically tracking the camera without using tennis balls to provide reliable features. Describe a method for automatically identifying good feature locations in an image for tracking purposes. Describe a method for recovering the motion of the camera given sequences of tracked features. In practice, a given feature can only be tracked for a certain number of frames before it disappears. How will your method handle the appearance and disappearance of feature tracks? Most techniques for recovering camera motion (egomotion) are designed to estimate all six DOF’s of a general camera displacement. Suppose that we are told that the camera can only translate forward (in the negative z axis direction) and not in the x or y directions. It still has three rotational DOF’s. How could you modify your motion estimation algorithm to enforce this additional constraint? Do you think this constraint would make much difference in the quality of the camera motion estimate? Explain.