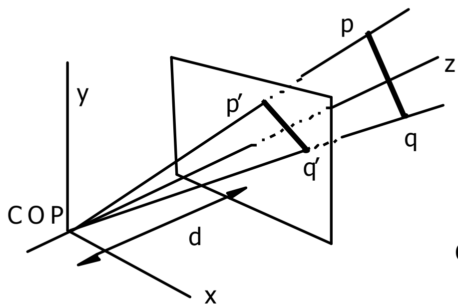
Delft University of Technology
Faculty Electrical Engineering, Mathematics,
and Computer Science
Mekelweg 4, Delft

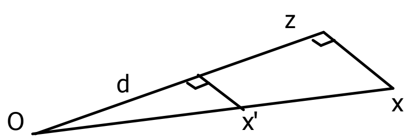
Answers for Examination IN4151

Wednesday, 24 January 2007

1. a. *For modelling and projective display of polygon models we apply several coordinate systems and transformations between these systems. Describe these coordinate systems and transformations. (3)*
The object is defined in the object coordinate space and transformed to world coordinate space. The viewing transformation transforms the object from world to viewing coordinate space, such that the viewpoint is positioned in the origin and the viewing direction is along the z-axis. After projection the image is scaled and translated to a viewport in screen coordinates.

   b. *Derive the perspective foreshortening of x and y as a function of z and the distance d of the center of projection (eye point) and the image plane. Give the matrix for the perspective transformation. (4)*





$$x' = \frac{d}{z} \cdot x = \frac{x}{z/d} \qquad x_s = x'/w'$$

$$y' = \frac{d}{z} \cdot y = \frac{y}{z/d} \qquad y_s = y'/w'$$

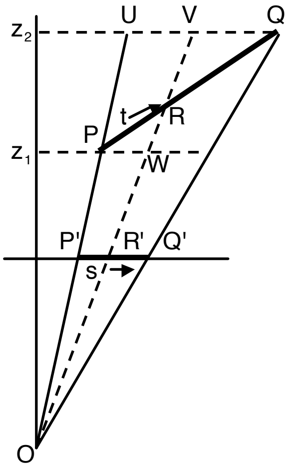
$$\begin{bmatrix} x' & y' & z' & w' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/d & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

The homogeneous divide is used to do the perspective transformation.
In the book of Shirley (par. 7.3) the following perspective matrix is derived

$$\mathbf{M}_p = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & n+f & -fn \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

for the points between the *near* and *far* clipping plane. The points at $z=n$ are not effected.

c. *Equal steps in image/screen space do not correspond with equal steps in object/texture space. Derive a relation between image/screen parameter s and parameter t in object/texture space. (3)*



$$\frac{PW}{UV} = \frac{z_1}{z_2}$$

$$\frac{PR}{RQ} = \frac{PW}{VQ}$$

$$\frac{PR}{RQ} = \frac{z_1}{z_2}\frac{UV}{VQ}$$

$$\frac{UV}{VQ} = \frac{P'R'}{R'Q'}$$

$$\frac{PR}{RQ} = \frac{z_1}{z_2}\frac{P'R'}{R'Q'}$$

$$\frac{t}{(1-t)} = \frac{z_1}{z_2}\frac{s}{(1-s)}$$

$$w_1 = 1/z_1 \qquad w_2 = 1/z_2$$

$$t = \frac{w_2 s}{w_1 + (w_1 - w_2)s}$$

In par. 7.4 of Shirley (p. 173) a similar relation is derived:

$$f(t) = \frac{h_R t}{h_r + t(h_R + h_r)}$$

2. a) *Describe the RGB and CIE xyY colour models. What is represented by Y and what do the x en y components represent? Mention two advantages of the xyY model compared to the RGB model. (6)*

The RGB model is based on the red, green and blue primary colours. The three basic colours are chosen such that they cover on basis of the tristimulus principle a large part of the high frequencies, middle frequencies and lower frequencies of the visual spectrum. In practice the RGB primaries correspond with the phosphors used in the video monitor. The video monitor can directly be driven with RGB values. The RGB model is an additive colour system, i.e. when all primaries are zero we have black and when all primaries are on we have white.
The CIE xyY model is derived form the CIE XYZ model that is based on three artificial saturated primary colours that are chosen such that all visible colours can be represented with positive coordinates. The x and y colour components are the normalized X and Y component normalizing them, i.e. scaling by (X+Y+Z). The Y component represents the luminance, the intensity as perceived by a human eye. The advantage of the xyY colour model is that colours can be interpolated and that the purity of the colour can be read from its position in the xy diagram.

b) *The gamut of a monitor can be represented by a triangular area in the xy-plane. What do the three vertices represent? (2)*

The three vertices are given by the three phosphors (red, green and blue) of the monitor. Different phosphors will give different gamuts.

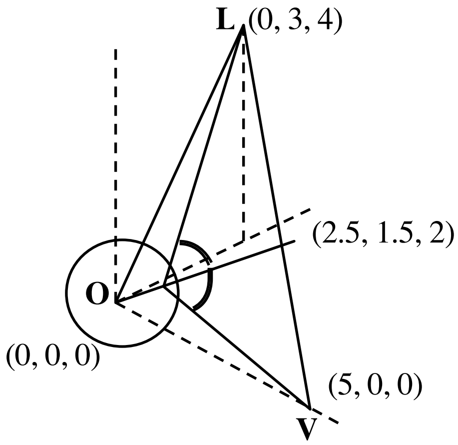
c) *The colours of these vertices are not saturated. How can we find the corresponding fully saturated colours in the xy plane? (2)*

Draw a line in the xyY diagram from the white point through the vertices to the boundary of the visible colour range to find the fully saturated colours.

3. a) *Describe the ray-sphere intersection algorithm for ray tracing. How is the normal at the intersection point calculated? (4)*

The vector equation of the ray is **r** = **e** + t.**f** , where **e** is the eye point, and **f** the direction of the ray. Substitute r_x, r_y and r_z in the quadratic equation of a sphere $(x-c_x)^2 + (y-c_y)^2 + (z-c_z)^2 = r^2$ where **c** is the centre of the sphere and solve for t. This gives a quadratic equation with two roots (two intersections), a double root (touching the sphere) or zero roots (no intersection). The intersection point nearest to the eye point (smallest t) is the first ray sphere intersection. The normal is the normalized vector from the center of the sphere to the intersection point.

b) *Given a sphere with radius 1 centered at the origin. A light source at position (0, 3, 4) and our viewpoint at (5, 0, 0). What are the coordinates of the highlight? (4)*



At the highlight the incoming light and the reflected light (in the direction of the view point) make equal angles with the normal. The normal is also in the same plane as the incoming and reflected light direction. This is the plane through V, L and O. The highlight is at the intersection of the bisection vector (2.5, 1.5, 2) and the sphere with radius 1.0. The length of the vector is $5\sqrt{2}$ thus the highlight is at position $(5\sqrt{2}/10, 3\sqrt{2}/10, 4\sqrt{2}/10)$.

c) *What are the coordinates for the point with the highest diffuse reflection? (2)*
The diffuse reflection is a function of the cosinus of the angle between the incoming light and the normal of the surface. The diffuse reflection is at a maximum where the angle is zero, i.e. where the normal is pointing to the light source and the surface is optimally facing the light. As the sphere is centered at (0, 0, 0), the maximum diffuse reflection is at the intersection of the vector (0, 3, 4) and the sphere, hence (0, 3/5, 4/5).

4. a) *A GPU contains vertex shaders and fragment/pixel shaders. For both types of shaders, describe which part of the viewing pipeline it accelerates. (3)*
The vertex shaders transform the vertices and do the lighting calculation. The fragment shaders do the scan interpolation and the pixel processing such as depth buffer test, texture mapping, filtering.

b) *For a complex rendering task, either the vertex or the fragment shader may become the limiting factor (bottleneck) for the rendering speed. In which type of scene would you expect the vertex shaders to be the bottleneck? (4)*
The vertex shaders do a constant number of operations per vertex, while the load of the fragment shaders is dependent on the number of pixels covered by the triangles. So for a scene with a few large objects the fragment shaders are more busy than the vertex shaders. The vertex shaders would become the bottleneck for models with a large number of polygons that are all very small.

c) *In a scene where the fragment shaders are the bottleneck, what would you expect to happen with the performance if you resize the viewport to half its size. How would this change be different if the vertex shaders were the bottleneck? (3)*
Reducing the viewport will accelerate the fragment shading, as less pixels have to be processed. For the vertex shaders there is no change (assuming the window is the same and the number of vertices does not change).

5. a) *In many VR systems, a stereoscopic effect is obtained through the use of two separate images on a single display and a pair of 'stereo' glasses. Describe passive and active stereo, including some differences, advantages and disadvantages of both. (4)*
Passive stereo uses light filtering through polarization or color filtering. Active uses shutters to consecutively close one eye, i.e. temporal filtering. Active stereo can be used on a normal CRT screen, but a high refresh rate is necessary to avoid flickering. Passive stereo needs two image sources, but gives steady image and, in the case of polarization needs a special screen.

b) *Tracking is used in VR systems to capture human actions. List at least three types of tracking methods, and describe the benefits of using tracking of a user's head in a VR system. (3)*
Mechanical, electromagnetic, inertial and camera tracking. Head tracking allows the generation of a correct perspective viewpoint of the world, and allows motion parallax effects.

c) *What is the influence of the tracking sampling frequency and tracking and rendering latency?* (3)
With high tracking-to-rendering latency, the measured actions suffer from long system delay before use in the graphics rendering. The input movements seem indirect or sluggish and, especially when headtracking is used in an immersive system, can lead to motion sickness. Generally, a lower bound for tracking-to-rendering latency is half of the tracker sampling time. A high frequency can be used to have fluent motion with a high latency/delay, e.g. when the tracking system buffers or data is sent over network.

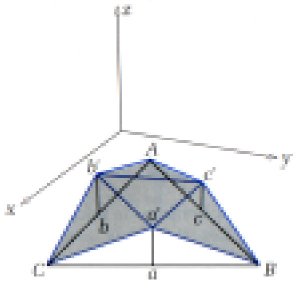
6. a) *Describe a procedural method to model a mountain landscape. How can snow be simulated? (3)*

Start with one triangle (or a set of triangles, being a rough approximation of the mountain landscape to be modelled). Apply the same process for every triangle in every iteration of the fractal generation process:

- displace the midpoints of all triangle edges over a random distance in z-direction (the random distances must be smaller for the next iteration than for the current generation).

- build new triangles from the old triangle vertices and the new displaced midpoints. See the figure below. Triangle ABC is replaced by a'b'c', Ab'c', a'Bc' and a'b'C in the next generation.

- repeat this process for every new generation of the mountain landscape.



Snow can be added by using different colours for different z-coordinates of triangles, i.e. different average z-coordinates of triangle vertices. When the z-coordinates are larger, the chance for a polygon to get the colour white becomes larger.

A similar process can be applied with quadrilaterals instead of triangles.

b) *Describe a procedural method to simulate a forest fire that proceeds over a mountain. (3)*

A forest fire can be modelled with a particle system. Rules are defined for the generation of so called explosion systems in the neighbourhood of existing explosion systems. An explosion system is a circular area on the mountain surface. Inside the circular area particles are born. The particles have an initial position inside the circular area, initial velocity, colour, and several other attributes. The attribute values of the particles change during their lifetimes. After some time particles die. Some attribute values of particles get random starting values (in a certain interval). Rules are defined for the attribute changes of particles during their life time. For instance their position, velocity and colour will change. For frames of the fire animation particles are drawn as points or short line segments at their current position with their current colour.

c) *What is an L-system? Which kinds of objects are modelled with L-systems? What is a context sensitive L-system? Explain context sensitivity with an example. (4)*

An L-system (Lindenmayer system) is a grammar based parallel rewrite system. Substitution rules are defined for non terminal symbols. The rules are applied in parallel to all non terminal symbols in a string. The result is a new string (the next generation). The strings can be converted to a graphical representation by converting the symbols to movements of a turtle that draws the fractal object on a screen. Graphical meaning of a symbol may for instance be: "Turn x degrees to the left" or "Move 1 unit forward, while drawing a line segment". A whole string can be processed in this way, resulting in a drawing of the fractal object.

An L-system is context sensitive when the conversion of a symbol to a string in the next generation is dependent of the context of the symbol (one or more symbols in front of or behind the symbol). Example:

□ : baaaaaaaa	initial string
p1 : b < a ⑧ b	if left context of a is b then a is substituted by b
p2 : b ⑧ a	b is substituted by a

Step	string
1	baaaaaaaa
2	abaaaaaaa
3	aabaaaaaa
4	aaabaaaaa