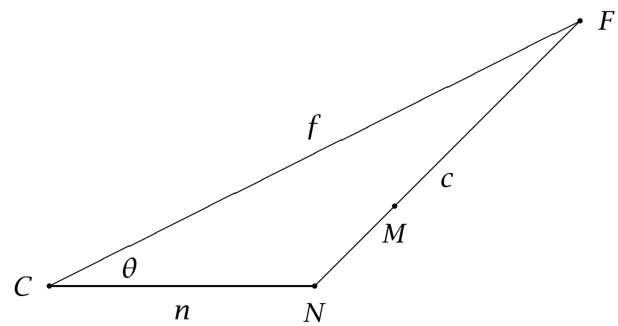
Camera Positions for Dialogue

Peter Grogono

We use the OpenGL coordinate system: Y is up; X is to the right; and Z is towards the camera (eye).

There are two people (i.e., heads): a nearer one at N and a further one at F. The camera is at C. We are interested in the triangle CNF. Note that n is the distance from the camera to the nearer head and f is the distance to the further head. The camera is aimed at M, which is roughly $1/3$ of the way from N to F.



Since three points define a plane, we need two dimensions only. We assume that N, F, and C are all at the same height and so have the same Y coordinate. Thus we work in the XZ plane.

The first thing we compute is the angle θ at C: this is the angle between the two heads as seen from the camera. Since the heads appear at positions $\frac{1}{3}$ and $\frac{5}{6}$ on the screen, the distance between them is $\frac{5}{6} - \frac{1}{3} = \frac{1}{2}$. Thus θ will be approximately half of the angle subtended by the entire screen width ("X angle").

Using a perspective projection, we have the Y angle β (called `fovy` in OpenGL) and the aspect ratio $r = w/h$. The relation between the X angle α and the Y angle β is

$$\frac{\tan(\alpha/2)}{\tan(\beta/2} = r$$

Consequently

$$\alpha = 2 \tan^{-1}(r \tan(\beta/2))$$

and

$$\begin{aligned} \theta &= \alpha/2 \\ &= \tan^{-1}(r \tan(\beta/2)) \end{aligned}$$

Applying the cosine rule to the triangle CNF gives

$$c^2 = f^2 + n^2 - 2 f n \cos\theta$$

Assume that the further head F is k times as far from the camera as the nearer head N. For most applications, $k \approx 3$. Then $f = k n$ and we have

$$\begin{aligned} c^2 &= k^2 n^2 + n^2 - 2 k n^2 \cos\theta \\ &= n^2(k^2 + 1 - 2 k \cos\theta) \end{aligned}$$

Let

$$s^2 = k^2 + 1 - 2k\cos\theta$$

Note that s is a constant that depends only on the perspective projection and the value chosen for k. We also know c, because it is the distance between the heads. We now have:

$$\begin{aligned} n &= c/s \\ f &= kc/s \end{aligned}$$

Next, assign coordinates in the XZ plane to each point:

$$\begin{aligned} C &\equiv (x_c, z_c) \\ N &\equiv (x_n, z_n) \\ F &\equiv (x_f, z_f) \end{aligned}$$

so that

$$\begin{aligned} n^2 &= (x_c - x_n)^2 + (z_c - z_n)^2 \\ f^2 &= (x_c - x_f)^2 + (z_c - z_f)^2 \end{aligned}$$

Then we have

$$\begin{aligned} (x_c - x_n)^2 + (z_c - z_n)^2 &= c^2/s^2 \\ (x_c - x_f)^2 + (z_c - z_f)^2 &= k^2c^2/s^2 \end{aligned}$$

We have to solve these equations to find the camera position (x_c, z_c). To simplify the equations, put the near person at $(0,0)$ and the far person at $(1,0)$. Then

$$\begin{aligned} x_n &= 0 \\ z_n &= 0 \\ x_f &= 1 \\ z_f &= 0 \end{aligned}$$

In this simplified coordinate system, we have $c = 1$ and $M \equiv (\frac{1}{3}, 0)$.

With these substitutions, the equations become:

$$\begin{aligned} x_c^2 + z_c^2 &= 1/s^2 \\ (x_c - 1)^2 + z_c^2 &= k^2/s^2 \end{aligned}$$

Subtracting eliminates z_c and gives

$$(x_c - 1)^2 - x_c^2 = (k^2 - 1)/s^2$$

which we can solve for $2x_c$ giving

$$\begin{aligned} 2x_c &= 1 - \frac{k^2 - 1}{s^2} \\ &= \frac{s^2 - k^2 + 1}{s^2} \\ &= \frac{k^2 + 1 - 2k\cos\theta - k^2 + 1}{s^2} \qquad \text{(using } s^2 = k^2 + 1 - 2k\cos\theta\text{)} \\ &= 2\left(\frac{1 - k\cos\theta}{s^2}\right) \end{aligned}$$

and so

$$x_c = \frac{1 - k\cos\theta}{s^2}$$

For z_c^2, we have:

$$\begin{aligned}
z_c^2 &= \frac{1}{s^2} - x_c^2 \\
&= \frac{1}{s^2} - \frac{(1 - k\cos\theta)^2}{s^4} \\
&= \frac{k^2 + 1 - 2k\cos\theta - 1 + 2k\cos\theta - k^2\cos^2\theta}{s^4} \\
&= \frac{k^2\sin^2\theta}{s^4}
\end{aligned}$$

and therefore

$$z_c = \pm\frac{k\sin\theta}{s^2}$$

The positive and negative square roots correspond to two possible camera positions. The diagram shows one position; the other position is obtained by reflecting C in the line NF.

We now have the camera position, (x_c, z_c), in the special coordinate system. To obtain the true camera position in the original coordinate system, we apply the following transformations to (x_c, y_c):

1. Rotate about the origin through an angle ϕ where

$$\begin{aligned}
\sin\phi &= \frac{x_f - x_n}{d} \\
\cos\phi &= \frac{z_f - z_n}{d} \\
\text{where}\quad d &= \sqrt{(x_f - x_n)^2 + (z_f - z_n)^2}
\end{aligned}$$

   (Check this!)

2. Scale (increase the distance NF from 1 to c and other distances in proportion):

$$\begin{aligned}
x' &= c\,x \\
z' &= c\,z
\end{aligned}$$

3. Translate (move N to its correct position (x_n, z_n)):

$$\begin{aligned}
x' &= x + x_n \\
z' &= z + z_n
\end{aligned}$$

When these transformations have been applied to (x_c, z_c), we should have the correct camera position. As a check, the same transformations applied to $(0, 0)$ and $(1, 0)$ should give the correct positions for $N \equiv (x_n, z_n)$ and $F \equiv (x_f, z_f)$, respectively. We can apply the same transformations to $(\frac{1}{3}, 0)$ to obtain the true coordinates of M.

Then we use C as the eye position and M as the model position in the `gluLookAt` call.