CS 465 Homework 4 (v.2): Transformation Matrices with Perspective

out: Sunday 21 September 2008
due: Friday 26 September 2008

In the last homework we discussed several categories of transformations and their compositions. In this problem set, we will further discuss the general affine transformations and also add perspective transformations. Before we start, let's quickly go over part of the course materials.

A 2D affine transformation can be written as a 3x3 matrix

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

which transforms a point (x, y) to (x', y')

$$M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

Similary a general 3D perspective transformation can be written as a 4x4 matrix. In the questions below, the points are lying on the ground so $z = 0$ for all the points; meanwhile, just by looking at the image we get we don't know the z values after transformation. Since z doesn't come into play at all, the perspective matrix degenerates into a 3x3 matrix

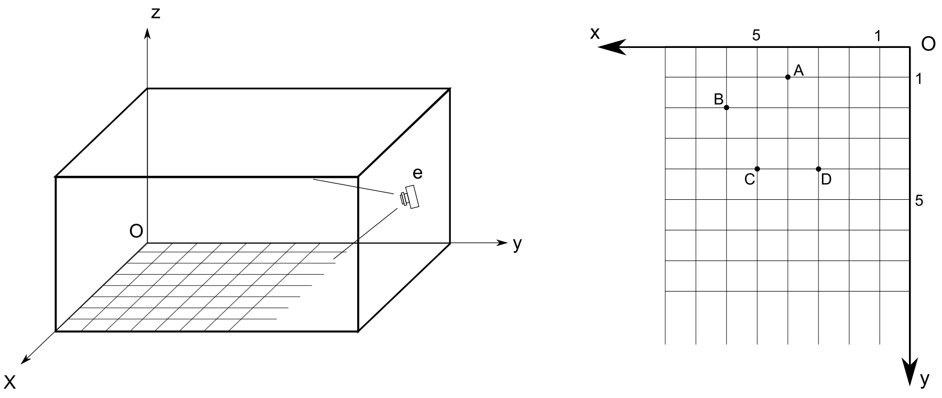
$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & 1 \end{bmatrix}$$

such that

$$M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \propto \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

Now consider the scene shown in figure 1(a), in which a camera at point e is taking pictures of a room. There are unit-size grid patterns on the floor. Points A, B, C and D are on the floor, and their cooridinates are $(4, 1, 0)$, $(6, 2, 0)$, $(5, 4, 0)$ and $(3, 4, 0)$ respectively.

1. Suppose the camera has a mysterious function that can take pictures under parallel projection, i.e. pictures without perspective (Fig. 2(a)). The coordinates of the images of the points A, B and C in the picture are $(0.3, 0.6)$, $(0.7, 0.65)$ and $(0.65, 0.85)$ respectively. Find out the transformation matrix M which transforms the points (x, y) on the floor to the points (x', y') in the picture. Note that figure 2(a) is just an illustration and it might not perfectly match your result. (Hint: the transformation is a 2D affine transformation.)



(a) Overview of the room

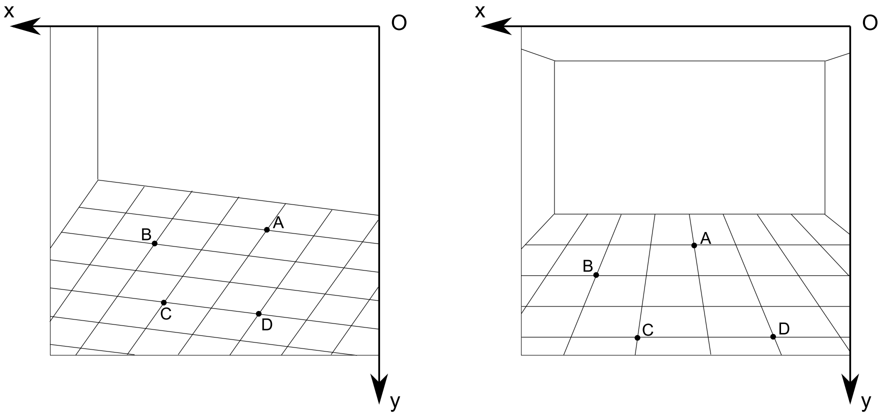
(b) Some points on the floor

Figure 1: A room with grids on the floor, and a camera at point e.

2. Generalize the above problem. Given three linearly independent points (x_i, y_i) on the floor and their coordinates (u_i, v_i) in the picture ($i = 1, 2, 3$), derive an affine transformation M that transforms (x_i, y_i) to (u_i, v_i). If you feel the closed form is too long to write down, you may leave your answer in terms of matrix multiplication and inversion.

3. Now turn off the mysterious function and the camera is in normal perspective mode. Figure 2(b) is a perspective view taken by it. The coordinates of the images of the points A, B, C and D in the pictures are $(0.45, 0.65)$, $(0.75, 0.7)$, $(0.65, 0.9)$ and $(0.25, 0.9)$. Find out the transformation matrix M which transforms the points (x, y) on the floor to the points (x', y') in the picture. Again, figure 2(b) is just an illustration and might not perfectly match your result. (Hint: the transformation is a 2D perspective transformation.)

4* Generalize the above problem. Given four points (x_i, y_i) in general position[1] on the floor and their coordinates (u_i, v_i) in the picture ($1 \leq i \leq 4$), derive a perspective transformation M that transforms (x_i, y_i) to (u_i, v_i). As before, you may leave matrix multiplication, inversion, or other intermediate steps in your formula.



(a) A picture without perspective

(b) A picture with perspective

Figure 2: The images of the room under parallel projection and perspective projection.

---

[1] General position in this context means no three points are collinear.