Perspective Projection
(Com S 477/577 Notes)

Yan-Bin Jia
Sep 5, 2019

1 Introduction

We now move on to visualization of three-dimensional objects, getting back to the use of homogeneous coordinates. Current display devices such as computer monitors, LCD screens, and printers are two-dimensional. Cameras also have 2D image screens. It is necessary to obtain a planar view of an object which gives three-dimensional realism. Visualization of an object is achieved by a sequence of operations called the viewing pipeline.

1. A projection is applied which maps the object to a new ‘flat’ object in a specified plane known as the viewplane.
2. A coordinate system is defined in the viewplane and the description of the ‘flat’ object in the coordinates is obtained.
3. The ‘flat’ object is mapped to the computer screen by means of a two-dimensional device coordinate transformation.

We will first look at projections of the plane onto a line, and subsequently projections of three-dimensional space onto a plane.

2 Projections of the Plane

Let us consider the problem of projecting the plane onto a line \( \ell \) contained in the plane. Let \( v \) be a point not on the line. The perspective projection from \( v \) onto \( \ell \) is the transformation which maps any point \( p \neq v \) to the point \( p' \) which is the intersection of the lines \( vp \) and \( \ell \), as illustrated in Figure 1. The point \( v \) is called the viewpoint or center of perspectivity, and the line \( \ell \) is called the viewline.

**Theorem 1** The perspective projection from a viewpoint \( v \) (in homogeneous coordinates) onto a viewline vector \( \ell \) is a two-dimensional transformation given by the matrix \( M = v\ell^T - (\ell \cdot v)I_3 \), where \( I_3 \) is the \( 3 \times 3 \) identity matrix.

![Figure 1: Perspective Projection](image-url)
Proof. The image point $p'$ of a point $p$ is the intersection of the viewline $\ell$ with the line $\ell'$ through $p$ and the viewpoint $v$. In homogeneous coordinates, the line $\ell'$ has the line vector $v \times p$, and therefore intersects $\ell$ at the point

\[
p' = \ell \times (v \times p) = (\ell \cdot p)v - (\ell \cdot v)p
\]

\[
= v(\ell \cdot p) - (\ell \cdot v)I_3p
\]

\[
= v\ell^T p - (\ell \cdot v)I_3p
\]

\[
= \left( v\ell^T - (\ell \cdot v)I_3 \right) p.
\]

Thus $p' = M p$ where

\[
M = v\ell^T - (\ell \cdot v)I_3.
\]

The matrix $M$ is called the projection matrix of the perspective projection from $v$ onto $\ell$. Lines through the viewpoint are called projectors. When the viewpoint $v$ is a point at infinity, the projection is called a parallel projection. As shown in Figure 2, a parallel projection has viewpoint $v = (v_1, v_2, 0)$, that is, the infinity point in the direction $(v_1, v_2)$. The projectors are parallel lines in the Cartesian plane with direction $(v_1, v_2)$. It is common practice to use the term ‘perspective projection’ to refer to a non-parallel projection.

Example 1. We determine the perspective projection of the triangle with vertices $(\frac{2}{3}, \frac{4}{3})$, $(\frac{4}{5}, \frac{2}{3})$, and $(\frac{3}{4}, \frac{1}{3})$ onto the line $5x + y - 4 = 0$ from the viewpoint $(\frac{10}{5}, \frac{2}{3})$. The homogeneous viewpoint is $v = (10, 2, 1)^T$, the line vector $\ell = (5, 1, -4)^T$ with $v \cdot \ell = 48$. We calculate the projection matrix as follows

\[
M = v\ell^T - (\ell \cdot v)I_3
\]

\[
= \begin{pmatrix} 50 & 10 & -40 \\ 10 & 2 & -8 \\ 5 & 1 & -4 \end{pmatrix} - 48 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\]

\[
= \begin{pmatrix} 2 & 10 & -40 \\ 10 & -46 & -8 \\ 5 & 1 & -52 \end{pmatrix}.
\]

The images of the vertices are obtained by multiplying $M$ to their homogeneous coordinates:

\[
\begin{pmatrix} 2 & 10 & -40 \\ 10 & -46 & -8 \\ 5 & 1 & -52 \end{pmatrix} \begin{pmatrix} 2 & 4 & 3 \\ 3 & 4 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -6 & 8 & -44 \\ -126 & -152 & 68 \\ -39 & -28 & -38 \end{pmatrix}.
\]

The Cartesian coordinates of the vertex images are $\left(\frac{2}{13}, \frac{4}{13}\right)$, $\left(-\frac{2}{7}, \frac{38}{7}\right)$, and $\left(-\frac{22}{19}, \frac{19}{19}\right)$. They are shown in Figure 3(a).
\( p_1 = (2, 3) \)
\( p_2 = (4, 4) \)
\( p_3 = (3, -1) \)
\( p_3' = (2, -1) \)
\( \ell : 5x + y - 4 = 0 \)

**Example 2.** Now let us determine the parallel projection of the same triangle in Example 1 onto the line \( 3x + 2y - 4 = 0 \) in the direction of the \( y \)-axis. The viewpoint is thus \( v = (0, 1, 0)^T \), the point at infinity in the direction of the \( y \)-axis. We have \( \ell = (3, 2, -4)^T \) and \( \ell \cdot v = 2 \). From (2) the projection matrix is determined to be

\[
M = \begin{pmatrix}
-2 & 0 & 0 \\
3 & 0 & -4 \\
0 & 0 & -2
\end{pmatrix}.
\]

And the homogeneous coordinates of the images of the triangle vertices are contained in the matrix product

\[
\begin{pmatrix}
-2 & 0 & 0 \\
3 & 0 & -4 \\
0 & 0 & -2
\end{pmatrix} \begin{pmatrix}
2 & 4 & 3 \\
3 & 4 & -1 \\
1 & 1 & 1
\end{pmatrix} = \begin{pmatrix}
-4 & -8 & -6 \\
2 & 8 & 5 \\
-2 & -2 & -2
\end{pmatrix}.
\]

The Cartesian coordinates are therefore \( \left( \frac{2}{-1} \right), \left( \frac{4}{-1} \right), \) and \( \left( \frac{3}{-5/2} \right) \), as shown in Figure 3(b).

### 3 Projections of Space

Analogous to a viewline in the projection of the plane, a viewplane is involved in the projection of three-dimensional space. Let \( n = (a, b, c, d) \) be the plane vector of a viewplane described by the general equation \( ax + by + cz + d = 0 \), and \( v \) be a point not on the viewplane. The **perspective projection** from \( v \) onto the viewplane \( n \) is a transformation that maps any point \( p \neq v \), onto the intersection point \( p' \) of the line \( vp \) and the plane. If \( v \) is a point at infinity then the projection is called a **parallel projection**. The two projections are illustrated in Figure 4.

**Theorem 2** The projection with homogeneous viewpoint \( v \) and viewplane with plane vector \( n \) is the three-dimensional transformation given by the matrix \( M = vn^T - (n \cdot v)I_4 \), where \( I_4 \) is the \( 4 \times 4 \) identity matrix.

**Proof** Let \( p \) with \( p \neq v \) be a point to be projected. It can be shown that every point on the line through \( p \) and \( v \) has the homogeneous coordinates of the form \( \alpha p + \beta v \) for some \( \alpha \) and \( \beta \) such as
that $\alpha \neq 0$ or $\beta \neq 0$. The line intersects the viewplane if $\alpha \mathbf{p} + \beta \mathbf{v}$ lies on the plane for some $\alpha$ and $\beta$, that is, when $\mathbf{n} \cdot (\alpha \mathbf{p} + \beta \mathbf{v}) = 0$. Thus

$$\alpha (\mathbf{n} \cdot \mathbf{p}) + \beta (\mathbf{n} \cdot \mathbf{v}) = 0. \quad (1)$$

The following two cases arise:

1. When $\mathbf{n} \cdot \mathbf{p} = 0$, the point $\mathbf{p}$ is on the viewplane and its projected image is itself. Meanwhile, we have that

$$M\mathbf{p} = \left(v\mathbf{n}^T - (\mathbf{n} \cdot \mathbf{v})I_4\right)\mathbf{p}$$

$$= (\mathbf{n} \cdot \mathbf{v})\mathbf{v} - (\mathbf{n} \cdot \mathbf{v})I_4\mathbf{p}$$

$$= -(\mathbf{n} \cdot \mathbf{v})\mathbf{p}.$$

Thus, $M\mathbf{p}$ is a multiple of $\mathbf{p}$, or equivalently, $\mathbf{p}$ itself in homogeneous coordinates.

2. When $\mathbf{n} \cdot \mathbf{p} \neq 0$, from (1) we obtain that $\alpha = -\beta (\mathbf{n} \cdot \mathbf{v})/(\mathbf{n} \cdot \mathbf{p})$. Substituting for $\alpha$, the point of intersection has homogeneous coordinates

$$\mathbf{p}' = \alpha \mathbf{p} + \beta \mathbf{v}$$

$$= \left(\frac{-\beta (\mathbf{n} \cdot \mathbf{v})}{\mathbf{n} \cdot \mathbf{p}}\right)\mathbf{p} + \beta \mathbf{v}.$$

Multiplying the coordinates by the scalar $\mathbf{n} \cdot \mathbf{p}$ and dividing by $\beta$ gives an alternative homogeneous coordinates in matrix form:

$$\mathbf{p}' = (\mathbf{n} \cdot \mathbf{v})\mathbf{v} - (\mathbf{n} \cdot \mathbf{v})\mathbf{p}$$

$$= \left(v\mathbf{n}^T - (\mathbf{n} \cdot \mathbf{v})I_4\right)\mathbf{p}.$$

Hence we also have $M = v\mathbf{n}^T - (\mathbf{n} \cdot \mathbf{v})I_4$. 

\[\square\]
Example 3. A prism shown in Figure 5(a) has vertices $(0, 0, 0)^T$, $(2, 0, 0)^T$, $(2, 3, 0)^T$, $(0, 3, 0)^T$, $(1, 2, 1)^T$, and $(1, 1, 1)^T$. Consider a parallel projection of the prism onto the plane in a direction parallel to the $z$-axis. The viewpoint is $v = (0, 0, 1, 0)^T$, and the plane vector is $n = (0, 0, 1, 0)^T$. We determine the projection matrix:

\[
M_1 = \begin{pmatrix}
0 \\
0 \\
1 \\
0
\end{pmatrix} (0 0 1 0) - 1 \cdot \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

\[
= \begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}
\]

Applying the projection to the vertices of the prism yields

\[
\begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix} \begin{pmatrix}
0 & 2 & 2 & 0 & 1 & 1 \\
0 & 0 & 3 & 3 & 2 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix} = \begin{pmatrix}
0 & -2 & -2 & 0 & -1 & -1 \\
0 & 0 & -3 & -3 & -2 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & -1 & -1 & -1 & -1
\end{pmatrix}
\]

Thus the images of the vertices under the projection are $(0, 0, 0)^T$, $(2, 0, 0)^T$, $(2, 3, 0)^T$, $(0, 3, 0)^T$, $(1, 2, 1)^T$, and $(1, 1, 1)^T$, as shown in Figure 5(b).

Next, we consider a perspective projection onto the plane $z = 0$ from the viewpoint $(1, 5, 3)^T$. Here $v = (1, 5, 3, 1)^T$ and $n = (0, 0, 1, 0)^T$. And the projection matrix is

\[
M_2 = \begin{pmatrix}
1 \\
5 \\
3 \\
1
\end{pmatrix} (0 0 1 0) - 3 \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

\[
= \begin{pmatrix}
-3 & 0 & 1 & 0 \\
0 & -3 & 5 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 1 & -3
\end{pmatrix}
\]

\[\text{Note that } \beta \neq 0 \text{ must hold, otherwise } n \cdot p = 0.\]
which, multiplied to the homogeneous coordinates of the prism vertices, yields

\[
\begin{pmatrix}
-3 & 0 & 1 & 0 \\
0 & -3 & 5 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 1 & -3
\end{pmatrix}
\begin{pmatrix}
0 & 2 & 2 & 0 & 1 & 1 \\
0 & 0 & 3 & 3 & 2 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}
= \begin{pmatrix}
0 & -6 & -6 & 0 & -2 & -2 \\
0 & 0 & -9 & -9 & -1 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 \\
-3 & -3 & -3 & -3 & -2 & -2
\end{pmatrix}
\]

Thus the image vertices are \((0,0,0)^T\), \((2,0,0)^T\), \((2,3,0)^T\), \((0,3,0)^T\), \((1,\frac{1}{2},0)^T\), and \((1,-1,0)^T\), as shown in Figure 5(c).

4 Viewplane Coordinate Mapping

So far the view of an object in the viewplane is expressed in homogeneous coordinates that correspond to the 3-dimensional world coordinates. The next stage is to define a coordinate system on the viewplane and represent the object in terms of these new coordinates. Figure 6 shows an image taken by a high speed camera of a robotic arm about to bat a flying dumbbell-shaped object. The camera’s image plane has a local coordinate system located at the upper left corner. What is transformation from a point, say, the center of the dumbbell, in the 3D world to image coordinates that is automatically carried out by the camera?

Generally, the viewplane \((u,v)\)-coordinate system is specified in world coordinates by an origin \(q = (q_1,q_2,q_3)^T\), and two unit vectors \(\hat{r} = (r_1,r_2,r_3)^T\) and \(\hat{s} = (s_1,s_2,s_3)^T\) which indicate the directions of the \(u\)- and \(v\)-axes, respectively. See Figure 7. Consider a point on the viewplane with homogeneous world coordinates \(p' = (x,y,z,t)^T\) and homogeneous viewplane coordinates \(p'' = (u,v,w)^T\). To obtain the Cartesian coordinates of \(p''\), we can simply project the vector \(qp'\)
onto the unit vectors $\hat{r}$ and $\hat{s}$ using inner products; namely,

\[
\frac{u}{w} = \left( \frac{x}{t} \right) - \left( \begin{array}{c} q_1 \\ q_2 \\ q_3 \end{array} \right) \cdot \left( \begin{array}{c} r_1 \\ r_2 \\ r_3 \end{array} \right),
\]

\[
\frac{v}{w} = \left( \frac{x}{t} \right) - \left( \begin{array}{c} q_1 \\ q_2 \\ q_3 \end{array} \right) \cdot \left( \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} \right).
\]

But let us try to solve the problem by a homogeneous transformation instead. More specifically, we would like to obtain $p''$ from $p'$ by a mapping $V$ such that $p'' = Vp'$, where $V$ is a $3 \times 4$ matrix. Rather than compute $V$ directly, we determine a $4 \times 3$ matrix $K$ such that $p' = Kp''$ and then express $V$ in terms of $K$.

The matrix $K$ can be determined, up to scaling, from four non-collinear points on the viewplane, using their homogeneous world coordinates and corresponding viewplane coordinates. We choose the following four points: (a) the origin $(q_1, q_2, q_3, 1)^T$ of the view plane; (b) the point at infinity $(r_1, r_2, r_3, 0)^T$ in the direction of the $u$-axis of the viewplane coordinate system; (c) the point at infinity $(s_1, s_2, s_3, 0)^T$ in the direction of the $v$-axis of the viewplane coordinate system; and (d) the point $(t_1, t_2, t_3, 1) = (q_1 + r_1 + s_1, q_2 + r_2 + s_2, q_3 + r_3 + s_3, 1)^T$ which is one-unit each in the $u$ and $v$ directions from the origin. The homogeneous viewplane coordinates of these points are $(0, 0, 1)^T$, $(1, 0, 0)^T$, $(0, 1, 0)^T$, and $(1, 1, 1)^T$, respectively. Then the corresponding points are mapped to each other as follows:

\[
\begin{pmatrix}
q_1 & r_1 & s_1 & t_1 \\
q_2 & r_2 & s_2 & t_2 \\
q_3 & r_3 & s_3 & t_3 \\
1 & 0 & 0 & 1
\end{pmatrix} =
\begin{pmatrix}
r_1 & s_1 & q_1 \\
r_2 & s_2 & q_2 \\
r_3 & s_3 & q_3 \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1
\end{pmatrix}
\]
Hence the transformation matrix is

\[ K = \begin{pmatrix} r_1 & s_1 & q_1 \\ r_2 & s_2 & q_2 \\ r_3 & s_3 & q_3 \\ 0 & 0 & 1 \end{pmatrix}. \]

Clearly, the column vectors of \( K \) are linearly independent. Therefore \( \text{rank}(K) = 3 \).

The viewplane coordinate mapping is an inverse of the transformation determined by the matrix \( K \). Since \( K \) is not a square matrix there is no matrix inverse \( K^{-1} \). We use a left inverse \( L \), for which \( LK = I_3 \). From \((K^TK)^{-1}K^TP = I_3\) we have a left inverse \( L = (K^TK)^{-1}K^T \). Here \((K^TK)^{-1}\) exists because \( \text{rank}(K) = 3 \). Since \( p' = Kp'' \), we see that

\[ Lp' = (K^TK)^{-1}K^T(Kp'') \]
\[ = (K^TK)^{-1}(K^TK)p'' \]
\[ = p''. \]

Hence the viewplane coordinate mapping is given by the matrix

\[ V = L = (K^TK)^{-1}K^T. \]

(2)

Observe that the viewplane coordinate matrix \( V \) is determined only by the choice of origin \((q_1, q_2, q_3)^T\), and the directions \((r_1, r_2, r_3)^T\) and \((s_1, s_2, s_3)^T\) of the \( u \)- and \( v \)-axes.

**Example 4.** Consider the perspective projection in Example 3 of a prism onto the plane \( z = 0 \) from the viewpoint \( v = (1, 5, 3)^T \). Let a \((u,v)\)-coordinate system on the viewplane be given by the origin \((1, 2, 0)^T\), \( u \)-axis direction \((3, 4, 0)^T\), and \( v \)-axis direction \((-4, 3, 0)^T\). The unit vectors in the axis directions are \( \hat{r} = (\frac{3}{5}, \frac{4}{5}, 0)^T \), and \( \hat{s} = (-\frac{4}{5}, \frac{3}{5}, 0)^T \). The viewplane coordinate matrix \( V = (K^TK)^{-1}K^T \) that maps a point in the space to a point in the viewplane is obtained in the following steps:

\[ K^TK = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} & 1 \\ \frac{4}{5} & \frac{3}{5} & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \]
In Example 3, we already determined the homogeneous world coordinates of the prism vertices. Their homogeneous viewplane coordinates are computed through a further multiplication by $V$:

$$
\begin{pmatrix}
\frac{3}{5} & \frac{4}{5} & 0 & -\frac{11}{5} \\
-\frac{4}{5} & \frac{4}{5} & 0 & -\frac{2}{5} \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
0 & -6 & -6 & 0 & -2 & -2 \\
0 & 0 & -9 & -9 & -1 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 \\
-3 & -3 & -3 & -3 & -2 & -2
\end{pmatrix}
= 
\begin{pmatrix}
\frac{33}{5} & 3 & -21 & -\frac{3}{5} & 12 & \frac{24}{5} \\
\frac{6}{5} & 6 & \frac{3}{5} & -\frac{21}{5} & 9 & \frac{18}{5} \\
0 & 0 & 0 & 0 & 0 & 0 \\
-3 & -3 & -3 & -3 & -2 & -2
\end{pmatrix}
$$

Hence the Cartesian viewplane coordinates of the vertices are \((-11/5), (-2/5), (7/5), (1/5), (-6/5), (-12/5), (-9/5)).

**References**