Perspective Projection

(Com S 477/577 Notes)

Yan-Bin Jia

Aug 27, 2020

1 Introduction

We look at the application of homogeneous coordinates to visualization of three-dimensional objects. 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 that yields three-dimensional realism. Visualization of the object is achieved via 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 threedimensional space onto a plane.

2 Projections of the Plane

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

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

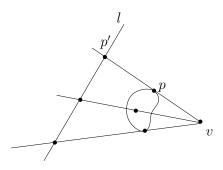


Figure 1: Perspective Projection

Proof The image point p' of a point p is the intersection of the viewline ℓ with the line ℓ' through p and the viewpoint v. In homogeneous coordinates, the line ℓ' has the line vector $v \times p$, and therefore intersects ℓ at the point

$$egin{array}{lll} oldsymbol{p}' &=& \ell imes (oldsymbol{v} imes oldsymbol{p}) \ &=& (\ell \cdot oldsymbol{p}) oldsymbol{v} - (\ell \cdot oldsymbol{v}) I_3 oldsymbol{p} \ &=& oldsymbol{v} (\ell \cdot oldsymbol{p}) - (\ell \cdot oldsymbol{v}) I_3 oldsymbol{p} \ &=& igg(oldsymbol{v} \ell^T oldsymbol{p} - (\ell \cdot oldsymbol{v}) I_3 igg) oldsymbol{p}. \end{array}$$

Thus $\mathbf{p}' = M\mathbf{p}$ where

$$M = \boldsymbol{v}\ell^T - (\ell \cdot \boldsymbol{v})I_3.$$

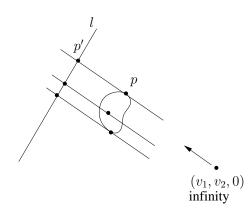


Figure 2: Parallel Projection

The matrix M is called the *projection matrix* of the perspective projection from \mathbf{v} onto ℓ . Lines through the viewpoint are called *projectors*. When the viewpoint \mathbf{v} is a point at infinity, the projection is called a *parallel projection*. As shown in Figure 2, a parallel projection has viewpoint $\mathbf{v} = (v_1, v_2, 0)$, 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 $\binom{2}{3}$, $\binom{4}{4}$, and $\binom{3}{-1}$ onto the line 5x+y-4=0 from the viewpoint $\binom{10}{2}$. The homogeneous viewpoint is $\boldsymbol{v}=(10,2,1)^T$, the line vector $\ell=(5,1,-4)^T$ with $\boldsymbol{v}\cdot\ell=48$. 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 $\binom{2/13}{42/13}$, $-\binom{2/7}{38/7}$, and $\binom{22/19}{-34/19}$. They are shown in Figure 3(a).

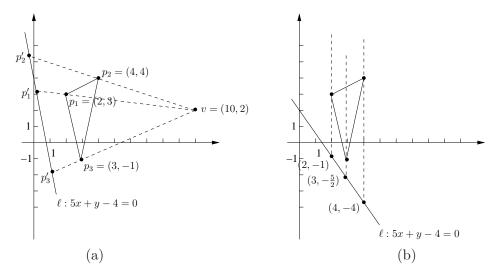


Figure 3: Perspective and parallel projections of a triangle.

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 $\mathbf{v}=(0,1,0)^T$, the point at infinity in the direction of the y-axis. We have $\ell=(3,2,-4)^T$ and $\ell\cdot\mathbf{v}=2$. From (2) the projection matrix is determined to be

$$M = \left(\begin{array}{rrr} -2 & 0 & 0\\ 3 & 0 & -4\\ 0 & 0 & -2 \end{array}\right).$$

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

$$\left(\begin{array}{ccc} -2 & 0 & 0 \\ 3 & 0 & -4 \\ 0 & 0 & -2 \end{array}\right) \left(\begin{array}{ccc} 2 & 4 & 3 \\ 3 & 4 & -1 \\ 1 & 1 & 1 \end{array}\right) = \left(\begin{array}{ccc} -4 & -8 & -6 \\ 2 & 8 & 5 \\ -2 & -2 & -2 \end{array}\right).$$

The Cartesian coordinates are therefore $\binom{2}{-1}$, $\binom{4}{-4}$, and $\binom{3}{-5/2}$, 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 $\mathbf{n}=(a,b,c,d)$ be the plane vector of a viewplane described by the general equation ax+by+cz+d=0, and \mathbf{v} be a point not on the viewplane. The perspective projection from \mathbf{v} onto the viewplane \mathbf{n} is a transformation that maps any point $\mathbf{p}\neq\mathbf{v}$, onto the intersection point \mathbf{p}' of the line $\overline{\mathbf{v}}$ and the plane. If \mathbf{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 \mathbf{v} and viewplane with plane vector \mathbf{n} is the three-dimensional transformation given by the matrix $M = \mathbf{v}\mathbf{n}^T - (\mathbf{n} \cdot \mathbf{v})I_4$, where I_4 is the 4×4 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 α and β such

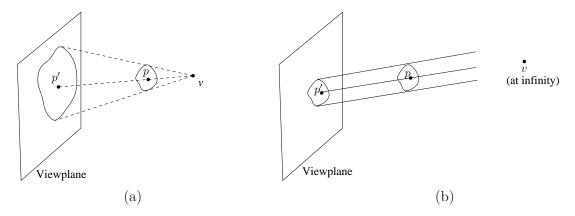


Figure 4: (a) Perspective projection and (b) parallel projection.

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 α and β , that is, when $\mathbf{n} \cdot (\alpha \mathbf{p} + \beta \mathbf{v}) = 0$. Thus

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

The following two cases arise:

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

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

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

Thus, Mp is a multiple of p, or equivalently, 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 α , the point of intersection has homogeneous coordinates

$$p' = \alpha p + \beta v$$

= $\left(-\beta (n \cdot v)/(n \cdot p)\right) p + \beta v.$

Multiplying the coordinates by the scalar $n \cdot p$ and dividing by β gives¹ the alternative homogeneous coordinates in matrix form:

$$p' = (n \cdot p)v - (n \cdot v)p$$

= $(vn^T - (n \cdot v)I_4)p$.

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

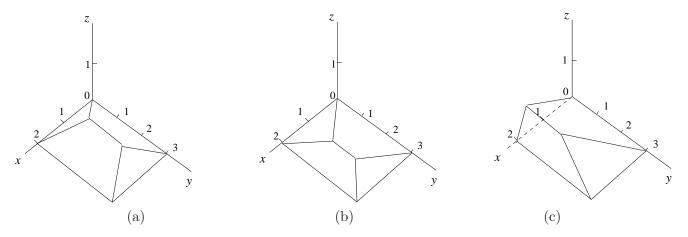


Figure 5: (a) A prism under (b) a parallel projection and (c) a perspective projection.

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 $\mathbf{v} = (0,0,1,0)^T$, and the plane vector is $\mathbf{n} = (0,0,1,0)^T$. We determine the projection matrix:

$$M_{1} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} (0 \quad 0 \quad 1 \quad 0) - 1 \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Applying the projection to the vertices of the prism yields

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

Thus the images of the vertices under the projection are $(0,0,0)^T$, $(2,0,0)^T$, $(2,3,0)^T$, $(0,3,0)^T$, $(1,2,0)^T$, and $(1,1,0)^T$, 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 $\mathbf{v} = (1,5,3,1)^T$ and $\mathbf{n} = (0,0,1,0)^T$. And the projection matrix is

$$M_{2} = \begin{pmatrix} 1 \\ 5 \\ 3 \\ 1 \end{pmatrix} (0 \quad 0 \quad 1 \quad 0) - 3 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} -3 & 0 & 1 & 0 \\ 0 & -3 & 5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -3 \end{pmatrix}$$

¹Note that $\beta \neq 0$ must hold, otherwise $\mathbf{n} \cdot \mathbf{p} = 0$.

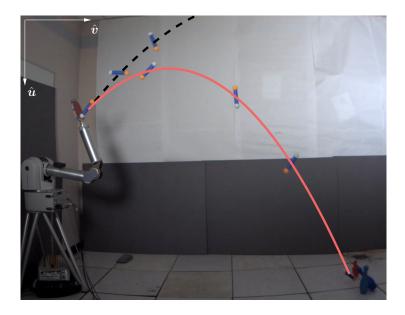


Figure 6: Image of a robotic arm batting an in-flight dumbbell-like object to hit three pins (courtsey of Matthew Gardner). Video on YouTube at https://youtu.be/dGBevZ54E3s.

which, multiplied to the homogeneous coordinates of the prism vertices, yields

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

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. Supermosed onto the image are some intermediate poses of the object along its pre- and post-batting trajectories. The camera's image plane has a local coordinate system located at the upper left corner. What is the 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 $\mathbf{q} = (q_1, q_2, q_3)^T$, and two unit vectors $\hat{\mathbf{r}} = (r_1, r_2, r_3)^T$ and $\hat{\mathbf{s}} = (s_1, s_2, s_3)^T$ which indicate the directions of the u- and v-axes, respectively. See Figure 7. Consider a point on the viewplane with homogeneous world coordinates $\mathbf{p}' = (x, y, z, t)^T$ and homogeneous viewplane coordinates $\mathbf{p}'' = (u, v, w)^T$. To obtain the Cartesian coordinates of \mathbf{p}'' , we can simply project the vector $\mathbf{q}\mathbf{p}'$

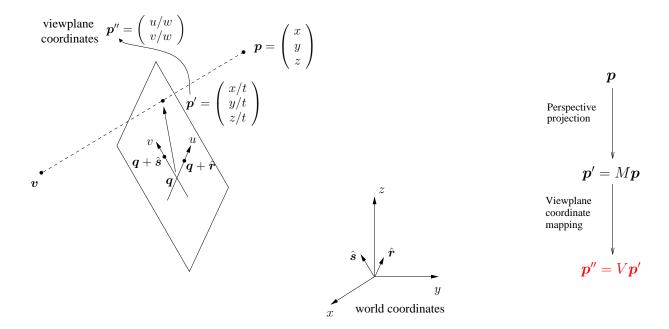


Figure 7: Viewplane and world coordinate systems.

onto the unit vectors \hat{r} and \hat{s} using inner products; namely,

$$\frac{u}{w} = \left(\begin{pmatrix} x/t \\ y/t \\ z/t \end{pmatrix} - \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \right) \cdot \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix},$$

$$\frac{v}{w} = \left(\begin{pmatrix} x/t \\ y/t \\ z/t \end{pmatrix} - \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \right) \cdot \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}.$$

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 3×4 matrix. Rather than compute V directly, we determine a 4×3 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 (4) 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:

$$\begin{pmatrix} q_1 & r_1 & s_1 & t_1 \\ q_2 & r_2 & s_2 & t_2 \\ q_3 & r_3 & s_3 & t_3 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} r_1 & s_1 & q_1 \\ r_2 & s_2 & q_2 \\ r_3 & s_3 & q_3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

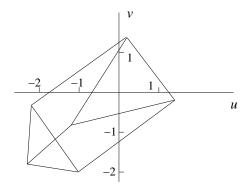


Figure 8: Projected prism in Figure 5 in viewplane coordinates.

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 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 $(K^TK)^{-1}K^TK = I_3$ we have a left inverse $L = (K^TK)^{-1}K^T$. Here $(K^TK)^{-1}$ exists because $\operatorname{rank}(K) = 3$. Since p' = Kp'', we see that

$$L\mathbf{p}' = (K^T K)^{-1} K^T (K\mathbf{p}'')$$

= $(K^T K)^{-1} (K^T K) \mathbf{p}''$
= \mathbf{p}'' .

Hence the viewplane coordinate mapping is given by the matrix

$$V = L = (K^T K)^{-1} K^T. (2)$$

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 $\mathbf{v}=(1,5,3)^T$. 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 $\hat{\mathbf{r}}=(\frac{3}{5},\frac{4}{5},0)^T$, and $\hat{\mathbf{s}}=(-\frac{4}{5},\frac{3}{5},0)^T$. 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:

$$K^{T}K = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} & 1 \\ \frac{4}{5} & \frac{3}{5} & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \frac{11}{5} \\ 0 & 1 & \frac{2}{5} \\ \frac{11}{5} & \frac{2}{5} & 6 \end{pmatrix},$$

$$(K^{T}K)^{-1} = \begin{pmatrix} \frac{146}{25} & \frac{22}{25} & -\frac{11}{5} \\ \frac{22}{25} & \frac{29}{25} & -\frac{2}{5} \\ -\frac{11}{5} & -\frac{2}{5} & 1 \end{pmatrix},$$

$$V = \begin{pmatrix} \frac{146}{25} & \frac{22}{25} & -\frac{11}{5} \\ \frac{22}{25} & \frac{29}{25} & -\frac{2}{5} \\ -\frac{11}{5} & -\frac{2}{5} & 1 \end{pmatrix} \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 & -\frac{11}{5} \\ -\frac{11}{5} & -\frac{2}{5} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 & -\frac{11}{5} \\ -\frac{4}{5} & \frac{3}{5} & 0 & -\frac{2}{5} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

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:

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

Hence the Cartesian viewplane coordinates of the vertices are $\binom{-11/5}{-2/5}$, $\binom{-1}{-2}$, $\binom{7/5}{-1/5}$, $\binom{1/5}{7/5}$, $\binom{-6/5}{-9/10}$, $\binom{-12/5}{-9/5}$.

References

[1] D. Marsh. Applied Geometry for Computer Graphics and CAD. Springer-Verlag, 1999.