Perspective Projection
Problem Set 1 Solutions

Computer Vision 2021
University of Bern

1 Image projections

1. Show that the perspective projection takes straight lines in 3D space and maps them to 2D straight lines on the image plane.

Solution A 3D line can be written as

$$P(t) = P_0 + tV \tag{1}$$

where $P_0, V \in \mathbf{R}^3$ are a 3D point on the line and the line direction respectively. $t \in \mathbf{R}$ is a scalar parameter that is used to generate all other 3D points on the line. Let us write the 3D coordinates explicitly, *i.e.*, $P(t) = [X(t)\ Y(t)\ Z(t)]^T$, $P_0 = [X_0\ Y_0\ Z_0]^T$ and $V = [X_v\ Y_v\ Z_v]^T$. The perspective projection of $P(t)$ is then

$$p(t) = [x(t)\ y(t)]^T = \frac{[X(t)\ Y(t)]}{Z(t)} = \frac{[X_0 + tX_v\ Y_0 + tY_v]^T}{Z_0 + tZ_v}. \tag{2}$$

To verify linearity we need to reparametrize $p(t)$ in the form $p(s) = p_0 + sv$ where $p_0, v \in \mathbf{R}^2$ and $s \in \mathbf{R}$. Let us define $p_0 = \frac{[X_0\ Y_0]^T}{Z_0}$ (obtained when $t = 0$), $v = [X_vZ_0 - X_0Z_v\ Y_vZ_0 - Y_0Z_v]^T$ and $s = \frac{t}{Z_0(Z_0+tZ_v)}$. Then, we have the following equality

$$p(s) = \frac{[X_0\ Y_0]^T}{Z_0} + s[X_vZ_0 - X_0Z_v\ Y_vZ_0 - Y_0Z_v]^T \tag{3}$$

which shows that $p(s)$ is also a line. Notice the nonlinear relationship between 3D points moving along the line and the corresponding movement of the 2D points on the line on the plane. Also, as $t \mapsto \infty$ we have that $s \mapsto \frac{1}{Z_0Z_v}$ which is a finite quantity.

2. Show that perspective projection preserves incidence in 3D space also onto the image plane.

   Solution The incidence of 2 lines in 3D space can be written as the solution of a system of 2 linear equations in 3D, *i.e.*,

$$P(t) = P_0 + tV \tag{4}$$
$$Q(n) = Q_0 + nW. \tag{5}$$

   If these two lines intercept in 3D space, then $P(t) = Q(n)$ for some $t = t^*$ and $n = n^*$. The projections $p(s)$ and $q(i)$ of $P(t^*) = Q(n^*)$ will also be identical because they are based on the same 3D point. Also, the projections form a system of 2 linear equations in 2D space because of the previous proof. Thus, they also represent the intersection of 2 straight lines.

3. Show that the opposite is not true in general: That is, if two lines intercept in 2D space then they may not do so in 3D space.

   Solution It suffices to show that there exist two non-intersecting lines in 3D space are intersecting in 2D space. Take for example (show drawing of lines)

$$P(t) = [0\ 0\ 1]^T + t[1\ 0\ 0]^T \tag{6}$$
$$Q(n) = [0\ 0\ 2]^T + n[0\ 1\ 0]^T. \tag{7}$$

   The projections are

$$p(s) = s[1\ 0]^T \tag{8}$$
$$q(i) = i[0\ 2]^T \tag{9}$$

   and they intersect at the origin on the plane $[0\ 0]^T$.

   Another important example is that of lines that are parallel in 3D and intercept on the image plane. In this case

$$P(t) = P_0 + tV \tag{10}$$
$$Q(n) = Q_0 + nV \tag{11}$$

   and $P_0 \neq Q_0$. Then, we have

$$p(s) = \frac{[X_{P0}\ Y_{P0}]^T}{Z_{P0}} + s[X_v Z_{P0} - X_{P0} Z_v\ \ Y_v Z_{P0} - Y_{P0} Z_v]^T \tag{12}$$
$$q(i) = \frac{[X_{Q0}\ Y_{Q0}]^T}{Z_{Q0}} + i[X_v Z_{Q0} - X_{Q0} Z_v\ \ Y_v Z_{Q0} - Y_{Q0} Z_v]^T. \tag{13}$$

Since these are 2 linear equations in 2 unknowns (s and i)

$$0 = \frac{1}{Z_{P0}} \begin{bmatrix} X_{P0} \\ Y_{P0} \end{bmatrix} - \frac{1}{Z_{Q0}} \begin{bmatrix} X_{Q0} \\ Y_{Q0} \end{bmatrix} + \begin{bmatrix} X_v Z_{P0} - X_{P0} Z_v & X_v Z_{Q0} - X_{Q0} Z_v \\ Y_v Z_{P0} - Y_{P0} Z_v & Y_v Z_{Q0} - Y_{Q0} Z_v \end{bmatrix} \begin{bmatrix} s \\ i \end{bmatrix} \tag{14}$$

then a solution is always possible (that is, the intercept exists) unless the 2×2 matrix is rank deficient (not invertible). Rank deficiency is achieved when the rows (or the columns) are linearly dependent. Without loss of generality we can pick $P_0 = [0\ 0\ 1]^T$ (translations do not change parallelism of 3D lines). Then, we obtain

$$0 = -\frac{1}{Z_{Q0}} \begin{bmatrix} X_{Q0} \\ Y_{Q0} \end{bmatrix} + \begin{bmatrix} X_v & X_v Z_{Q0} - X_{Q0} Z_v \\ Y_v & Y_v Z_{Q0} - Y_{Q0} Z_v \end{bmatrix} \begin{bmatrix} s \\ i \end{bmatrix} \tag{15}$$

so that rank deficiency is achieved when

$$\begin{bmatrix} X_v \\ Y_v \end{bmatrix} = s \begin{bmatrix} X_{Q0} \\ Y_{Q0} \end{bmatrix}. \tag{16}$$

4. Show that angles are not preserved in the perspective projection.

Solution As in the previous case, we could pick 1 example where the claim is correct and be done. The simplest example is one where the 2 lines in 3D space lie on a plane orthogonal to the image plane. Then, the projection of this plane onto the image plane is a 2D line. Thus any angle between two 3D lines in this plane becomes zero on the image plane.

To better understand the process however, we also examine a more general solution.

Two intersecting lines form an angle θ between them. We can write them as

$$P(t) = P_0 + tV \tag{17}$$

$$Q(n) = P_0 + nW \tag{18}$$

so that $< V, W >= |V||W| \cos\theta$, and let us choose $P_0 = [0\ \ 0\ \ 1]$ for simplicity (we can always rotate and translate the 3D space –this does not change angles and lengths in 3D space– such that this is the case). Let us also choose unitary vectors V and W (*i.e.*, such that $|V| = |W| = 1$). Then, $< V, W >= X_v X_w + Y_v Y_w + Z_v Z_w = \cos\theta$, that is, the inner product of the two directions corresponds to the angle between the lines. We repeat the same procedure for the corresponding 2D projections on the image plane. Then, we have

$$p(s) = p_0 + sv \tag{19}$$

$$q(i) = p_0 + iw \tag{20}$$

where $p_0 = [0\ 0]^T$ and $v = [X_v Z_0 - X_0 Z_v\ Y_v Z_0 - Y_0 Z_v]^T = [X_v\ Y_v]^T$ and $w = [X_w Z_0 - X_0 Z_w\ Y_w Z_0 - Y_0 Z_w]^T = [X_w\ Y_w]^T$. The inner product between v and w becomes

$$< v, w >= X_v X_w + Y_v Y_w = \cos\theta - Z_v Z_w \tag{21}$$

so that it is clear that the inner product on the image plane depends on the Z coordinates of the directions of the lines in 3D space.

5. Show that lengths are not preserved in the perspective projection.

Solution The solution follows a path similar to the previous ones.

One example is sufficient to prove the claim. Take the line $P(t) = [0\ 0\ 1]^T + t[0\ 0\ 1]^T$. This projects to the 2D point $p = [0\ 0]^T$ for all t. So any segment of $P(t)$ projects to a segment of length 0 on the image plane.

More in general, we can ignore lines and just consider 3D points P and Q and their distance $|P - Q|$. One point can always be fixed to some arbitrary location (except for the origin, where the projection is undefined) without loss of generality (distance is invariant to rotations and translations). Let us choose $P = [0\ 0\ 1]^T$. Then we have

$$|P - Q| = \sqrt{X_q^2 + Y_q^2 + (Z_q - 1)^2}. \tag{22}$$

Since $p = [0\ 0]$, the corresponding distance on the image plane is

$$|p - q| = |q| = \frac{\sqrt{X_q^2 + Y_q^2}}{|Z_q|} \neq \sqrt{X_q^2 + Y_q^2 + (Z_q - 1)^2} = |P - Q|. \tag{23}$$

6. Determine the 3D plane that corresponds to the horizon.

Solution The horizon is defined as the set of points at the same height as the camera, that is, with coordinate $Y = 0$. Then, the 3D points with this constraint are all points

$$P = \begin{bmatrix} X \\ 0 \\ Z \end{bmatrix} \tag{24}$$

for arbitrary coordinates X and Z. The horizon plane is also spanned by the vectors $[1\ 0\ 0]^T$ and $[0\ 0\ 1]^T$.

7. Determine the projection of an opaque sphere onto the image plane.

Solution Let us consider all points on the surface of the sphere. Each point is projected to the camera center via a straight line. The 3D points whose lines are tangential to the sphere define the boundary of the projection. These lines and points do not change as we rotate the sphere around its center (a sphere does not change its shape due to rotation); similarly, any rotation of the camera center with respect to the sphere center will not change these lines and points (with respect to the camera center). By symmetry, the 3D points whose lines are tangential to the sphere, form a circle. Thus, this circle and all the lines going through the camera center form a cone. The projection on the image plane of these lines is then obtainable as the section of a cone. We know that cone sections can be one of the following: circle, ellipse, hyperbole, parabola.

Let us define S as the sphere center, the camera center is the origin, ρ is the sphere radius and N is the normal to the image plane (at a distance 1 from the origin). Some useful calculations:

$$|S| = \text{distance sphere center to camera center} \tag{25}$$

$$\theta = \text{half angle of the cone} \tag{26}$$

$$\rho = |S| \sin\theta \tag{27}$$

$$L = |S|\cos\theta = \sqrt{|S|^2 - \rho^2} \quad \text{length of cone side} \tag{28}$$

$$H = L\cos\theta = \frac{L^2}{|S|} = |S|\left(1 - \frac{\rho^2}{|S|^2}\right) \quad \text{cone height} \tag{29}$$

$$D = L\sin\theta = L\frac{\rho}{|S|} = \rho\sqrt{1 - \frac{\rho^2}{|S|^2}} \quad \text{radius of cone base.} \tag{30}$$

The last quantities H and D are all we need to define the cone. Let us rotate the cone and the image plane around the camera center until the base of the cone lies on the $X - Y$ plane. 3D points at the surface of the cone can be written as

$$P(t,\phi) = tD \begin{bmatrix} \cos\phi \\ \sin\phi \\ H/D \end{bmatrix} \tag{31}$$

where ϕ determines a point on the base circle and t the normalized distance from the cone tip ($t = 1$ reaches the base and $t = 0$ is at the tip). The rotated image plane will have normal $\tilde{N}$. 3D points Q lying on the image plane can be written then as

$$0 = < Q - \tilde{N}, \tilde{N} > = Q^T\tilde{N} - 1. \tag{32}$$

To find the intersection between the cone and the image plane we put the 2 equations in a system and obtain

$$\tilde{N}^T \begin{bmatrix} tD\cos\phi \\ tD\sin\phi \\ tH \end{bmatrix} = 1 \tag{33}$$

or, equivalently,

$$X_{\tilde{N}} t \cos\phi + Y_{\tilde{N}} t \sin\phi + Z_{\tilde{N}} t H/D = 1. \tag{34}$$

We can write $t = \frac{1}{X_{\tilde{N}} \cos\phi + Y_{\tilde{N}} \sin\phi + Z_{\tilde{N}} H/D}$. We can substitute in $P(t, \phi)$ and obtain all 3D points at the intersection as

$$P(\phi) = \frac{D}{X_{\tilde{N}} D \cos\phi + Y_{\tilde{N}} D \sin\phi + Z_{\tilde{N}} H} \begin{bmatrix} D\cos\phi \\ D\sin\phi \\ H \end{bmatrix}. \tag{35}$$

The coordinates of $P(\phi) = [X \ Y \ Z]^T$ are then related by

$$X^2 + Y^2 = \frac{D^2}{H^2} Z^2 \tag{36}$$

$$X_{\tilde{N}} X + Y_{\tilde{N}} Y + Z_{\tilde{N}} Z = 1. \tag{37}$$

If $Z_{\tilde{N}} \neq 0$ we can obtain

$$Z = \frac{1}{Z_{\tilde{N}}} \left(1 - X_{\tilde{N}} X - Y_{\tilde{N}} Y\right) \tag{38}$$

and substitute

$$X^2 + Y^2 - \frac{D^2}{H^2 Z_{\tilde{N}}^2} \left(1 - X_{\tilde{N}} X - Y_{\tilde{N}} Y\right)^2 = 0 \tag{39}$$

and finally

$$\left(\frac{H^2 Z_{\tilde{N}}^2}{D^2} - X_{\tilde{N}}^2\right) X^2 + \left(\frac{H^2 Z_{\tilde{N}}^2}{D^2} - Y_{\tilde{N}}^2\right) Y^2 - 2X_{\tilde{N}} Y_{\tilde{N}} XY + 2X_{\tilde{N}} X + 2Y_{\tilde{N}} Y - 1 = 0. \tag{40}$$

This last equation represents a quadric, that is, any among circle, ellipse, hyperbola, and parabola.

8. What is the orthographic projection of an opaque sphere?

Solution The orthographic projection consists of simply taking the first two coordinates of a 3D point. Geometrically it amounts to projecting 3D points along the normal of the image plane. In the case of a sphere the projection will therefore always be a circle whose radius is the radius of the sphere.

The 3D points on the sphere that project on the boundary of the image plane projection can be found by imposing that the normal vector going through them must be tangential to the sphere. The calculations are analogous to those presented above in the case of perspective projection.