Perspective projections

1 Coordinate systems

The analysis of perspective projections involves three systems of coordinates:

- The 3D world coordinates, denoted by (X, Y, Z).
- The camera coordinates, denoted by (x, y).
- The image coordinates, denoted by (u, v).

1.1 World coordinates (X, Y, Z)

The camera is positioned at the point $(0, 0, 0)$ in world coordinates, and it is pointing at the positive Z direction. The image plane is parallel to the (X, Y) plane, at distance f from the camera. The constant f is called **the focal length**. Thus, the point $(0, 0, f)$ is the projection of the camera on the image plane.

1.2 Camera coordinates (x, y)

Coordinates measured *on the image plane* with $(0, 0, f)$ as the origin are called camera coordinates, and we denote them by (x, y). (The point (x, y) on the image plane is the point (x, y, f) in world coordinates.) A point (X, Y, Z) in world coordinates in projected on the point $(x, y))$ in camera coordinates, satisfying:

$$x = f\frac{X}{Z}, \qquad y = f\frac{Y}{Z}$$

1.3 Image coordinates (u, v)

Point location in the image, with respect to an *arbitrarily chosen* origin is denoted by (u, v). For example, the most common choice is to place the origin at the upper left corner of the image. The projection of the camera on the image plane, the point $(x = 0, y = 0)$ in the camera coordinates, is denoted by the point (u_0, v_0) in the image coordinate system. The point $(u = u_0, v = v_0)$ is called **the principal point**. This gives the following relation between camera coordinates and image coordinates:

$$x = u - u_0, \qquad y = v - v_0$$

The relation between world coordinates and image coordinates is given by:

$$u = u_0 + f\frac{X}{Z}, \qquad v = v_0 + f\frac{Y}{Z}$$

Observe that this mapping involves three constant values: f, u_0, v_0. These are called **the camera calibration parameters**.

2 Parametric representation of lines

2.1 2D

The simplest representation of a line in 2D is:

$$y = ax + b$$

But this representation is not symmetric in x and y, and cannot be used to represent vertical lines. A more general form is a parametric representation of the form:

$$x = a_1 t + b_1, \qquad y = a_2 t + b_2$$

In vector notation this can be written as:

$$\begin{pmatrix} x \\ y \end{pmatrix} = At + B, \qquad \text{where} \quad A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

where A is a vector in the direction of the line. Notice that the value of A is not uniquely determined by the line. For example, the line $y = 1 - x$ can be written as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} t + \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix} t + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Since the multiplication of A by a scalar doesn't change the line, we can normalize A by assuming with out loss of generality that as a vector it has a size of 1. This means that:

$$a_1^2 + a_2^2 = 1$$

Therefore, all lines in the same direction (parallel lines) have the same normalized A but may have different B.

The lines $\begin{pmatrix} x \\ y \end{pmatrix} = A_1 t + B_1$ and $\begin{pmatrix} x \\ y \end{pmatrix} = A_2 t + B_2$ are orthogonal if the dot product of A_1, A_2 is zero. With $A_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ and and $A_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ the orthogonality of these lines can be expressed as:

$$a_{11}a12 + a_{21}a_{22} = 0$$

2.2 3D

The exact same relations hold for 3D lines. The parametric representation is:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = At + B, \qquad \text{where} \quad A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

where A is a vector in the direction of the line. The value of A is the direction of the line in 3D, and it is determined uniquely with the additional condition:

$$a_1^2 + a_2^2 + a_3^2 = 1$$

The lines in the directions A_1, A_2 are orthogonal if:

$$a_{11}a_{12} + a_{21}a_{22} + a_{31}a_{32} = 0$$

3 Vanishing points

It is very easy to compute the 3D direction of 3D lines from the projection of two or more parallel 3D lines. We begin by showing that all projections of parallel 3D lines meet at a single point in 2D, called *the vanishing point*. Given a 3D line

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = At + B, \qquad \text{where} \quad A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

it has the following projection in 2D:

$$x = f\frac{a_1 t + b_1}{a_3 t + b_3}, \quad y = f\frac{a_2 t + b_2}{a_3 t + b_3}$$

The vanishing point is the value of (x, y) when $t \to \infty$. It is given by:

$$x_\infty = f\frac{a_1}{a_3}, \quad y_\infty = f\frac{a_2}{a_3}$$

Since these values are independent of B, they are the same for all lines sharing the same value of A (all lines parallel to the given line). In the special case where $a_3 = 0$ the projections of the lines do not meet in the finite plane. In this special case we say that the vanishing point is at infinity.

To compute the 3D direction (the value of A) from the vanishing point (x_∞, y_∞) we need to solve the following 3 equations with 3 unknowns:

$$x_\infty = f\frac{a_1}{a_3}, \quad y_\infty = f\frac{a_2}{a_3}, \quad a_1^2 + a_2^2 + a_3^2 = 1$$

The solution is:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \frac{1}{\sqrt{x_\infty^2 + y_\infty^2 + f^2}} \begin{pmatrix} x_\infty \\ y_\infty \\ f \end{pmatrix}$$

This solution uses the camera coordinates. To compute an answer in terms of the image coordinates we need all three camera calibration parameters u_0, v_0, f:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \frac{1}{\sqrt{(u_\infty - u_0)^2 + (v_\infty - v_0)^2 + f^2}} \begin{pmatrix} u_\infty - u_0 \\ v_\infty - v_0 \\ f \end{pmatrix}$$

4 Vanishing points of orthogonal lines

Given the projections of two sets of parallel 3D lines that are known to be orthogonal in 3D we can compute two vanishing points in image coordinates: (u_1, v_1) for the first set and (u_2, v_2) for the second set. Therefore, up to a constant the 3D direction of the first set is $\begin{pmatrix} u_1 - u_0 \\ v_1 - v_0 \\ f \end{pmatrix}$, and the 3D direction of the second set is $\begin{pmatrix} u_2 - u_0 \\ v_2 - v_0 \\ f \end{pmatrix}$ The orthogonality implies that:

$$(u_1 - u_0)(u_2 - u_0) + (v_1 - v_0)(v_2 - v_0) + f^2 = 0$$

This can be used to compte one of the camera calibration parameters from the other two.

5 Cross Ratio

3D world coordinates: X, Y, Z. Picture coordinates: x, y.

Let p_1, p_2, p_3, p_4 be 4 points on the 3D line:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} t + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Without loss of generality p_i is the point for which $t = t_i$ $(i = 1, 2, 3, 4)$. Let q_i be the 2D projection of p_i.

$$q_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \quad x_i = f\frac{a_1 t_i + b_1}{Z_i}, \quad y_i = f\frac{a_2 t_i + b_2}{Z_i}, \quad \text{where: } Z_i = a_3 t_i + b_3$$

The cross ratio of the four 2D points is defined as:

$$\mathrm{CrossRatio}(q_1, q_2, q_3, q_4) = \frac{|q_1 - q_2| \cdot |q_3 - q_4|}{|q_1 - q_3| \cdot |q_2 - q_4|}$$

where $|q_i - q_j|$ is the 2D distance between q_i and q_j.

Theorem: The Cross Ratio of four collinear points is invariant to the perspective projection. (In other words, the cross ratio gives a value that depends on the relative location of the points in 3D. This value would be the same for all perspective projections.)

Proof: $|q_i - q_j|^2 = (x_i - x_j)^2 + (y_i - y_j)^2$. We begin by computing $x_i - x_j$.

$$x_i - x_j = f\frac{(a_1 t_i + b_1)(a_3 t_j + b_3) - (a_1 t_j + b_1)(a_3 t_i + b_3)}{(a_3 t_i + b_3)(a_3 t_j + b_3)} = f\frac{(a_1 b_3 - a_3 b_1)(t_i - t_j)}{Z_i Z_j}$$

Similarly:

$$y_i - y_j = f\frac{(a_2 b_3 - a_3 b_2)(t_i - t_j)}{Z_i Z_j}$$

Define: $K_1 = f(a_1 b_3 - a_3 b_1)$, $K_2 = f(a_2 b_3 - a_3 b_2)$. Notice that K_1, K_2 depend on the line parametrization and the perspective projection parameters, but *not* on the points. We have:

$$(x_i - x_j)^2 + (y_i - y_j)^2 = \frac{(K_1^2 + K_2^2)(t_i - t_j)^2}{Z_i^2 Z_j^2}$$

and by putting $K = \sqrt{K_1^2 + K_2^2}$:

$$|q_i - q_j| = K\frac{|t_i - t_j|}{|Z_i||Z_j|}$$

Therefore,

$$\begin{aligned} |q_1 - q_2||q_3 - q_4| &= K^2\frac{|t_1 - t_2||t_3 - t_4|}{|Z_1||Z_2||Z_3||Z_4|} \\ |q_1 - q_3||q_2 - q_4| &= K^2\frac{|t_1 - t_3||t_2 - t_4|}{|Z_1||Z_3||Z_2||Z_4|} \end{aligned}$$

So that:

$$\mathrm{CrossRatio}(q_1, q_2, q_3, q_4) = \frac{|t_1 - t_2| \cdot |t_3 - t_4|}{|t_1 - t_3| \cdot |t_2 - t_4|}$$

Clearly, this expression is independent of the perspective projection or the line parametrization.

6 Summary of formulas

3D (world coordinates)	2D (image coordinates)	relation
a point $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$	a point $\begin{pmatrix} u \\ v \end{pmatrix}$	$u = u_0 + f\frac{X}{Z}$, $v = v_0 + f\frac{Y}{Z}$
a line	a line	
parallel lines in the direction $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$	lines meet at vanishing point $\begin{pmatrix} u_\infty \\ v_\infty \end{pmatrix}$	$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = K \begin{pmatrix} u_\infty - u_0 \\ v_\infty - v_0 \\ f \end{pmatrix}$ $K = \frac{1}{\sqrt{(u_\infty - u_0)^2 + (v_\infty - v_0)^2 + f^2}}$
two sets of parallel lines in directions $\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}$, $\begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}$ these directions are orthogonal in 3D	two vanishing points $\begin{pmatrix} u_1 \\ v_1 \end{pmatrix}$ and $\begin{pmatrix} u_2 \\ v_2 \end{pmatrix}$	$(u_1 - u_0)(u_2 - u_0) + (v_1 - v_0)(v_2 - v_0) + f^2 = 0$
p_1, p_2, p_3, p_4 four points on a 3D line	q_1, q_2, q_3, q_4 four points on a 2D line	$\frac{\lvert q_1 - q_2 \rvert \cdot \lvert q_3 - q_4 \rvert}{\lvert q_1 - q_3 \rvert \cdot \lvert q_2 - q_4 \rvert} = \frac{\lvert p_1 - p_2 \rvert \cdot \lvert p_3 - p_4 \rvert}{\lvert p_1 - p_3 \rvert \cdot \lvert p_2 - p_4 \rvert}$