The Geometry of Perspective Projection

- Pinhole camera and perspective projection

  - This is the simplest imaging device which, however, captures accurately the geometry of perspective projection.
  
  - Rays of light enters the camera through an infinitesimally small aperture.
  
  - The intersection of the light rays with the image plane form the image of the object.
  
  - Such a mapping from three dimensions onto two dimensions is called *perspective projection*. 
- A simplified geometric arrangement

- In general, the world and camera coordinate systems are not aligned.

- To simplify the derivation of the perspective projection equations, we will make the following assumptions:

  (1) the center of projection coincides with the origin of the world.

  (2) the camera axis (optical axis) is aligned with the world’s z-axis.
(3) avoid image inversion by assuming that the image plane is in front of the center of projection.

- **Some terminology**

  - The model consists of a plane (image plane) and a 3D point \( O \) (**center of projection**).

  - The distance \( f \) between the image plane and the center of projection \( O \) is the **focal length** (e.g., the distance between the lens and the CCD array).

  - The line through \( O \) and perpendicular to the image plane is the **optical axis**.

  - The intersection of the optical axis with the image plane is called **principal point** or **image center**.

  (note: the principal point is not always the "actual" center of the image)
• The equations of perspective projection

(notation: \((x, y, z) \rightarrow (X, Y, Z), r \rightarrow R, (x', y', z') \rightarrow (x, y, z), r' \rightarrow r)\)

- Using the following similar triangles:

1. from \(OA'B'\) and \(OAB\):

\[
\frac{f}{Z} = \frac{r}{R}
\]

2. from \(A'B'C'\) and \(ABC\):

\[
\frac{x}{X} = \frac{y}{Y} = \frac{r'}{R}
\]

**perspective proj. eqs:**

\[
x = \frac{Xf}{Z}, \quad y = \frac{Yf}{Z}, \quad z = f
\]

- Using matrix notation:

\[
\begin{bmatrix}
x_h \\
y_h \\
z_h \\
w
\end{bmatrix}
= \begin{bmatrix}
f & 0 & 0 & 0 \\
0 & f & 0 & 0 \\
0 & 0 & f & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]

- Verify the correctness of the above matrix (homogenize using \(w = Z\)):

\[
x = \frac{x_h}{w} = \frac{fX}{Z}, \quad y = \frac{y_h}{w} = \frac{fY}{Z}, \quad z = \frac{z_h}{w} = f
\]
• Properties of perspective projection

Many-to-one mapping

- The projection of a point is not unique (any point on the line $OP$ has the same projection).

Scaling/Foreshortening

- The distance to an object is inversely proportional to its image size.
- When a line (or surface) is parallel to the image plane, the effect of perspective projection is *scaling*.

- When a line (or surface) is not parallel to the image plane, we use the term *foreshortening* to describe the projective distortion (i.e., the dimension parallel to the optical axis is compressed relative to the frontal dimension).

![Effect of focal length](image)

*Effect of focal length*

- As \( f \) gets smaller, more points project onto the image plane (*wide-angle camera*).

- As \( f \) gets larger, the field of view becomes smaller (more *telescopic*).

*Lines, distances, angles*

- Lines in 3D project to lines in 2D.

- Distances and angles are *not* preserved.

- Parallel lines *do not* in general project to parallel lines (unless they are parallel to the image plane).
Vanishing point

* parallel lines in space project perspectively onto lines that on extension intersect at a single point in the image plane called *vanishing point* or *point at infinity*.

* (alternative definition) the vanishing point of a line depends on the orientation of the line and not on the position of the line.

* the vanishing point of any given line in space is located at the point in the image where a parallel line through the center of projection intersects the image plane.

Vanishing line

* the vanishing points of all the lines that lie on the same plane form the *vanishing line*.

* also defined by the intersection of a parallel plane through the center of projection with the image plane.
Orthographic Projection

- It is the projection of a 3D object onto a plane by a set of parallel rays orthogonal to the image plane.

- It is the limit of perspective projection as $f \rightarrow \infty$ (i.e., $f/Z \rightarrow 1$)

orthographic proj. eqs: $x = X$, $y = Y$ (drop $Z$)

- Using matrix notation:

$$
\begin{bmatrix}
  x_h \\
y_h \\
z_h \\
w
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z \\
  1
\end{bmatrix}
$$

- Verify the correctness of the above matrix (homogenize using $w=1$):

$$
x = \frac{x_h}{w} = X \quad y = \frac{y_h}{w} = Y
$$

• Properties of orthographic projection

- Parallel lines project to parallel lines.

- Size does not change with distance from the camera.
**Weak Perspective Projection**

- Perspective projection is a non-linear transformation.

- We can approximate perspective by scaled orthographic projection (i.e., linear transformation) if:

  1. the object lies close to the optical axis.

  2. the object’s dimensions are small compared to its average distance $\bar{Z}$ from the camera (i.e., $\delta z < \bar{Z}/20$)

  \[
  \text{weak perspective proj. eqs:} \quad x = \frac{Xf}{Z} \approx \frac{Xf}{\bar{Z}} \quad y = \frac{Yf}{Z} \approx \frac{Yf}{\bar{Z}} \quad \text{drop } Z
  \]

  - The term $\frac{f}{\bar{Z}}$ is a scale factor now (e.g., every point is scaled by the same factor).

  - Using matrix notation:

    \[
    \begin{bmatrix}
    x_h \\
    y_h \\
    z_h \\
    w
    \end{bmatrix} =
    \begin{bmatrix}
    f & 0 & 0 & 0 \\
    0 & f & 0 & 0 \\
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & \bar{Z}
    \end{bmatrix}
    \begin{bmatrix}
    X \\
    Y \\
    Z \\
    \bar{Z}
    \end{bmatrix}
    \]

  - Verify the correctness of the above matrix (homogenize using $w = \bar{Z}$):

    \[
    x = \frac{x_h}{w} = \frac{fX}{\bar{Z}} \quad y = \frac{y_h}{w} = \frac{fY}{\bar{Z}}
    \]