1 几何变换

1.1 几何变换基础

定理 1.1 平移 平移对应的数学操作为向量加法. 设物体上任意一点 P 的坐标为 $\boldsymbol{p}$, 平移向量为 $\boldsymbol{t}$, 则平移后的点 P' 的坐标 $\boldsymbol{p}' = \boldsymbol{p} + \boldsymbol{t}$. 例如, 对于三维空间而言有

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ z + t_z \end{bmatrix}$$

定理 1.2 旋转 旋转对应的数学操作为矩阵乘法. 设物体上任意一点 P 的坐标为 $\boldsymbol{p}$, 旋转矩阵为 $\boldsymbol{R}$, 则旋转后的点 P' 的坐标 $\boldsymbol{p}' = \boldsymbol{R}\boldsymbol{p}$. 在二维平面中, 绕原点逆时针旋转 θ 角度的旋转矩阵为

$$\boldsymbol{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

在三维空间中, 绕 x, y, z 轴分别旋转 α, β, γ 角度的旋转矩阵分别为

$$\boldsymbol{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}, \quad \boldsymbol{R}_y = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}, \quad \boldsymbol{R}_z = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

定理 1.3 绕任意轴的旋转 在三维空间中, 绕单位向量 $\boldsymbol{a} = (a_x, a_y, a_z)$ 定义的轴逆时针旋转 θ 角度的旋转矩阵 $\boldsymbol{R}$ 为

$$\boldsymbol{R} = \cos\theta \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (1 - \cos\theta) \begin{bmatrix} a_x^2 & a_x a_y & a_x a_z \\ a_x a_y & a_y^2 & a_y a_z \\ a_x a_z & a_y a_z & a_z^2 \end{bmatrix} + \sin\theta \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$$

证明. 考虑三维空间中的点 $\boldsymbol{v}$, 旋转轴 $\boldsymbol{a}$ 和旋转角度 θ. 从 $\boldsymbol{v}$ 旋转至 $\boldsymbol{v}'$ 的操作可以分解为平行于 $\boldsymbol{a}$ 的分量 $\boldsymbol{r}$ 和垂直于 $\boldsymbol{a}$ 的分量 $\boldsymbol{u}$.

根据投影向量的定义, $\boldsymbol{r} = (\boldsymbol{v} \cdot \boldsymbol{a})\boldsymbol{a}$, $\boldsymbol{u} = \boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{a})\boldsymbol{a}$. 旋转过程中, $\boldsymbol{r}$ 保持不变, 而 $\boldsymbol{u}$ 则绕 $\boldsymbol{a}$ 旋转 θ 角度, 这可以由下面的方式得到

$$\boldsymbol{u}' = \boldsymbol{u}\cos\theta + \boldsymbol{u}^{\perp}\sin\theta$$

注意到 $\boldsymbol{u}^{\perp} = \boldsymbol{a} \times \boldsymbol{v}$, 于是可得

$$\begin{aligned}\boldsymbol{v}' &= [(\boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{a})\boldsymbol{a}) \cos\theta] + (\boldsymbol{a} \times \boldsymbol{v}) \sin\theta + (\boldsymbol{v} \cdot \boldsymbol{a})\boldsymbol{a} \\ &= \boldsymbol{v} \cos\theta + (1 - \cos\theta)(\boldsymbol{v} \cdot \boldsymbol{a})\boldsymbol{a} + (\boldsymbol{a} \times \boldsymbol{v}) \sin\theta\end{aligned}$$

定义 $\boldsymbol{a}$ 的反对称矩阵为

$$\boldsymbol{a}^* = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$$

不难根据矩阵乘法和向量叉乘的定义得出 $\boldsymbol{a} \times \boldsymbol{v} = \boldsymbol{a}^* \boldsymbol{v}$.
根据点积的定义有

$$(\boldsymbol{a} \cdot \boldsymbol{x})\boldsymbol{a} = \left(\boldsymbol{a}^{\mathrm{t}} \boldsymbol{x}\right) \boldsymbol{a} = \boldsymbol{a} \left(\boldsymbol{a}^{\mathrm{t}} \boldsymbol{x}\right) = \left(\boldsymbol{a}\boldsymbol{a}^{\mathrm{t}}\right) \boldsymbol{x}$$

于是将上述 $\boldsymbol{v}'$ 的式子写作矩阵乘法即有

$$\boldsymbol{v}' = \left[\boldsymbol{I} \cos\theta + \boldsymbol{a}\boldsymbol{a}^{\mathrm{t}}(1 - \cos\theta) + \boldsymbol{a}^* \sin\theta\right] \boldsymbol{v}$$

于是旋转矩阵即为

$$\boldsymbol{R} = \boldsymbol{I} \cos\theta + \boldsymbol{a}\boldsymbol{a}^{\mathrm{t}}(1 - \cos\theta) + \boldsymbol{a}^* \sin\theta$$

□

1.2 齐次坐标系及其中的几何变换

我们在上一节中简单介绍了几何变换对应的数学操作. 例如, 平移可以视作向量的加法, 旋转可以视作旋转矩阵与向量的乘法. 但是, 这些操作并不能统一表示, 因而带来了额外的麻烦. 例如, 平移无法表示为矩阵与向量的乘法. 为了解决这个问题, 我们可以引入齐次坐标系的概念.

1.2.1 齐次坐标系的定义

齐次坐标系的基本思想是, 通过引入额外的维度, 将所有的几何变换都表示为矩阵与向量的乘法.

定义 1.4 齐次坐标系 在 n 维空间 R^n 中, 点 $\boldsymbol{v} = (x_1, \cdots, x_n)$ 的齐次坐标表示为 $(x_1, \cdots, x_n, 1)$, 即在原有坐标的基础上增加一个分量恒为 1 的维度.
相应地, 齐次坐标系中的向量 $\boldsymbol{v} = (x_1, \cdots, x_n)$ 表示为 $(x_1, \cdots, x_n, 0)$, 即在原有坐标的基础上增加一个分量恒为 0 的维度.

如此, 我们可以将所有的几何变换都表示为 $(n+1) \times (n+1)$ 的矩阵与齐次坐标向量的乘法.

1.2.2 齐次坐标系中几何变换的表示方式

这里以三维空间为例介绍齐次坐标系中几何变换的表示方式.

定理 1.5 齐次坐标系中的平移矩阵 平移向量 $\boldsymbol{t} = (t_x, t_y, t_z)$ 的平移矩阵为

$$\boldsymbol{T} = \boldsymbol{I} + \begin{bmatrix} \boldsymbol{t} \\ 0 \end{bmatrix}$$

对于空间中任意一点 $\boldsymbol{p}$, 在齐次坐标系下总有

$$\begin{bmatrix} \boldsymbol{p} + \boldsymbol{t} \\ 1 \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} \boldsymbol{p} \\ 1 \end{bmatrix}$$

证明. 简单地运用矩阵乘法的定义就能得到. □

定理 1.6 齐次坐标系中的旋转矩阵 齐次坐标系中的旋转矩阵为

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}' & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix}$$

其中 $\boldsymbol{R}'$ 为三维空间中的旋转矩阵, 在前面已经推出其具体形式.

于是在齐次坐标系下, 任何平移与旋转变换都可以写作矩阵

$$\begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0} & 1 \end{bmatrix}$$

与齐次坐标向量的乘法形式.

1.3 几何变换在模型渲染中的应用

对于渲染这一过程来说, 将三维模型投影到二维平面上是一个重要的步骤. 观察者的位置和视角决定了投影的结果. 通常, 我们需要观察者的位置 $\boldsymbol{p}$, 观察平面的法向量 $\boldsymbol{n}$ 和视角正上方的方向 $\boldsymbol{v}$ 决定.

1.3.1 坐标变换

在观察者的坐标中,$\boldsymbol{n}$ 通常指向 z 轴正方向,$\boldsymbol{v}$ 指向 y 轴正方向. 而 x 轴方向根据左右手坐标系而确定. 这里假定 x 轴的方向为 $\boldsymbol{u} = \boldsymbol{v} \times \boldsymbol{n}$. 因此, 我们先需要把模型从世界坐标系 $\mathcal{C}_{\text{world}}$ 变换到观察坐标系 $\mathcal{C}_{\text{obs}}$ 上.

二维坐标变换 为了研究坐标系变换问题, 我们从二维坐标系变换开始.

考虑 $\boldsymbol{p}$ 在 uv 坐标和 xy 坐标下的表示分别为

$$(p_u, p_v) = \boldsymbol{o}_{uv} + p_u\boldsymbol{u} + p_v\boldsymbol{v}$$

$$(p_x, p_y) = \boldsymbol{o}_{xy} + p_x\boldsymbol{x} + p_y\boldsymbol{y}$$

其中 $\boldsymbol{o}_{uv}$ 和 $\boldsymbol{o}_{xy}$ 分别为两个坐标系的原点位置. 并且有

$$\boldsymbol{u} = u_x\boldsymbol{x} + u_y\boldsymbol{y}$$

$$\boldsymbol{v} = v_x\boldsymbol{x} + v_y\boldsymbol{y}$$

于是可以列出方程

$$\boldsymbol{o}_{uv} + p_u(u_x\boldsymbol{x} + u_y\boldsymbol{y}) + p_v(v_x\boldsymbol{x} + v_y\boldsymbol{y}) = \boldsymbol{o}_{xy} + p_x\boldsymbol{x} + p_y\boldsymbol{y}$$

即

$$\boldsymbol{o}_{uv} - \boldsymbol{o}_{xy} = (p_x - p_u u_x - p_v v_x)\boldsymbol{x} + (p_y - p_u u_y - p_v v_y)\boldsymbol{y}$$

设 uv 坐标的原点在 xy 坐标系下的表示为 (e_x, e_y), 则有

$$\begin{cases} e_x = p_x - p_u u_x - p_v v_x \\ e_y = p_y - p_u u_y - p_v v_y \end{cases}$$

于是可得

$$\begin{cases} p_x = e_x + p_u u_x + p_v v_x \\ p_y = e_y + p_u u_y + p_v v_y \end{cases}$$

我们发现上述结果恰好对应于下面的几何变换:

$$\begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & e_x \\ 0 & 1 & e_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_x & v_x & 0 \\ u_y & v_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_u \\ p_v \\ 1 \end{bmatrix}$$

需要注意的是, 以矩阵乘法表示的几何变换只能在同一坐标系下才具有几何意义. 因此, 上述等式左右两边的点都应在 xy 坐标系下描述其坐标. 也即, 我们把 xy 坐标系中的点 $(p_u, p_v)_{xy}$ 先旋转, 使得旋转后的 $\boldsymbol{x}$, $\boldsymbol{y}$ 轴分别与 $\boldsymbol{u}$, $\boldsymbol{v}$ 轴对齐, 再平移使得 xy 坐标系的原点与 uv 坐标系的原点重合. 此时,$(p_u, p_v)_{xy}$ 将被变换到 $(p_x, p_y)_{xy}$, 这一点的即 uv 坐标系中的点 $(p_u, p_v)_{uv}$ 的位置. 于是我们就完成了坐标系变换. 上述过程如下图所示.

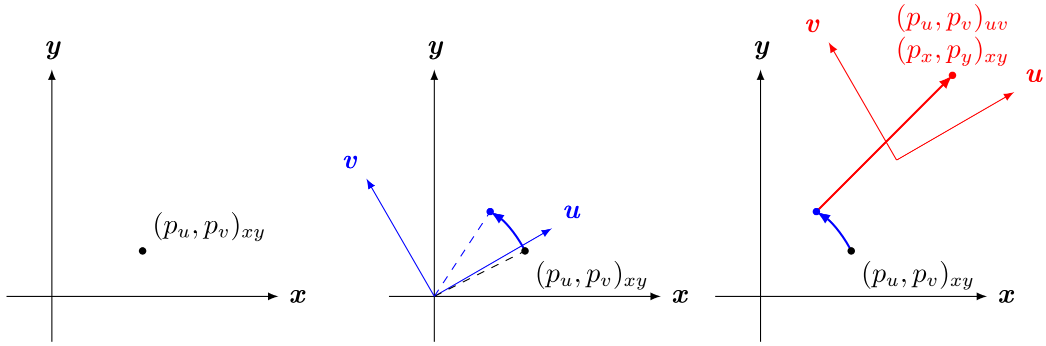

图 1: 二维坐标系变换的示意图

三维坐标变换 同样地, 可以推出三维坐标系的矩阵表示为

$$\begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & e_x \\ 0 & 1 & 0 & e_y \\ 0 & 0 & 1 & e_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_x & v_x & w_x & 0 \\ u_y & v_y & w_y & 0 \\ u_z & v_z & w_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_u \\ p_v \\ p_w \\ 1 \end{bmatrix}$$

1.3.2 投影变换

在解决坐标变换的问题后, 我们需要将物体投影到观察平面上. 这里我们介绍两种常见的投影方式: 正交和透视投影. 在此之前, 需要明确一些投影中的概念.

定义 1.7 投影线 连接对象点和投影点的直线称作**投影线**.

定义 1.8 平行投影 投影线相互平行的投影方式称作**平行投影**

定义 1.9 (投影) 观察体 观察得到的图像对应在三维空间中的区域称作 **(投影) 观察体**.

定义 1.10 裁剪平面 观察体在 z 方向 (也就是投影平面法向) 的边缘通过选取平行于投影平面的两个平面决定, 这两个平面分别称作**近裁剪平面**和**远裁剪平面**, 分别记作 z_{near} 和 z_{far}.

有时, z_{near} 和 z_{far} 也指两个裁剪平面的 z 轴坐标. 这需要依据上下文确定其含义.

正交投影 我们先来介绍比较简单的投影方式.

定义 1.11 正交投影 **正交投影**属于平行投影的一种, 它的投影线全部与投影平面垂直.

不难看出正交投影是保长度的, 因此工程和建筑测绘常用正交投影.

正交投影从观察坐标到观察平面的变换很简单, 任意一点 $\boldsymbol{v}=(x,y,z)$ 的投影点就是 $\boldsymbol{i}=(x,y)$. 通常, 需要将对象描述转化到规范化坐标系, 即建立 $\boldsymbol{v}\to[-1,1]^3$ 的映射 (即将观察体映射到边长为 2, 范围从 $(-1,-1,-1)$ 到 $(1,1,1)$ 的立方体中). 这对应于一个放缩操作, 因此在齐次坐标系下正交投影的变换矩阵为

$$\boldsymbol{M}_{\mathrm{ortho}}=\begin{bmatrix}\frac{2}{x_{\max}-x_{\min}} & 0 & 0 & -\frac{x_{\max}+x_{\min}}{x_{\max}-x_{\min}}\\ 0 & \frac{2}{y_{\max}-y_{\min}} & 0 & -\frac{y_{\max}+y_{\min}}{y_{\max}-y_{\min}}\\ 0 & 0 & \frac{2}{z_{\max}-z_{\min}} & -\frac{z_{\max}+z_{\min}}{z_{\max}-z_{\min}}\\ 0 & 0 & 0 & 1\end{bmatrix}$$

其中 $x_{\max},x_{\min},y_{\max},y_{\min}$ 分别为观察体在 x 和 y 方向上的上下界, $z_{\min}$ 和 $z_{\max}$ 即为 z_{near} 和 z_{far}. 这样, 在规范化坐标系中的坐标 (x',y',z') 和规范化前的坐标 (x,y,z) 就满足

$$\begin{bmatrix}x' & y' & z' & 1\end{bmatrix}^{\mathrm{t}}=\boldsymbol{M}_{\mathrm{ortho}}\begin{bmatrix}x & y & z & 1\end{bmatrix}^{\mathrm{t}}$$

透视投影 尽管正交投影容易生成, 且可以保持对象的比例不变, 但它的成像缺发真实感. 人眼观察和相机拍摄到的图像通常符合透视投影规律.

定义 1.12 透视投影 透视投影的投影线投影线汇聚于投影中心 C, 投影中心到投影平面的距离称为焦距, 记作 f.

透视投影观察体是棱台形状, 近剪切面 z_{near} 小, 远剪切面 z_{far} 大. 我们需要将观察体映射到一个适合正交投影的区域内 (即将棱台映射到一个长方体内). 具体而言, 是把观察体挤压到一个以 z_{near} 为底面的长方体内. 设观察体内一点 $\boldsymbol{v}=(x,y,z)$ 变换到正交投影区域内的点 $\boldsymbol{u}=(x',y',z')$. 根据相似三角形的性质, 有

$$\frac{x'}{z_{\mathrm{near}}}=\frac{x}{z},\quad \frac{y'}{z_{\mathrm{near}}}=\frac{y}{z}$$

即

$$x'=\frac{xz_{\mathrm{near}}}{z},\quad y'=\frac{yz_{\mathrm{near}}}{z}$$

我们现在求矩阵 $\boldsymbol{M}_{\mathrm{persp}\to\mathrm{ortho}}$ 使得在齐次坐标系下有

$$\begin{bmatrix}\boldsymbol{u}\\ 1\end{bmatrix}=\boldsymbol{M}_{\mathrm{persp}\to\mathrm{ortho}}\begin{bmatrix}\boldsymbol{v}\\ 1\end{bmatrix}$$

注意到 x',y' 的表达式中 z 在分母, 这意味着映射可能是非线性的. 为了避免这一情况, 注意到对齐次坐标系

的各个分量同乘一数不改变其含义, 于是将两边同乘 z 可得 (这里把 z 合并进矩阵中)

$$\begin{bmatrix} xz_{\text{near}} \\ yz_{\text{near}} \\ z'z \\ z \end{bmatrix} = \boldsymbol{M}_{\text{persp}\to\text{ortho}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

我们可以逆推出矩阵除第三行以外的形式. 于是有

$$\begin{bmatrix} xz_{\text{near}} \\ yz_{\text{near}} \\ z'z \\ z \end{bmatrix} = \begin{bmatrix} z_{\text{near}} & 0 & 0 & 0 \\ 0 & z_{\text{near}} & 0 & 0 \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

由于 z 坐标的变换与 x, y 坐标无关, 因此 $m_{31} = m_{32} = 0$. 自然, 我们希望映射是线性的, 不应该改变两个裁剪平面的 z 坐标值. 于是分别将 $z = z' = z_{\text{near}}$ 和 $z = z' = z_{\text{far}}$ 代入矩阵的第三行可得

$$\begin{cases} z_{\text{near}} m_{33} + m_{34} = z_{\text{near}}^2 \\ z_{\text{far}} m_{33} + m_{34} = z_{\text{far}}^2 \end{cases}$$

解得 $m_{33} = z_{\text{near}} + z_{\text{far}}, m_{34} = -z_{\text{near}} z_{\text{far}}$. 这样, 观察体就被映射到一个以近裁剪平面为底面,z 坐标取值为 $[z_{\text{near}}, z_{\text{far}}]$ 的长方体内. 再对此长方体应用前面求出的正交投影的变换矩阵, 可知投射投影的变换矩阵为

$$\begin{aligned} \boldsymbol{M}_{\text{persp}} &= \boldsymbol{M}_{\text{ortho}} \boldsymbol{M}_{\text{persp}\to\text{ortho}} \\ &= \begin{bmatrix} \frac{2}{x_{\max}-x_{\min}} & 0 & 0 & -\frac{x_{\max}+x_{\min}}{x_{\max}-x_{\min}} \\ 0 & \frac{2}{y_{\max}-y_{\min}} & 0 & -\frac{y_{\max}+y_{\min}}{y_{\max}-y_{\min}} \\ 0 & 0 & \frac{2}{z_{\text{far}}-z_{\text{near}}} & -\frac{z_{\text{far}}+z_{\text{near}}}{z_{\text{far}}-z_{\text{near}}} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_{\text{near}} & 0 & 0 & 0 \\ 0 & z_{\text{near}} & 0 & 0 \\ 0 & 0 & z_{\text{near}} + z_{\text{far}} & -z_{\text{near}} z_{\text{far}} \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{aligned}$$

这就完成了把观察体映射到规范化坐标系的立方体内的过程.

需要注意的是, 在透视投影后的第四个分量即为深度值 z, 这在后续的深度测试中会用到. 此外, 根据齐次坐标的定义, 在使用透视投影矩阵后需要再做一次齐次除法, 即将前三个分量分别除以第四个分量, 才能得到最终的规范化坐标系下的坐标值.

视口映射 在完成将观察体映射到规范化坐标系的工作后, 由于不同设备的屏幕分辨率大小不同, 我们需要将规范化坐标系内的物体再映射到屏幕坐标上 (这里不对 z 坐标做变换, 在以后的步骤中另有他用). 假定屏幕的宽度和高度分别为 w, h, 那么视口变换就是把 xy 坐标 $[-1, 1] \times [-1, 1]$ 映射到 $[0, w] \times [0, h]$ 上. 这也是一个放

缩操作, 因此不难写出视口变换的矩阵为

$$\boldsymbol{M}_{\text{view}} = \begin{bmatrix} w/2 & 0 & 0 & w/2 \\ 0 & h/2 & 0 & h/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$