

# 1 几何变换

## 1.1 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.

### 1.1.1 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 1.1 齐次坐标系** 在  $n$  维空间  $R^n$  中, 点  $\mathbf{v} = (x_1, \dots, x_n)$  的齐次坐标表示为  $(x_1, \dots, x_n, 1)$ , 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  $\mathbf{v} = (x_1, \dots, x_n)$  表示为  $(x_1, \dots, x_n, 0)$ , 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. 例如, 在二维空间中的一些几何操作对应的矩阵如下:

$$\begin{array}{cc} \text{平移} & \text{旋转} \\ \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} & \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \end{array}$$

### 1.1.2 旋转矩阵

在上一节中我们不加证明地给出了旋转矩阵的定义. 在此我们详细地推导旋转矩阵的形式.

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

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

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

注意到  $\mathbf{u}^\perp = \mathbf{a} \times \mathbf{v}$ , 于是前面的推导可以整理成

$$\mathbf{v}' = [(\mathbf{v} - (\mathbf{v} \cdot \mathbf{a})\mathbf{a}) \cos \theta] + (\mathbf{a} \times \mathbf{v}) \sin \theta + (\mathbf{v} \cdot \mathbf{a})\mathbf{a}$$

□

## 1.2 几何变换在铰接刚体中的应用

### 1.2.1 前向运动学

以人体为例, 我们可以把各个关节视作节点, 各个骨骼视作连接节点的边. 这样, 人体就可以表示为一个带有根节点的树形结构. 我们只需要知道根节点的位置和各个关节的旋转角度, 就可以计算出各个节点的位置.

以刚体铰链模型  $P_0 \cdots P_n$  为例展示前向运动学的过程. 节点  $P_0, \dots, P_n$  的初始位置记作  $\mathbf{p}_0^{\text{st}}, \dots, \mathbf{p}_n^{\text{st}}$ , 向下一个关节的位移即为  $\mathbf{l}_i = \mathbf{p}_i^{\text{st}} - \mathbf{p}_{i-1}^{\text{st}} (i = 1, \dots, n)$ , 每个节点上都带有一个局部坐标系  $\mathcal{C}_i$ .

现在, 我们为除去  $P_n$  外的每个关节  $P_j$  指定一个旋转矩阵  $\mathbf{R}_j^{\text{loc}} (j = 1, \dots, n-1)$  (这里的旋转矩阵是基于该关节关联的局部坐标系  $\mathcal{C}_j$  所定义的), 每个关节  $P_j$  处的旋转带动其后的所有关节 (包括其上带的局部坐标系) 和骨骼旋转. 因此, 节点  $P_j$  的总的旋转矩阵  $\mathbf{R}_j^{\text{tot}}$  即为它的局部坐标系  $\mathcal{C}_j$  相对于根节点的坐标系  $\mathcal{C}_0$  的旋转与它本身的旋转的复合. 而  $\mathcal{C}_j$  的旋转事实上就是由  $P_{j-1}$  节点的旋转定义的, 因此  $\mathbf{R}_j^{\text{tot}} = \mathbf{R}_{j-1}^{\text{tot}} \mathbf{R}_j^{\text{loc}}$ .

现在, 我们就可以从根节点开始计算每个子节点的位置  $\mathbf{p}_j^{\text{ed}} (j = 1, \dots, n)$ . 根节点的位置  $\mathbf{p}_0^{\text{ed}} = \mathbf{p}_0^{\text{st}}$  已知, 而每个节点的位置可以通过其父节点的位置加上旋转后的位移向量得到, 即

$$\mathbf{p}_j^{\text{ed}} = \mathbf{p}_{j-1}^{\text{ed}} + \mathbf{R}_{j-1}^{\text{tot}} \mathbf{l}_j$$

## 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  $\mathbf{p}$ , 观察平面的法向量  $\mathbf{n}$  和视角正上方的方向  $\mathbf{v}$  决定.

### 1.3.1 坐标变换

在观察者的坐标中,  $\mathbf{n}$  通常指向  $z$  轴正方向,  $\mathbf{v}$  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  $\mathbf{u} = \mathbf{v} \times \mathbf{n}$ . 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  $\mathbf{p}$  在  $uv$  坐标和  $xy$  坐标下的表示分别为

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

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

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

$$\mathbf{u} = x_u \mathbf{x} + y_u \mathbf{y}$$

$$\mathbf{v} = x_v \mathbf{x} + y_v \mathbf{y}$$

不妨设  $uv$  坐标的原点在  $xy$  坐标中的表示为  $(o_x, o_y)$ . 那么可以把坐标变换视作平移和旋转的复合, 即

$$\begin{bmatrix} p_u \\ p_v \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_u & y_u & 0 \\ x_v & y_v & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix}$$

### 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 1.2 投影线** 连接对象点和投影点的直线称作**投影线**.

**定义 1.3 平行投影** 投影线相互平行的投影方式称作**平行投影**

**定义 1.4 (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 1.5 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  $z_{\text{near}}$  和  $z_{\text{far}}$ .

有时,  $z_{\text{near}}$  和  $z_{\text{far}}$  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 1.6 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  $\mathbf{v} = (x, y, z)$  的投影点就是  $\mathbf{i} = (x, y)$ . 通常, 需要将对象描述转化到规范化坐标系, 即建立  $\mathbf{v} \rightarrow [-1, 1]^3$  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

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

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  $z_{\text{near}}$  和  $z_{\text{far}}$ . 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

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

**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, 但它的成像缺乏真实感. 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 1.7 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  $z_{\text{near}}$  小, 远剪切面  $z_{\text{far}}$  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  $z_{\text{near}}$  为底面的长方体内. 设观察体内一点  $\mathbf{v} = (x, y, z)$  变换到正交投影区域内的点  $\mathbf{u} = (x', y', z')$ . 根据相似三角形的性质, 有

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

即

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

我们现在求矩阵  $\mathbf{M}_{\text{persp} \rightarrow \text{ortho}}$  使得在齐次坐标系下有

$$\begin{bmatrix} \mathbf{u} \\ 1 \end{bmatrix} = \mathbf{M}_{\text{persp} \rightarrow \text{ortho}} \begin{bmatrix} \mathbf{v} \\ 1 \end{bmatrix}$$

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)

$$\begin{bmatrix} x z_{\text{near}} \\ y z_{\text{near}} \\ z' z \\ z \end{bmatrix} = \mathbf{M}_{\text{persp} \rightarrow \text{ortho}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

我们可以逆推出矩阵除第三行以外的形式. 于是有

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

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

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

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.

**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

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