

# CV Problem\_Set\_1

王泽恺 2400013155

## 1

### 1.1

1. 齐次坐标的核心是比例不变性：N 维笛卡尔坐标系中的同一个点，在 N+1 维齐次坐标系中可以用无数组坐标表示，但这些坐标满足“对应分量成比例”的关系。

已知条件：笛卡尔坐标系中点  $P = (x_1, x_2, \dots, x_N)$ ；

齐次坐标系中该点的一般形式为  $P' = (x_1, x_2, \dots, x_N, w)$  ( $w \neq 0$ , w 为权重)；

齐次坐标的“规范形式”为  $(x'_1, x'_2, \dots, x'_N, 1)$  (即令最后一维坐标为 1)。

由于齐次坐标的比例不变性，一般形式与规范形式表示同一个点，因此两组坐标的对应分量成比例，比例系数为 w (或  $1/w$ )：

$$\frac{x_1}{x'_1} = \frac{x_2}{x'_2} = \dots = \frac{x_N}{x'_N} = \frac{w}{1} = w$$

由此可解出笛卡尔坐标  $x_i$  与齐次规范形式坐标  $x'_i$  的关系：若已知齐次规范形式  $(x'_1, \dots, x'_N, 1)$ ，则笛卡尔坐标为： $x_i = w \cdot x'_i$  (其中 w 是齐次坐标的权重因子)。

2.

- 统一表示“有限点”与“无穷远点”，解决“平行线透视相交”问题

对于笛卡尔坐标系中的有限点  $(x_1, \dots, x_N)$ ，齐次坐标为  $(x_1, \dots, x_N, w)$  ( $w \neq 0$ )，规范形式为  $(x_1/w, \dots, x_N/w, 1)$ ；

对于“无穷远点”，齐次坐标可表示为  $(a_1, a_2, \dots, a_N, 0)$  ( $w = 0$ )，其中  $(a_1, \dots, a_N)$  代表该无穷远点的“方向”。

- 将非线性透视投影转化为线性矩阵乘法，简化计算流程

若使用齐次坐标，透视投影关系可表示为线性矩阵乘法： $x \sim P \cdot X$  (符号 “~” 表示齐次坐标的比例等价，即  $x = k \cdot P \cdot X$ ,  $k \neq 0$ )。

若用笛卡尔坐标直接表示，透视投影含  $1/Z$  的除法，是非线性变换，无法用矩阵乘法统一描述

-利用齐次坐标，通过简化透视投影和统一有限点和无穷远点，摄影计算能够更好的将三维场景映射到二维屏幕中，因此被采用。

## 1.2

推拉变焦的第一个核心目标是主体在画面中尺寸不变，对应物理量是“主体在传感器上的像高 $h'$ ”恒定  
因为  $\frac{h}{h'} = \frac{z}{z'}$

又由高斯透镜公式  $\frac{1}{f} = \frac{1}{z} + \frac{1}{z'}$

所以  $h' = \frac{(h \cdot f)}{z - f}$

设  $h' = C$

则不难推出  $f = \frac{C}{C+h} \cdot z$

这即是希区柯克变焦的重要数学约束，即焦距和物距始终成正比。

推拉变焦的视觉特效（背景拉伸 / 压缩），本质是视场角（FoV）随 $z$ 和 $f$ 的联动发生剧烈变化

由  $\tan(\theta/2) = D/2z'$  和高斯透镜公式

得  $\tan(\theta/2) = D(z - f)/2zf$

又因为  $f = kz$

所以  $\tan(\theta/2) = D(1 - k)/2kz$

由此公式不难得出当相机靠近主体（ $z$ 减小）时， $\tan(\frac{\theta}{2})$ 增大， $\theta$ 减小（视场角变窄，背景范围缩小）；

当相机远离主体（ $z$ 增大）时， $\tan(\frac{\theta}{2})$ 减小， $\theta$ 增大（视场角变宽，背景范围扩大）。

## 2

1. 相机高度  $H$  是世界坐标系中相机的  $z$  坐标（因地面为  $z = 0$ ）。根据透视投影的相似三角形原理：参考物体的实际高度（0.76 米）与相机高度  $H$  的比值，等于“参考物体上、下边缘在图像中的像素距离 ( $|y_{top} - y_{bottom}|$ )”与“参考物体下边缘到消失线的像素距离 ( $|y_{horizon} - y_{bottom}|$ )”的比值。公式可表示为： $\frac{0.76}{H} = \frac{|y_{top} - y_{bottom}|}{|y_{horizon} - y_{bottom}|}$ ，整理后得  $H = 0.76 \times \frac{|y_{horizon} - y_{bottom}|}{|y_{top} - y_{bottom}|}$ 。

经过测算后  $y_{horizon} = 1777, y_{top} = 1952, y_{bottom} = 2133$

得出  $H \approx 1.49$

2.

$$K = \begin{pmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

将世界点  $A_{world}$  通过外参  $R$  和  $T$  转换为相机坐标系下的坐标  $A_{cam} = (X_c, Y_c, Z_c)^T$ ，公式为：

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = R \cdot \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + T$$

第二步：内参矩阵  $K$  对相机坐标  $A_{cam}$  的作用可拆解为  $u$ 、 $v$  两个方向的投影，取  $v$  方向根据相似三角形写出方程  $v = f \cdot \frac{Y_c}{Z_c} + v'$  ( $v'$  是花坛底部的点)。

第三步：整理方程，解出  $f$ 。

整理得： $f = \frac{(v-v') \cdot Z_c}{Y_c}$  再将“相机坐标 $Y_c, Z_c$ 与外参 $R, T$ 的关系”代入（即 $Y_c = R_{21}X_w + R_{22}Y_w + R_{23}Z_w + T_x$ ,  $Z_c = R_{31}X_w + R_{32}Y_w + R_{33}Z_w + T_z$ ），最终得到：

$$f = \frac{(v - v') \cdot (R_{31}X_w + R_{32}Y_w + R_{33}Z_w + T_z)}{R_{21}X_w + R_{22}Y_w + R_{23}Z_w + T_x}$$

最终算出来在相机坐标系中 $A_{cam}$ 坐标为 $(-0.76, -0.73, 13.4)$ ,  $v = 1952, v' = 2133$ , 代入解得 $f = 3191 pixel$ 。