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$P = (x_1, x_2, ..., x_N)$；
齐次坐标系中该点的一般形式为$P' = (x_1, x_2, ..., x_N, w)$（$w \neq 0$，w 为权重）；
齐次坐标的 “规范形式” 为$(x'_1, x'_2, ..., x'_N, 1)$（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$1/w$）：
$\frac{x_1}{x'_1} = \frac{x_2}{x'_2} = ... = \frac{x_N}{x'_N} = \frac{w}{1} = w$
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$(x'_1, ..., x'_N, 1)$，则笛卡尔坐标为：$\boxed{x_i = w \cdot x'_i}$（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$(x_1, ..., x_N)$，齐次坐标为$(x_1, ..., x_N, w)$（$w \neq 0$），规范形式为$(x_1/w, ..., x_N/w, 1)$；
  对于 “无穷远点”，齐次坐标可表示为$(a_1, a_2, ..., a_N, 0)$（$w = 0$），其中$(a_1, ..., a_N)$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，$\tan\left(\frac{\theta}{2}\right)$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，$\tan\left(\frac{\theta}{2}\right)$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.

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

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$A_{\text{cam}} = (X_c, Y_c, Z_c)^T$，
第二步：相机坐标→像素坐标（内参作用）忽略缩放因子s（投影的 “比例一致性” 可消去s），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u方向（v方向推导逻辑一致）：$u = f \cdot \frac{X_c}{Z_c} + c_x$该公式的物理意义：$\frac{X_c}{Z_c}$：将 3D 相机坐标的 “水平偏移” 转换为 2D 归一化坐标（无单位，仅表示方向）；$f \cdot \frac{X_c}{Z_c}$：通过焦距f将归一化坐标缩放为 “像素偏移量”(单位：像素)；$+c_x$：将像素偏移量的原点从 “相机坐标系原点” 平移到 “图像主点”(对齐像素坐标系)。
第三步：整理方程，解出f将$c_x = \frac{W}{2}$（由图像宽度W决定）代入上式，整理得：$f = \frac{(u-c_x) \cdot Z_c}{X_c} =$

$\frac{\left(u-\frac{W}{2}\right)\cdot Z_c}{X_c}$再将 “相机坐标$X_c, Z_c$与外参$R, T$的关系” 代入（即$X_c = R_{11}X_w + R_{12}Y_w + R_{13}Z_w + T_x$，$Z_c = R_{31}X_w + R_{32}Y_w + R_{33}Z_w + T_z$），最终得到：

$$\boxed{f = \frac{\left(u-\frac{W}{2}\right)\cdot\left(R_{31}X_w + R_{32}Y_w + R_{33}Z_w + T_z\right)}{R_{11}X_w + R_{12}Y_w + R_{13}Z_w + T_x}}$$

最终算出来在相机坐标系中A_{cam}坐标为$(-0.76, -0.73, 13.4)$,$u = 1892$,,代入解得$f = 2750 pixel$
。