

# HW1

刘智琦

2024 年 9 月 23 日

## 1 Short questions

### 1.1 Homogeneous coordination

#### 1.1.1

$$\begin{aligned}\because P' &= (x_1, x_2, \dots, x_n, \omega), \quad P' = (x'_1, x'_2, \dots, x'_n, 1) \\ \therefore x'_i &= x_i / \omega\end{aligned}$$

#### 1.1.2

1. **统一表示平移和其他变换：**在二维或三维空间中，平移变换不能用矩阵乘法直接表示，而需要加上一个平移向量。通过引入齐次坐标，平移变换可以与旋转、缩放等其他线性变换统一表示为矩阵乘法，这简化了变换的组合和计算。

2. **处理无穷远点：**齐次坐标允许表示无穷远点，这在计算机图形学和计算摄影中非常重要。例如，在透视投影中，平行线在无穷远处相交，通过齐次坐标可以方便地处理这种情况。

### 1.2 Dolly zoom

在相机移动的同时调节镜头的焦距，比如向前移动相机的同时调小焦距，可以使前景保持不变，而背景变大，产生视觉上的扭曲效果。

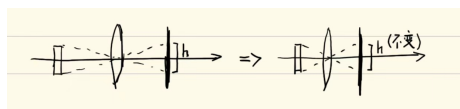


图 1: Dolly zoom

## 2 Camera parameters from the image

### 2.1

取  $x$  轴，屋檐作为一组平行线，中间花坛的两条边作为另一组平行线，画出 Vanishing Line。过点  $A$  作  $PH \perp y$  轴于点  $H$  交 Vanishing Line 于点  $P$ 。

经测量

$$\frac{PH}{AH} \approx 1.9$$

$$\therefore Height \approx 0.76 \times 1.9 = 1.44m$$

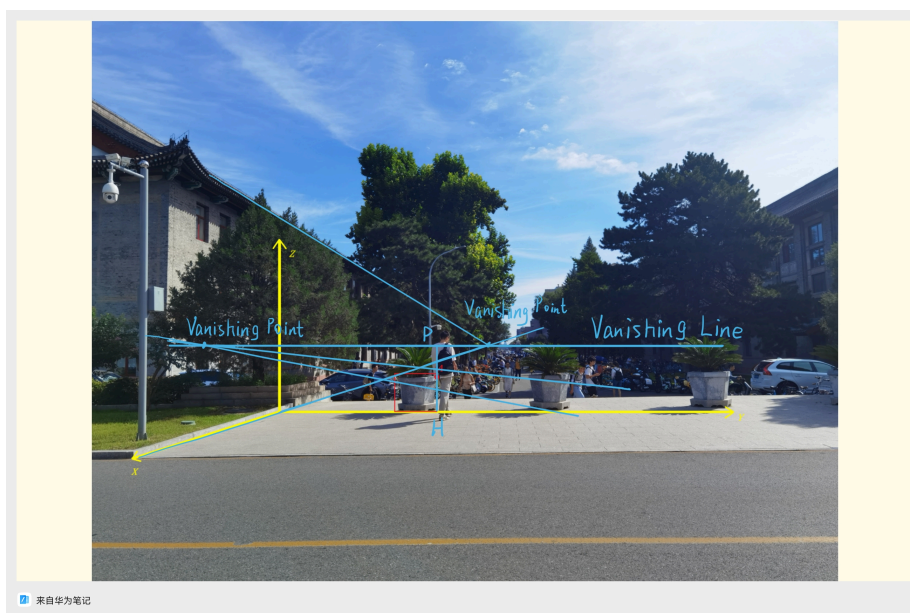


图 2: Camera parameters

## 2.2

$$\mathbf{\Pi} = \underbrace{\begin{bmatrix} f & s & c_x \\ 0 & \alpha f & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{intrinsics}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{projection}} \underbrace{\begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}}_{\text{rotation}} \underbrace{\begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}}_{\text{translation}}$$

identity matrix

图 3: Camera parameters

由 (1) 可得, 相机的世界坐标为 (13.4, 4.5, 1.44)

$$\therefore T = \begin{bmatrix} 1 & 0 & 0 & -13.4 \\ 0 & 1 & 0 & -4.5 \\ 0 & 0 & 1 & -1.44 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

经测量, 世界坐标系的 y 轴, z 轴分别平行于图像坐标系的 x 轴, y 轴。

又  $\because$  世界坐标系的 x 轴  $\perp$  yz 平面, 图像坐标系的 z 轴  $\perp$  xy 平面

$\therefore$  世界坐标系的 x 轴与图像坐标系的 z 轴平行

$\therefore$  相机的旋转矩阵为

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

经查表, 图像平面上 1 米对应 2800 个像素点

$$\therefore = \begin{bmatrix} f & 0 & 0.731 \\ 0 & f & 0.549 \\ 0 & 0 & 1 \end{bmatrix}$$

经测量, 点 A 的像素坐标为 (1890, 1950), 原点 O 的像素坐标为 (1040, 2140)

则点 A 的图像坐标为 (0.675, 0.696), 原点 O 的图像坐标为 (0.371, 0.764)

$$\therefore \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -13.4 \\ 0 & 1 & 0 & -4.5 \\ 0 & 0 & 1 & -1.44 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 3.74 \\ 0.76 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.76 \\ 0.68 \\ 13.4 \end{bmatrix} = 13.4 \begin{bmatrix} -0.0567 \\ 0.0507 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -13.4 \\ 0 & 1 & 0 & -4.5 \\ 0 & 0 & 1 & -1.44 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -4.5 \\ 1.44 \\ 13.4 \end{bmatrix} = 13.4 \begin{bmatrix} -0.336 \\ 0.107 \\ 1 \end{bmatrix}$$

$\therefore$  点 A, 点 O 的相机齐次坐标分别为  $(-0.0567, 0.0507, 1)^T$ ,  $(-0.336, 0.107, 1)^T$

$$\begin{bmatrix} 0.675 \\ 0.696 \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0.731 \\ 0 & f & 0.549 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.0567 \\ 0.0507 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0.371 \\ 0.764 \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0.731 \\ 0 & f & 0.549 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.336 \\ 0.107 \\ 1 \end{bmatrix}$$

考虑到误差, 取平均值作为 f 的值, 得到  $f=1.74\text{m}$

### 3 Image filtering and subsampling

通过一系列函数逐步实现了高斯金字塔图像处理算法。以下是各个部分的简要说明:

导入库:

numpy 用于数组和矩阵操作。PIL (Python Imaging Library) 用于图像处理。

函数 `cross_correlation_2d`:

计算二维图像和二维核的交叉相关。对图像进行填充以处理边界。遍历图像的每个像素，计算局部区域与核的乘积和。

函数 `convolve_2d`:

对核进行翻转(水平和垂直),然后调用 `cross_correlation_2d` 进行卷积操作。

函数 `gaussian_blur_kernel_2d`:

生成一个二维高斯模糊核。根据给定的大小和标准差计算高斯分布值，并进行归一化。

函数 `low_pass`:

对图像应用低通滤波(高斯模糊)。将图像分解为红、绿、蓝三个通道，分别进行卷积操作。将模糊后的通道重新组合成一个图像。

函数 `image_subsampling`:

对图像进行下采样，将图像尺寸减半。

函数 `gaussian_pyramid`:

构建高斯金字塔。对图像进行多次模糊和下采样，并保存每个层级的图像。

主程序:

从用户输入获取图像文件名。打开图像并调用 `gaussian_pyramid` 函数处理图像。这个文件的主要功能是对输入图像进行高斯模糊和下采样，生成一系列不同分辨率的图像，并保存这些图像。