



无人机赛课第2讲

--视觉传感器和二维码

安琪儿

联系方式: shenyuan_ee@tsinghua.edu.cn

时间: 2019年11月3日





目 录

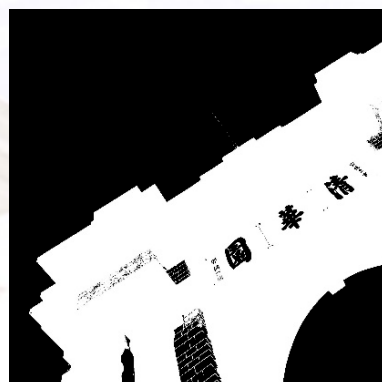
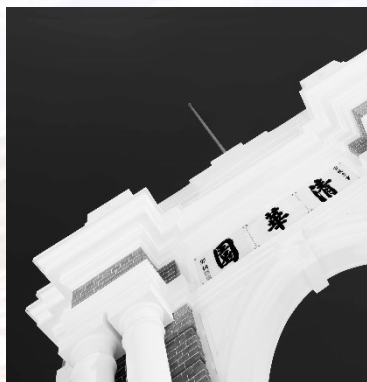
- 1 图像处理相关知识
- 2 二维码相关知识
- 3 基于二维码的机器人视觉定位



1 图像处理相关知识

- 图像灰度化、二值化

- 彩色图像：R, G, B三通道，每个通道取值0~255
- 灰度图像：单通道，取值0~255
- 二值图像：单通道黑白两色，取值0, 255



- 图像灰度化方法：

- 最大值法： $Y = \max(R, G, B)$
- 平均值法： $Y = (R + G + B)/3$
- 加权平均法： $Y = (W_R R + W_G G + W_B B)/(W_R + W_G + W_B)$



1 图像处理相关知识

- 图像灰度化、二值化

- 图像二值化方法:

- 根据阈值对像素点分类, 将灰度图像转为黑白二值图像

$$I(x, y) = \begin{cases} 1, & i(x, y) \geq thresh \\ 0, & i(x, y) < thresh \end{cases}$$

- 全局阈值:

- e.g. Otsu算法 (最大化类间方差)

(其中 u_0, u_1 分别为黑白两类像素的类内均值, u 为总体均值, w_0, w_1 为两类像素所占比例)

$$t^* = \operatorname{argmax}(g(t))$$

$$g(t) = w_0(u_0 - u)^2 + w_1(u_1 - u)^2$$

- 局部阈值: 目标像素与邻域对比确定取0或255

- e.g. Niblack算法

(其中 $m(x, y)$ 为邻域均值, $s(x, y)$ 为邻域标准差, $t(x, y)$ 为像素处的阈值)

$$t(x, y) = m(x, y) + k * s(x, y)$$

- 自适应阈值: 根据不同的局部特征选择不同的二值化方法



1 图像处理相关知识

- 图像滤波

- 在保留图像细节的同时抑制噪声
- 图像卷积：利用卷积核对图像邻域进行加权

$$O(i, j) = \sum_{m, n} I(i + m, j + n) * K(m, n)$$

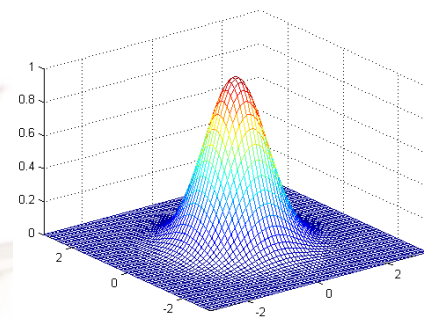
- 线性滤波器：

- 每个像素的输出值是输入像素的线性加权和
- 均值滤波：将卷积核覆盖区域的像素均值作为输出
- 高斯滤波：

- 卷积核 $K(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$

- 非线性滤波器：

- 中值滤波：选取卷积核覆盖区域图像像素的中值作为输出，常用以去除椒盐噪声

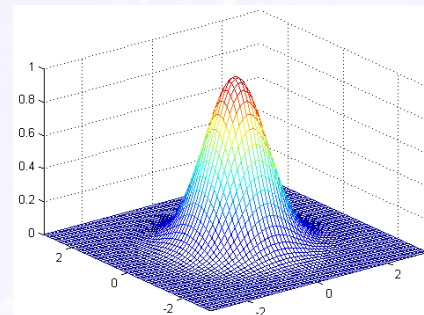




1 图像处理相关知识

- 图像滤波
 - 高斯滤波

- 卷积核 $K(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$
- σ 越大/卷积核尺寸越大，平滑效果越好，图像越模糊
- 不同 σ 、不同卷积核尺寸下滤波效果：



3×3 高斯核, $\sigma^2 = 1$



3×3 高斯核, $\sigma^2 = 3$



9×9 高斯核, $\sigma^2 = 1$



9×9 高斯核, $\sigma^2 = 3$



原图



1 图像处理相关知识

- 图像滤波

- 中值滤波

- 选取卷积核覆盖区域图像像素的中值作为输出，常用以去除椒盐噪声





1 图像处理相关知识

- 图像膨胀、腐蚀

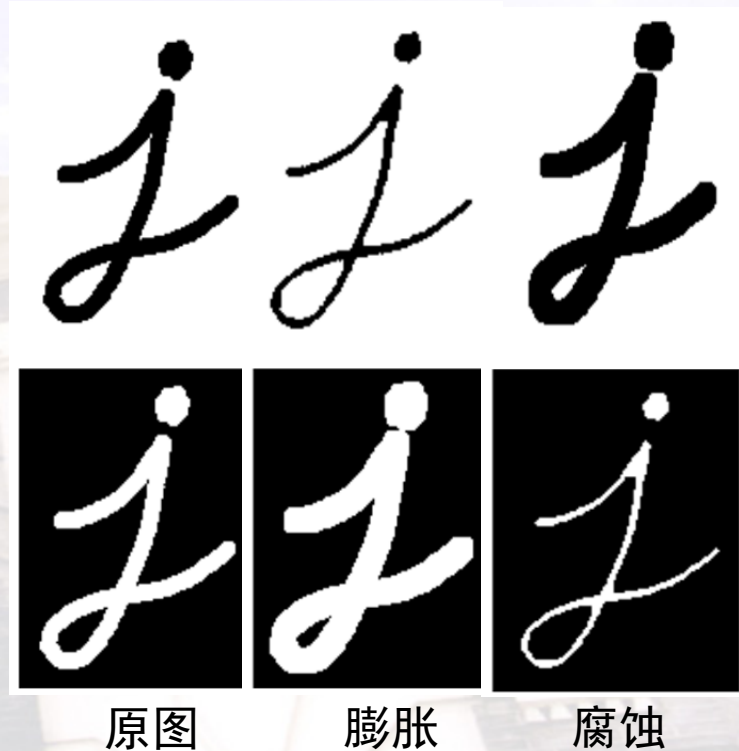
- 形态学中的两种对偶操作，是针对像素值大的区域而言

- 膨胀：

- 用卷积核扫描图像，将叠加区域的最大像素值作为中心点输出值
→ 图片中亮的区域“膨胀”

- 腐蚀：

- 用卷积核扫描图像，将叠加区域的最小像素值作为中心点输出值
→ 图片中暗的区域“膨胀”





1 图像处理相关知识

- 图像膨胀、腐蚀
 - 不同尺寸卷积核下膨胀、腐蚀效果



膨胀



10 × 10卷积核



腐蚀

50 × 50卷积核

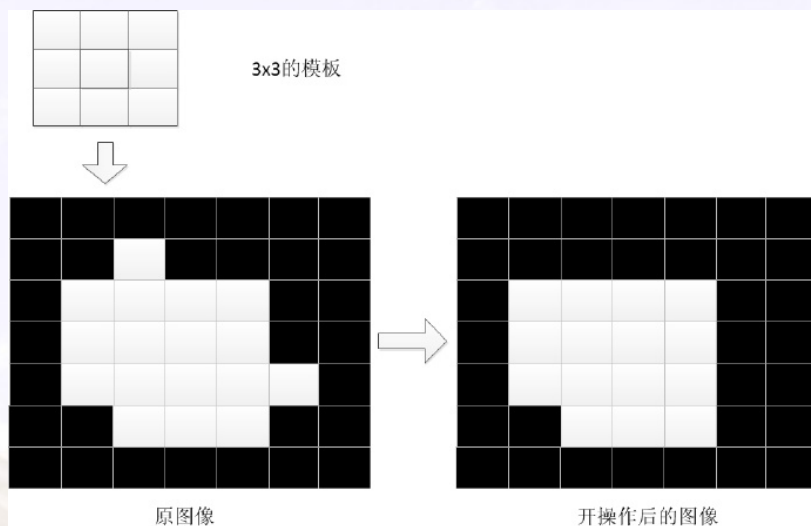


1 图像处理相关知识

- 图像膨胀、腐蚀

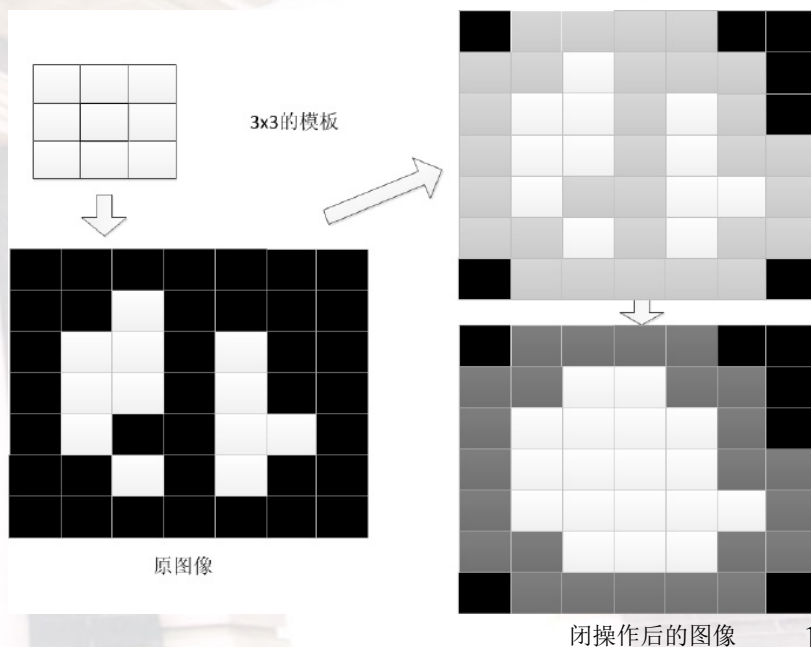
- 开运算

- 先腐蚀，再膨胀
 - 消除高亮小物体，平滑物体边界



- 闭运算

- 先膨胀，再腐蚀
 - 消除小的黑洞，连通相邻高亮区域





1 图像处理相关知识

- 图像边缘检测

- 检测图像中亮度变化明显的点，常用以轮廓提取等
- Sobel算子：

- x, y 两个方向的卷积核，分别计算横向、纵向图像梯度

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I, \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I$$

- 边缘判断：

$$G_x + G_y > thresh$$

- 边缘方向：

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

- 对噪声有平滑作用，精度不高，容易检测到伪边缘



1 图像处理相关知识

- 图像边缘检测
 - Sobel算子:



原图



X
方向
梯度
图



Y
方向
梯度
图



整体
梯度
图



1 图像处理相关知识

• 图像边缘检测

– Canny算子:

- Step 1: 图像灰度化
- Step 2: 高斯滤波
- Step 3: 计算梯度值和方向

$$G = \sqrt{G_x^2 + G_y^2}, \quad \theta = \arctan\left(\frac{G_y}{G_x}\right)$$

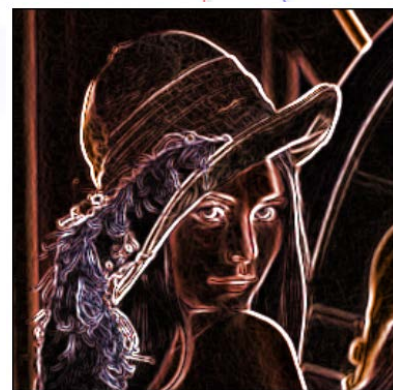
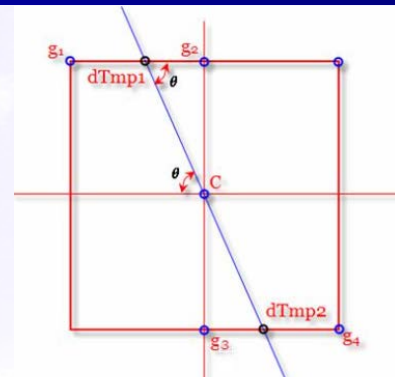
- Step4: 非极大值抑制 (NMS)

(保留梯度局部最大值)

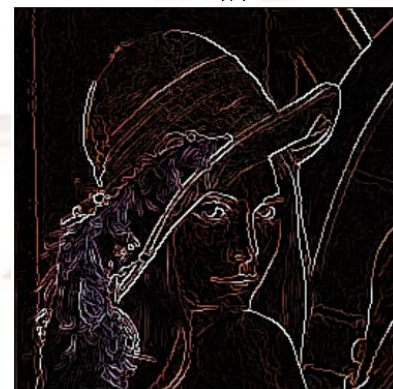
- 蓝色线为边缘方向, g_1, g_2, g_3, g_4 为 c 的相邻像素点, 插值得到 $G(dTmp1), G(dTmp2)$
- $G(c) > G(dTmp1), G(c) > G(dTmp2)$

- Step 5: 双阈值选取、边缘连接

- $I(x, y) > T_H$, 赋1; $I(x, y) < T_L$, 赋0
- $T_L < I(x, y) < T_H$ & 与赋值1像素点连通, 赋1



NMS前

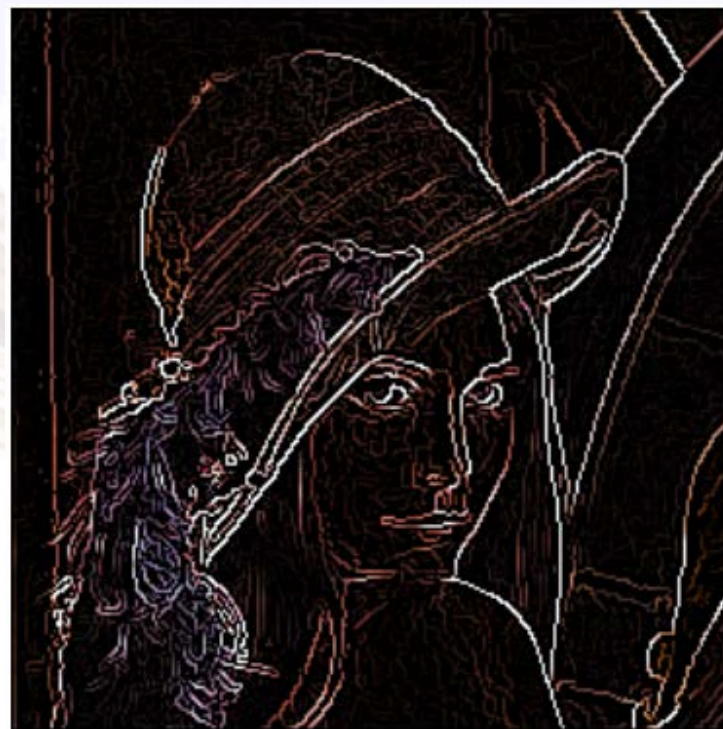


NMS后



1 图像处理相关知识

- 图像边缘检测
 - Sobel算子 vs Canny算子:





2 二维码相关知识

- 基本概念

- 二维码：又称二维条码，通过二维平面上黑白相间的几何图形记录数据符号信息
- 分类：
 - 行列式二维码：由两行或多行一维码组成



PDF417



Ultracode

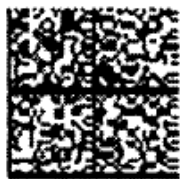


Code 49



Code 16K

- 矩阵式二维码：通过矩阵空间中黑白像素的排布进行编码
 - 常见：Data Matrix, **QR Code (最常见)**, 汉信码（我国自主研发）等



Data Matrix



Maxi Code



Aztec Code



QR Code



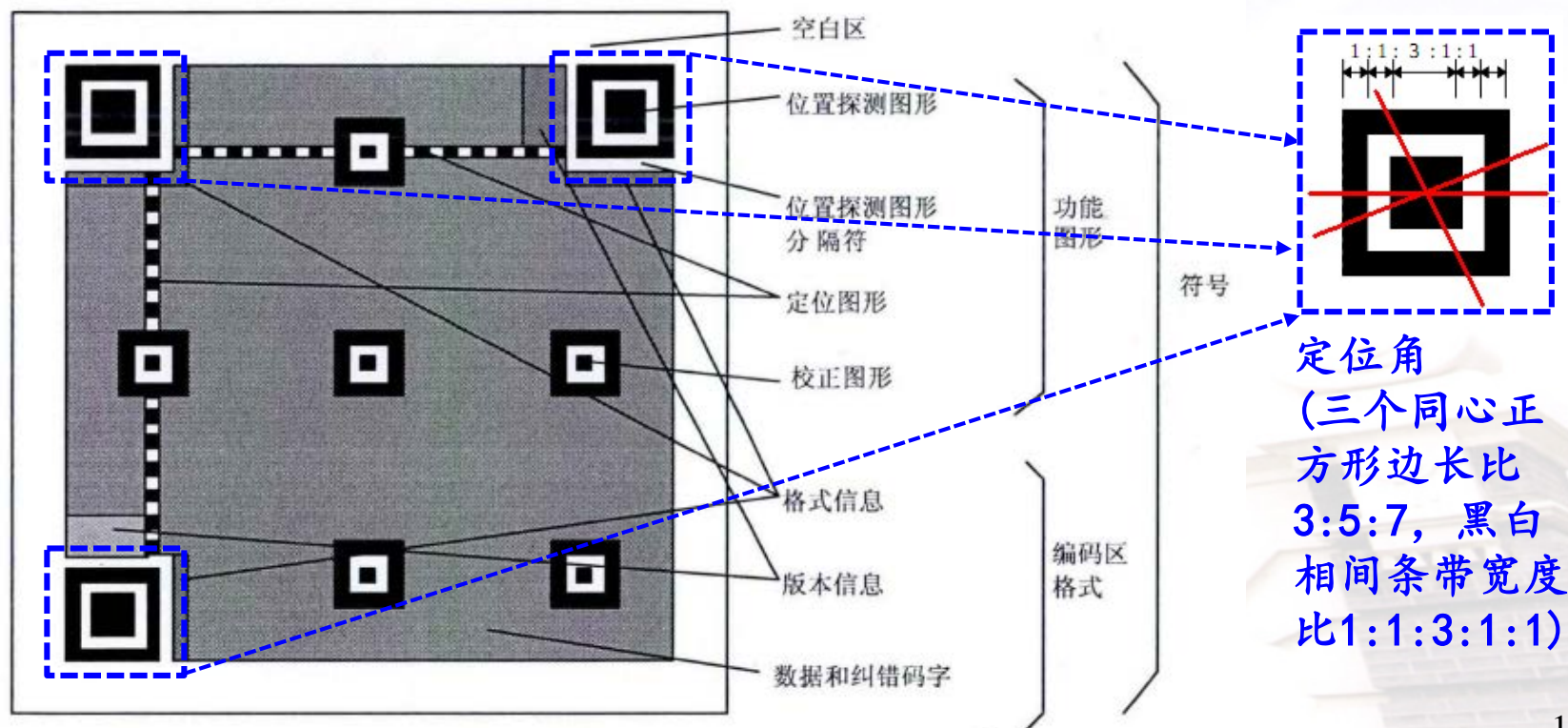
Vericode



2 二维码相关知识

• QR码

- “Quick Response”的缩写
- 40个版本，码元数目为 $21 \times 21, 25 \times 25, \dots, 177 \times 177$
- 符号结构：





2 二维码相关知识

- QR码译码

- 输入：单目相机拍摄到的RGB图片
- 输出：图片中所含二维码的译码数据
- 译码步骤：

- 图像预处理：

- 图像灰度化、二值化：降低数据量
 - 滤波：消除高频噪声
 - 图像腐蚀、膨胀：消除局部反光等细小区域影响

- QR码译码：

- 定位：根据三个定位角的图形特点，确定图中二维码区域位置
 - 解码：对定位得到的二维码区域进行解码



解码结果：

Tsinghua University-QR code 4



3 基于二维码的机器人视觉定位

- 基本原理

- 预先在环境中贴上一系列二维码，并精确测量其在世界坐标系中的坐标
- 机器人搭载的单目相机对环境拍照，通过识别图像中的二维码，结合成像原理，估计机器人在世界坐标系中的位置坐标
- 优点：定位精度高、抗环境干扰、运算量小、成本低
- 关键技术：
 - 二维码的定位与识别
 - 基于成像原理的相机位置估计

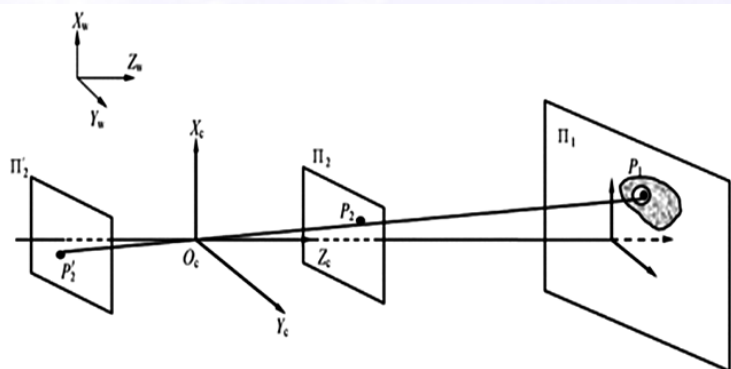




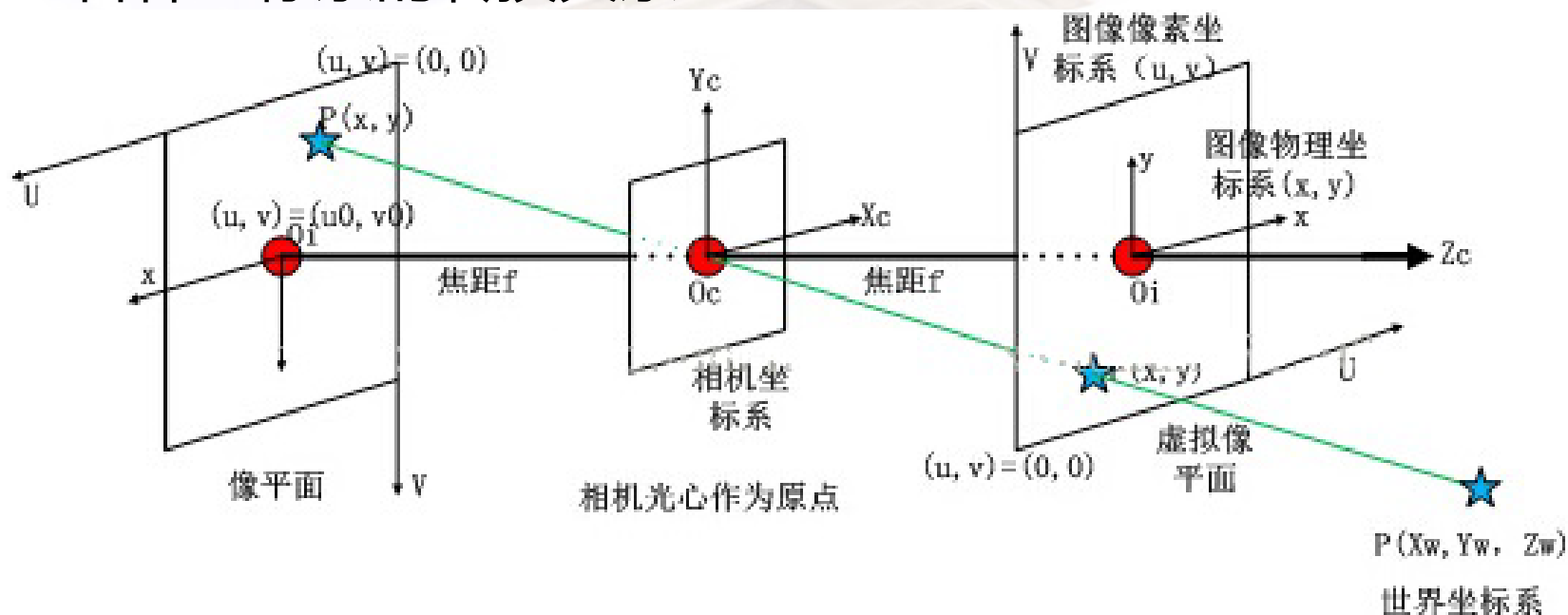
3 基于二维码的机器人视觉定位

- 基于成像原理的相机位置估计

- 小孔成像示意图



- 各种坐标系的转换关系





3 基于二维码的机器人视觉定位

- 基于成像原理的相机位置估计

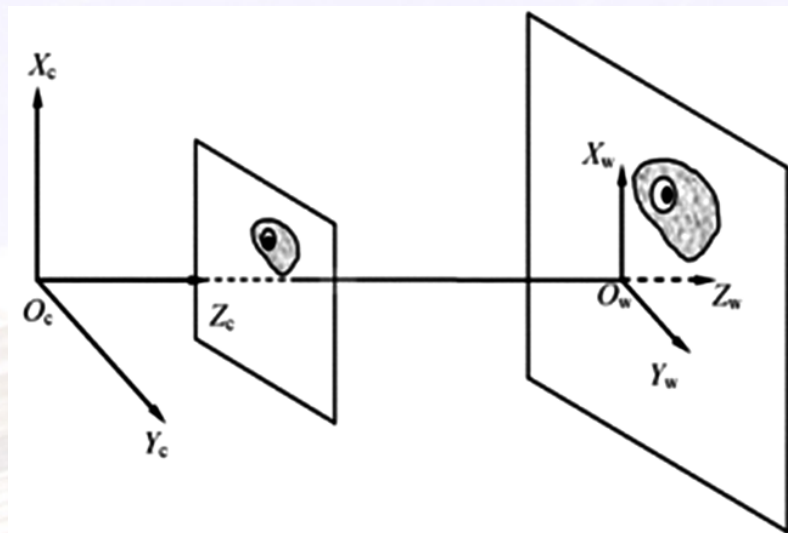
- 针孔相机投影模型

$$x = \frac{1}{z_c} * K * T * X + n$$

$$T = [R \quad t]_{3 \times 4}$$

$$x = (u, v, 1)^T$$

$$X = (x_w, y_w, z_w, 1)^T$$



- 相机外参: $T = [R \quad t]_{3 \times 4}$, $t = -R * c$, 其中 c 为相机中心

- 相机内参: $K = \begin{bmatrix} \alpha_x & 0 & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$

- 三维空间点的坐标 (世界坐标系中): $X = (x_w, y_w, z_w, 1)^T$

- 投影像素点坐标: $x = (u, v, 1)^T$



3 基于二维码的机器人视觉定位

- 基于成像原理的相机位置估计

- 针孔相机投影模型

- 当相机正视二维码中心，且点 \mathbf{X} 位于世界坐标系 $z_w = 0$ 平面时， $\mathbf{R} = \mathbf{I}_{3 \times 3}$ ， $\mathbf{t} = (0, 0, d)^T$ ， $\mathbf{X} = (x_w, y_w, 0, 1)^T$

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{d} \begin{pmatrix} \alpha_x x_w + u_0 d \\ \alpha_y y_w + v_0 d \\ d \end{pmatrix}$$

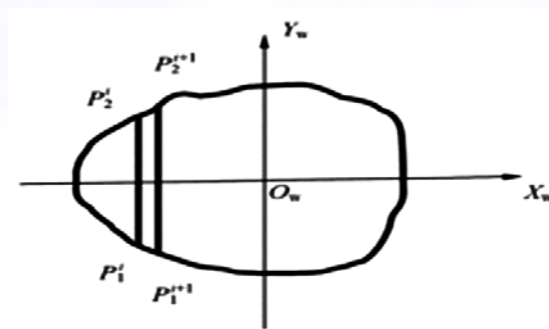
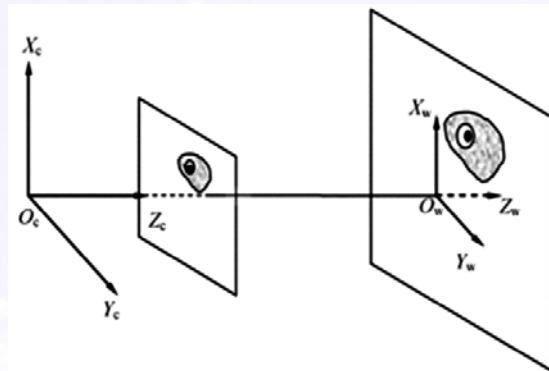
- 二维码实际面积 S 与图像中面积 s_0 的关系：

$$S = \sum_{i=1}^N (P_{2y}^i - P_{1y}^i)(P_{1x}^{i+1} - P_{1x}^i) = \sum_{i=1}^N \frac{(v_2^i - v_1^i)(u_1^{i+1} - u_1^i)d^2}{\alpha_x \alpha_y} = \frac{s_0 * d^2}{\alpha_x \alpha_y}$$

$$\rightarrow d = \sqrt{\alpha_x \alpha_y \frac{S}{s_0}}$$

→ 相机相对于二维码的位置坐标

→ 由于二维码坐标已知，从而得到相机在世界坐标系中的位置坐标





谢谢!