

一、相机模型

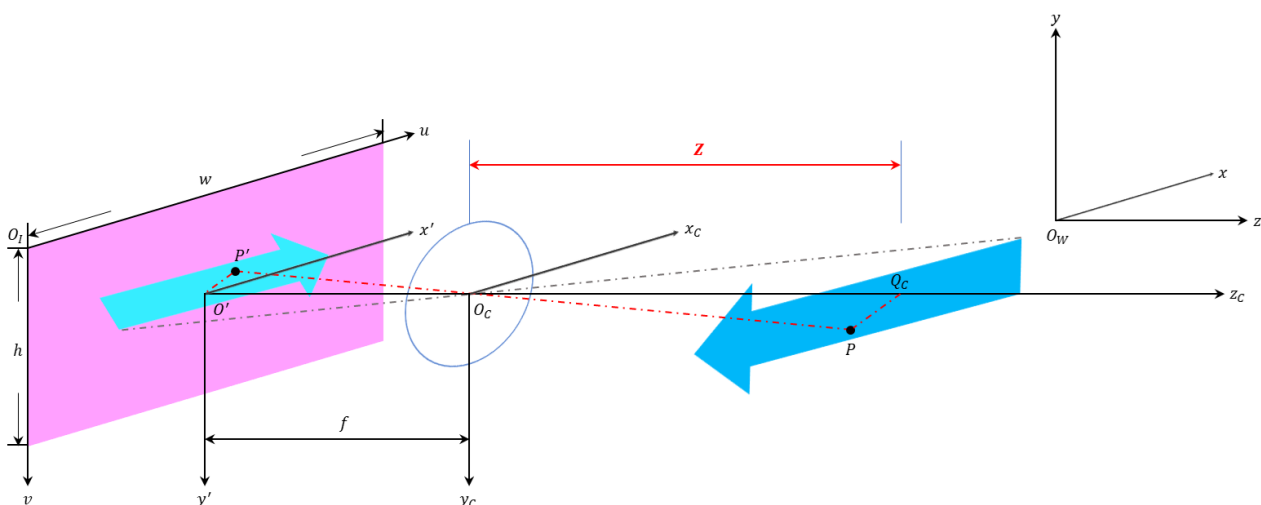


Figure 1 相机模型

相机模型根本上是凸透镜成像模型。而对于计算机视觉而言，能够将数字图像中每一个特定像素与其对应真实物体联系清晰准确地描述是我们通过图像了解和认识世界的基础，也是图像处理和研究的意义所在。否则我们对图像处理获得的信息和结论将对我们认识世界毫无帮助。

为了阐明这一关系，我们引入坐标这一数学工具帮助我们解释。假设我们有一台固定相机、一个处在特定时空的物体和某一时刻用相机拍摄的物体的图像。

我们从图像中物体入手一步一步反推至对应特定时空下原物体的位置信息。我们随机选取数字图像中一点 $P'(v, u)$ ， (v, u) 为其在图像像素坐标系下的坐标值。

1.1 图像像素坐标系→图像物理坐标系

由于数字图像在像平面上生成，因此 O_Ivu 图像像素坐标系平面与 $O'y'x'$ 图像物理坐标系平面共面。那么根据平面坐标变换可得如下关系：

$$\begin{cases} v = \frac{y}{dy} + v_0 \\ u = \frac{x}{dx} + u_0 \end{cases}$$

(dx, dy) 为单位物理长度对应像素数， (v_0, u_0) 为 O' 点在图像像素坐标系下的坐标

将上式用齐次坐标表示即为：

$$\begin{bmatrix} v \\ u \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dy} & & v_0 \\ & \frac{1}{dx} & u_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} y \\ x \\ 1 \end{bmatrix} \dots \textcircled{1}$$

1.2 图像物理坐标系→相机坐标系

由于相机的镜头的主光轴(图中 z_c 轴)穿过像平面几何中心(即 O')， $O'O_c = f$ ，再加上一般而言我们认为相机平面与像平面是平行的，因此 $O'y'x'$ 图像物理坐标系平面与 $O_cy_cx_c$ 相机平面平行且量纲相同。

那么在 $O_cy_cx_cz_c$ 相机坐标系中， $P'(y, x, -f)$ ， $P(y_c, x_c, Z)$ ，根据 $\Delta P'O'O_c \sim \Delta PO_cQ_c$ ，可以得到如下数

量关系：

$$\begin{aligned}\frac{P'O'}{O'O_c} &= \frac{PO_c}{O_cQ_c} \\ \left(\frac{P'O'}{O'O_c}\right)^2 &= \left(\frac{PO_c}{O_cQ_c}\right)^2 \\ \frac{y^2}{f^2} + \frac{x^2}{f^2} &= \frac{y_c^2}{z^2} + \frac{x_c^2}{z^2}\end{aligned}$$

但这个式子无法得到确切的坐标关系，不过却给了我们提示。将两个三角形分别投影到 Oyz 和 Oxz 平面上，得到的两对新的三角形分别相似。则有

$$\begin{cases} \frac{y}{f} = \frac{y_c}{z} \\ \frac{x}{f} = \frac{x_c}{z} \end{cases}$$

同样用齐次坐标表示如下：

$$Z \begin{bmatrix} y \\ x \\ 1 \end{bmatrix} = \begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix} \begin{bmatrix} y_c \\ x_c \\ 1 \end{bmatrix} \dots \textcircled{2}$$

1.3 相机坐标系→世界坐标系

这里我们遇到了一个挑战，就是相对于前三个坐标系而言，世界坐标系更加“自由独立”，世界坐标系可以与相机坐标系有更多的差异，比如旋转的差异(假设都是三维直角坐标系)。

那么首先我们就要简要介绍一下欧式变换公式：

$$C = R_{3 \times 3}W + T_{3 \times 1}$$

其中 $R_{3 \times 3}$ 为旋转矩阵， $T_{3 \times 1}$ 为平移矩阵。

平移很容易理解，那么旋转矩阵是如何达到其效果，这里我参考学习了一片博客：

[双目立体视觉系统——by AndyCheng hgcc](#)

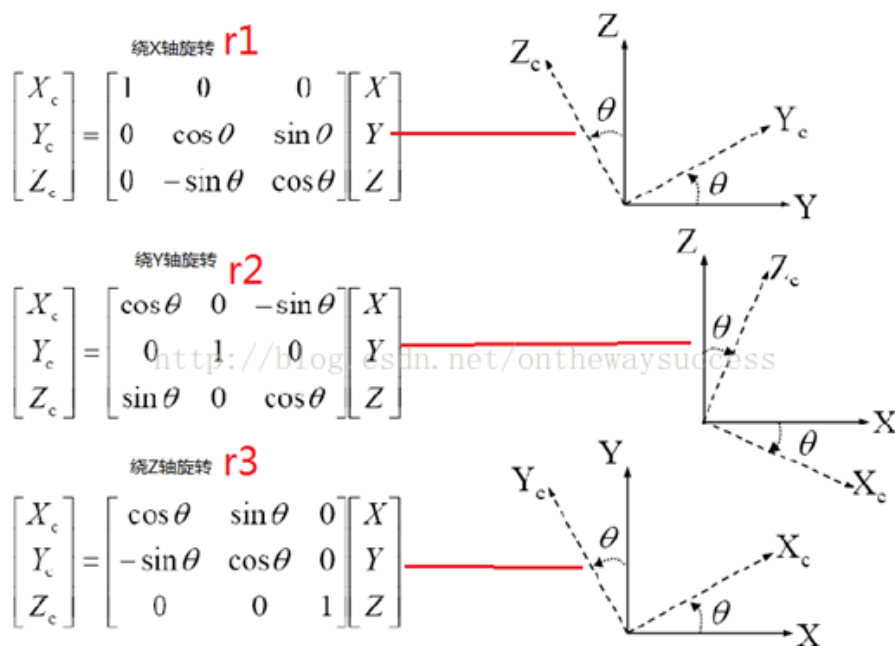


Figure 2 旋转矩阵及效果图示

旋转矩阵 $R_{3 \times 3} = r_1 r_2 r_3$ ，认识了欧式变换之后，我们可以完成相机坐标系→世界坐标系的坐标变换了：

$$\begin{bmatrix} y_C \\ x_C \\ z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} y_W \\ x_W \\ z_W \\ 1 \end{bmatrix} \dots \textcircled{3}$$

1.4 图像像素坐标系→世界坐标系

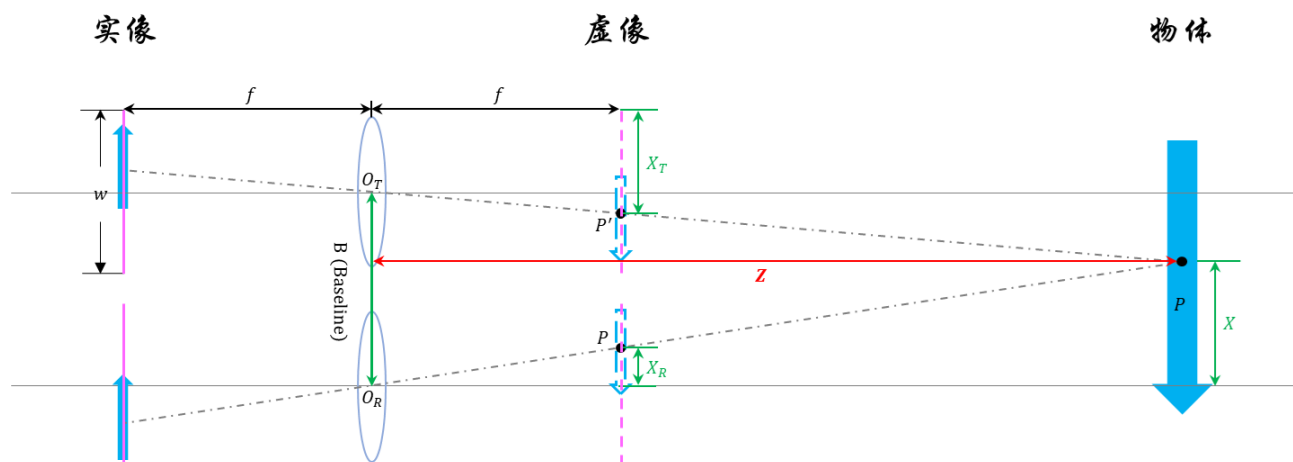
至此，根据①②③式，我们可以得到：

$$Z \begin{bmatrix} v \\ u \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dy} & v_0 \\ & \frac{1}{dx} & u_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} f & & \\ & f & \\ & & 1 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} y_W \\ x_W \\ z_W \\ 1 \end{bmatrix} \dots \textcircled{4}$$

我们已经成功地将图片中反应现实物体的某些信息通过数学上的直角坐标系联系了起来，然而，我们发现④式中右侧世界坐标系下的物体远近距离信息在经过变换的过程中消失了，以至于在目前条件下似乎无法判断图像中不同物体的“深度”方面的相对位置信息。

不过不要慌，我们可以通过下面的双目系统，实现“深度”信息的获得。

二、双目系统



相机利用凸透镜成像原理会在镜片另一侧成倒立缩小的实像，实像关于光心中心对称操作可在透镜和物体之间形成一个正立的虚像(像平面亦可同样中心对称称为虚平面)。

我们的目标是求得物体(Object)与基线(Baseline)之间距离 Z 值。

已知：

物理量	像平面尺寸	Baseline 长度	焦距	P'到像平面边距	P 到光轴距离
值	w	B	f	X_T	X_R

根据 $\triangle PO_T O_R \sim \triangle OP' P$ 得到如下数量关系：

$$\frac{P'P}{B} = \frac{Z-f}{Z}$$

$$\frac{B - \left(\frac{w}{2} - X_R\right) - \left(X_T - \frac{w}{2}\right)}{B} = 1 - \frac{f}{Z}$$

$$\frac{X_R - X_T}{B} = \frac{f}{Z}$$

$$Z = \frac{B \times f}{X_R - X_T}$$

令 $D = X_R - X_T$

$$Z = \frac{B \times f}{D}$$

D 即为视差(disparity)



我们看到正是由于视差存在让生物和具有“双目”系统的器械能够从接收到的图像中“感知”到不同物体“深度”方面的相对位置信息，而且这种感知力与视差大小反相关。