Camera Model Note

FanmingL

2018 年 11 月 28 日

图 1: 示例图

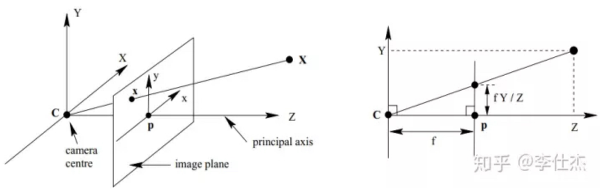

图 2: 相机坐标系下的坐标到像素坐标的转换，图源水印

1 相机模型

1.1 问题的提出

如图1，当我们拍摄地面时，若我们知道图中白线的规格（长度以及间隔），如何推算图中所有像素点在地面上的坐标（当然这要先假设图中所有物体都在地面上）?

要解决这个问题，还得首先从相机模型开始理解。

1.2 相机模型

1.2.1 相机内参

就假设我们的成像满足小孔成像模型，见图2。由这个模型很容易推出，在相机坐标系下坐标为 $[x, y, z]^T$ 的点，是如何映射到像素坐标上的。

见图2的右图，其中右上角的黑点表示我们观测的点，经过 p 点的竖线表示我们的感光片，我们的光线打到感光片上就能成像，在感光片上的位置可以由这个模型去确认，图2的右图中标出了 Y, Z 轴的方向，这就是相机坐标系的 Y, Z 轴，cp 的长度为焦距，通过相似三角形原则我们可以确定，该物体在像素平面上的纵的像素坐标 v 为

$$v = \frac{fy}{z} \tag{1}$$

同理，横的像素坐标 u

$$u = \frac{fx}{z} \tag{2}$$

上面求出的是以相机正中心为原点求出的坐标，而我们的图像的原点在图像的左上角，所以还需要加个偏移量，使我们求出图像原点在左上角的坐标，假设偏移分别为 u_0, v_0。

$$v = \frac{fy}{z} + v_0 \tag{3}$$

$$u = \frac{fx}{z} + u_0 \tag{4}$$

但是到这里还不是我们想要的，因为以上坐标的单位是厘米或者毫米，我们需要的是像素坐标，要转化为像素坐标，就需要知道感光片上的传感器的间隔，即传感器之间的的实际间隔，假设分别为 d_u, d_v。

$$v = \frac{\frac{fy}{z} + v_0}{d_v} \tag{5}$$

$$u = \frac{\frac{fx}{z} + u_0}{d_u} \tag{6}$$

稍作整理

$$v = \frac{\frac{f}{d_v}}{z} y + \frac{v_0}{d_v} \tag{7}$$

$$u = \frac{\frac{f}{d_u}}{z} x + \frac{u_0}{d_u} \tag{8}$$

令

$$\begin{gathered} f_y = \frac{f}{d_v} \\ f_x = \frac{f}{d_u} \\ c_y = \frac{v_0}{d_v} \\ c_x = \frac{u_0}{d_u} \end{gathered} \tag{9}$$

得到

$$v = \frac{f_y}{z} y + c_y \tag{10}$$

$$u = \frac{f_x}{z} x + c_x \tag{11}$$

其实这个能够写成矩阵形式

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{12}$$

令

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{13}$$

写得更原始点就是

$$
\mathbf{K}=\begin{bmatrix} \frac{f}{d_u} & 0 & \frac{u_0}{d_u} \\ 0 & \frac{f}{d_v} & \frac{v_0}{d_v} \\ 0 & 0 & 1 \end{bmatrix} \tag{14}
$$

因为 $\mathbf{K}$ 只与相机本身的性质（焦距、主点偏移量、传感器间隔）有关，我们称 $\mathbf{K}$ 为相机内参。

在我们调整相机焦距时 f 会发生改变，调整相机分辨率时 d_u, d_v 会发生改变，这点希望能够记住，调整相机分辨率之后实际上不需要重新标定相机，只需要对内参矩阵做一些小修改即可。

1.2.2 相机外参

上一节我们知道只要对相机坐标系下的坐标 $[x, y, z]$ 左乘一个相机内参矩阵，再把坐标转为齐次坐标（即把结果向量的第三项归一化），就能够得到像素坐标。但是，相机总是运动的，我们很难确定某个物体在相机坐标系下的坐标，但是我们能够给定一个世界坐标系，一个静止的物体在世界坐标系下的坐标肯定是固定的。

为了方便区分，令相机坐标系下的坐标为 $[x_c, y_c, z_c]^T$，世界坐标系下的坐标为 $[x_w, y_w, z_w]^T$。那么式12应当写为

$$
z_c\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}=\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \tag{15}
$$

我们知道在三维世界中，一个坐标系到另一个坐标系的变换可以由一个旋转和一个平移表示，即

$$
\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}=\mathbf{R}\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}+\mathbf{t} \tag{16}
$$

其中 $\mathbf{R}$ 为旋转矩阵，$\mathbf{t}$ 为平移向量，$\mathbf{R}$ 有以下性质

$$
\begin{aligned} \mathbf{R}^T\mathbf{R} &= I \\ |\mathbf{R}| &= 1 \end{aligned} \tag{17}
$$

式16其实可以这样理解：先把世界坐标系做个旋转，使得三个基矢和相机坐标系的平行，再直接将世界坐标系的原点平移至相机坐标系处。为了方便，式16可写成矩阵形式。

$$
\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}=[\mathbf{R} \quad \mathbf{t}]\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{18}
$$

上式也可以写成齐次的形式

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{19}$$

令

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{20}$$

由于 $\mathbf{M}$ 只和相机的位姿有关，我们称 $\mathbf{M}$ 为相机外参，由于矩阵 $\mathbf{R}$ 的特殊正交性，$\mathbf{M}$ 的逆可以很快写出

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{R}^T & \mathbf{R}^T\mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{21}$$

至此我们可以将相机成像的全部流程用一个公式写出

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \left[\mathbf{R} \quad \mathbf{t} \right] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{22}$$

上述模型是很多单目视觉定位的基础，比如视觉 SLAM。

1.3 单应性矩阵

考虑一个特殊情况，我们拍摄的所有东西都在一个平面上，然后我们以这个平面为世界坐标系的 xy 平面，那么这个平面上所有的点的 z_w 都为零。代入至式22得到

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \left[\mathbf{R} \quad \mathbf{t} \right] \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} \tag{23}$$

我们把旋转矩阵分块写，令 $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3]$，上式可写成如下

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} \tag{24}$$

因为有个零存在，上式可继续化简

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \tag{25}$$

令

$$\mathbf{H} = \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \tag{26}$$

遂式25可写为

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \tag{27}$$

我们称 $\mathbf{H}$ 为单应性矩阵。单应性变换通常描述从一个平面到另一个平面的变换，它的特点是，一个直线变换过去之后仍然是一根直线。

可以看出，该矩阵的尺度的变化（$\mathbf{H}$ 乘上一个标量）并不会影响这个变换关系，所以 $\mathbf{H}$ 有八个自由度，要解出它来，至少需要八个方程。而解出它之后有两种使用方法。

第一种是我们在开篇设定的问题，匹配到若干特征点之后如何将其他像素点也映射到目标平面。

第二种是估计相机的位置，对式26两边同时乘以 $\mathbf{K}^{-1}$

$$\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} = \mathbf{K}^{-1}\mathbf{H} \tag{28}$$

由于 $\mathbf{R}$ 的正交特性，我们可以很快由 $\mathbf{r}_1, \mathbf{r}_2$ 求出 $\mathbf{r}_3$

$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2 \tag{29}$$

其中 $\times$ 代表叉乘（外积）。

求得旋转矩阵 $\mathbf{R}$ 后，我们可以得到其旋转对应的四元数。

$$\begin{aligned} q_0 &= \frac{\sqrt{1+Tr(\mathbf{R})}}{2} \\ q_1 &= \frac{R_{32}-R_{23}}{4q_0} \\ q_2 &= \frac{R_{13}-R_{31}}{4q_0} \\ q_3 &= \frac{R_{21}-R_{12}}{4q_0} \end{aligned} \tag{30}$$

相机相对于世界坐标系的旋转、平移便可由此估计出来！

1.4 单应性矩阵的最小二乘解

这部分主要推导当已经有若干点对之时，如何估计单应性矩阵 $\mathbf{H}$

1.4.1 优化目标

假设我们有 m 个点对 $(m >= 4)$，我们可以得到 m 个式27，写成式子的形式即

$$\begin{aligned} u_i &= \frac{H_{11}x_i + H_{12}y_i + H_{13}}{H_{31}x_i + H_{32}y_i + H_{33}} \\ v_i &= \frac{H_{21}x_i + H_{22}y_i + H_{23}}{H_{31}x_i + H_{32}y_i + H_{33}} \end{aligned} \tag{31}$$

分母乘到左边，并移项。

$$
\begin{aligned}
H_{31}x_iu_i + H_{32}y_iu_i + H_{33}u_i - H_{11}x_i - H_{12}y_i - H_{13} = 0 \\
H_{31}x_iv_i + H_{32}y_iv_i + H_{33}v_i - H_{21}x_i - H_{22}y_i - H_{23} = 0
\end{aligned} \tag{32}
$$

稍作整理

$$
\begin{aligned}
-H_{11}x_i - H_{12}y_i - H_{13} + H_{31}x_iu_i + H_{32}y_iu_i + H_{33}u_i = 0 \\
-H_{21}x_i - H_{22}y_i - H_{23} + H_{31}x_iv_i + H_{32}y_iv_i + H_{33}v_i = 0
\end{aligned} \tag{33}
$$

又可以写成矩阵的形式

$$
\begin{aligned}
\mathbf{a}_{xi}^T\mathbf{h} = 0 \\
\mathbf{a}_{yi}^T\mathbf{h} = 0
\end{aligned} \tag{34}
$$

其中

$$
\begin{aligned}
\mathbf{h} &= (H_{11}, H_{12}, H_{13}, H_{21}, H_{22}, H_{23}, H_{31}, H_{32}, H_{33})^T \\
\mathbf{a}_{xi} &= (-x_i, -y_i, -1, 0, 0, 0, x_iu_i, y_iu_i, u_i)^T \\
\mathbf{a}_{yi} &= (0, 0, 0, -x_i, -y_i, -1, x_iv_i, y_iv_i, v_i)^T
\end{aligned} \tag{35}
$$

我们的优化对象即

$$
\min \quad \sum_{i=1}^{m} (\mathbf{a}_{xi}^T\mathbf{h})^2 + (\mathbf{a}_{yi}^T\mathbf{h})^2 \tag{36}
$$

将上式写成矩阵形式即

$$
\min \quad \|\mathbf{A}\mathbf{h}\|_2^2 \tag{37}
$$

其中

$$
A = \begin{bmatrix} \mathbf{a}_{x1}^T \\ \mathbf{a}_{y1}^T \\ \vdots \\ \mathbf{a}_{xm}^T \\ \mathbf{a}_{ym}^T \end{bmatrix} \tag{38}
$$

式37就是我们的优化对象。

1.4.2 用 SVD 求解问题

把式37写成矩阵乘法的形式

$$
\min \quad J(\mathbf{h}) = \mathbf{h}^T\mathbf{A}^T\mathbf{A}\mathbf{h} \tag{39}
$$

令 $J(\mathbf{h})$ 对 $\mathbf{h}$ 求偏导。

$$
\frac{\partial J(\mathbf{h})}{\partial \mathbf{h}} = 2\mathbf{A}^T\mathbf{A}\mathbf{h} \tag{40}
$$

令偏导为零

$$
\mathbf{A}^T\mathbf{A}\mathbf{h} = 0 \tag{41}
$$

这实际上是一个求解齐次方程组的问题，$\mathbf{A}^T\mathbf{A}$ 为 9×9 的矩阵，若 $\mathbf{A}^T\mathbf{A}$ 的秩为 8，那么我们可以得到一个 $\mathbf{h}$ 的非平凡解，这样就再好不过了。若 $\mathbf{A}^T\mathbf{A}$ 的秩小于 8，这种情况下，我们不能得到一个准确的答案。但是若我们提供 4 对以上不共线的点对构成 $\mathbf{A}$，$\mathbf{A}^T\mathbf{A}$ 往

往是满秩的，$\mathbf{h}$ 只有零解，所以我们需要对 $\mathbf{A}^T\mathbf{A}$ 进行些改造，让其不满秩。如何改造呢?首先观察 $\mathbf{A}$ 的奇异值分解的形式。

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T \tag{42}$$

$\mathbf{A}^T\mathbf{A}$ 便可化为

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Sigma^T\Sigma\mathbf{V}^T \tag{43}$$

设 Σ 的第 i 行 i 列的值（即，$\mathbf{A}$ 的第 i 个奇异值) 为 σ_i。$\mathbf{v}_i$ 为 $\mathbf{V}$ 的第 i 列。上式化为

$$\mathbf{A}^T\mathbf{A} = \sum_{i=1}^{9}\sigma_i^2\mathbf{v}_i\mathbf{v}_i^T \tag{44}$$

$\mathbf{v}_i\mathbf{v}_i^T$ 的秩必然为 1，上式的含义为 $\mathbf{A}^T\mathbf{A}$ 由 9 个秩为 1 的矩阵相加得到，所以使 $\mathbf{A}^T\mathbf{A}$ 不满秩的最简单的方法即，去掉其中一个矩阵，使新的矩阵由剩下的 8 个矩阵相加得到。8 个秩为 1 的矩阵之和的矩阵，其秩必然小于等于 8，而 $\mathbf{A}^T\mathbf{A}$ 为 9×9 的矩阵，此时 $\mathbf{A}^T\mathbf{A}$ 必不满秩。

假设我们去掉第 ℓ 个矩阵，式41化为

$$\begin{aligned}(\mathbf{A}^T\mathbf{A} - \sigma_\ell^2\mathbf{v}_\ell\mathbf{v}_\ell^T)\mathbf{h} &= 0\\ \mathbf{A}^T\mathbf{A}\mathbf{h} &= \sigma_\ell^2\mathbf{v}_\ell\mathbf{v}_\ell^T\mathbf{h}\end{aligned} \tag{45}$$

现在问题的解转化为矩阵 $\mathbf{A}^T\mathbf{A}, \mathbf{v}_\ell\mathbf{v}_\ell^T$ 的泛化特征值的问题，不过，这里情况特殊，我们可以直接写出一个解。即

$$\mathbf{h} = \mathbf{v}_\ell \tag{46}$$

我们可以验证一下这个解的正确性

$$\begin{aligned}\mathbf{A}^T\mathbf{A}\mathbf{h} &= \sum_{i=1}^{9}\sigma_i^2\mathbf{v}_i\mathbf{v}_i^T\mathbf{h}\\ &= \sigma_\ell^2\mathbf{v}_\ell\\ \sigma_\ell^2\mathbf{v}_\ell\mathbf{v}_\ell^T\mathbf{h} &= \sigma_\ell^2\mathbf{v}_\ell\mathbf{v}_\ell^T\mathbf{v}_\ell\\ &= \sigma_\ell^2\mathbf{v}_\ell\end{aligned} \tag{47}$$

上式用到了 $\mathbf{V}$ 的正交性，即

$$\mathbf{v}_i^T\mathbf{v}_j = \delta_{ij} \tag{48}$$

当且仅当 $i=j$ 时 $\delta_{ij}=1$，其余 $\delta_{ij}=1$。

那么应该怎么取 ℓ 呢？我们把式46代回到式39中，得到

$$\begin{aligned}&\min \quad \mathbf{v}_\ell^T\mathbf{A}^T\mathbf{A}\mathbf{v}_\ell\\ &\min \quad \mathbf{v}_\ell^T\sigma_\ell^2\mathbf{v}_\ell\\ &\min \quad \sigma_\ell^2\end{aligned} \tag{49}$$

所以我们要取最小的奇异值对应的下标作为 ℓ，得益于 SVD 算法的高效性，在 Σ 中，奇异值已经按从大到小的顺序排列好了，σ_9 就是最小的奇异值，所以 $\ell=9$，所以，

$$\mathbf{h} = \mathbf{v}_9 \tag{50}$$

即 $\mathbf{V}$ 的第 9 列。

1.4.3 总结流程

所以我们拿到 $m(m >= 4)$ 对不共线的点对之后，首先将其整理成式38的形式，然后对 $\mathbf{A}$ 进行奇异值分解，取 $\mathbf{V}$ 的最后一列作为 $\mathbf{h}$ 的估计量即可！