

附录二

摄像机成像模型

一点说明

从图像恢复空间结构以及视觉定位等, 都需要利用摄像机成像模型. 所以对图像的成像有一定的了解是计算机视觉研究人员所必备的知识. 本节是国科大研究生课程:《计算机视觉》讲义的附录二, 旨在介绍摄像机成像的几何投影过程.

附录中的很多内容均是中科院自动化研究所 吴福朝 研究员撰写的. 尽管笔者也参与了一些编写, 但主体工作都是吴福朝研究员完成的, 在此特别强调说明, 并对吴老师致以谢意.

胡占义

2019-8-23

本节的主要内容有:

第一节, 介绍作为中心投影的摄像机模型, 这种模型可由投影矩阵(摄像机矩阵)得到完全描述. 摄像机矩阵实现三维空间到二维空间(摄像机像平面)的一个射影变换, 具有 11 个自由度. 在本节, 还介绍中心投影摄像机的几何结构(即摄像机矩阵元素所具有的几何意义)、从摄像机矩阵分解出内参数与外参数的方法以及估计摄像机矩阵的线性算法与非线性迭代算法。

第二节, 着重介绍中心投影摄像机对某些空间几何元素(点、线、面、二次曲线与二次曲面)的投影性质, 以及对像平面的几何元素(点、线、二次曲线)的反投影性质。

摄像机模型

2.1.1 摄像机模型

基本针孔模型

考虑空间到平面 π 的中心投影(Central Projection)。假定投影中心在 O_c 点, 它到平面 π 的距离为 f 。空间点 X_c 在平面 π 上的投影点(或像点) m 是以点 O_c 为端点且通过点 X_c 的射线与平面 π 的交点, 如图 2.1.1(a)所示。平面 π 称为摄像机的像平面(Image Plane), 点 O_c 称为摄像机的中心(或光心), f 称为摄像机的焦距(Focal Length), 以点 O_c 为端点且垂直于像平面的射线称为光轴或主轴(Principal Axis), 主轴与像平面的交点 p 称为摄像机的主点(Principal Point)。

为了在代数上描述这种投影几何关系, 需要建立摄像机(欧氏)坐标系和像平面(欧氏)坐标系。以主点 p 为像平面的坐标原点 o , 以水平线与铅直线分别为 x 轴和 y 轴, 建立图像坐标系 $o-xy$ 。以摄像机中心 O_c 为坐标原点, 以主轴为 z_c 轴, 以平行于 x 轴且通过摄像机中心 O_c 的直线为 x_c 轴, 以平行于 y 轴且通过摄像机中心 O_c 的直线为轴 y_c , 建立摄像机坐标系 $O_c-x_c y_c z_c$, 如图 2.1.1(b)所示。

空间点 X_c 在摄像机坐标系中的欧氏坐标记为 $\tilde{X}_c = (x_c, y_c, z_c)^T$, 它的像点 m 在图像坐标系中的坐标记为 $\tilde{m} = (x, y)^T$ 。从图 2.1.1(b)不难看出空间点 X_c 与它的像点 m 满足下述关系:

$$\begin{cases} x = \frac{fx_c}{z_c} \\ y = \frac{fy_c}{z_c} \end{cases} \quad (2.1.1)$$

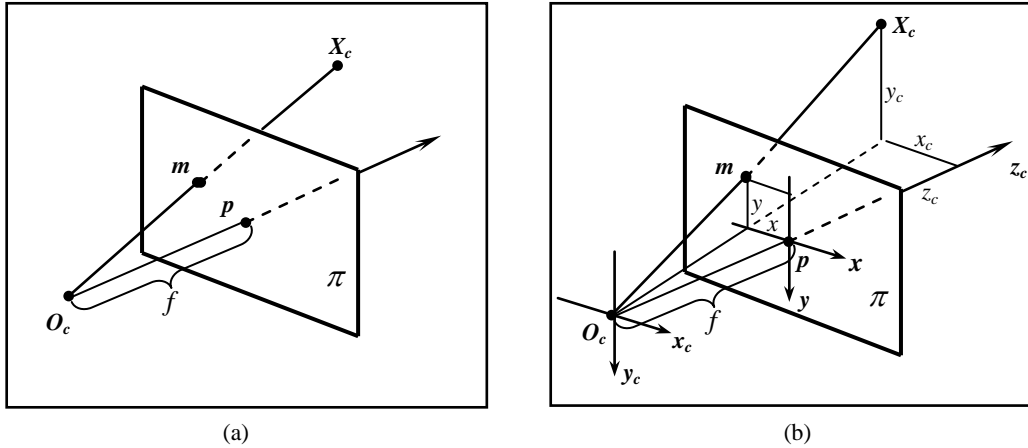


图 2.1.1: (a)基本针孔模型; (b) 根据三角形相似原理有 $x = \frac{fx_c}{z_c}, y = \frac{fy_c}{z_c}$

如果空间点与图像点都用齐次坐标来表示, 即 $X_c = (x_c, y_c, z_c, 1)^T$, $m = (x, y, 1)^T$, 则上述关系表示为

$$z_c \mathbf{m} = \begin{pmatrix} fx_c \\ fy_c \\ z_c \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{X}_c \quad (2.1.2)$$

这是从空间点到图像点的齐次线性变换。记

$$P = \text{diag}(f, f, 1)(I, \mathbf{0}) \quad (2.1.3)$$

则齐次线性变换(2.1.2)可表示为

$$\mathbf{m} = P\mathbf{X}_c \quad (2.1.4)$$

矩阵 P 是一个 3×4 矩阵，通常称它为摄像机矩阵(Camera Matrix)。

注意，(2.1.4)是一个齐次等式，表示在相差一个非零常数因子的意义下相等。

注：一般情况下图像点与空间点都使用齐次坐标，因此所涉及到的等式大多数都是齐次等式，齐次等式仍用等号“=”来表示。在某些情况下，例如：利用齐次等式进行计算和推理，如果齐次等式中的常数因子有特别的作用时，将显式地写出这个常数因子而成为常规的等式。对此，读者在阅读本书时应特别注意。

主点不在图像中心

由于事先不知道主点的确切位置，通常无法建立上述图像平面坐标系。在实际中，都是以图像中心或者图像的左上角作为图像坐标系的原点来建立图像坐标系的（在本书中除特别说明外都是以图像中心作为图像坐标系的原点），在此坐标系下主点可能不在图像中心，因此摄像机矩阵可能不再具有式 (2.1.3)的形式。

假定主点在像平面的坐标为 $\mathbf{p} = (x_0, y_0, 1)^T$ ，则摄像机的投影关系满足：

$$\begin{cases} x = \frac{fx_c}{z_c} + x_0 \\ y = \frac{fy_c}{z_c} + y_0 \end{cases} \quad (2.1.5)$$

使用齐次坐标，上式可表示为

$$z_c \mathbf{m} = \begin{pmatrix} fx_c \\ fy_c \\ z_c \end{pmatrix} = \begin{pmatrix} f & 0 & x_0 & 0 \\ 0 & f & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{X}_c \quad (2.1.6)$$

则摄像机矩阵具有下述形式：

$$P = K(I, \mathbf{0}) \quad (2.1.7)$$

其中

$$K = \begin{pmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.1.8)$$

并称它为摄像机内参数矩阵(Intrinsic Parameters Matrix)。摄像机的投影关系可写成下面的形式：

$$\mathbf{m} = \mathbf{P}\mathbf{X}_c \quad (2.1.9)$$

CCD 摄像机

用于计算机处理的图像通常都是用 CCD 摄像机所获取的数字图像,或者是由模拟信号摄像机获取的图像再经过特别设备进行数字离散化后的数字图像。一般地, CCD 摄像机内参数矩阵不具有(2.1.8)的形式。为了得到 CCD 摄像机的模型,必须刻划 CCD 摄像机的数字离散化过程。

假定 CCD 摄像机数字离散化后的像素是一个矩形,矩形的长与宽分别为 d_x, d_y , 如图 2.1.2 所示。

设图像点 $(x, y, 1)^T$ 在离散化后的坐标为 $(u, v, 1)^T$, 则必有

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} 1/d_x & 0 & 0 \\ 0 & 1/d_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

将上式代入(2.1.6)式,可以得到

$$\mathbf{z}_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} 1/d_x & 0 & 0 \\ 0 & 1/d_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & x_0 & 0 \\ 0 & f & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{X}_c = \begin{pmatrix} f/d_x & 0 & x_0/d_x & 0 \\ 0 & f/d_y & y_0/d_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{X}_c$$

如果离散化后的图像坐标仍用 \mathbf{m} 表示,则上式可写成

$$\mathbf{m} = \mathbf{K}(\mathbf{I}, 0)\mathbf{X}_c = \mathbf{P}\mathbf{X}_c \quad (2.1.10)$$

其中

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.1.11)$$

矩阵 \mathbf{K} 称为 CCD 摄像机的内参数矩阵, $f_x = f/d_x, f_y = f/d_y$ 称为 CCD 摄像机在 u 轴和 v 轴方向上的尺度因子(Measure Factors), $(u_0, v_0)^T = (x_0/d_x, y_0/d_y)^T$ 称为 CCD 摄像机的主点(Principal Point)。

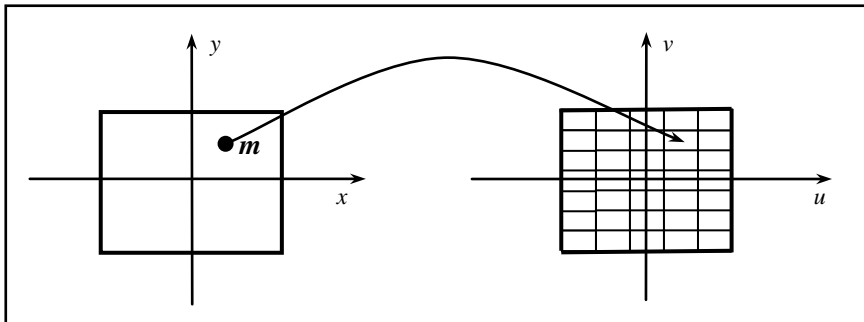


图 2.1.2: 图像平面的数字离散化变换 (矩形像素)

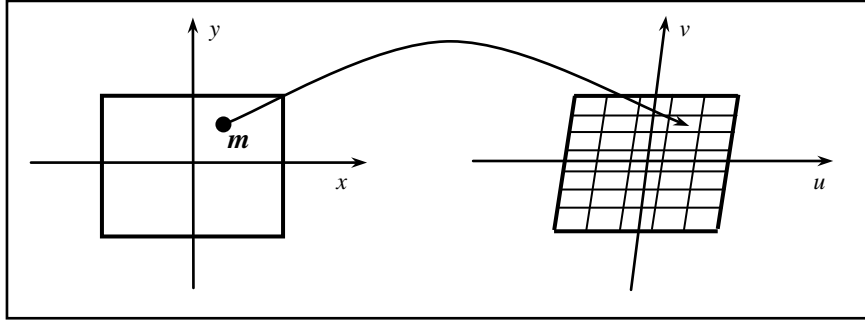


图 2.1.3: 图像平面的数字离散化变换 (平行四边形像素)

由于制造工艺的限制, 一般情况下, CCD 摄像机数字离散化后的像素不是一个矩形而是一个平行四边形, 四边形的-边平行于 u 轴, 而另一边与 u 轴构成一个 $\theta (\neq \frac{\pi}{2})$ 角。令平行四边形两边长分别为 d_x, d_y , 如图 2.1.3 所示。

设图像点 $(x, y, 1)^T$ 离散化后的坐标为 $(u, v, 1)^T$, 则必有

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} 1/d_x & -ctg\theta/d_x & 0 \\ 0 & 1/d_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

结合(2.1.6)式, 可得到

$$z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f/d_x & -fctg\theta/d_x & (x_0 - y_0ctg\theta)/d_x & 0 \\ 0 & f/d_y & y_0/d_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} X_c$$

如果离散化后的图像坐标仍用 m 表示, 则上式可写成

$$m = K(I, 0)X_c = PX_c$$

其中

$$K = \begin{pmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.1.12)$$

这里 $f_x = f/d_x, f_y = f/d_y$ 称为一般 CCD 摄像机在 u 轴和 v 轴方向上的尺度因子, $(u_0, v_0)^T = ((x_0 - y_0ctg\theta)/d_x, y_0/d_y)^T$ 称为一般 CCD 摄像机的主点, 而 $s = -fctg\theta/d_x$ 称为一般 CCD 摄像机的畸变因子或倾斜因子(Skew Factor)。

在本书中, 所涉及的摄像机均指 CCD 摄像机, 除特别说明外均假定摄像机内参数矩阵具有 (2.1.12) 的形式。

摄像机矩阵的一般形式

上面所介绍的摄像机矩阵是在摄像机坐标系下的结果。由于摄像机的中心和主轴等事先都是未知的, 这个坐标系不能给出空间点的具体坐标值, 另外摄像机可安放在环境中的任何位置。因此, 需要一个基准坐标系来描述空间点和摄像机的位置。这个基准坐标系通常称为世界坐标

系(World System)。

世界坐标系与摄像机坐标系之间的关系可以用一个旋转矩阵和一个平移向量来描述，如图 2.1.4 所示。令空间点在世界坐标系与摄像机坐标系的坐标分别为 $\mathbf{X} = (x, y, z, 1)^T$ ， $\mathbf{X}_c = (x_c, y_c, z_c, 1)^T$ ，则它们之间的关系为

$$\mathbf{X}_c = \begin{pmatrix} R & -R\tilde{C} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{X} \quad (2.1.13)$$

其中 \tilde{C} 表示摄像机中心在世界坐标系中的非齐次坐标，即摄像机中心的齐次坐标为 $\mathbf{C} = (\tilde{C}^T, 1)^T$ 。以后，点 \mathbf{X} 的非齐次坐标总用 $\tilde{\mathbf{X}}$ 来表示。

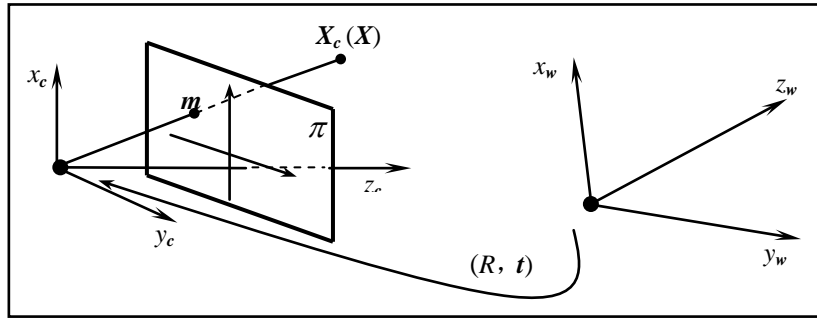


图 2.1.4: 世界坐标系与摄像机坐标系之间的欧氏变换

将(2.1.13)代入式(2.1.10)，则有

$$\mathbf{m} = K(I, 0) \begin{pmatrix} R & -R\tilde{C} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{X} = KR(I, -\tilde{C})\mathbf{X} \quad (2.1.14)$$

这样，就得到了摄像机矩阵的一般形式：

$$\mathbf{P} = KR(I, -\tilde{C}) \quad (2.1.15)$$

矩阵 $R(I, -\tilde{C})$ 称为摄像机的外参数矩阵(Extrinsic Parameters Matrix)。

有时，也用 $\tilde{\mathbf{X}}_c = R\tilde{\mathbf{X}} + \mathbf{T}$ 来描述世界坐标系与摄像机坐标系之间的关系，此时摄像机矩阵为

$$\mathbf{P} = K(R, \mathbf{T}) \quad (2.1.16)$$

其中， $\mathbf{T} = -R\tilde{C}$ 。

由于摄像机矩阵是一个齐次矩阵，所以在实际中所估计的摄像机矩阵 \mathbf{P} 通常与式(2.1.15)相差一个非零常数因子(关于摄像机矩阵的估计，见第 2.1.3 节)。事实上，若两个摄像机矩阵 $\mathbf{P}_1, \mathbf{P}_2$ 满足 $\mathbf{P}_1 = s\mathbf{P}_2$ ，令 $(u, v, w)^T = \mathbf{P}_1 \mathbf{X}$ ， $(u', v', w')^T = \mathbf{P}_2 \mathbf{X}$ ，则必有 $(u/w, v/w, 1)^T = (u'/w', v'/w', 1)^T$ ，即 $\mathbf{P}_1, \mathbf{P}_2$ 作用在同一空间点上的图像是相同的。另外值得注意的是，空间点在一个固定位置上的摄像机下的图像只与内参数矩阵有关，而与世界坐标系选择无关。也就是说，不论如何选择世界坐标系，空间点在像平面上的投影点的位置是不变的。

摄像机矩阵是一个秩 3 的 3×4 矩阵，因为它的前三列所构成的子矩阵是一个可逆矩阵。另

外，由于摄像机矩阵的齐次性，所以它仅有 11 个独立元素。

2.1.2 摄像机矩阵元素的几何意义

摄像机中心(Camera Centre)

考虑摄像机中心在世界坐标系中的坐标。由于

$$PC = KR(I, -\tilde{C}) \begin{pmatrix} \tilde{C} \\ 1 \end{pmatrix} = KR(\tilde{C} - \tilde{C}) = 0$$

所以，摄像机中心 $C = (\tilde{C}^T, 1)^T$ 是方程 $PC=0$ 的一个解。

由于摄像机矩阵 P 的秩等于 3，这导致 P 仅有一维右零空间。因此，摄像机中心的齐次坐标构成 P 的右零空间。在已知摄像机矩阵 P 的情况下，可以通过求解方程 $PX=0$ 而得到摄像机中心在世界坐标系中的坐标。

事实上，如果令 $P = (H, p_4)$ ，其中 H 为 P 的前三列所构成的 3×3 矩阵， p_4 是 P 的第四列向量，则 $PX=0$ 的解为

$$X = s \begin{pmatrix} -H^{-1}p_4 \\ 1 \end{pmatrix}$$

于是，摄像机中心在世界坐标系中的坐标为

$$C = \begin{pmatrix} -H^{-1}p_4 \\ 1 \end{pmatrix} \quad (2.1.17)$$

世界坐标系的原点与坐标轴方向

记摄像机矩阵为 $P = (p_1, p_2, p_3, p_4)$ ，其中 p_j 为 P 的第 j 列向量。

世界坐标系的原点 由于世界坐标系的原点坐标为 $X = (0, 0, 0, 1)^T$ ，所以它的图像点坐标为

$$s_0 m_0 = PX = (p_1, p_2, p_3, p_4) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = p_4$$

即摄像机矩阵的第 4 列向量是世界坐标原点的图像的齐次坐标。

坐标轴方向 下面考虑世界坐标系三个坐标轴方向的图像，即三个坐标轴与无穷远平面交点的图像。显然，三个坐标轴的方向为

$$X = (1, 0, 0, 0)^T, Y = (0, 1, 0, 0)^T, Z = (0, 0, 1, 0)^T$$

因此它们的图像分别为

$$s_1 m_1 = PX = p_1$$

$$s_2 m_2 = PY = p_2$$

$$s_3 m_3 = PZ = p_3$$

即摄像机矩阵的前三个列向量分别是世界坐标系三个坐标轴方向的图像点的齐次坐标。

如果已知世界坐标原点以及三个坐标轴方向的图像，则摄像机矩阵可以表示为

$$P = (s_1 \mathbf{m}_1, s_2 \mathbf{m}_2, s_3 \mathbf{m}_3, s_4 \mathbf{m}_4) \quad (2.1.18)$$

其中 s_j 是未知参数。由于 P 的齐次性，这些未知参数中只有三个是独立的，也就是说可以将其中的任一个置为 1。

主平面与轴平面

记摄像机矩阵为 $P = \begin{pmatrix} p^{1T} \\ p^{2T} \\ p^{3T} \end{pmatrix}$ ，其中 p^{1T}, p^{2T}, p^{3T} 分别为 P 的三个行向量。

主平面 (Principal Plane) 摄像机的坐标平面 $O_c - x_c y_c$ ，即与像平面平行的坐标平面，称为主平面。主平面在世界坐标系中可用摄像机矩阵的第 3 行向量 p^{3T} 来表示。因为主平面与像平面平行，所以它们的交线是一条无穷远直线，即主平面的像是像平面上的无穷远直线。令 X 为主平面上的任一点，则它在摄像机下的图像必为

$$\begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = PX = \begin{pmatrix} p^{1T} X \\ p^{2T} X \\ p^{3T} X \end{pmatrix}$$

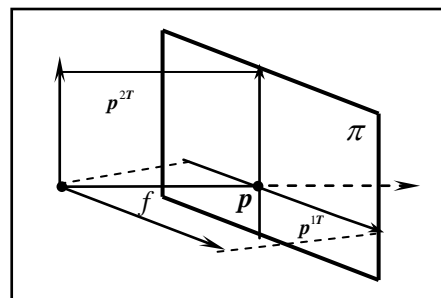
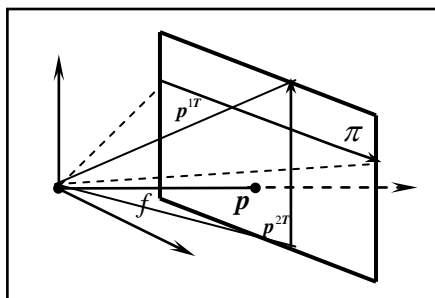
所以有 $p^{3T} X = 0$ ，于是主平面在世界坐标系中的坐标为摄像机矩阵的第 3 行 p^{3T} 。

轴平面 (Axis Plane) 考虑由方程 $p^{1T} X = 0$ 所确定的平面，即在世界坐标系中坐标为 p^{1T} 的平面。它是图像平面的 v 轴与摄像机中心所构成的平面，通常称它为轴平面。令 X 为这个轴平面上的任一点，则它在摄像机下的图像必为

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = PX = \begin{pmatrix} p^{1T} X \\ p^{2T} X \\ p^{3T} X \end{pmatrix} = \begin{pmatrix} 0 \\ p^{2T} X \\ p^{3T} X \end{pmatrix}$$

因此，这个轴平面的图像点集合是 $\{(0, v, 1)^T\}$ ，即像平面上的 v 轴。同理，由 $p^{2T} X = 0$ 所确定的另一个轴平面是图像平面的 u 轴与摄像机中心所构成的平面。

如图 2.1.5 所示，轴平面与主平面的不同之处在于它依赖于图像坐标系的选择，即对于不同的图像坐标系，对应的轴平面是不同的。如图 2.1.5(b) 所示。如果主点在图像中心，则轴平面 p^{1T} 正好是摄像机坐标轴 x_c, z_c 所构成的坐标平面，轴平面 p^{2T} 正好是 y_c, z_c 轴所构成的坐标平面。



(a)

(b)

图 2.1.5: 在 (a)、(b) 中, 图像坐标轴不相同, 对应的轴平面也不相同。

主轴与主点

主轴 主轴与主平面是正交的, 因此主轴必为主平面的法线。一般地, 一个平面 $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T$ 的法线有两个方向(正向与负向), 它们是 $\hat{n} = \pm(\pi_1, \pi_2, \pi_3)^T$, 在无穷远平面上它们表示同一个点, 即法线与无穷远平面的交点。由于主平面是 $\mathbf{p}^{3T} = (p_{31}, p_{32}, p_{33}, p_{34})^T$, 所以主轴两个方向为 $\hat{\mathbf{p}}^3 = \pm(p_{31}, p_{32}, p_{33})^T$ 。通常所讲的主轴方向是它的正方向, 即指向摄像机前方的方向, 由于摄像机矩阵可以相差一个常数, 因此 $\hat{\mathbf{p}}^{3T}$ 的正号并不代表主轴的正向。如果摄像机矩阵 $P = (H, \mathbf{p}_4)$ 与标准摄像机矩阵 $K(R, \mathbf{t})$ 相差一个正常数, 必有 $\det(H) > 0$, 否则 $\det(H) < 0$ 。因此主轴的正向是

$$\mathbf{v} = \det(H) \mathbf{h}^3 \quad (2.1.19)$$

其中 \mathbf{h}^{3T} 是矩阵 H 的第三行向量。

主点 主点是主轴与像平面的交点, 由于主轴过摄像机中心, 因此主点必为主轴方向的图像点, 所以主点坐标为

$$\mathbf{p} = P(\mathbf{h}^{3T}, 0)^T = H\mathbf{h}^3 \quad (2.1.20)$$

空间点的深度(Depth of Space Points)

令空间点 \mathbf{X} 在摄像机坐标系下的坐标为 $\mathbf{X}_c = (x_c, y_c, z_c, 1)^T$, 则该点的深度定义为

$$d(P, \mathbf{X}) = z_c$$

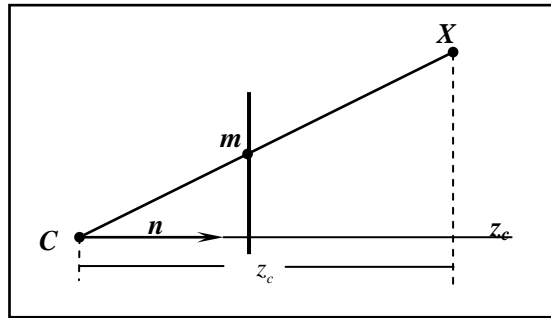
其中 P 表示摄像机矩阵。

下面给出深度与摄像机矩阵元素之间的关系。令空间点 \mathbf{X} 在世界坐标系中的坐标为 $\mathbf{X} = (x, y, z, 1)^T$, 它在摄像机矩阵下的图像点为 $\mathbf{m} = w(u, v, 1)^T$, 则必有

$$\mathbf{w} = \mathbf{p}^{3T} \mathbf{X} = \mathbf{p}^{3T} (\mathbf{X} - \mathbf{C}) = \mathbf{h}^{3T} (\tilde{\mathbf{X}} - \tilde{\mathbf{C}}) \quad (2.1.21)$$

上式中的第二个等式是因为摄像机光心 \mathbf{C} 满足 $\mathbf{P}\mathbf{C} = 0$ 。另一方面, 假定 \mathbf{n} 是主轴正向的单位向量(如图 2.1.6 所示), 则深度可表示为

$$d(P, \mathbf{X}) = z_c = \mathbf{n}^T (\tilde{\mathbf{X}} - \tilde{\mathbf{C}}) \quad (2.1.22)$$

图 2.1.6: 空间点的深度 z_c

由式(2.1.19)，可以得到

$$\mathbf{n} = \frac{\det(H)}{\|\det(H)\mathbf{h}^3\|} \mathbf{h}^3 = \frac{\text{sign}(\det(H))}{\|\mathbf{h}^3\|} \mathbf{h}^3$$

因此，

$$d(P, X) = \frac{\text{sign}(\det(H))w}{\|\mathbf{h}^3\|}$$

总结上述讨论，有下述命题：

命题 2.1.1 令空间点 X 在世界坐标系中的坐标为 $X = (x, y, z, 1)^T$ ，它在摄像机矩阵 $P = (H, \mathbf{p}_4)$ 下的图像点为 $\mathbf{m} = w(u, v, 1)^T$ ，则 X 的深度为

$$d(P, X) = \frac{\text{sign}(\det(H))w}{\|\mathbf{h}^3\|} \quad (2.1.23)$$

如果空间点使用一般化的齐次坐标 $X = (x, y, z, t)^T$ ，则深度可表示为

$$d(P, X) = \frac{\text{sign}(\det(H))w}{t \|\mathbf{h}^3\|} \quad (2.1.24)$$

摄像机矩阵的分解

给定摄像机矩阵 P ，通过矩阵分解的方法可以求出摄像机的内参数与外参数。令摄像机内参数矩阵为 K ，外参数矩阵为 (R, \mathbf{t}) ，则摄像机矩阵可以写成下述形式

$$P = \alpha K(R, \mathbf{t})$$

将 P 写成 $P = (H, \mathbf{p}_4)$ ，则有

$$H = \alpha KR, \quad \mathbf{p}_4 = \alpha K\mathbf{t}$$

对 H 作 RQ 分解： $H = \overline{K}\overline{R}$ ，其中 \overline{K} 是对角元均为正数的上三角矩阵， \overline{R} 为旋转矩阵，这种分解是唯一的。由于摄像机内参数矩阵最后一个元素为 1，所以，

$$K = \overline{k}_{33}^{-1} \overline{K}, \quad R = \overline{R}, \quad \mathbf{t} = \overline{k}_{33}^{-1} \mathbf{p}_4 \quad (2.1.25)$$

其中 \overline{k}_{33} 是矩阵 \overline{K} 的第(3, 3)元素。

2.1.3 摄像机矩阵的估计

估计摄像机矩阵的前提，是需要已知一些空间点在世界坐标系下的坐标和它们的图像坐标。

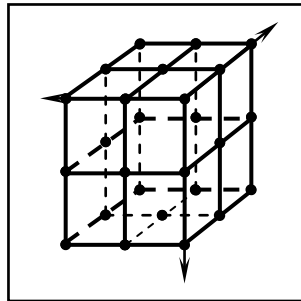


图 2.1.7：用于求解摄像机投影矩阵的立方体：正交的三条棱作为世界坐标系的三个坐标轴，各面上直线的交点的三维欧氏坐标已精确测定，根据这些特征点与其图像点的对应可以求解摄像机投影矩阵。

例如：在实践中可以使用一个标定物和它的图像求解摄像机矩阵。由于标定参考物上的每一个特征点相对于世界(欧氏)坐标系的位置在制作时都已经过精确测定(世界坐标系可以选为参考物的物体坐标系,如图 2.1.7 中参考物立方体的相互正交的三条棱可以作为世界坐标系的三个坐标轴),所以根据标定参考物上的特征点与其图像点的对应关系可以建立关于摄像机矩阵的约束方程,从而求解摄像机矩阵。

线性算法

线性算法有时又称为直接线性变换算法 (DLT Algorithm, [2]), 下面给出这种算法的基本思想。记摄像机矩阵为

$$P = \begin{pmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \\ \mathbf{p}^{3T} \end{pmatrix}$$

其中 \mathbf{p}^{jT} 为矩阵 P 的第 j 行向量。

令 $\mathbf{X}_j = (x_j, y_j, z_j, 1)^T$ 是特征点在世界坐标系下的坐标, 对应的图像点坐标为 $\mathbf{m}_j = (u_j, v_j, 1)^T$, 于是根据摄像机的投影关系, 得到

$$s_j \mathbf{m}_j = P \mathbf{X}_j = \begin{pmatrix} \mathbf{p}^{1T} \mathbf{X}_j \\ \mathbf{p}^{2T} \mathbf{X}_j \\ \mathbf{p}^{3T} \mathbf{X}_j \end{pmatrix} \quad (2.1.26)$$

因此, 消去上式中的常数因子后, 可得到下述方程:

$$\begin{pmatrix} \mathbf{p}^{1T} \mathbf{X}_j \\ \mathbf{p}^{2T} \mathbf{X}_j \\ \mathbf{p}^{3T} \mathbf{X}_j \end{pmatrix} \times \begin{pmatrix} u_j \\ v_j \\ 1 \end{pmatrix} = 0^*)$$

上式可以写成下述线性方程组的形式:

$$\begin{cases} \mathbf{p}^{2T} \mathbf{X}_j - v_j \mathbf{p}^{3T} \mathbf{X}_j = 0 \\ \mathbf{p}^{1T} \mathbf{X}_j - u_j \mathbf{p}^{3T} \mathbf{X}_j = 0 \\ v_j \mathbf{p}^{1T} \mathbf{X}_j - u_j \mathbf{p}^{2T} \mathbf{X}_j = 0 \end{cases}$$

在这个方程组中, 第三个方程可由前两个方程线性表示, 因此只有两个方程是线性独立的。即上述方程组与下面的方程组是等价的:

$$A_j \mathbf{p} = \begin{pmatrix} 0 & \mathbf{X}_j^T & -v_j \mathbf{X}_j^T \\ \mathbf{X}_j^T & 0 & -u_j \mathbf{X}_j^T \end{pmatrix} \begin{pmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{pmatrix} = 0 \quad (2.1.27)$$

*) 这里符号 “ \times ” 表示两个三维向量的叉积。关于它的定义及运算性质, 请参考附录 A2.1。本书其它地方, 除特别说明外 “ \times ” 均表示叉积。

因此，给定 $n(>5)$ 个以上的特征点与其图像点的对应，可求解摄像机矩阵 P 。

具体计算方法如下：

对每一个点对应，得到一个 2×12 的矩阵 A_j ，再将这 n 个矩阵组合起来得到一个 $2n \times 12$ 的矩阵 $A = (A_1^T \cdots A_n^T)^T$ ，对 A 作奇异值分解 $A = UDV^T$ ，则 $p = v_{12}$ (V 的最后一个列向量) 是方程 $Ap = 0$ 的最小二乘解^{*)}，再将 p 写成矩阵的形式就得到摄像机矩阵 P 。该算法概括在下述算法 2.1.1 中。

算法 2.1.1：求解摄像机矩阵的线性算法 (DLT 算法)

目标： 给定空间点与其图像点的 $N(\geq 6)$ 个对应 $\{X_j \leftrightarrow m_j\}$ ，计算摄像机矩阵 P

算法：

- 对每一个点对应 $X_j \leftrightarrow m_j$ ，构造矩阵 $A_j = \begin{pmatrix} 0 & X_j^T & -v_j X_j^T \\ X_j^T & 0 & -u_j X_j^T \end{pmatrix}$ ，再将所有的 A_j 排成一个 $2N \times 12$ 的矩阵 A ；
 - 对 A 进行奇异值分解： $A = UDV^T$ ；
 - 由矩阵 V 的最后一个列向量 v_{12} 得到摄像机矩阵 $P = \begin{pmatrix} v_{1,12} & v_{2,12} & v_{3,12} & v_{4,12} \\ v_{5,12} & v_{6,12} & v_{7,12} & v_{8,12} \\ v_{9,12} & v_{10,12} & v_{11,12} & v_{12,12} \end{pmatrix}$
-

非线性算法(Non-Linear Algorithm)

如果空间点和图像点都没有测量误差时，上述线性算法能对摄像机矩阵作出很好的估计。当空间点或图像点有测量误差时，就需要优化算法来估计摄像机矩阵。使用优化算法，首先要建立优化目标函数或称代价函数。

图像点误差 当空间点的测量数据无误差(例如利用精确的标定物体)时，数据误差的来源仅是图像点的数据测量误差。此时，可以将图像点的几何距离(Geometric Distance)的平方和：

$$\sum_j d^2(m_j, \hat{m}_j)$$

作为优化目标函数。其中： m_j 是空间点 X_j 的图像点的测量数据， \hat{m}_j 是由摄像机矩阵作用在空间点 X_j 上的理想图像点，即 $\hat{m}_j = PX_j$ 。注意，在这里 $m_j, \hat{m}_j (PX_j)$ 代表的是图像点，在计算距离时应该使用它的欧氏坐标。于是，估计摄像机矩阵问题就归结到求解下述优化问题：

$$\min_P \sum_j d^2(m_j, PX_j) \quad (2.1.27)$$

当图像点的测量误差服从高斯分布时，上述最小化的 P 是摄像机矩阵的最大似然估计(Maximum Likelihood Estimation (MLE))。

空间点误差 数据误差的来源仅是空间点的测量数据时，即图像点的数据无测量误差时，可以使用三维几何距离

^{*)} 参见附录 B1.4，推论 B.1.6。

$$\min \sum_j d^2(\mathbf{X}_j, \hat{\mathbf{X}}_j) \quad (2.1.28)$$

作为优化目标函数。这里 $\hat{\mathbf{X}}_j$ 是使得 $\mathbf{m}_j = P\hat{\mathbf{X}}_j$ 成立的空间点。与式(2.1.27)不同的是(2.1.28)是一个约束优化问题，因为要优化的对象是使得 $\mathbf{m}_j = P\hat{\mathbf{X}}_j$ 成立并最接近于测量点 \mathbf{X}_j 的空间点 $\hat{\mathbf{X}}_j$ 。

一般情况 在更一般的情况下，当空间点和图像点都存在测量误差时，需要同时优化(2.1.27)和(2.1.28)。由于图像点与空间点使用的不是同一个物理度量单位，所以不能简单地将(2.1.27)与(2.1.28)的和作为目标函数，此时需要对图像距离的平方与空间距离的平方进行加权求和，以此作为目标函数，或者更一般地使用 Mahalanobis(Mah) 距离^{*)}进行度量，即使用下述目标函数：

$$\min_P \sum_j (d_{Mah}^2(\mathbf{m}_j, P\hat{\mathbf{X}}_j) + d_{Mah}^2(\mathbf{X}_j, \hat{\mathbf{X}}_j)) \quad (2.1.29)$$

在求解优化问题时，还需要一种迭代算法和开始迭代的初值。对于迭代算法，可以使用 L-M 算法^{**)；}对于初值，一般可以将线性算法的估计结果作为初值。

2.2 摄像机对基本几何元素的作用

2.2.1 对点的作用

正向投影(Forward Projection)

正如上节所述，一个摄像机可以用它的矩阵 P 来表示。空间点 \mathbf{X} 被它作用到(通常说投影到)图像平面上的一个像点 \mathbf{m} ：

$$\mathbf{m} = P\mathbf{X}$$

这种投影关系，称为摄像机的正向投影，有时简称为投影。因此，一个空间点在摄像机的正向投影下仍然是(图像平面的)一个点。

在正向投影中，正如以后将要看到的那样，无穷远点的投影是非常重要的。无穷远点的齐次坐标可以表示为

$$\mathbf{X}_\infty = (\mathbf{d}^T, 0)^T$$

其中 \mathbf{d} 是一个三维向量，它代表了通过点 \mathbf{X}_∞ 的直线方向。在空间中通过同一无穷远点 $\mathbf{X}_\infty = (\mathbf{d}^T, 0)^T$ 的所有直线是相互平行的，因此它们有相同的方向。记摄像机矩阵为 $P = (H, \mathbf{p}_4)$ ，则无穷远点在摄像机下的投影为

$$P\mathbf{X}_\infty = (H, \mathbf{p}_4)\mathbf{X}_\infty = H\mathbf{d} \quad (2.2.1)$$

因此，无穷远点 $\mathbf{X}_\infty = (\mathbf{d}^T, 0)^T$ 的投影仅与摄像机矩阵的前三列有关，而与其第 4 列无关。

反向投影(Back Projection)

^{*)} Mahalanobis(Mah) 距离定义为： $d_M(\mathbf{X}, \mathbf{Y}) = (\mathbf{X} - \mathbf{Y})^T \Sigma^{-1} (\mathbf{X} - \mathbf{Y})$ ，其中： Σ 是正定矩阵，在实际应用中通常取它为数据的协方差矩阵。

^{**) 见附录 B2.6}

反向投影是针对图像平面而言的，图像平面上点 m 的反投影是指在一个摄像机的作用下具有像点 m 的所有空间点的集合，即

$$l_b = \{X \mid m = PX\}$$

从几何上不难看出，图像点的反投影是从摄像机中心出发并通过图像点 m 的一条射线。此时，将图像平面视为三维空间中的一张平面，即将图像点看作三维空间中的一个点，如图 2.2.1 所示。

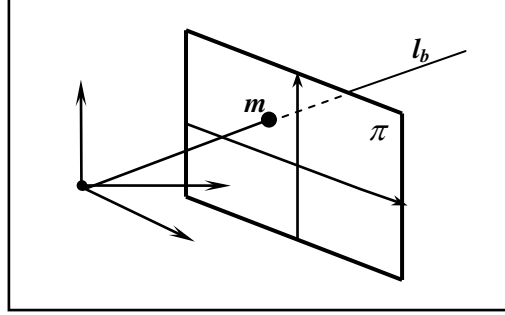


图 2.2.1: 图像点 m 的反投影是射线 l_b

考虑图像点 m 的反投影 l_b 在世界坐标系中的方程是非常重要的，因为三维计算机视觉的主要研究对象是从图像恢复三维几何信息。

如果能确定摄像机中心 C 和 l_b 上另一个点的三维坐标，就可以确定 l_b 在三维空间中的方程，因为两点可以确定一条直线。从式(2.1.17)，知道摄像机中心 C 在世界坐标系中的坐标为

$$C = \begin{pmatrix} -H^{-1}p_4 \\ 1 \end{pmatrix}$$

再考虑另一个空间点 P^+m ，它是由图像点 m 和 P 的广义逆 $P^+ = P^T(PP^T)^{-1}$ 所定义的。因此，必有

$$P(P^+m) = PP^T(PP^T)^{-1}m = m$$

所以， P^+m 是 l_b 上的点。于是，根据 l_b 上的两个点 C ， P^+m ，得到 l_b 的参数方程：

$$X(u) = u(P^+m) + C \quad (2.2.2)$$

利用 l_b 的无穷远点也可以确定它的参数方程。令 $X_\infty = (d^T, 0)^T$ 是这条射线上的无穷远点，则 $PX_\infty = m$ 。从式(2.2.1)，得到 $d = H^{-1}m$ ，于是 l_b 的参数方程可写成

$$X(u) = u \begin{pmatrix} H^{-1}m \\ 0 \end{pmatrix} + \begin{pmatrix} -H^{-1}p_4 \\ 1 \end{pmatrix} = \begin{pmatrix} H^{-1}(um - p_4) \\ 1 \end{pmatrix} \quad (2.2.3)$$

2.2.2 对线的作用

正向投影

空间直线 L 可由它上面的两个点来表示，即 L 的参数方程可写成

$$X(u) = X_1 + uX_2$$

其中 X_1, X_2 是 L 上的两个点。在摄像机矩阵 P 的作用下,

$$m(u) = P(X_1 + uX_2) = PX_1 + uPX_2 = m_1 + um_2 \quad (2.2.4)$$

于是, 直线 L 的图像是连结两个像点 m_1, m_2 的一条直线。因此, 空间直线在摄像机下的投影仍是一条直线。

下面应用空间直线的 Plucker 矩阵和 Plucker 坐标来描述空间直线的投影。

命题 2.2.1 如果空间直线 L 使用 Plucker 矩阵表示, 即 $L = X_1X_2^T - X_2X_1^T$, 其中 X_1, X_2 是 L 上的两个点, 则它的像直线 l 的坐标满足:

$$[l]_x = PLP^T \quad (2.2.5)$$

证明 记 $PX_1 = m_1, PX_2 = m_2$, 则 $l = m_1 \times m_2$ 。于是,

$$PLP^T = m_1m_2^T - m_2m_1^T = [m_1 \times m_2]_x = [l]_x$$

证毕。

对于像直线 l 的坐标, 也可以象点的投影一样由直线投影矩阵来给出。由摄像机矩阵 P 定义一个秩 3 的 3×6 矩阵:

$$P_L = \begin{pmatrix} P^2 \wedge P^3 \\ P^3 \wedge P^1 \\ P^1 \wedge P^2 \end{pmatrix} \quad (2.2.6)$$

其中 p^{jT} 为摄像机矩阵 P 的第 j 行向量, 它表示空间中的一张平面; $P^i \wedge P^j$ 是平面 $\{P^i, P^j\}$ 交线的 Plucker 坐标。根据下述命题 2.2.2, P_L 是空间直线的投影矩阵。

命题 2.2.2 如果空间直线 L 使用 Plucker 坐标, 则它在摄像机 P 下的像直线 l 是:

$$l = P_L L = \begin{pmatrix} (P^2 \wedge P^3 / L) \\ (P^3 \wedge P^1 / L) \\ (P^1 \wedge P^2 / L) \end{pmatrix} \quad (2.2.7)$$

证明 记 X_1, X_2 是 L 上的两个点, 其像点 $PX_1 = m_1, PX_2 = m_2$, 则

$$l = m_1 \times m_2 = (PX_1) \times (PX_2)$$

于是,

$$(P^i \wedge P^j / L) = (P^{iT} X_1)(P^{jT} X_2)^T - (P^{jT} X_2)(P^{iT} X_1)^T$$

所以,

$$l = m_1 \times m_2 = \begin{pmatrix} (P^2 \wedge P^3 / L) \\ (P^3 \wedge P^1 / L) \\ (P^1 \wedge P^2 / L) \end{pmatrix} = P_L L$$

证毕。

反向投影

从几何上，不难看出像平面上一条直线的反投影是空间中通过摄像机中心的一张平面，如图 2.2.2 所示。

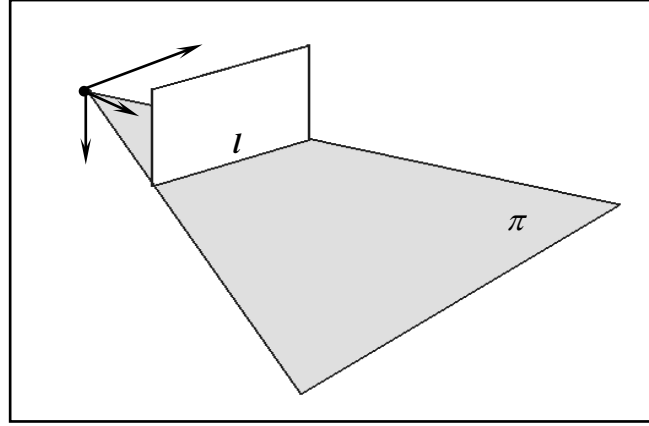


图 2.2.2: 图像直线 l 的反投影平面是 $\pi = P^T l$ 。

在代数上，可用下述命题来描述：

命题 2.2.3 在摄像机 P 下的像直线 l 的反投影是空间平面 $\pi = P^T l$ 。

证明 令 X 是摄像机 P 将它投影到直线 l 上的任一空间点，则必有 $X^T (P^T l) = 0$ 。因此，像直线 l 的反投影是空间平面 $\pi = P^T l$ 。证毕。

2.2.3 对平面的作用

考虑摄像机关于空间平面 π 的投影关系。空间点 X 在摄像机 P 的作用下，其图像点 $m = PX$ ，并且图像点 m 在图像平面中的坐标仅依赖于图像坐标系的选择而与世界坐标系无关。因此，我们可以自由地选择世界坐标系。

以空间平面 π 为世界坐标系的 O - xy 平面，如图 2.2.3 所示，则平面 π 上点的坐标为 $X = (x, y, 0, 1)^T$ ，则必有

$$m = PX = (p_1, p_2, p_3, p_4) \begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = (p_1, p_2, p_4) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

记 $H = (p_1, p_2, p_4)$ ， $X_\pi = (x, y, 1)^T$ ，则上式可以写成

$$m = HX_\pi \quad (2.2.8)$$

在一般情况下，矩阵 H 的秩为 3。 $\text{rank}(H)=2$ 的充要条件是摄像机中心在平面 π 上。事实上， $C = (x, y, z, 1)^T$ 为摄像机光心当且仅当 $PC = 0$ 。若 $\text{rank}(H)=2$ ，则 H 有一维右零空间，于是存在 $C_\pi = (x, y, 1)^T \in \pi$ 使得 $HC_\pi = P(x, y, 0, 1)^T = 0$ ，因此摄像机光心必在平面 π 上。若摄像机光心在平面 π 上，则 H 有一维右零空间，于是 $\text{rank}(H)=2$ 。

当摄像机光心在平面 π 上时，平面 π 的投影是一条直线，这条直线是平面与像平面的交线

(如图 2.2.3(b)所示)。当摄像机光心不在平面 π 上时, H 是一个秩 3 的矩阵, 表示平面 π 到像平面的二维射影变换(如图 2.2.3(a)所示)。以后在讨论平面投影时, 除特别说明外均假定摄像机光心不在这个平面上。

通常称矩阵 H 为平面 π 到像平面的单应矩阵, 或简称单应, 它实现平面 π 到像平面的一个二维射影变换。由于单应矩阵是齐次的, 因此它有 8 个自由度。

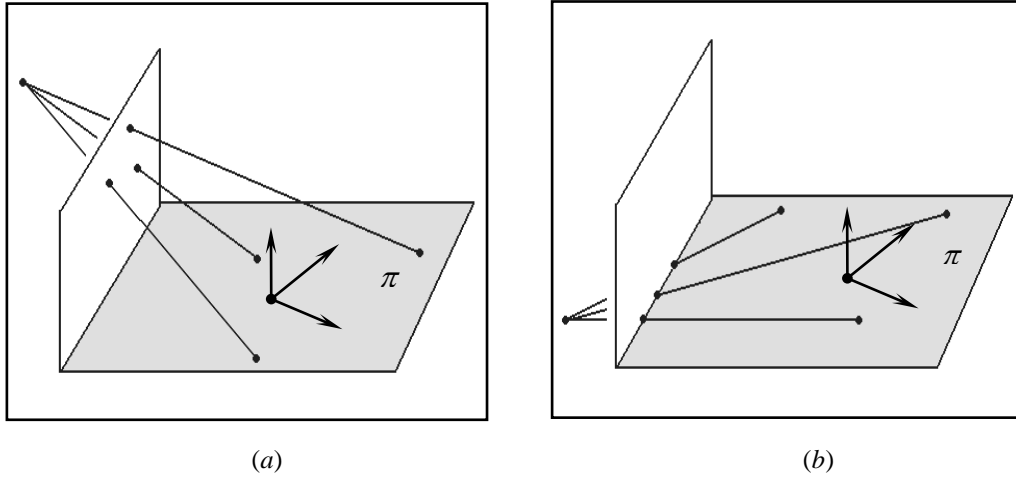


图 2.2.3: (a)空间平面到图像平面的单应变换 H ; (b)退化情况: 摄像机中心在空间平面上。此时空间平面的像是该平面与像平面的交线, $\text{rank}(H)=2$ 。

无穷远平面到像平面的单应矩阵 假定无穷远点 $X = (x, y, z, 0)^T$ 的图像为 m , 则有

$$sm = PX = (p_1, p_2, p_3)X_\infty$$

其中 $X_\infty = (x, y, z)^T$ 。所以, 无穷远平面到像平面的单应矩阵可表示为

$$H_\infty = (p_1, p_2, p_3) = KR \quad (2.2.9)$$

式(2.2.9)同时也说明摄像机矩阵前三列所构成的子矩阵是无穷远平面到像平面的单应矩阵, 简称无穷远平面的单应矩阵(Infinite Homography)。这样, 摄像机矩阵可表示为 $P = (H_\infty, p_4)$ 。

单应矩阵的估计

下面主要讨论如何估计单应矩阵。在实际中, 常用的方法主要有基于点对应(Point Correspondences)和线对应(Line Correspondences)的两种估计方法。

基于点对应 给定一个平面点与其像点的点对应: $X = (x, y, 1)^T \leftrightarrow m = (u, v, 1)^T$, 从式 $sm = HX$ 中消去未知常数因子 s , 可得到关于 H 的两个线性约束方程。因此, 在一般情况下至少需要 4 组以上的点对应才能在相差一个常数因子的意义下唯一求解单应矩阵。在仅有 4 个点的情况下, 有下述命题:

命题 2.2.4 给定 4 组点对应 $X_j \leftrightarrow m_j, (1 \leq j \leq 4)$, 则在相差一个常数因子的意义下可唯一确定单应矩阵 H 的充要条件是 $X_j (1 \leq j \leq 4)$ 中任意三点均不共线。并且可以通过下述公式计算单应矩阵:

$$H = (m_1, m_2, m_3) \text{diag} \left(\frac{p'_1}{p_1}, \frac{p'_2}{p_2}, \frac{p'_3}{p_3} \right) (X_1, X_2, X_3)^{-1} \quad (2.2.10)$$

其中, $(p_1, p_2, p_3)^T = (X_1, X_2, X_3)^{-1} X_4$, $(p'_1, p'_2, p'_3)^T = (m_1, m_2, m_3)^{-1} m_4$ 。

证明: 见附录一命题 1.1.6。

由于图像噪声的影响, 仅用 4 个点对应所计算单应矩阵通常是很不理想的。因此, 在实际中通常使用 4 个以上点对应来估计单应矩阵。

最小二乘估计 令 $X_j \leftrightarrow m_j, (1 \leq j \leq n)$ 为给定的 n 个点对应。不难计算:

$$m_j \times HX_j = [m_j]_{\times} HX_j = [m_j]_{\times} \begin{pmatrix} h_1^T X_j \\ h_2^T X_j \\ h_3^T X_j \end{pmatrix} = \begin{pmatrix} 0 & -1 & v_j \\ 1 & 0 & -u_j \\ -v_j & u_j & 0 \end{pmatrix} \begin{pmatrix} X_j^T h_1 \\ X_j^T h_2 \\ X_j^T h_3 \end{pmatrix}$$

其中 $X_j = (x_j, y_j, 1)^T$ 。由于 $s_j m_j = HX_j, (1 \leq j \leq n)$, 所以 $m_j \times HX_j = 0$ 。于是, 有下述约束线性方程组:

$$\begin{pmatrix} 0 & -X_j^T & v_j X_j^T \\ X_j^T & 0 & -u_j X_j^T \\ -v_j X_j^T & u_j X_j^T & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = 0, \quad j = 1, 2, \dots, n \quad (2.2.11)$$

实际上, 对于每一个对应点所导出的上述三个方程中的第 3 个方程可由前两个方程线性表示。记

$$A_j = \begin{pmatrix} 0 & -X_j^T & v_j X_j^T \\ X_j^T & 0 & -u_j X_j^T \end{pmatrix}, \quad j = 1, 2, \dots, n$$

这样, 方程组(2.2.11)的最小二乘解是矩阵 $A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}$ 的最小奇异值的右奇异向量。于是, 我们有

下述基于点对应的最小二乘估计算法:

算法 2.2.1: 估计单应矩阵的最小二乘算法

目标: 给定空间平面点与其图像的 $N(\geq 4)$ 个点对应 $\{X_j \leftrightarrow m_j\}$, 计算单应矩阵 H

算法:

- 对每一个点对应 $X_j \leftrightarrow m_j$, 构造矩阵 $A_j = \begin{pmatrix} 0 & X_j^T & -v_j X_j^T \\ X_j^T & 0 & -u_j X_j^T \end{pmatrix}$, 再将所有的 A_j 排成一个 $2N \times 9$ 的矩阵 A ;
 - 对 A 进行奇异值分解: $A = UDV^T$;
-

- 由矩阵 V 的最后一个列向量 \mathbf{v}_9 得到摄像机矩阵 $H = \begin{pmatrix} v_{1,9} & v_{2,9} & v_{3,9} \\ v_{4,9} & v_{5,9} & v_{6,9} \\ v_{7,9} & v_{8,9} & v_{9,9} \end{pmatrix}$

基于线对应 空间直线 L 与它的像直线 l 构成一个线对应，记为 $L \leftrightarrow l$ 。利用线对应，同样可估计出平面的单应矩阵。

对于线对应 $L \leftrightarrow l$ ，根据第 1 章命题 1.1.7，有

$$s\mathbf{l} = H^{-T}\mathbf{L} \quad (2.2.12)$$

利用式(2.2.12)，可得到与命题 2.2.4 类似的命题：

命题 2.2.5 给定 4 组线对应 $L_j \leftrightarrow l_j, (1 \leq j \leq 4)$ ，则在相差一个常数因子的意义下可唯一确定单应矩阵 H 的充要条件是 $L_j (1 \leq j \leq 4)$ 中任意三直线均不共点。

与基于点对应的估计类似，也可以得到基于线对应估计单应矩阵的最小二乘算法。

2.2.4 对二次曲线的作用

正向投影

考虑空间中一条二次曲线在摄像机下的投影。二次曲线是平面曲线，将二次曲线所在的平面记为 π ，并如上节那样建立世界坐标系，则摄像机关于平面 π 的作用可由一个单应矩阵 H 来描述，即 $\mathbf{m} = H\mathbf{X}_\pi$ 。一条二次曲线可以用一个 3×3 对称矩阵 C 来表示，即它的方程可写成

$$\mathbf{X}_\pi^T C \mathbf{X}_\pi = 0, \quad (\mathbf{X}_\pi = (x, y, 1)^T \in \pi)$$

由于对任意 $\mathbf{X}_\pi \in C$ ，其图像点 $\mathbf{m} = H\mathbf{X}_\pi$ ，所以

$$\mathbf{m}^T H^{-T} C H^{-1} \mathbf{m} = \mathbf{X}_\pi^T C \mathbf{X}_\pi = 0 \quad (2.2.13)$$

由于 $C_m = H^{-T} C H^{-1}$ 仍为一个 3×3 对称矩阵，所以它表示图像平面上的一条二次曲线。因此，二次曲线的图像仍是一条二次曲线。

命题 2.2.6 二次曲线 C 的图像仍是一条二次曲线 C_m 。如果二次曲线 C 所在平面到像平面的单应矩阵为 H ，则有

$$C_m = H^{-T} C H^{-1} \quad (2.2.14)$$

绝对二次曲线的图像(Image of the Absolute Conic (IAC)) 绝对二次曲线 Ω_∞ 是无穷远平面 π_∞ 上的一条(点)二次曲线。在 π_∞ 上， Ω_∞ 可由下述方程来描述(见第 1 章第 1.2.5 节)：

$$(x, y, z)I(x, y, z)^T = 0 \quad (2.2.15)$$

它是 π_∞ 上一条虚二次曲线。

下面考虑绝对二次曲线在摄像机下的像曲线。

令摄像机矩阵为 $P = (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4)$ 。从式(2.2.9)，无穷远平面到像平面的单应矩阵为 $H_\infty = (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ 。于是，由命题 2.2.6， Ω_∞ 的图像是二次曲线：

$$C_m = H_\infty^{-T} I H_\infty^{-1}$$

因 $H_\infty = (p_1, p_2, p_3) = \alpha KR$ ，所以，

$$C_m = K^{-T} K^{-1} \quad (2.2.16)$$

这里略去了常数因子 α^2 ，因为二次曲线的矩阵表示与常数因子无关，即 C 与 sC 表示同一条二次曲线)。这样，就得到了下述命题：

命题 2.2.7 绝对二次曲线在摄像机下的像曲线为 $C = K^{-T} K^{-1}$ 。

这是一个非常重要的命题，因为它指出绝对二次曲线的图像与世界坐标系的选择无关(或者说与摄像机位置无关)，仅与摄像机内参数有关。以后，将会看到它构成了摄像机自标定的理论基础。

反向投影

令 C_m 是图像平面上的一条二次曲线，考虑它在摄像机 P 下的反投影。在几何上不难看出， C_m 的反投影是顶点在摄像机中心并通过二次曲线 C_m 的一个锥面 Q ，它是一个退化的二次曲面(如图 2.2.4 所示)。

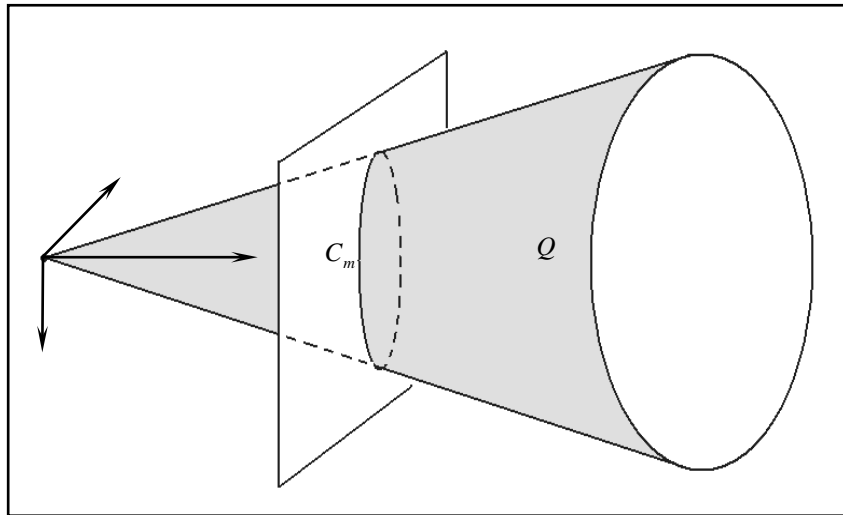


图 2.2.4：二次曲线的反投影是一个锥面

在代数上，可以用下述命题来描述：

命题 2.2.8 设摄像机矩阵为 P ，则二次曲线 C_m 的反投影 Q 为

$$Q = P^T C_m P \quad (2.2.17)$$

证明 图像点 m 在二次曲线 C_m 上，当且仅当 $m^T C_m m = 0$ 。由于空间点 X 的投影是 $m = PX$ ，所以 X 被投影到二次曲线 C_m 上当且仅当 $X^T P^T C_m P X = 0$ ，这就是说二次曲线的反投影是(退化)二次曲面 $Q = P^T C_m P$ 。证毕。

2.2.5 对二次曲面的作用

二次曲面 Q 可用一个 4×4 的对称矩阵来表示，它的方程为 $\mathbf{X}^T Q \mathbf{X} = 0$ 。当 $\text{rank}(Q)=4$ 时， Q 表示一个非退化的二次曲面；当 $\text{rank}(Q)=3$ 时， Q 是一个锥面；当 $\text{rank}(Q)=2$ 时， Q 表示两张不重合的平面；当 $\text{rank}(Q)=1$ 时， Q 表示两张重合的平面。在考虑摄像机对二次曲面的作用时，一般假定 Q 是非退化的或是一个锥面。

二次曲面的轮廓线 令 \mathbf{X}_0 是空间中的一个点， Q_c 是以 \mathbf{X}_0 为顶点且与 Q 相切的所有射线组成的集合，它是一个以 \mathbf{X}_0 为顶点的锥面，通常称 Q_c 为二次曲面 Q 的视锥面。锥面 Q_c 与二次曲面 Q 相切于一条二次曲线 Γ ，它是 Q_c 中所有母线与 Q 的切点的集合，如图 2.2.5 所示。称 Γ 是 Q 的一条轮廓线。显然， Q 的轮廓线与视锥面的顶点有关。

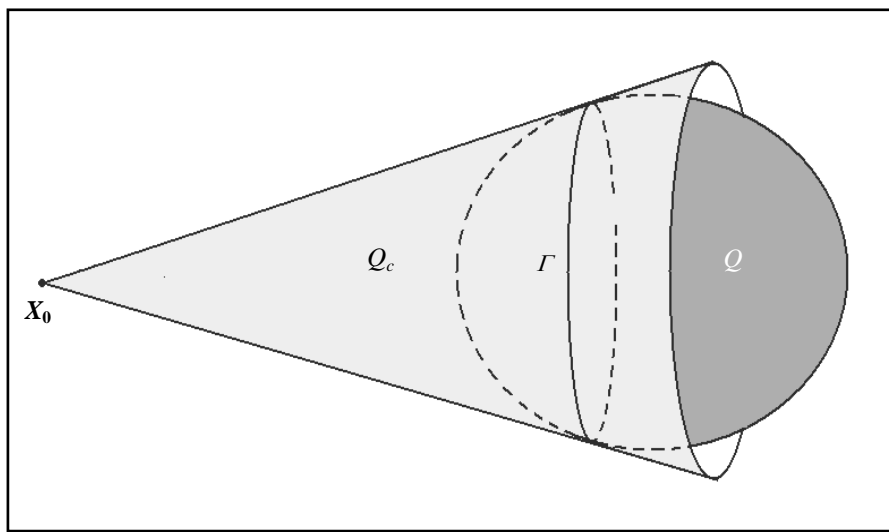


图 2.2.5：二次曲面的轮廓线 Γ 是一条二次曲线

二次曲面的投影

下面考虑二次曲面的投影。在几何上，二次曲面 Q 的轮廓 Γ 在图像平面上的投影 C 是顶点在摄像机光心的视锥面与像平面的交线。二次曲面 Q 的图像是交线 C 所包含的区域，如图 2.2.6 所示。

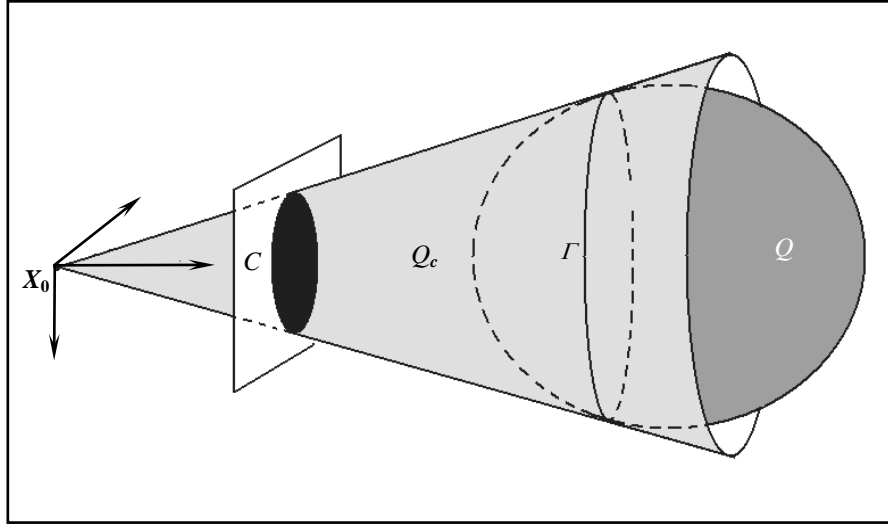


图 2.2.6: 二次曲面的投影

通常称 C 是二次曲面图像的轮廓线。因 Γ 是一条二次曲线，所以 C 也是一条二次曲线。如果二次曲面上没有纹理，它的轮廓线 Γ 的图像 C 是唯一可以得到的信息。因此，以后称 C 是二次曲面 Q 的图像。

在代数上，我们有下述命题：

命题 2.2.9 (1) 令摄像机矩阵为 P ，二次曲面 Q 的对偶为 Q^* ，它的图像 C 的对偶为 C^* ，则有

$$C^* = PQ^*P^T \quad (2.2.18)$$

(2) 二次曲面的轮廓线 Γ 所在的平面为 $\pi_\Gamma = QO$ ，其中 O 是摄像机的光心坐标。

证明 结论(2)可从二次曲面对极关系推出，因为摄像机的中心 C 关于二次曲面的极面是 π_Γ 。下面仅证明(1)。令 l 是 C 的任一条切线，则必有

$$l^T C^* l = 0 \quad (2.2.19)$$

由命题 2.2.3， l 的反投影面为 $\pi = P^T l$ ，并且它与二次曲面 Q 相切。因此，有

$$l^T PQ^*P^T l = 0$$

结合(2.2.21)式，就得到 $C^* = PQ^*P^T$ 。证毕。

关于二次曲面的反投影锥面 Q_c ，由式(2.2.17) 直接给出。

绝对二次曲面的图像 绝对二次曲线 Ω_∞ 的对偶是三维空间中一个退化的对偶二次曲面，称为绝对二次曲面并记为 Q_∞^* 。在几何上， Q_∞^* 由 Ω_∞ 的切平面所组成。在代数上 Q_∞^* 由一个秩为 3 的 4×4 的齐次矩阵来表示（见第 1 章第 1.2.5 节）：

$$Q_\infty^* = \begin{pmatrix} I & 0 \\ 0^T & 0 \end{pmatrix} \quad (2.2.20)$$

绝对二次曲面是绝对二次曲线的对偶，所以它在摄像机 P 下的图像是绝对二次曲线的对偶

图像(Dual Image of the Absolute Conic (DIAC))。因此, 绝对二次曲面 Q_{∞}^* 的图像是

$$C^* = C^{-1} = KK^T \quad (2.2.21)$$

式(2.2.21)也可以从式(2.2.18) 直接导出: 因为, 从式(2.2.18), 我们有

$$C^* = PQ_{\infty}^*P^T = H_{\infty}H_{\infty}^T = KK^T$$

这样, 就得到了下述命题:

命题 2.2.10 绝对二次曲面在摄像机下的图像 (DIAC)为 $C^* = KK^T$ 。

与 IAC 一样, DIAC 也是图像平面上的一条(对偶)二次曲线, 与摄像机位置无关而仅与内参数有关。因此, 它也构成了摄像机自标定的理论基础。

2.3 欧氏空间与射影空间

在以上各节中, 我们都假定世界坐标系是欧氏坐标系, 并使用了射影几何中的一些概念, 例如: 无穷远平面 π_{∞} , 并用无穷远点来表示一条直线的方向, 以及齐次坐标等等。

在本书的后续章节中, 将要更进一步地使用射影坐标系。如果世界坐标系是射影坐标系, 摄像机矩阵又具有什么样的形式呢? 这是不难回答的。因为摄像机坐标系和射影坐标系之间的变换仍然可以用一个 4×4 的齐次矩阵 H 表示, 即

$$X_c = HX$$

并且由三维射影空间到图像平面的映射仍然可以用一个秩为 3 的 3×4 矩阵 P 表示。事实上, 对于最一般的情形, 摄像机是一个从三维射影空间到二维射影平面的映射, 可以用矩阵的合成来表示这个映射:

$$P = K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} H \quad (2.3.1)$$

它仍是一个 3×4 的矩阵。

注意: 摄像机坐标系总是使用欧氏坐标系, 因为摄像机本身是一个欧氏装置。这一点是非常重要的。也就是说, 不应该因为有一个摄像机投影模型, 我们就可以回避欧氏几何的概念。在数学上, 对于一个 3×4 矩阵, 总能象 2.1.2 节所讨论的那样进行 RQ 分解而获得一个旋转矩阵和一个上三角矩阵 K 。这样获得的参数作为欧氏解释, 当且仅当它与空间坐标系相匹配时才有意义, 即世界坐标系是欧氏坐标系时才有意义。如果世界坐标系是仿射的或射影的, 这种分解的结果不能作为摄像机的欧氏解释。

(1) 如果 $H = \begin{pmatrix} A & b \\ 0^T & 1 \end{pmatrix}$, 其中 $\text{rank}(A)=3$, 它表示三维空间的一个仿射变换, 式(2.3.1)变为

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} A & b \\ 0^T & 1 \end{pmatrix}$$

它是世界坐标系为仿射坐标系的摄像机矩阵，称它为仿射空间中的摄像机矩阵。

(2) 如果 $H = \begin{pmatrix} sR & t \\ 0^T & 1 \end{pmatrix}$ ，其中 R 是旋转矩阵， s 为非零常数，它表示三维空间的一个相似变

换，此时式(2.3.1)变为

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} sR & t \\ 0^T & 1 \end{pmatrix}$$

它是世界坐标系为欧氏坐标系(但度量单位为 s)的摄像机矩阵，称它为相似空间中的摄像机矩阵。

(3) 如果 $H = \begin{pmatrix} R & t \\ 0^T & 1 \end{pmatrix}$ ，其中 R 是旋转矩阵，它表示三维空间的一个欧氏变换，此时式(2.3.1)

变为

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R & t \\ 0^T & 1 \end{pmatrix}$$

它是世界坐标系为欧氏坐标系(其度量是绝对度量)的摄像机矩阵，称它为欧氏空间中的摄像机矩阵，即在以前各节所介绍的摄像机矩阵。

(4) 如果 $H=(I, 0)$ ，此时式(2.3.1)变成

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

它是以摄像机坐标系为世界坐标系的摄像机矩阵。

参考文献

1. R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge , 2000.