

附录 B 摄象机模型和外极线几何

B.1 摄象机模型

B.1.1 针孔模型和透视投影

B.1.2 摄象机外参数

B.1.3 摄象机内参数和归一化摄象机

B.1.4 透视矩阵的一般形式

B.2 透视投影的各种线性近似

B.2.1 正投影 (orthographic projection)

B.2.2 弱透视 (weak perspective)

B.2.3 平行透视 (paraperspective projection)

B.2.4 仿射摄象机

B.3 透视投影下的外极线几何

B.3.1 外极线几何中的概念

B.3.2 归一化坐标系中的外极线方程

B.3.3 像素坐标系中的外极线方程

B.3.4 投影矩阵下的外极线方程

B.3.5 基础矩阵和外极几何变换

B.1 摄象机模型

在大部分应用环境中可以用理想的针孔模型来近似实际摄象机。针孔模型的几何关系就是透视投影。下面我们先介绍透视投影的几何关系。

B.1.1 针孔模型和透视投影

针孔摄象机的模型在第五章中已有介绍。在那里我们使用了矢量代数的表示方法，下面我们坐标变换的方法来推导之。

我们定义的第一个坐标系是摄象机坐标系。该坐标系的原点在焦心 C ， X_c 、 Y_c 、 Z_c 轴由 A 、 H 和 V 决定，其中 A 为光轴方向， H 和 V 是正交的方向，三者组成右手直角坐标系。三维点在该坐标系中的坐标 M_c 记为 $(X_c, Y_c, Z_c)^T$ 。

为表示透视模型我们还需要在图象平面中建立图象坐标系。这是一个二维坐标系，其原点位于光轴和图象平面的交点 c （称为主点，principal point），两坐标轴与 H 和 V 平行且反向。在该坐标系中像点 m 的坐标表示为 $(u, v)^T$ 。

在定义了这两个坐标系后，投影模型可表示为

$$\frac{u}{X_c} = \frac{v}{Y_c} = \frac{f}{Z_c} \quad (B.1)$$

其中 f 为焦心到图象平面的距离，即焦距。

摄象机坐标系和图象坐标系如图 B.1 所示。

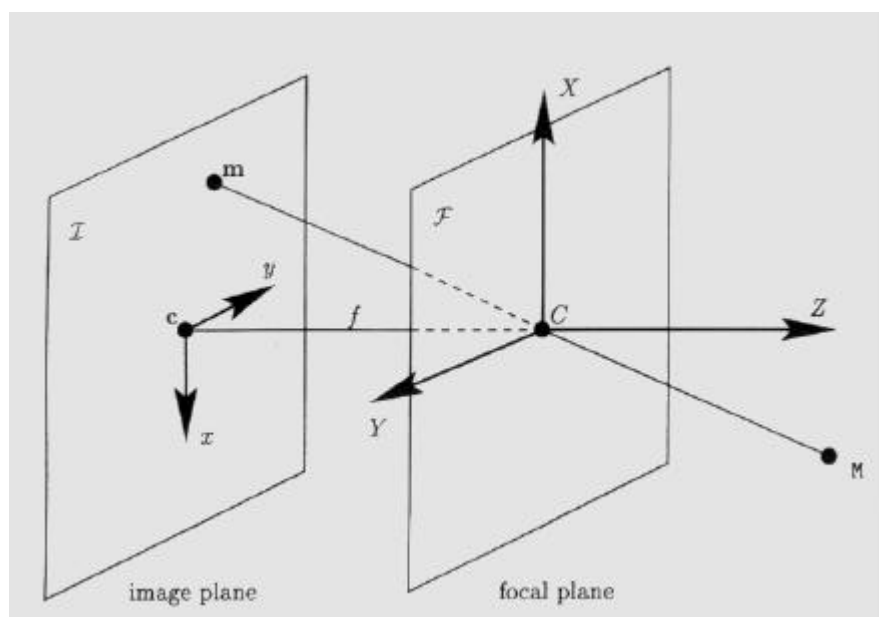


图 B.1 针孔模型，其中过焦点 C 和图象平面平行的平面称为焦平面

在实际应用中物体离焦点的距离一般都远大于焦距。因此我们常在光轴上和实际图象平面关于焦点对称的位置上设置一个虚拟图象平面，如图 B.2。并在该虚拟平面上建立二维坐标系。原点在光轴和该平面的交点，两坐标轴与 H 和 V 平行并且方向相同。把像平面上的点经焦点作中心对称映射到该虚拟平面上。这样该平面上点的二维坐标与三维点的摄像机坐标系坐标同样满足上面的投影模型。以后我们一般把此虚拟平面称为像平面。

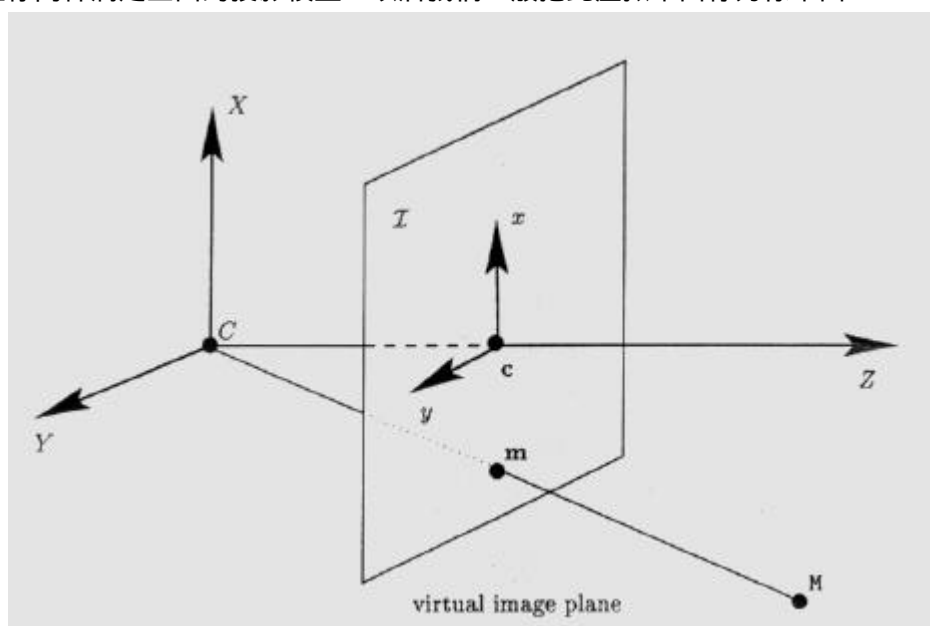


图 B.2 使用虚拟图象平面的针孔模型

对任一坐标值 $\mathbf{x}=(x_1, x_2, \dots)^T$ ，我们用 $\tilde{\mathbf{x}}$ 表示 \mathbf{x} 的齐次坐标。它由 \mathbf{x} 的所有分量加上一个为 1 的元素生成，即 $\tilde{\mathbf{x}}=(\mathbf{x}^T, 1)^T$ 。

在投影模型中如果我们使用 \mathbf{m} 点的齐次坐标 $(U, V, S)^T$ ，则有

$$\begin{pmatrix} U \\ V \\ S \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (\text{B.2})$$

其中 U, V, S 满足 $S \neq 0$ 且 $u=U/S$, $v=V/S$ 。另外, 我们用 \mathbf{P}_c 表示 3×4 矩阵 $\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$,

则上式可写成线性形式

$$s\tilde{\mathbf{m}} = \mathbf{P}_c \tilde{\mathbf{M}}_c \quad (\text{B.3})$$

其中 $s=S$ 为一比例因子。

B.1.2 摄象机外参数

上面的讨论都是在摄象机坐标系 (C, X_c, Y_c, Z_c) 中进行的, 不过在实际应用中, 摄象机的位置和方向并不总是固定不变的, 因此我们需要用固定的世界坐标系 (O, X, Y, Z) 来表示三维点。记 \mathbf{P}_c 在世界坐标系中的坐标为 $\mathbf{M}=(X, Y, Z)^T$, 则两坐标系的关系可用 $\mathbf{M}_c = \mathbf{R}\mathbf{M} + \mathbf{t}$ 来表示, \mathbf{R} 是旋转矩阵, 它表示摄象机的方向; \mathbf{t} 则与摄象机的位置有关, 它实际是世界坐标系原点在摄象机坐标系中的坐标。这两者被称为摄象机的外参数。

如果我们使用齐次坐标, 上面的坐标系间的关系可写成

$$\tilde{\mathbf{M}}_c = \mathbf{D}\tilde{\mathbf{M}} \quad (\text{B.4})$$

$$\text{其中 } \mathbf{D} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix}, \quad \mathbf{0}_3 = [0, 0, 0]^T \quad (\text{B.5})$$

结合(B.3)和(B.4)式, 我们有

$$s\tilde{\mathbf{m}} = \mathbf{P}_c \tilde{\mathbf{M}}_c = \mathbf{P}_c \mathbf{D} \tilde{\mathbf{M}} = \mathbf{P} \tilde{\mathbf{M}} \quad (\text{B.6})$$

其中 $\mathbf{P} = \mathbf{P}_c \mathbf{D}$ 为世界坐标系中的投影矩阵。

B.1.3 摄象机内参数和归一化摄象机

下面我们考虑图象平面上的坐标变换。它在实际应用中十分重要, 这是因为:

- 像素坐标系的原点不一定与光轴和图象平面的交点(即主点)重合;
- 像素坐标系中两坐标轴的单位由实际设备的采样率决定, 它们不一定相同;
- 像素坐标系中两坐标轴不一定成直角。

为了处理这些问题, 我们需要建立图象坐标系和像素坐标系间的仿射变换关系。如图 B.3 所示, (c, x, y) 是上一节介绍的图象坐标系。它是直角坐标系, 并且两个轴向上的单位相

同。 (o,u,v) 则是像素坐标系，其原点常位于图象的左上角而不是主点 c ，像素也常常不是方的。设 k_u 、 k_v 是 u,v 轴上的单位在图象坐标系中的度量值， θ 是 u,v 两轴的夹角， (u_0, v_0) 是 c 在像素坐标系中的坐标。这五个参数就是摄像机的内参数。

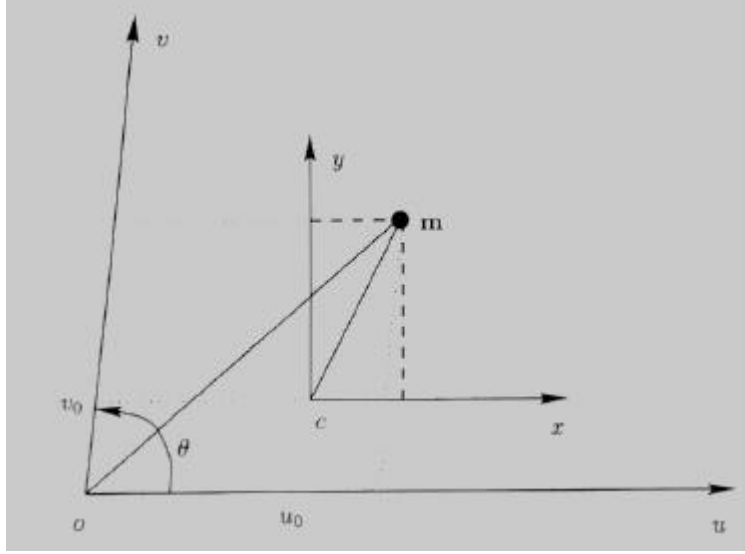


图 B.3 摄像机的内参数，图象平面上的坐标变换

令 $\mathbf{m}_{old}=[x,y]^T$ 为图象坐标系中的坐标值， $\mathbf{m}_{new}=[u,v]^T$ 则是像素坐标。显然有

$$\tilde{\mathbf{m}}_{new} = H\tilde{\mathbf{m}}_{old},$$

其中 $H = \begin{bmatrix} k_u & k_u \cot \theta & u_0 \\ 0 & k_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ 。另外，根据(B.3)式，我们有 $s\tilde{\mathbf{m}}_{old} = P_{old}\tilde{\mathbf{M}}_c$ ，代入上式，

得

$$s\tilde{\mathbf{m}}_{new} = HP_{old}\tilde{\mathbf{M}}_c = P_{new}\tilde{\mathbf{M}}_c$$

$$\text{其中 } P_{new} = HP_{old} = H \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} fk_u & fk_u \cot \theta & u_0 & 0 \\ 0 & fk_v / \sin \theta & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (\text{B.7})$$

P_{new} 即像素坐标表示的投影矩阵。从中我们可以看到真正起作用的是 $\alpha_u = fk_u$ 和 $\alpha_v = fk_v$ ，即焦距的变化和像素尺度的变化在最终的图象上是不可区分的。

在实际应用中为简化公式，我们常常使用所谓的归一化坐标系。该坐标系也是定义在图象平面上的二维坐标系，如果用该坐标系表示图象平面上的点，则投影矩阵 P_N 有非常简单

$$\text{的形式 } P_N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (\text{B.8})$$

对摄像机坐标系中的点 (X_c, Y_c, Z_c) ，其像点的归一化坐标 (x_N, y_N) 满足

$$x_N = X_c / Z_c, \quad y_N = Y_c / Z_c \quad (\text{B.9})$$

根据 P_N 和 P_{new} 的定义，我们可以看到 P_N 和 P_{new} 满足

$$P_{\text{new}} = AP_N \quad (\text{B.10})$$

$$\text{其中, } A = \begin{bmatrix} \mathbf{a}_u & \mathbf{a}_u \cot \mathbf{q} & u_0 \\ 0 & \mathbf{a}_v / \sin \mathbf{q} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{B.11})$$

而归一化坐标 $(x_N, y_N)^T$ 和像素坐标 $(u, v)^T$ 满足

$$\begin{bmatrix} x_N \\ y_N \\ 1 \end{bmatrix} = A^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (\text{B.12})$$

使用归一化坐标系实际是把摄像机的内外参数分解开来考虑, 这样当我们关注于摄像机的外参数变化的情况(例如运动视觉)时就不必考虑使用的是什么样的摄像机。

B.1.4 透视矩阵的一般形式

同时考虑摄像机的内外参数, 即建立像素坐标系中点 \mathbf{m} 与三维对应点在世界坐标系中的坐标值 \mathbf{M}_w 之间的映射关系, 我们有

$$s\tilde{\mathbf{m}} = P\tilde{\mathbf{M}}_w \quad (\text{B.13})$$

$$\text{其中 } P = AP_N D = A \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \quad (\text{B.14})$$

其中 A , R , \mathbf{t} 如上两节所述。

透视矩阵 P 为 3×4 矩阵, 由于可相差一个尺度, 因此只有 11 个独立参量。

记 p_{ij} 为 P 中 i 行 j 列的元素, 消去尺度因子 s , 我们得到

$$u = \frac{p_{11}X_w + p_{12}Y_w + p_{13}Z_w + p_{14}}{p_{31}X_w + p_{32}Y_w + p_{33}Z_w + p_{34}} \quad (\text{B.15})$$

$$v = \frac{p_{21}X_w + p_{22}Y_w + p_{23}Z_w + p_{24}}{p_{31}X_w + p_{32}Y_w + p_{33}Z_w + p_{34}} \quad (\text{B.16})$$

上两式可用于标定摄像机。如果有六个以上的像点和三维点的对应就可求出 P , 然后从 P 中分解出内外参数。这方面的内容就不详述了。

根据 P 我们还可以求出焦心 C 在世界坐标系中的坐标。令 $P = [B \ b]$, 其中 B 为 3×3 矩阵,

b 为三维列向量。显然有 $P\tilde{C}_w = 0$, 即 $\begin{bmatrix} B & b \end{bmatrix} \begin{bmatrix} C_w \\ 1 \end{bmatrix} = 0$, 故

$$C_w = -B^{-1}b \quad (\text{B.17})$$

根据 P 和一个像点 \mathbf{m} 我们可以得到空间中一条射线, 该射线由焦心和 \mathbf{m} 决定, 并且其上的所有点都透视投影为 \mathbf{m} 。这条射线的参数方程为 $\mathbf{M} = B^{-1}(-b + \lambda \tilde{\mathbf{m}})$, 其中 λ 为任一正实数。

B.2 透视投影的各种线性近似

从上节的讨论可以知道，透视投影实际是一个非线性映射。这在实际求解时可能需要大的计算量；更重要的是，如果透视效果并不明显，直接使用该模型可能会使实际问题称为病态问题。另外，在某些条件下，例如，摄像机的视场很小，并且物体的尺寸相对于到观察者的距离也很小，透视模型可以很好地用线性模型近似。这种近似可大大简化推导和计算。

为简单起见，如果不作特别说明的话，下面的讨论都认为像点用其归一化坐标表示，三维点用其在摄像机坐标系中的坐标表示。

B.2.1 正投影（orthographic projection）

最简单的线性近似称为正投影。这种近似完全忽略了深度信息。在这种投影方式下，物体到摄像机的垂直距离（深度信息）和物体到光轴的距离（位置信息）都完全丢失了。因此，它只在这两种信息确实可以忽略时才可使用。

正投影的公式为 $x=X, y=Y$ 。

B.2.2 弱透视（weak perspective）

如果物体的尺寸相对其到摄像机的距离很小的话，物体上各点的深度可以用一共同的深度值 Z_0 近似，这个值一般取物体质心的深度。这样透视模型可近似为

$$\begin{aligned} x &= \frac{X}{Z_0} \\ y &= \frac{Y}{Z_0} \end{aligned} \quad (\text{B.18})$$

这种近似可以看作两阶段投影的合成。第一步，整个物体按平行于光轴的方向正投影到经过物体质心并与图象平面平行的平面上；第二步，再按透视模型投影到图象平面上，这一步实际是全局的放缩。因此，弱透视也被称为放缩正投影（scaled orthographic projection）。

$$\text{令 } P_{wp} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & Z_0 \end{bmatrix} \quad (\text{B.19})$$

则弱透视模型可写成与透视投影类似的形式

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = P_{wp} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \text{将摄像机的内外参数都考虑进来，我们有}$$

$s\tilde{\mathbf{m}} = AP_{wp}D\tilde{\mathbf{M}}_w$ ，其中 s 为一比例因子， A 和 D 如上节定义。消去比例因子，我们看

到二维点和三维点的对应关系确实是线性的。

下面我们来推导这种近似带来的误差。现在我们又回到归一化坐标系和摄像机坐标系。设三维点 \mathbf{M} 的真正深度值为 $Z=Z_0+\Delta Z$ 。该点按透视模型投影为 \mathbf{m}_p ，而弱透视的结果为 \mathbf{m}_{wp} ，我们用泰勒公式把 Z 在 Z_0 处展开并略去高阶项，得到两者的差 \mathbf{m}_{error} 为

$$\begin{aligned}\mathbf{m}_{error} &= \mathbf{m}_p - \mathbf{m}_{wp} = \frac{1}{Z_0 + \Delta Z} \begin{bmatrix} X \\ Y \end{bmatrix} - \frac{1}{Z_0} \begin{bmatrix} X \\ Y \end{bmatrix} = \frac{1}{Z_0} \left(1 - \frac{\Delta Z}{Z_0} + \left(\frac{\Delta Z}{Z_0} \right)^2 - \dots \right) \begin{bmatrix} X \\ Y \end{bmatrix} - \frac{1}{Z_0} \begin{bmatrix} X \\ Y \end{bmatrix} \\ &= -\frac{\Delta Z}{Z_0^2} \left(1 - \frac{\Delta Z}{Z_0} + \left(\frac{\Delta Z}{Z_0} \right)^2 - \dots \right) \begin{bmatrix} X \\ Y \end{bmatrix} = -\frac{\Delta Z}{Z_0^2} \frac{1}{1 + \frac{\Delta Z}{Z_0}} \begin{bmatrix} X \\ Y \end{bmatrix} = -\frac{1}{Z} \frac{\Delta Z}{Z_0} \begin{bmatrix} X \\ Y \end{bmatrix}\end{aligned}$$

我们可以看到有两种原因带来误差，一是 $\frac{\Delta Z}{Z_0}$ ，即物体的深度信息；二是 $\begin{bmatrix} X \\ Y \end{bmatrix}$ ，即位置信息。

误差对物体的不同部分是不同的。在实际应用中为使用弱透视模型，一般要求 $Z_0 > 10 * |\Delta Z|$ 。

B.2.3 平行透视 (paraperspective projection)

在弱透视投影中，三维点先被正投影到过物体质心并与图象平面平行的平面上。这一过程中丢失了物体的位置信息。如果物体离光轴较远，弱透视带来的误差是很大的。在平行透视中，投影过程仍可分为两步，第一步仍是把物体平行投影到过质心且与象平面平行的平面上，不过这次的投影线不是平行于光轴，而是平行于质心 G 和焦心 C 的连线 CG 。

容易得到平行透视的公式为

$$\begin{aligned}x &= \frac{1}{Z_0} \left(X - \frac{X_0}{Z_0} Z + X_0 \right) \\ y &= \frac{1}{Z_0} \left(Y - \frac{Y_0}{Z_0} Z + Y_0 \right)\end{aligned}\tag{B.20}$$

其中 (X_0, Y_0, Z_0) 为质心的三维坐标。

为将该模型写成与透视投影类似的形式，我们令

$$P_{pp} = \begin{bmatrix} 1 & 0 & -\frac{X_0}{Z_0} & X_0 \\ 0 & 1 & -\frac{Y_0}{Z_0} & Y_0 \\ 0 & 0 & 0 & Z_0 \end{bmatrix}\tag{B.21}$$

则 $s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = P_{pp} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$ 。即使考虑了摄像机的内外参数，二维点与三维点间的对应关系也

仍是线性的。

最后我们也来看看平行透视引入的误差。令 $(X,Y,Z)^T = (X_0 + \Delta X, Y_0 + \Delta Y, Z_0 + \Delta Z)^T$ ，我们用与上节类似的方法得到

$$\begin{aligned} \mathbf{m}_{error} &= \mathbf{m}_p - \mathbf{m}_{pp} = \frac{1}{Z_0 + \Delta Z} \begin{bmatrix} X_0 + \Delta X \\ Y_0 + \Delta Y \end{bmatrix} - \frac{1}{Z_0} \begin{bmatrix} X - \frac{X_0}{Z_0} Z + X_0 \\ Y - \frac{Y_0}{Z_0} Z + Y_0 \end{bmatrix} \\ &= \frac{1}{Z_0} \left(1 - \frac{\Delta Z}{Z_0} + \left(\frac{\Delta Z}{Z_0} \right)^2 - \dots \right) \begin{bmatrix} X_0 + \Delta X \\ Y_0 + \Delta Y \end{bmatrix} - \frac{1}{Z_0} \begin{bmatrix} X - \frac{X_0}{Z_0} Z + X_0 \\ Y - \frac{Y_0}{Z_0} Z + Y_0 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{\Delta X \Delta Z}{Z_0^2} + \frac{X}{Z_0} \left(\frac{\Delta Z}{Z_0} \right)^2 + O\left(\left(\frac{\Delta Z}{Z_0} \right)^3 \right) \\ -\frac{\Delta Y \Delta Z}{Z_0^2} + \frac{Y}{Z_0} \left(\frac{\Delta Z}{Z_0} \right)^2 + O\left(\left(\frac{\Delta Z}{Z_0} \right)^3 \right) \end{bmatrix} \end{aligned}$$

可以看到，像点误差是三维点误差的二阶无穷小。而在弱透视的情况下，像点误差是三维点误差的一阶无穷小。这说明，平行透视确实是比较弱透视更好的近似。

图 B.4 是各种线性近似和透视模型的比较。从图中可看到，各模型的近似程度和我们前面的分析是一致的。

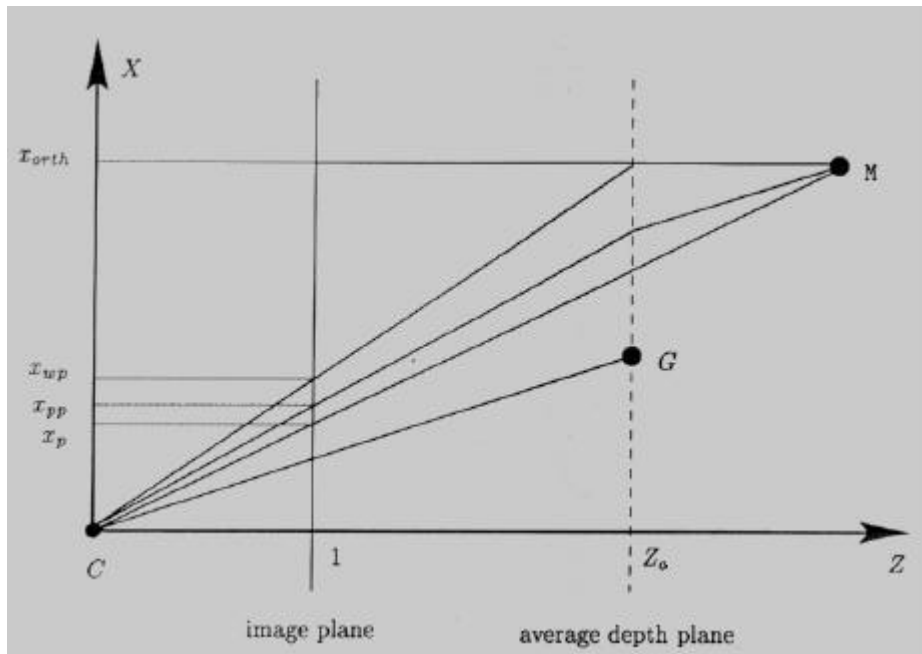


图 B.4 透视模型及其各种线性近似的比较

图中 x_{orth} 、 x_{wp} 、 x_{pp} 和 x_p 分别是物体上的三维点 M 在正投影、弱透视、平行透视和透视投

影下的投影点。G 为物体的质心。

透视投影还有一种误差更小的线性近似，即正透视（orthoperspective）。同样从两阶段过程来考虑，正透视与平行透视的区别在于第一阶段的投影平面不是与图象平面平行，而是经过质心和 CG（即焦心和质心的连线）垂直。它也可被看作以下三步操作的合成：

1. 摄象机绕焦心 C 旋转，直至光轴与 CG 重合；
2. 进行弱透视投影；
3. 将摄象机旋转回原位置，这引起图象平面上的一个仿射变换。

正透视的数学表示形式很复杂，在这里就不给出其公式了。

B.2.4 仿射摄象机

观察在正投影、弱透视和平行透视下的投影矩阵，我们发现它们都具有如下的形式：

$$P_A = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{34} \end{bmatrix} \quad (B.22)$$

P_A 是一 3×4 矩阵，它决定了一个三维空间到二维平面的线性映射（用齐次坐标表示），所以我们把 P_A 称为仿射摄象机。与透视投影矩阵类似， P_A 也可相差一个尺度因子，因此它只有 8 个自由参量。它可由 4 组二维点和三维点的对应决定。

如果用非齐次坐标，仿射摄象机可表示为

$$m = T_A M + t_A \quad (B.23)$$

其中 T_A 为一 2×3 矩阵，其元素 $T_{ij} = p_{ij}/p_{34}$ ，而 t_A 为二维向量 $[p_{14}/p_{34}, p_{24}/p_{34}]^T$ 。

仿射摄象机的一个重要性质是保平行性：三维空间的平行线投影为二维空间的平行线。这和透视投影是不同的。证明很简单：设 $M_1(\lambda) = M_a + \lambda \cdot \mathbf{u}$ 和 $M_2(\mu) = M_b + \mu \cdot \mathbf{u}$ 是三维空间的两平行线，其中 \mathbf{u} 是三维方向向量， λ 和 μ 是直线参数。投影结果为 $m_1(\lambda) = (T_A M_a + t_A) + \lambda T_A \mathbf{u}$ 和 $m_2(\mu) = (T_A M_b + t_A) + \mu T_A \mathbf{u}$ ，显然它们都平行于方向向量 $T_A \mathbf{u}$ 。

仿射摄象机的另一个重要性质是它把三维点集的质心投影为对应二维投影点的质心。这也是透视投影不具有的性质。

仿射摄象机的缺点是几何意义不明显。它当然是前几节介绍的各种透视投影线性近似的推广。这种推广可按下面两种方式来理解：

- (1) 允许三维物体作某种非刚性变形。实际上，如果对 P_A 右乘 $D = \begin{bmatrix} M & t \\ 0_3^T & 1 \end{bmatrix}$ 这样的三

维仿射变换，其中 M 为 3×3 矩阵，而 t 为三维列向量，相乘的结果仍是仿射摄象机。

- (2) 无需标定摄象机内参数。实际上，对 P_A 左乘 $B = \begin{bmatrix} B_1 & b \\ 0_2^T & 1 \end{bmatrix}$ 这样的二维仿射变换，

其中 B_1 为 2×2 矩阵，而 b 为二维列向量，相乘的结果也仍是仿射摄象机。

即使不标定摄象机的内参数，我们仍能从图象中提取出如平行性、定长度比这样的仿射度量。对于某些视觉任务来说，这样的仿射度量就足够了。

最后需要指出的是仿射摄象机是实际摄象机的近似。它只在感兴趣目标的深度变化相对

其深度而言可忽略不计才适用。

B.3 透视投影下的外极线几何

本节我们介绍透视投影下的外极线几何，并讨论其概念和数学表示。

B.3.1 外极线几何中的概念

外极线几何讨论的是两摄象机图象平面间的关系。考虑图 B.5 中的两个摄象机。设 C 为第一个摄象机的焦心， C' 为第二个摄象机焦心； m 为第一个摄象机图象平面 I 上的点， m' 为其在第二个摄象机图象平面 I' 上的对应点（即两者是同一三维点 M 分别在两个像平面上的投影）。则 m' 必然位于该像平面内由 m 决定的一条直线 l_m' 上。其中 l_m' 是 I' 和由三维点 M 、两焦心 C 、 C' 决定的平面 π （称为外极平面）的交线。这是因为在 I 上的投影点是 m 的三维点必然在射线 CM 上，而 CM 在 I' 上的投影为 l_m' 。 l_m' 称为 m 决定的外极线。不仅如此，我们还发现， I' 上所有的外极线交于同一点 e' ，此点称为外极点。它是 CC' 和 I' 的交点。这可以这样解释：任给 I 上一点 m ，它在 I' 上决定的外极线 l_m' 是 I' 与 π_m 的交线，其中 π_m 是由 m 决定的外极平面。这样所有的 π_m 形成过 CC' 的平面族，这个平面族与 CC' 的交点位于所有外极线上，因而是外极点 e' 。由于两摄象机的地位是完全对称的，因而类似的分析可发现， I' 上任一点 m' 在 I 上决定一条外极线 $l_{m'}$ ， I 上所有外极线交于外极点 e 。 e 和 e' 都在直线 CC' 上。 m 、 m' 、 C 和 C' 四点共面（其中 m 和 m' 是对应点），这被称为共面约束。它在已知摄象机内参数的情况下从运动求解结构中有重要作用。

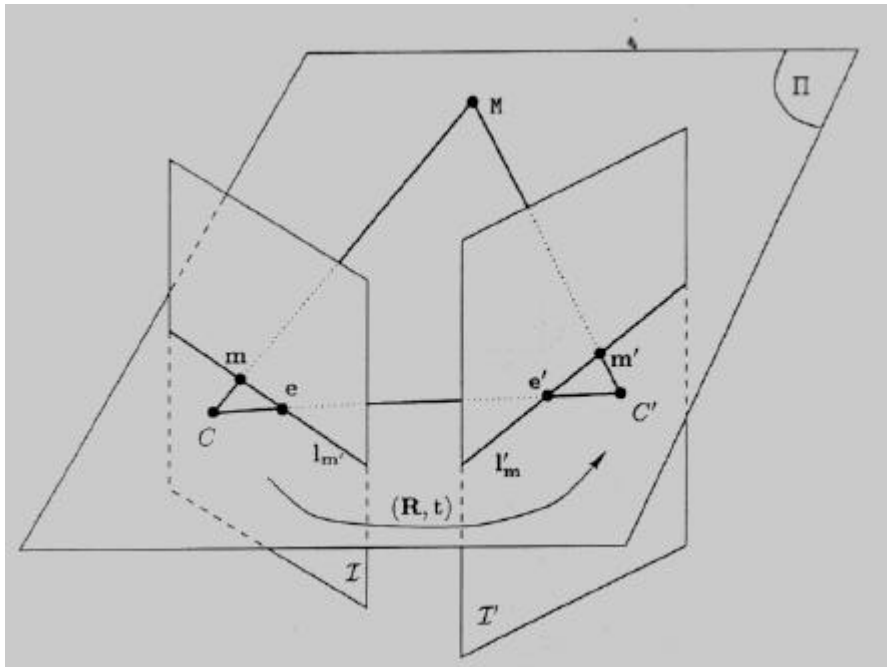


图 B.5 外极线几何

利用外极线的概念可大大减少在两幅图象间匹配对应点的计算量。由于 I 中点 m 在 I' 中的对应点在由 m 决定的外极线上，因此搜索空间的维数由二维降为一维。该约束称为外

极线约束。

如果两焦心的连线 CC' 与一个或两个摄像机的像平面平行，那么该像平面上的外极点就位于无穷远处。这说明该平面上的外极线彼此平行。另外，如果还有 CC' 与摄像机的水平扫描线平行，则外极线也是水平的。实际上，很多立体视觉算法都假设外极线与水平扫描线平行。

下面我们考虑两摄像机间的方向和位置关系。设第二个摄像机坐标系可由第一个摄像机坐标系经由旋转 R 和平移 t 得到。这样同一三维点在第一个坐标系中的坐标 (X,Y,Z) 和其在第二个摄像机坐标系中的坐标 (X',Y',Z') 满足

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + t \quad (B.24)$$

$$\text{其中 } R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}.$$

R 中只有三个自由参量，因为旋转矩阵 R 满足 $RR^T=I$ 且 $\det(R)=1$ ，这将提供六个约束。

从下节开始我们将首先推导归一化坐标系中的外极线方程，然后把它扩展到像素坐标系的情形，最后给出直接利用投影矩阵表示的形式。

B.3.2 归一化坐标系中的外极线方程

设 $M=(X,Y,Z)^T$ 在两摄像机像平面中的投影分别为 $m=(x,y)^T$ 和 $m'=(x',y')^T$ 。像平面上的点用其归一化坐标表示， M 是三维点在第一个摄像机坐标系中的坐标。再设 $M'=(X',Y',Z')^T$ 是同一三维点在第二个摄像机坐标系中的坐标。根据针孔摄像机的透视模型，我们有

$$\tilde{m} = M / Z, \tilde{m}' = M' / Z'.$$

利用(B.24)消去 M 和 M' ，则有

$$\tilde{m} = \frac{1}{Z} (Z' R \tilde{m}' + t),$$

其中仍有两个与三维点有关的参数 Z 和 Z' 。为此我们用 t 对上式两边作叉乘，得到

$$t \times \tilde{m} = \frac{Z'}{Z} t \times (R \tilde{m}'), \text{ 然后再两边都与 } \tilde{m} \text{ 作点积 (即内积), 最后得到}$$

$$\tilde{m}^T t \times (R \tilde{m}') = 0 \quad (B.25)$$

上式中不再含有 Z 和 Z' 。

(B.25)式在根据运动求解结构中十分重要。它的几何意义是十分明显的： CC' 、 $C\tilde{m}$ 和 $C'\tilde{m}'$ 三向量共面。如果用第一个摄像机坐标系中的坐标表示这三个向量，它们分别是 t 、 \tilde{m} 和 $R\tilde{m}'$ 。上节提到的共面约束决定了这三个向量的混合积为零，这就是(B.25)式。

下面我们定义一个从三维向量到 3×3 反对称矩阵的变换：它将向量 $(x_1, x_2, x_3)^T$ 映射为矩

$$\text{阵} \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}, \text{ 记该反对称矩阵为 } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{\times}.$$

利用上面的定义我们可以将两向量的叉乘表示成矩阵与向量的乘积，容易验证 $a \times b = [a]_{\times} b$ ，其中 a, b 为任意两个三维向量。因此(B.25)式可写成

$$\tilde{m}^T E \tilde{m}' = 0 \quad (\text{B.26})$$

$$\text{其中 } E = [t]_{\times} R \quad (\text{B.27})$$

(B.26)式称为外极线方程。其中 E 称为本质矩阵 (Essential Matrix)。它最早由 Longuet-Higgins 在 1981 年由从运动到结构的求解中导出。该矩阵仅由两摄象机间的旋转和平移决定。由于 $[t]_{\times}$ 是反对称矩阵，因此 $\det(E)=0$ 。实际上， E 的秩为 2。

在具体分析(B.26)式前，我们首先来讨论二维平面中直线的表示方法。二维空间里的任何直线可用方程 $ax+by+c=0$ 来表示。这就是说，直线可用一个三维向量 $l=(a,b,c)^T$ 表示，该直线上的所有点 $\tilde{m}=(x,y,1)^T$ 满足 $\tilde{m}^T l=0$ 。当然 l 可相差一个比例因子。实际上， l 乘以任一非零的比例因子 λ 后得到的向量 λl 表示的是同一条直线。如果一条直线过 m_1 和 m_2 两点，则我们有 $\tilde{m}_1^T l=0$ 且 $\tilde{m}_2^T l=0$ ，因此

$$l = \tilde{m}_1 \times \tilde{m}_2 \quad (\text{B.28})$$

下面我们讨论在第一个摄象机像平面内外极线的表示。设 $m'=(x',y')^T$ 为第二个摄象机像平面的一点，其对应的三维点 M' 必在由 C' 和 M'_∞ 决定的射线 $C'M'_\infty$ 上，其中 M'_∞ 是该射线上的无穷远点。 \tilde{m}' 也在该射线上。根据透视模型我们有，

$$M' = \mathbf{I} \tilde{m}' = \mathbf{I} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}, \mathbf{I} \in (0, \infty) \quad (\text{B.29})$$

这实际上就是射线 $C'M'_\infty$ 的参数表示。而该三维点 M' 在第一个摄象机坐标系中的坐标

$M = R M' + t = \mathbf{I} R \tilde{m}' + t$ 。记 l_m 为 $C'M'_\infty$ 在第一个摄象机图象平面 I 中的投影，即 m' 在 I 中决定的外极线。

这条直线可由两个点确定。为求得第一个点，我们用 $\lambda=0$ 代入 $C'M'_\infty$ 的参数方程，它

的投影点为 $\tilde{e} = \frac{1}{t_z} t$ ，其中 t_z 为 t 的第三个分量。这是第二个摄象机焦心在 I 上的投影，即

外极点。为得到第二个点，我们令 $\mathbf{I} = \infty$ ，这样我们得到 $\tilde{m}_\infty = \frac{1}{r_3^T \tilde{m}'} R \tilde{m}'$ ，其中 r_3 为 R 的

第三行。根据前面对直线表示方法的讨论 (B.28)式), 我们有

$$l_{m'} = \tilde{e} \times \tilde{m}_{\infty} = t \times R\tilde{m}' = E\tilde{m}' \quad (\text{B.30})$$

由于可忽略比例因子, 所以我们丢掉了向量前的系数 $\frac{1}{t_z}$ 和 $\frac{1}{r_3^T \tilde{m}'}$ 。

同样的推导过程可以得出 $l'_m = E^T \tilde{m}$ 。即外极线可以由 E 表出。外极线方程实际是外极线约束的代数表示。

下面我们来讨论 e 和 e' 的性质。由于 e 是第一个摄象机像平面上所有外极线的交点, 因此对任何 \tilde{m}' 有 $\tilde{e}^T E \tilde{m}' = 0$ 成立, 这就是说 $\tilde{e}^T E = 0^T$, 即 $\tilde{e}^T [t]_{\times} R = 0^T$, 由于 R 可逆, 故

$$\tilde{e}^T [t]_{\times} = 0^T, \quad \tilde{e} = \begin{pmatrix} t_x & t_y & 1 \\ t_z & t_z & 1 \end{pmatrix}^T。这里我们又得到了外极点的坐标表示。$$

同样的推导, $0 = E \tilde{e}' = [t]_{\times} R \tilde{e}'$, 故 $R \tilde{e}' = \lambda t$, λ 为某个非零比例因子。

$$\tilde{e}' = \frac{1}{r_3^T t} R^T t = \begin{bmatrix} \frac{r_1^T t}{r_3^T t} & \frac{r_2^T t}{r_3^T t} & 1 \end{bmatrix}^T, \quad \text{其中 } r_i' = [r_{1i} \quad r_{2i} \quad r_{3i}]^T, \quad i=1,2,3 \text{ 是 } R \text{ 的列标号。}$$

e 是无穷远点对应于 $t_z=0$; e' 是无穷远点则对应于 $r_3^T t=0$, 即第一个摄象机的焦心位于第二个摄象机的焦平面 (即过焦心和图象平面平行的平面) 内。如果还要求两个焦平面重

合, 那么除了 $t_z=0$ 外还有 $r_{13}=r_{23}=0$, $r_{33}=1$, 故 R 形如 $\begin{bmatrix} \cos \mathbf{q} & \sin \mathbf{q} & 0 \\ -\sin \mathbf{q} & \cos \mathbf{q} & 0 \\ 0 & 0 & 1 \end{bmatrix}$, 即两坐标系间

只有绕光轴的旋转。

将上面的 t 及 R 的特殊形式代入(B.27), 我们有

$$E = [t]_{\times} R = \begin{bmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y \cos \mathbf{q} - t_x \sin \mathbf{q} & -t_y \sin \mathbf{q} + t_x \cos \mathbf{q} & 0 \end{bmatrix} \quad (\text{B.31})$$

如果将 E 的具体形式代入(B.26), 我们发现外极线方程对所有的图象坐标值都是线性的。这说明两个像平面上的外极线都是平行的, 并且方向和像点没有关系。

B.3.3 像素坐标系中的外极线方程

现在 m 和 m' 是两摄象机像平面上用像素坐标表示的一组对应点。显然有

$$\tilde{m}^T F \tilde{m}' = 0 \quad (\text{B.32})$$

$$\text{其中, } F = A^{-T} E A'^{-1} \quad (\text{B.33})$$

A 和 A' 为两摄象机的内参数矩阵。

上式的推导十分简单。只要把像素坐标和归一化坐标的关系(B.12)代入(B.26)式即得。方程(B.32)是两帧图象间对应像素满足的基本约束。上式也可直接用透视模型来从头推导。我们设世界坐标系和第二个摄象机坐标系重合，则

$$s\tilde{m} = A[R \quad t] \begin{bmatrix} M' \\ 1 \end{bmatrix}, \quad s'\tilde{m}' = A'[I \quad 0] \begin{bmatrix} M' \\ 1 \end{bmatrix}$$

用与上节类似的方法消去 M' ， s 和 s' ，我们就可得到(B.32)式。

3×3 矩阵 F 称为基础矩阵。由于 E 秩为 2， F 的秩也为 2。另外，和 E 一样， F 也可相差一比例因子。因此 F 也只有 7 个自由参量。

下面我们推导这时外极点 e 和 e' 的形式。显然 e 是 C' 的投影，而 $C' = 0$ ，故我们有

$$s_e \tilde{e} = A[R \quad t] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = At, \quad \text{故 } \tilde{e} \text{ 即为向量 } At \text{ 被它的第三个分量除的结果。}$$

同理， e' 是 C 的投影，而由(B.17)知 $C = -R^{-1}t$ ，代入 $s_{e'} \tilde{e}' = A'[I \quad 0] \begin{bmatrix} C \\ 1 \end{bmatrix}$ ，得

$s_{e'} \tilde{e}' = -A'R^{-1}t$ ，即 \tilde{e}' 为向量 $-A'R^{-1}t$ 被它的第三个分量除的结果。

现在我们来推导外极线的表示。为求出第一个摄象机图象平面上由 m' 决定的外极线 $l_{m'}$ ，我们需要 $l_{m'}$ 上的两个点。一个点是已求出的 e ，另一个点可以是射线 $C'\tilde{m}'$ 上的任一点，我们选择使 $\tilde{m}' = A'[I \quad 0] \begin{bmatrix} M'_0 \\ 1 \end{bmatrix}$ ，即该点的比例因子 s 为 1。这样 $M'_0 = A'^{-1} \tilde{m}'$ ，该点在第一

个摄象机图象平面 I 上的投影点 m_0 满足 $s_0 \tilde{m}_0 = A[R \quad t] \begin{bmatrix} M'_0 \\ 1 \end{bmatrix} = A(RA'^{-1} \tilde{m}' + t)$ 。

根据 \tilde{e} 和 \tilde{m}_0 ，我们得到

$$l_{m'} = \tilde{e} \times \tilde{m}_0 = \frac{1}{s_e s_0} (s_e \tilde{e}) \times (s_0 \tilde{m}_0) = \frac{1}{s_e s_0} (At) \times (ARA'^{-1} \tilde{m}' + At) = \frac{1}{s_e s_0} (At) \times (ARA'^{-1} \tilde{m}')$$

易证 $(Ax) \times (Ay) = \det(A)A^{-T}(x \times y)$ 对所有的可逆阵 A 成立，再加上 $l_{m'}$ 可相差一比例因子，故 $l_{m'} = A^{-T}(t) \times (RA'^{-1} \tilde{m}') = F\tilde{m}'$ ，其中 F 为基础矩阵。

类似地，我们有 $l'_m = F^T \tilde{m}$ ，推导如下：

为求出 l'_m 我们同样需要 e' 之外的另一点。这次我们选择射线 $C\tilde{m}$ 上的点 M_0 满足 $\tilde{m} = A[R \quad t] \begin{bmatrix} M_0 \\ 1 \end{bmatrix}$ ，同样比例因子为 1。这样 $M_0 = (AR)^{-1}(\tilde{m} - At)$ ，该点在第二个摄象

机像平面 I' 上的投影满足 $s'_0 \tilde{m}'_0 = A'[I \quad 0] \begin{bmatrix} M_0 \\ 1 \end{bmatrix} = A'R^{-1}A^{-1}\tilde{m} - A'R^{-1}t$ ，因此有

$$l'_m = (s'_e \tilde{e}') \times (s'_0 \tilde{m}'_0) = -(A'R^{-1}t) \times (A'R^{-1}A^{-1}\tilde{m}) = -(A'R^{-1})^{-T}(t) \times (A^{-1}\tilde{m}) = F^T \tilde{m}$$

推导过程中我们已在 l'_m 上添加了比例因子 $s'_e s'_0$ 。

外极点的位置也可根据外极线方程求出。由于外极点是所有外极线的交点，故对所有 m 和 m' ，有 $\tilde{e}^T F \tilde{m}' = 0$ 和 $\tilde{m}^T F \tilde{e}' = 0$ 成立。此即 $F^T \tilde{e} = 0$ 且 $F \tilde{e}' = 0$ 。将 F 的具体形式代入，我们最终得到 $s'_e \tilde{e}' = -A'R^{-1}t$ 和 $s_e \tilde{e} = At$ ，和以前的结论相同。

B.3.4 投影矩阵下的外极线方程

在某些应用（例如立体视觉）中，摄像机投影矩阵 P 是我们的已知量。因而我们希望直接用 P 表示外极线几何，而不再使用 A, R, t 等量。

设 P 和 P' 分别为一、二台摄像机的投影矩阵。我们记 $P = [B \ b]$ ，其中 B 为 3×3 矩阵， b 为三维向量。类似地有 $P' = [B' \ b']$ 。

根据透视模型，我们有

$$s\tilde{m} = [B \ b] \begin{bmatrix} M \\ 1 \end{bmatrix}, \quad s'\tilde{m}' = [B' \ b'] \begin{bmatrix} M \\ 1 \end{bmatrix}, \quad \text{其中 } M \text{ 是三维点在世界坐标系中的坐标。}$$

如果 B 和 B' 可逆，则

$$M = sB^{-1}\tilde{m} - B^{-1}b = s'B'^{-1}\tilde{m}' - B'^{-1}b',$$

从中消去 M ，并且两边都左乘 B ，有

$$s\tilde{m} = s'BB'^{-1}\tilde{m}' + b - BB'^{-1}b',$$

再两边都与 $b - BB'^{-1}b'$ 作叉乘，则

$$s(b - BB'^{-1}b') \times \tilde{m} = s'(b - BB'^{-1}b') \times BB'^{-1}\tilde{m}',$$

又与 \tilde{m} 作点积，我们得到

$$\tilde{m}^T F \tilde{m}' = 0, \quad \text{其中 } F = [b - BB'^{-1}b']_{\times} BB'^{-1}, \quad \text{这就是用投影矩阵表示的外极线方程。}$$

下面我们来推导外极点的表示。根据 (B.17)，我们有 $C' = -B'^{-1}b'$ ，故

$$s_e \tilde{e} = [B \ b] \begin{bmatrix} C' \\ 1 \end{bmatrix} = b - BB'^{-1}b'; \quad \text{同理，} \quad s'_e \tilde{e}' = b' - B'B^{-1}b.$$

外极线方向的推导与上两节完全类似，最后的结论 $l_{m'} = F\tilde{m}'$ ， $l'_m = F^T \tilde{m}$ 也与上节相同。

注意，本节我们将 P 分解为 B 和 b ，并且要求 B 可逆。实际上可将这一表示法推广，改用 P 和 P' 来直接表示外极线方程，这方面的内容我们就不详述了。

B.3.5 基础矩阵和外极几何变换

本节我们讨论基础矩阵在外极几何变换中的作用。所谓外极几何变换，是指两个像平面间外极点和外极线的几何变换关系。

设 m' 为第二个摄象机像平面上的任一点，它在第一个摄象机的像平面上决定了一条外极线 $l_{m'} = F\tilde{m}'$ 。它必然经过外极点 $\tilde{e} = [e_1 \ e_2 \ e_3]^T$ 和 m' 的对应点 $\tilde{m} = [u \ v \ s]^T$ ，故 $l_{m'}$ 有两种表示方法 $l_{m'} = F\tilde{m}' = \tilde{e} \times \tilde{m}$ 。现在我们都使用像点的齐次坐标。同样， m 决定的外极线 l_m 过外极点 $\tilde{e}' = [e_1' \ e_2' \ e_3']^T$ 和 m 的对应点 $\tilde{m}' = [u' \ v' \ s']^T$ ， l_m 同样有两种表示方法 $l_m = F^T \tilde{m} = \tilde{e}' \times \tilde{m}'$ 。

为建立外极线间的关系，我们首先确定外极点的坐标。根据 $F^T \tilde{e} = 0$ 和 $F\tilde{e}' = 0$ ，我们可以求得

$$\begin{aligned} e_1' &= F_{23}F_{12} - F_{22}F_{13} \\ e_2' &= F_{13}F_{21} - F_{11}F_{23} \\ e_3' &= F_{22}F_{11} - F_{21}F_{12} \end{aligned} \quad (B.34)$$

和

$$\begin{aligned} e_1 &= F_{32}F_{21} - F_{22}F_{31} \\ e_2 &= F_{31}F_{12} - F_{11}F_{32} \\ e_3 &= F_{22}F_{11} - F_{21}F_{12} \end{aligned} \quad (B.35)$$

e 和 e' 的三坐标值可相差一个比例因子。

确定了外极点，外极线的方向可由其方向决定。考虑 $l_{m'} = \tilde{e} \times \tilde{m}'$ ，设其方向向量为

$u' = [1 \ t' \ 0]^T$ ，故 $l_{m'}$ 可表示成参数形式 $\tilde{m}' = \tilde{e}' + \lambda' u'$ ，其中 λ' 为参数。 m' 决定的外极线

$$l_{m'} = F\tilde{m}' = F\tilde{e}' + \lambda' F u' = \mathbf{I}' \begin{bmatrix} F_{11} + F_{12}t' \\ F_{21} + F_{22}t' \\ F_{31} + F_{32}t' \end{bmatrix} \quad (B.36)$$

另外，我们设 l_m 的方向向量为 $u = [1 \ t \ 0]^T$ ，利用 l_m 的另一种表示方法，则有

$$l_{m'} \cong \tilde{e} \times \tilde{m} = \tilde{e} \times (\tilde{e} + \mathbf{I}u) = \mathbf{I}\tilde{e} \times u = \mathbf{I} \begin{bmatrix} -(F_{11}F_{22} - F_{21}F_{12})t \\ (F_{11}F_{22} - F_{21}F_{12}) \\ (F_{32}F_{21} - F_{22}F_{31})t - F_{31}F_{12} + F_{11}F_{32} \end{bmatrix} \quad (B.37)$$

其中 \cong 表示在相差一个比例因子的意义下相等。由于(B.36)和(B.37)是同一直线的两种表示，令它们相等。得到

$$\mathbf{t} = \frac{a\mathbf{t}'+b}{c\mathbf{t}'+d} \quad (\text{B.38})$$

其中 $a=F_{12}$, $b=F_{11}$, $c=-F_{22}$, $d=-F_{21}$ 。

上式写成矩阵形式如下：

$$\mathbf{r} \begin{bmatrix} \mathbf{t} \\ 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \mathbf{t}' \\ 1 \end{bmatrix} \quad (\text{B.39})$$

ρ 为比例因子。这种关系被称为 τ 和 τ' 的同构关系。我们也可以说两幅图象间的外极线有同构关系。

当然上面的推导只在外极线方向的第一个分量不为零时才成立。如果第一个分量为零，我们只要设该方向为 $[\mathbf{t} \ 1 \ 0]^T$ ，就能得到类似的结果。

从上面的推导可以看到，外极几何变换由外极点 $\tilde{e} = [e_1 \ e_2 \ e_3]^T$ 、

$\tilde{e}' = [e_1' \ e_2' \ e_3']^T$ 、以及同构中的四个参数 a 、 b 、 c 、 d 确定。外极点坐标可相差一比例因子，同构参数也可相差一比例因子，因此一共有 7 个自由参数。这刚好是基础矩阵中自由参数的数量。

如果我们确定了外极几何变换，即求出了两个外极点和四个同构参数，则能用它们表出基础矩阵。关系如下：

$$\begin{aligned} F_{11} &= be_3 e_3' \\ F_{12} &= ae_3 e_3' \\ F_{13} &= -(ae_2' + be_1')e_3 \\ F_{21} &= -de_3 e_3' \\ F_{22} &= -ce_3 e_3' \\ F_{23} &= (ce_2' + de_1')e_3 \\ F_{31} &= (de_2 - be_1)e_3' \\ F_{32} &= (ce_2 - ae_1)e_3' \\ F_{33} &= -(ce_2' + de_1')e_2 + (ae_2' + be_1')e_1 \end{aligned} \quad (\text{B.40})$$

参考文献：

[Xug 96] Xu, G. and Zhang, Z. Y., Epipolar geometry in stereo, motion and object recognition, a unified approach, Kluwer Academic Publishers, 1996.