

# 计算机视觉的 PNP 问题的最优解<sup>1)</sup>

徐文立

(清华大学自动化系, 北京 100084)

## 摘 要

本文讨论了计算机视觉的 PNP (Perspective-N-Point) 问题: 根据观察到的  $n$  个已知特征点的象点求解被观察物体相对于相机的三维运动参数。由于象噪声, 该问题本质上是非线性最优化问题。本文导出一个闭式解, 并提出若干克服象噪声影响的方法。仿真试验的结果表明本文的方法有很好的应用前景。

**关键词** 计算机视觉, PNP 问题, 运动参数, RS 分解, 相机阵列。

## 一、前 言

设被观察物体上有  $n$  个特征点, 且已知它们之间的相对位置。所谓 PNP (Perspective-N-Point) 问题就是要从透视投影得到的  $n$  个象点中, 求解摄象时刻物体相对于相机的位姿(位置和姿态)。显然, 如果分别求得两个不同摄象时刻下物体相对于相机的位姿, 则可以非常容易地求得这两个时刻间物体相对于相机的运动(平移和旋转)。本文用“运动参数”来统一描述“位置和姿态”以及“平移和旋转”。由于 PNP 问题在目标跟踪、零件装配、物体识别、视觉导航等许多方面有着广阔的应用背景, 因而一直是计算机视觉研究中的一个重要问题。在假定象点位置没有误差的理想情况下, Fischler 等<sup>[1]</sup>用投影几何<sup>[2]</sup>的方法证明了为保证 PNP 问题解的唯一性, 在特征点共面的情形下需要四个点, 在非共面情形下需要六个点。Hung 等<sup>[3]</sup>从另一角度讨论了四点共面的情形, 所得的解和文献[2]给出的解尽管形式上不同, 实质上是一致的。最近, Horaud 等<sup>[4]</sup>又导出了在四个非共面点情形下的解析解, 关键是求解一个双二次多项式方程, 不保证解的唯一性。事实上, 基于不存在象噪声假设下导出的准确解对于象点的位置误差是十分敏感的, 为克服象噪声的影响, PNP 问题应按最优化问题来处理。Hung 等<sup>[3]</sup>假定四个特征点构成一个已知的平面四角形, 利用迭代的方法将观察到的四个象点在象平面上作微小的移动, 根据新的象点位置求得对应特征点在空间构成的四角形, 直至所得四角形和已知四角形的形状大小一致。该方法的缺点是合理性值得怀疑, 且难以推广到多于四点的情形。Lowe<sup>[5,6]</sup>利用迭代法不断修正运动参数, 使得根据运动参数所计算出来的象点位置 and 实际观察到的象点位置尽可能接近。该方法的合理性是显然的, 缺点是没有给出迭代过程所需要的初始解, 也没有讨论唯一性问题。

本文于 1991 年 7 月 5 日收到。

1) 获国家教委和清华大学资助。

本文讨论在有象噪声的情形下如何求解 PNP 问题, 提出了行之有效的方法。

## 二、估计特征点的空间位置

### 1. 特征点共面的情形

在此情形下, 总可以设置物体坐标系  $\{B\}$  使得所有特征点位于  $X-Y$  平面内, 设  $\mathbf{p}_i^b = (x_i^b, y_i^b, 0)^T$  是点  $p_i$  相对于  $\{B\}$  的位矢。相机坐标系  $\{C\}$  的设置如图 1 所示, 不失一般性, 令焦距  $f=1$ 。设  $\mathbf{p}_i^c = (X_i^c, Y_i^c, Z_i^c)^T$  和  $\mathbf{q}_i = (x_i, y_i, 1)^T$  分别是点  $p_i$  以及它的对应象点  $q_i$  相对于  $\{C\}$  的位矢。假定  $\{B\}$  相对于  $\{C\}$  的位姿由旋转矩阵  $R$  和平移向量  $\mathbf{t}$  来描述, 即

$$\mathbf{p}_i^c = Z_i^c \mathbf{q}_i = R \mathbf{p}_i^b + \mathbf{t}, \quad j = 1, \dots, n. \quad (1)$$

由上式可得

$$T \mathbf{p}_i = k_i \mathbf{q}_i, \quad j = 1, \dots, n, \quad (2)$$

符号的定义如下:

$$\begin{aligned} \mathbf{p}_i &= (X_i^b, Y_i^b, 1)^T, \quad \mathbf{q}_i = (x_i, y_i, 1)^T, \\ T &= [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}], \quad k_i = Z_i^c. \end{aligned} \quad (3)$$

这里  $\mathbf{r}_j$  是  $R$  的第  $j$  个列向量。若记  $T = [\mathbf{t}_1 \ \mathbf{t}_2 \ \mathbf{t}_3]^T$ , 由式 (2) 得

$$\mathbf{t}_1^T \mathbf{p}_i = k_i x_i, \quad \mathbf{t}_2^T \mathbf{p}_i = k_i y_i, \quad \mathbf{t}_3^T \mathbf{p}_i = k_i. \quad (4)$$

将上面的第三个方程代入前两个, 得

$$\begin{aligned} \mathbf{t}_1^T \mathbf{p}_i &= x_i \mathbf{t}_3^T \mathbf{p}_i, \\ \mathbf{t}_2^T \mathbf{p}_i &= y_i \mathbf{t}_3^T \mathbf{p}_i, \\ j &= 1, \dots, n. \end{aligned} \quad (5)$$

由此可见, 每个特征点提供二个方程。由于  $T$  有九个元素, 为了确定  $T$  (不计比例因子), 至少需要四个特征点。定义

$$\begin{aligned} P &= [\mathbf{p}_1, \dots, \mathbf{p}_n], \\ X &= \text{diag}(x_1, \dots, x_n), \\ Y &= \text{diag}(y_1, \dots, y_n), \end{aligned} \quad (6)$$

式 (5) 可改写成

$$\begin{aligned} P^T \mathbf{t}_1 &= X P^T \mathbf{t}_3, \\ P^T \mathbf{t}_2 &= Y P^T \mathbf{t}_3. \end{aligned} \quad (7)$$

注意,  $P$  反映了特征点集的结构, 是事先知道的, 可认为不含误差。而  $X$  和  $Y$  是观察到的象点坐标, 必定含有误差。因此, 为了从式 (7) 确定  $T$ , 可建立最小化目标函数

$$J = \|P^T \mathbf{t}_1 - X P^T \mathbf{t}_3\|^2 + \|P^T \mathbf{t}_2 - Y P^T \mathbf{t}_3\|^2, \quad (8)$$

服从  $\|\mathbf{t}_3\| = 1$ 。这是线性最小二乘问题。给定  $\mathbf{t}_3$ , 最优的  $\mathbf{t}_1$  和  $\mathbf{t}_2$  应是<sup>[7]</sup>

$$\mathbf{t}_1 = (P^+)^T X P^T \mathbf{t}_3, \quad \mathbf{t}_2 = (P^+)^T Y P^T \mathbf{t}_3, \quad (9)$$

这里,  $P^+ = P^T (P P^T)^{-1}$ 。将式 (9) 代入 (8) 得

$$J = \mathbf{t}_3^T (B^T B) \mathbf{t}_3, \quad (10)$$

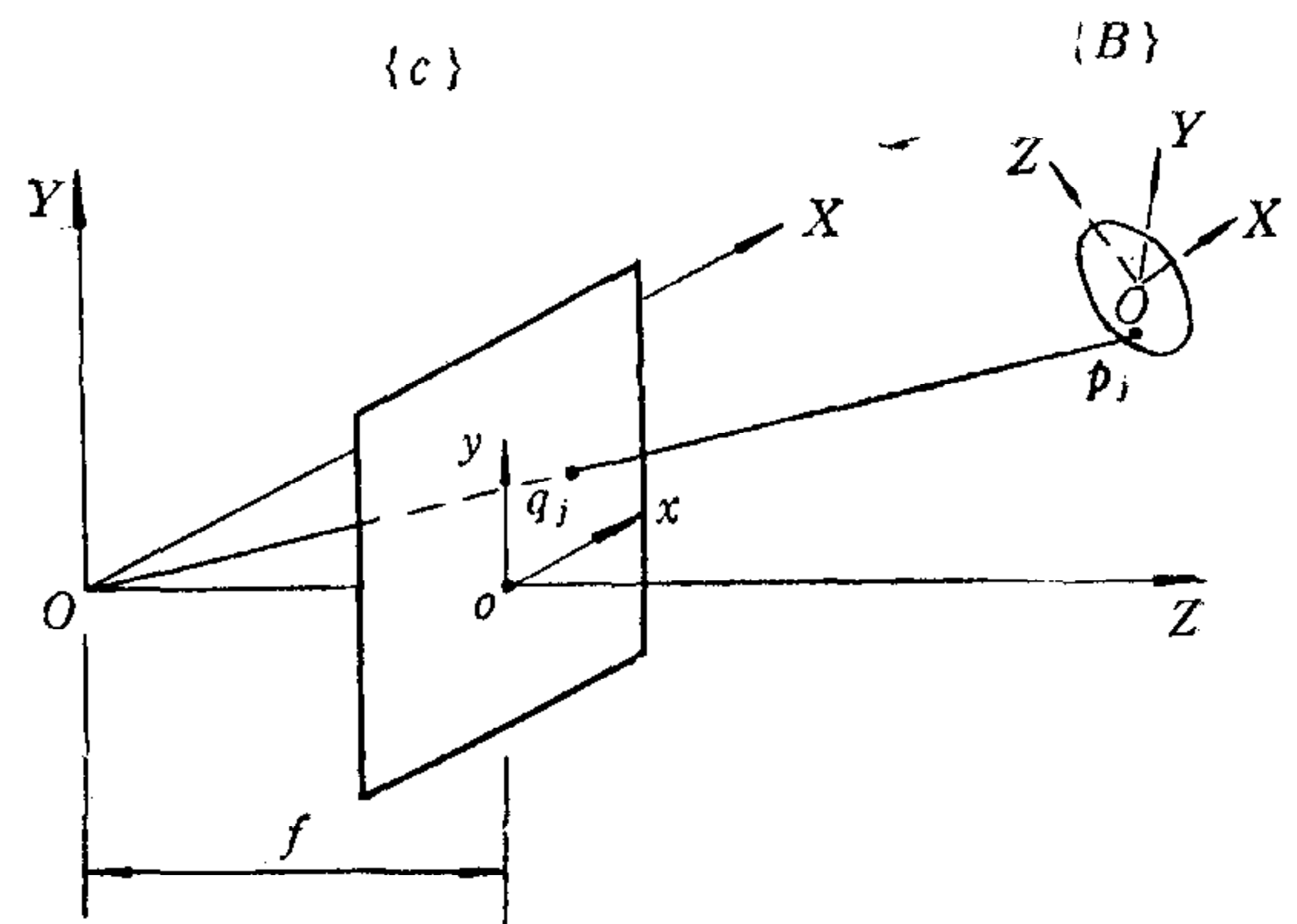


图 1 相机坐标系  $\{C\}$  的设置

这里,  $B^T = [PXU PYU]$ ,  $U = I - P^+P$ . 不难证明  $U^T U = U^2 = U$ , 于是有

$$B^T B = P W P^T, \quad (11)$$

这里,  $W = XUX + YUY$ , 或写成元素形式

$$W = (w_{ij})_{n \times n}, \quad w_{ij} = (x_i x_j + y_i y_j) u_{ij}. \quad (12)$$

由二次型最优化理论知, 使式(10)的二次型最小且模为1的  $t_3$  应是和  $3 \times 3$  对称矩阵  $B^T B$  的最小特征值相对应的单位特征向量. 记  $t_3^*$  为所得的最优解, 由式(1), (3)和(4)得

$$p_i^c = \lambda \tilde{q}_i, \quad \tilde{q}_i = t_3^{*T} p_i q_i, \quad j = 1, \dots, n. \quad (13)$$

至此, 除比例因子  $\lambda$  未定外, 已得到特征点  $p_i$  在相机坐标系中的位矢  $p_i^c$ . 如何确定  $\lambda$ , 将在下节讨论.

## 2. 特征点非共面的情形

一般地设定  $p_i^b = (X_i^b, Y_i^b, Z_i^b)^T$ . 此时, 方程(2)仍成立, 只要将式(3)中的  $p_i$  和  $T$  重新定义如下:

$$p_i = (X_i^b, Y_i^b, Z_i^b, 1)^T, \quad T = [r_1 \ r_2 \ r_3 \ t]. \quad (14)$$

和式(3)相比,  $p_i$  由  $3 \times 1$  向量变成了  $4 \times 1$  向量, 而  $T$  则由  $3 \times 3$  矩阵变成了  $3 \times 4$  矩阵. 不难看出, 前面提出的共面点情形下估计特征点空间位置的方法同样适用于非共面的一般情形, 只要注意上述的维数变化即可. 由于在一般情形下,  $T$  有 12 个元素, 而每个特征点仅提供 2 个求解  $T$  的方程, 见式(5), 因此, 在一般情形下确定  $T$  及点的空间位置至少需要 6 个点.

## 三、估计运动参数

本节讨论如何估计运动参数  $R$  和  $t$ , 使其导致的特征点的空间位置  $p_i^c = R p_i^b + t$  和前节得到的估计值  $p_i^c = \lambda \tilde{q}_i$  尽可能接近. 确切地说, 是要求一个旋转阵  $R$ 、一个向量  $t$  和一个正数  $\lambda$  使下述目标函数最小:

$$J_1 = \sum_{i=1}^n \|\lambda \tilde{q}_i - (R p_i^b + t)\|^2. \quad (15)$$

不难看出, 给定  $\lambda$  和  $R$ , 最优的  $t$  应为

$$t = \lambda \tilde{q}_c - R p_c^b. \quad (16)$$

这里,  $\tilde{q}_c = \frac{1}{n} \sum_{i=1}^n \tilde{q}_i$ ,  $p_c^b = \frac{1}{n} \sum_{i=1}^n p_i^b$ . 式(15)可改写为

$$J_1 = \sum_{i=1}^n \|\lambda \bar{q}_i - R \bar{p}_i\|^2 = \|\lambda \bar{Q} - R \bar{P}\|^2, \quad (17)$$

这里,

$$\begin{aligned} \bar{q}_i &= \tilde{q}_i - \tilde{q}_c, \quad \bar{Q} = [\bar{q}_1, \dots, \bar{q}_n], \\ \bar{p}_i &= p_i^b - p_c^b, \quad \bar{P} = [\bar{p}_1, \dots, \bar{p}_n]. \end{aligned} \quad (18)$$

下面, 引用作者在另一篇文章<sup>[8]</sup>中的结果.

**定理 1.** 对任一  $n \times n$  实矩阵  $A$ , 存在一个旋转阵  $R$  和一个对称阵  $S$ , 使得  $A =$

$RS$ . 而且,  $S$  总可以选择为相似于对角阵  $\text{diag}(s_1, \dots, s_n)$  使得

$$s_1 \geq \dots \geq s_{n-1} \geq |s_n|.$$

满足该相似条件的  $RS$  分解称作规格化  $RS$  分解.

**定理 2.** 给定一对  $3 \times n$  矩阵  $\bar{P}$  和  $\bar{Q}$ , 建立目标函数  $J = \|\bar{Q} - R\bar{P}\|^2$  ( $R$  为旋转阵). 若  $\bar{R}\bar{S}$  是矩阵  $\bar{P}\bar{Q}^T$  的规格化  $RS$  分解, 则使  $J$  取最小值的最优旋转阵为  $R^* = \bar{R}^T$ ,  $J$  的最小值为  $J^* = \|\bar{P}\|^2 + \|\bar{Q}\|^2 - 2\text{tr}(\bar{S})$ .

以上两定理的证明以及规格化  $RS$  分解的方法和唯一性问题参考文献[8].

关于式(17)的最优化问题. 设  $\bar{R}\bar{S}$  是矩阵  $\bar{P}\bar{Q}^T$  的规格化  $RS$  分解, 则  $\bar{R}(\lambda\bar{S})$  是矩阵  $\bar{P}(\lambda\bar{Q})^T$  的规格化  $RS$  分解. 根据定理 2, 使式(17)的  $J_1$  取最小值的旋转阵为  $R^* = \bar{R}^T$ , 相应的  $J_1$  值为:  $J_1 = \|\bar{P}\|^2 + \|\lambda\bar{Q}\|^2 - 2\text{tr}(\lambda\bar{S})$ , 它是  $\lambda$  的函数. 令

$$\frac{\partial J_1}{\partial \lambda} = 0,$$

可得使  $J_1$  取最小值的  $\lambda$  值:  $\lambda^*$ . 有了  $R^*$  和  $\lambda^*$ , 最优的  $t$  可由式(16)确定. 综上, 有

$$\lambda^* = \text{tr}(\bar{S})/\|\bar{Q}\|^2, R^* = \bar{R}^T, t^* = \lambda^* \bar{q}_c - R^* p_c^b. \quad (19)$$

## 四、迭代修正

如图 1 所示,  $q_i$  是点  $p_i$  的象点. 理想情况下, 点  $p_i$  必定在射线  $Oq_i$  上, 但当  $q_i$  有观察误差时, 该结论就不成立了. 本节讨论如何估计运动参数  $R$  和  $t$ , 使其导致的特征点空间位置  $p_i^c = Rp_i^b + t$  尽可能接近所对应的射线  $Oq_i$ , 即希望找到一个旋转阵  $R$ 、一个向量  $t$  和一组正数  $\lambda_j, j = 1, \dots, n$ , 使得下述目标函数最小:

$$J_2 = \sum_{i=1}^n \|\lambda_i q_i - (Rp_i^b + t)\|^2. \quad (20)$$

比较这里的  $J_2$  和式(15)的  $J_1$  可以看出, 二者在形式上是一样的, 不同之处仅在于  $J_1$  中的  $\lambda$  是不依赖于各别特征点的公因子, 而  $J_2$  中的  $\lambda_j$  是和特征点有关的. 对于  $J_1$ , 导出了一个闭式解; 而对于  $J_2$ , 却求不出闭式解, 下面给出一种迭代法.

由式(20)知, 对于给定的  $R$  和  $t$ , 最优的  $\lambda_j$  应为  $\lambda_j = (q_i^T q_i)^{-1} q_i^T (Rp_i^b + t)$ , 代入式(20)得

$$J_2 = \sum_{i=1}^n \|g_i\|^2, g_i = A_i (Rp_i^b + t). \quad (21)$$

这里,  $A_i = (q_i^T q_i)^{-1} q_i q_i^T - I$ , 是一个常数矩阵, 与  $R$  和  $t$  无关, 而  $g_i$  是  $R$  和  $t$  的函数. 用  $\text{Rot}(\mathbf{k}, \phi)$  来表示旋转阵, 其对应的转轴为  $\mathbf{k}$  (单位向量), 转角(按右手惯例)为  $\phi$ , 于是有<sup>[9]</sup>

$$\text{Rot}(\mathbf{k}, \phi) = \cos \phi I + (1 - \cos \phi) \mathbf{k} \mathbf{k}^T + \sin \phi \mathbf{k}^X, \quad (22)$$

这里,  $\mathbf{k}^X$  定义为对应于向量  $\mathbf{k}$  的反对称矩阵

$$\mathbf{k} = \begin{bmatrix} k_x \\ k_y \\ k_z \end{bmatrix}, \quad \mathbf{k}^X = \begin{bmatrix} 0 & -k_z & k_y \\ k_z & 0 & -k_x \\ -k_y & k_x & 0 \end{bmatrix}. \quad (23)$$

定义

$$\Delta R = \text{Rot}(\mathbf{k}, \Delta\phi), \quad \Delta\phi = (\Delta\phi)\mathbf{k}. \quad (24)$$

当  $\Delta\phi$  足够小时,由式(22),(23)和(24)可知下式近似成立:

$$\Delta R = I + (\Delta\phi)^X. \quad (25)$$

定义  $\Delta(R\mathbf{p}) = (\Delta R)R\mathbf{p} - R\mathbf{p}$ , 由式(23), (25)得

$$\Delta(R\mathbf{p}) = (\Delta R - I)R\mathbf{p} = (\Delta\phi)^X R\mathbf{p} = -(R\mathbf{p})^X \Delta\phi. \quad (26)$$

上式导致

$$\frac{\partial(R\mathbf{p})}{\partial\phi} = -(R\mathbf{p})^X. \quad (27)$$

由式(21),(27)得

$$\frac{\partial \mathbf{g}_i}{\partial \mathbf{t}} = A_i, \quad \frac{\partial \mathbf{g}_i}{\partial \phi} = -A_i(R\mathbf{p}_i^b)^X. \quad (28)$$

有了这两个雅可比矩阵,高斯-牛顿迭代法<sup>[10]</sup>可有效地用于式(21)的非线性最小二乘问题,即寻找  $R$  和  $\mathbf{t}$  使  $J_2$  最小,具体算法如下:

- 1) 用第三节的闭式解求得的  $R$  和  $\mathbf{t}$  作初始解.
- 2) 由目前的  $R$  和  $\mathbf{t}$  计算  $\mathbf{g}_i$ ,  $\partial \mathbf{g}_i / \partial \mathbf{t}$  和  $\partial \mathbf{g}_i / \partial \phi$ .
- 3) 求下列线性方程组的最小二乘解  $\Delta \mathbf{t}$  和  $\Delta \phi$ .

$$\frac{\partial \mathbf{g}_i}{\partial \mathbf{t}} \Delta \mathbf{t} + \frac{\partial \mathbf{g}_i}{\partial \phi} \Delta \phi = -\mathbf{g}_i, \quad j = 1, \dots, n.$$

4) 计算  $\Delta\phi = \|\Delta\phi\|$ ,  $\mathbf{k} = \Delta\phi / \Delta\phi$ ,  $\Delta R = \text{Rot}(\mathbf{k}, \Delta\phi)$ , 令  $\mathbf{t} \leftarrow \mathbf{t} + \Delta \mathbf{t}$ ,  $R \leftarrow (\Delta R)R$ .

5) 对于给定的小的正数  $\varepsilon_t$  和  $\varepsilon_\phi$ , 若  $\|\Delta \mathbf{t}\| \leq \varepsilon_t$  且  $\Delta\phi \leq \varepsilon_\phi$ , 则停止迭代,否则转步骤 2).

## 五、相机阵列

本节提出的相机阵列是一种多相机系统. 在该阵列中, 每个相机具有相同的焦距以及和图 1 相同的坐标系设置. 相机的安排使得各个相机坐标系的原点(光心)彼此共面, 相应的坐标轴互相平行. 假定一个相机阵列由  $m$  个相机组成, 它们的坐标系分别记作  $\{C_k\}$ ,  $k = 1, \dots, m$ , 则可以想象在这个阵列中存在有一个虚构的“中心相机”, 它的坐标系, 记作  $\{C_o\}$ , 的原点(光心)是所有  $m$  个坐标系原点(光心)的中点. 设特征点  $p_i$  在相机  $\{C_k\}$  上的象点为  $q_{ik}$ ,  $k = 1, \dots, m$ , 在虚构的“中心相机”  $\{C_o\}$  上的虚构象点为  $q_{io}$ . 用  $\mathbf{q}_{ik}$  和  $\mathbf{q}_{io}$  分别表示点  $q_{ik}$  和  $q_{io}$  在坐标系  $\{C_k\}$  和  $\{C_o\}$  中的位矢. 根据前述相机阵列的几何结构, 则有

$$\mathbf{q}_{io} = \frac{1}{m} \sum_{k=1}^m \mathbf{q}_{ik}, \quad j = 1, \dots, n. \quad (29)$$

假定  $\mathbf{q}_{ik}$  具有独立的随机误差, 其均值为 0, 方差为  $\sigma^2$ . 由式(29)知,  $\mathbf{q}_{io}$  具有的随机

误差,其均值为 0, 方差为  $\sigma^2/m$ . 这说明,在统计意义上,  $q_{j_0}$  较  $q_{i_k}$  误差小. 基于  $q_{j_0}$ ,  $j = 1, \dots, n$ , 用前面提出的闭式解和迭代修正方法可以估计出物体  $\{B\}$  相对于“中心相机”  $\{C_0\}$  的运动参数  $R_0$  和  $t_0$ . 有时需要知道  $\{B\}$  相对于  $\{C_k\}$ ,  $k \in \{1, \dots, m\}$  的运动参数  $R_k$  和  $t_k$ . 如果已知  $\{C_k\}$  和  $\{C_0\}$  的相对位移是  $d_{0k}$ , 则

$$R_k = R_0, \quad t_k = t_0 + d_{0k}. \quad (30)$$

显然,若已知各相机之间准确的相对位置,则  $d_{0k}$  可事先算出. 事实上,即使不知道相机间的相对位置,  $d_{0k}$  也可按下式在线估计出来:

$$d_{0k} = \frac{1}{n} \sum_{j=1}^n (q_{i_k} - q_{j_0}) Z_j^0. \quad (31)$$

这里,  $Z_j^0$  是点  $p_j$  在  $\{C_0\}$  中的  $Z$  向坐标,由向量  $R_0 p_j^b + t_0$  的第三元素确定.

利用相机阵列估计运动参数的方法有如下特点: 1) 对阵列中每个相机而言所要完成的工作是相同的,因而可以平行处理. 2) 相机之间距离的大小并不影响结果的精度,因而相机阵列可以安排得十分紧凑. 3) 估计物体相对于阵列中任一相机的运动参数并不需要事先知道各个相机之间的相对位置.

## 六、仿真和结果

作者对本文提出的各种方法进行了计算机仿真试验. 试验的基本步骤如下:

- 1) 给定各特征点在物体坐标系  $\{B\}$  中的位矢以及  $\{B\}$  相对于参考相机坐标系  $\{C\}$  的运动参数  $R$  和  $t$ , 由此求得特征点在  $\{C\}$  中的位矢.
- 2) 给定相机焦距  $f$ , 根据透视投影以及特征点在  $\{C\}$  中的位置, 确定出相应的象点.
- 3) 在准确象点的  $x$  和  $y$  坐标上分别加上独立的随机噪声, 以模拟实际观察到的象点.
- 4) 基于加噪声后的象点, 利用第三节的闭式解求得运动参数的估计值  $R^*$  和  $t^*$ , 并用第四节的迭代法进行修正.
- 5) 计算运动参数的估计值和给定值之间的差异  $\Delta t = t - t^*$  以及  $\Delta R = RR^{*T}$ , 记  $\Delta R$  所对应的旋转角为  $\Delta\phi$ .

6) 重复执行步骤 3)——5), 共 1000 次, 统计出  $\|\Delta t\|$  和  $|\Delta\phi|$  的均值和方差.

表 1—4 给出了大量试验数据中的一组典型例子. 在这组例子中采用的设定值如下: 四个基本特征点在  $\{B\}$  中的坐标为  $(-10, 20, -30)$ ,  $(10, 20, -30)$ ,  $(-10, 95, -30)$ ,  $(10, 95, -30)$ , 单位为  $\text{cm}$ . 这四个点构成了一个矩形, 在此矩形内, 可以任意增加特征点的个数. 物体  $\{B\}$  对于参考相机  $\{C\}$  的运动参数给定为  $R = \text{Rot}(y, 10^\circ)$ ,  $t = (0, 0, t_z)^T$ ,  $t_z$  的取值为 250, 500, 750, 1000, 单位为  $\text{cm}$ . 相机的焦距设为  $f = 12 \text{ mm}$ . 对于相机阵列, 相邻相机间的距离取为 5  $\text{cm}$ , 这个值仅用来产生象点, 在估计运动参数时认为是未知的. 加在象点上的噪声设为  $\mathcal{N}(0, \sigma^2)$ ,  $\sigma$  选为 0.3, 0.6, 0.9, 1.2, 单位为 0.01  $\text{mm}$ . 迭代修正算法中的容差  $\varepsilon_t$  和  $\varepsilon_\phi$  分别取 0.1  $\text{cm}$  和 0.01 弧度.

表 1 单个相机 4 个特征点

$t_z$	$\sigma$	$\ \Delta t\ $		$ \Delta\phi $	
		均值	方差	均值	方差
250	.3	1.68	1.21	1.18	.375
		.797	.282	.546	.080
250	.6	3.27	5.19	2.30	1.41
		1.61	1.19	1.12	.336
250	.9	5.19	12.4	3.62	3.62
		2.37	2.50	1.67	.722
250	1.2	6.46	21.0	4.67	6.11
		3.09	4.79	2.21	1.52
500	.3	7.32	29.1	5.26	7.52
		2.72	3.43	1.88	1.06
750	.3	17.8	332	12.2	35.7
		6.05	21.2	5.97	50.3
1000	.3	45.6	2636	20.9	95.9
		8.72	28.1	9.56	70.0

表 2 单个相机 16 个特征点

$t_z$	$\sigma$	$\ \Delta t\ $		$ \Delta\phi $	
		均值	方差	均值	方差
250	.3	.624	.115	.732	.220
		.413	.060	.340	.035
250	.6	1.23	.468	1.42	.758
		.813	.256	.677	.139
250	.9	1.83	.988	2.23	2.03
		1.17	.498	.974	.268
250	1.2	2.50	1.91	2.86	3.45
		1.63	1.03	1.35	.504
500	.3	2.81	2.34	3.33	4.69
		1.53	.970	1.15	.371
750	.3	6.67	16.2	7.63	23.4
		3.63	11.5	3.40	26.7
1000	.3	12.3	80.0	13.6	70.4
		6.66	21.6	7.96	71.3

在表 1—4 中, 对应于每一组  $t_z$  和  $\sigma$ , 上行给出的是利用闭式解得到的结果, 下行给出的是迭代修正后的结果。为方便起见, 称  $\|\Delta t\|$  以及  $|\Delta\phi|$  的均值为平均估计误差, 观察每张表列出的结果, 不难看出, 平均估计误差随象噪声的标准离差  $\sigma$  的增加按线性关系增长, 随物体离相机的纵深  $t_z$  的增加按平方关系增长; 迭代修正可以使平均估计

表 3 4 个相机 4 个特征点

$t_z$	$\sigma$	$\ \Delta t\ $		$ \Delta\phi $	
		均值	方差	均值	方差
250	.3	.810	.296	.573	.093
		.393	.070	.280	.023
250	.6	1.60	1.13	1.18	.365
		.796	.285	.575	.090
250	.9	2.41	2.77	1.74	.850
		1.20	.603	.879	.193
250	1.2	3.22	4.82	2.35	1.46
		1.53	.979	1.12	.335
500	.3	3.63	5.80	2.62	1.79
		1.25	.723	.915	.239
750	.3	8.31	46.1	6.20	10.9
		2.48	4.14	1.88	5.75
1000	.3	17.9	390	11.0	28.3
		4.21	12.1	4.11	34.4

表 4 9 个相机 36 个特征点

$t_z$	$\sigma$	$\ \Delta t\ $		$ \Delta\phi $	
		均值	方差	均值	方差
250	.3	.142	.006	.163	.016
		.088	.003	.049	.004
250	.6	.284	.023	.363	.058
		.166	.010	.138	.010
250	.9	.418	.054	.521	.111
		.267	.026	.228	.017
250	1.2	.539	.091	.690	.220
		.339	.041	.305	.031
500	.3	.644	.115	.812	.256
		.310	.039	.250	.022
750	.3	1.51	.707	1.87	1.44
		.581	.142	.448	.055
1000	.3	2.92	2.80	3.44	4.86
		.867	.316	.651	.112

误差减小为闭式解的一半。将表 1 和表 2、表 3 比较可以看出,增加特征点的个数或增加相机的个数均可减小平均估计误差。将表 1 和表 4 比较可以看出,多特征点、相机阵列以及迭代修正三种方法的组合使用可导致相当高的估计精度。



## 七、结 束 语

本文研究了计算机视觉中的 PNP 问题,这个问题本质上是非线性最优化问题,本文提出了两种目标函数  $J_1$  和  $J_2$ 。后者较前者更合理,因而有更好的估计精度;但前者可导出一个闭式解,后者则需用迭代法。利用闭式解作为迭代的初始解,并采用本文所设计的对  $J_2$  求雅可比阵的方法,可使迭代过程很快结束,除迭代修正外,为了提高估计精度,一个可行的办法是增加特征点的个数,这将增加象点提取和对应等问题的复杂性以及计算的工作量,在实际应用中增加特征点数目有时是不现实的。本文提出的采用相机阵列的设想既可以有效地提高估计精度,又不增加问题的复杂性和计算工作量。计算机仿真试验的结果表明,本文提出的各种方法以及它们的相互组合对于克服象噪声对估计精度的影响是行之有效的。

## 参 考 文 献

- [1] Fischler, M. A. and Bolles, R. C., Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Proc of the DARPA Image Understanding Workshop, College Park, MD, 1980, 71—85.
- [2] Wylie, C. R., Introduction to Projective Geometry, McGraw-Hill Book Company, New York, 1970.
- [3] Hung, Y., Yeh, P. and Harwood, D., Passive Ranging to Known Planar Point Sets, Proc. of the IEEE International Conference on Robotics and Automation, St. Louis, Missouri, 1985, 80—85.
- [4] Horaud, R., Conio, B., Leboulloux, O. and Lacolle, B., An Analytic Solution for the Perspective 4-Point Problem, *Computer Vision, Graphics, and Image Processing*, 47(1989), 1, 33—44.
- [5] Lowe, D. G., Solving for the Parameters of Object Models from Image Descriptions, Proc. of the DARPA Image Understanding Workshop, College Park, MD, 1980, 121—127
- [6] Lowe, D. G., Three-Dimensional Object Recognition from Single Two-Dimensional Images, *Artificial Intelligence*, 31(1987), 355—395
- [7] Lawson, C. L. and Hanson, R. J., Solving Least-Squares Problems, Prentice-Hall Inc., Englewood Cliffs, New Jersey, (1974).
- [8] Su, R. J. and Xu, W. L., Matrix Decomposition and Optimal Estimation of Rigid-Body Displacement, Proc. of the International Symposium MTNS-89, Amsterdam, The Netherland 3(1989), 513—520.
- [9] Bottema, O. and Roth, B., Theoretical Kinematics, North Holland Publishing Company, New York, 1979.
- [10] Dennis, J. E. and Schnabel, R. B., Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1983.

## OPTIMAL SOLUTIONS OF THE PNP PROBLEM IN COMPUTER VISION

XU WENLI

(Dept. of Automation, Tsinghua University, Beijing 100084)

### ABSTRACT

This paper gives a comprehensive treatment of the PNP (Perspective-N-Point) problem in computer vision: find the 3-D motion parameters of a viewed object relative to a camera from the observed image points of  $n$  known feature points. This problem is essentially one of non-linear optimization due to the image noise. In this paper, a closed-form solution is developed, and several methods to combat the image noise are presented. The simulation results show that the approaches given in this paper hold great promise in application.

**Key words:** Computer vision; PNP problem; motion parameters; RS decomposition; camera array.



**徐文立** 1947年生。1970年毕业于清华大学电机工程系。1981年于清华大学自动化系获硕士学位。1990年于美国科罗拉多大学获博士学位。现任清华大学自动化系副教授。目前研究领域有智能控制、计算机视觉、计算机控制与管理系统等。