1000-9825/2003/14(03)0682 ©2003 Journal of Software 软 件 学 报 Vol.14, No.3

P*n*P 问题的线性求解算法*

吴福朝+, 胡占义

(中国科学院 自动化研究所 模式识别国家重点实验室,北京 100080)

A Linear Method for the P*n*P Problem

WU Fu-Chao+, HU Zhan-Yi

(National Laboratory of Pattern Recognition, Institute of Automation, The Chinese Academy of Sciences, Beijing 100080, China)

+ Corresponding author: Phn: 86-10-62616540, E-mail: fcwu@nlpr.ia.ac.cn

http://nlpr-web.ia.ac.cn/english/rv

Received 2001-11-20; Accepted 2002-06-10

Wu FC, Hu ZY. A linear method for the P*n*P problem. *Journal of Software*, 2003,14(3):682~688.

Abstract: The classical P*n*P problem (3≤*n*≤5) is inherently non-linear and generally of multiple solutions and sensitive to errors associated with image points. In the classical P*n*P problem, only a single image is involved. Inspired by the problems in robot navigation, the P*n*P problem is extended to the two-image case where the camera undergoes a pure translation between the two images. The obtained main results are: given two images of *n* known control points captured by a translating camera, (1) When *n*=3, the camera's pose and two scaling parameters of the two image axes can be linearly determined; (2) When *n*≥4, the camera's pose and all its 5 intrinsic parameters can be linearly determined. In other words, the P*n*P problem with an uncalibrated camera can be linearly solved in this case. The results presented in this paper appears instructive from both theoretical and practical points of view.

Key words: P*n*P problem; camera intrinsic parameter; camera pose

摘 要: 经典的P*n*P(3≤*n*≤5)问题从本质上来说是非线性的,不但具有多解性而且对图像点的位置误差极为敏感.经典 P*n*P 问题仅涉及一幅图像,针对机器人导航中的实际需求,将 P*n*P 问题扩展到了摄像机作平移运动下的两幅图像的情况,并研究探讨了在这种情况下 P*n*P 问题的线性求解方法,主要结果有:给定平移运动下 n 个控制点的两幅图像,(1) 当 $n=3$ 时,可线性求解摄像机的方位以及摄像机内参数的两个尺度因子;(2) 当 $n\geq 4$ 时,不仅可以线性求解摄像机的方位,而且能够确定摄像机的所有内参数.也就是说,给定平移视点下的两幅控制点图像,可以线性求解摄像机未标定的 P*n*P 问题.结果具有一定的理论意义和应用价值.

关键词: P*n*P 问题;摄像机内参数;摄像机方位

中图法分类号: TP391 **文献标识码**: A

P*n*P 问题是计算机视觉、射影测量学乃至数学领域的一个经典问题.自从该问题首次于 1981 年提出之后[1],由于其在物体定位方面的重要应用价值,引起了人们的广泛重视.而后涉及该领域的大量文章的相继问世充分

* Supported by the National Natural Science Foundation of China under Grant Nos.60375009, 69975021 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2002AA422230 (国家高技术研究发展计划)

第一作者简介: 吴福朝(1957－),男,安徽安庆人,教授,博士生导师,主要研究领域为计算机视觉,人工智能.

© 中国科学院软件研究所 http://www.jos.org.cn

说明了这一点[1~14].文献中对 PnP 问题的研究集中在对 P3P,P4P 与 P5P 问题的研究上.这是因为,P2P 问题总有无限多组解,当 n>5 时,PnP 问题均可以线性求解.目前,对 P3P,P4P 与 P5P 问题的研究,除了无穷多解的情况,有如下结果:P3P 问题最多有 4 个解且解的上限可以达到[1,2,4,5];对于 P4P 问题,当 4 个控制点共面时有惟一解[1,7,12],当 4 个控制点不共面时,最多可能有 4 个解且解的上限可以达到[13];P5P 问题最多可能有两个解,且解的上限可以达到[14],在 PnP 问题的研究中大多都假定摄像机内参数是已知的,即摄像机是标定的,关于摄像机未标定的 PnP 问题的研究请参考文献[15].

对于 3≤n≤5 的情况,PnP 问题不但具有多解性,而且所有的求解算法均是非线性的.PnP 问题的非线性是其本身所固有的,也就是说,如果没有其他信息可以利用,PnP 问题不可能有线性求解算法.由于控制点非常少,非线性算法对图像点的位置误差极为敏感.在实际应用中,例如机器人定位问题,机器人在移动过程中,运动姿态并非每时每刻都在变化,它的运动轨迹通常是一条折线,这说明我们可以得到平移运动下控制点的图像序列.因此,研究以下问题就具有特别重要的实际意义:给定平移视点下的两幅控制点图像,能否线性求解摄像机的方位?这即是本文所要讨论的问题.

本文讨论的主要结果是:(1) 当 n=3 时,可线性求解摄像机的方位以及摄像机内参数中的两个尺度因子;(2)当 n≥4 时,不但可以线性求解摄像机的方位,而且能够确定摄像机的所有内参数.这就是说,给定平移视点的两幅控制点图像,我们可以线性地惟一求解摄像机未标定的 PnP 问题.因此,本文的结果不但具有重要的应用价值,同时也具有一定的理论意义.

在本文中,对符号的使用作如下规定:点与其坐标向量(列向量)用同一个小写黑体字母表示,如 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\ldots$;矩阵用大写黑体字母表示,如 $\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z},\ldots$;$\boldsymbol{X}^T(\boldsymbol{x}^T)$ 表示矩阵 $\boldsymbol{X}$(向量 $\boldsymbol{x}$)的转置;标量用小写字母表示,如 $x,y,z,\ldots$

1 问题陈述

在本文中,记摄像机内参数矩阵为 $\boldsymbol{K}$,摄像机坐标系 O_c-$x_cy_cz_c$ 关于物体坐标系 O-xyz 的方位(变换)记为 $(\boldsymbol{R}\quad \boldsymbol{t})$,即 $\boldsymbol{x}_c=\boldsymbol{R}\boldsymbol{x}+\boldsymbol{t}$.其中 $\boldsymbol{R}$ 是旋转矩阵,表示摄像机的方向;$\boldsymbol{t}$ 是平移向量,表示摄像机的位置.令 $\boldsymbol{x}=(x,y,z)^T\in R^3$ 是控制点在物体坐标系中的坐标,它在摄像机坐标系中的坐标记为 $\boldsymbol{x}_c=(x_c,y_c,z_c)^T\in R^3$,对应的齐次坐标分别记为 $\tilde{\boldsymbol{x}}=(\boldsymbol{x}^T,1)^T$ 与 $\tilde{\boldsymbol{x}}_c=(\boldsymbol{x}_c{}^T,1)^T$,其图像点坐标记为 $\boldsymbol{u}=(u,v)^T$,对应的齐次坐标记为 $\tilde{\boldsymbol{u}}=(\boldsymbol{u}^T,1)^T$,则有

$$s\tilde{\boldsymbol{u}}=\boldsymbol{K}(\boldsymbol{R}\boldsymbol{x}+\boldsymbol{t})=\boldsymbol{K}(\boldsymbol{R}\quad \boldsymbol{t})\tilde{\boldsymbol{x}}, \tag{1}$$

其中 $s=z_c>0$,这是因为空间点 $\tilde{\boldsymbol{x}}$ 总位于摄像机的前方,即 $z_c>0$.

已知 n 个控制点在物体坐标系中的坐标 $\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_2,\ldots,\tilde{\boldsymbol{x}}_n$,摄像机在初始位置关于 n 个控制点的图像坐标为 $\tilde{\boldsymbol{u}}_1^{(1)},\tilde{\boldsymbol{u}}_2^{(1)},\ldots,\tilde{\boldsymbol{u}}_n^{(1)}$,摄像机作平移运动后关于 n 个控制点的图像坐标为 $\tilde{\boldsymbol{u}}_1^{(2)},\tilde{\boldsymbol{u}}_2^{(2)},\ldots,\tilde{\boldsymbol{u}}_n^{(2)}$,求摄像机内参数 $\boldsymbol{K}$ 与平移运动前、后的方位 $(\boldsymbol{R}\quad \boldsymbol{t}^{(1)})$ 与 $(\boldsymbol{R}\quad \boldsymbol{t}^{(2)})$.

2 n=3 的情况

在本节中,我们作下述基本假设:

(1) 3 个图像点不共线;

(2) 每一个控制点都不在平移运动前、后两光心的连线上;

(3) 摄像机内参数矩阵 $\boldsymbol{K}=\mathrm{diag}(f_u,f_v,1)$,或等价地,已知摄像机的主点并且畸变因子等于 0.

命题 1. 若上述基本假设成立*,则从 3 个控制点可线性惟一求解摄像机内参数矩阵 $\boldsymbol{K}$ 以及方位 $(\boldsymbol{R}\quad \boldsymbol{t}^{(1)})$ 和 $(\boldsymbol{R}\quad \boldsymbol{t}^{(2)})$ 的充要条件是,3 个控制点所确定的平面与摄像机坐标系的 3 个坐标轴均不平行.

下面我们给出命题 1 的构造性证明,即证明过程也是求解过程.

☛ 如果基本假设中任何一条不成立,则问题的解不惟一;当(3)不成立时,问题有无穷多解.

© 中国科学院软件研究所 http://www.jos.org.cn

以 3 个控制点 $\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_2,\tilde{\boldsymbol{x}}_3$ 所在的平面π为坐标平面 $O\text{-}xy$,以垂直于平面π的直线为 z 轴建立物体坐标系,则 $\tilde{\boldsymbol{x}}_i=(x_i,y_i,0,1)^T$.摄像机在初始位置关于3个控制点的图像坐标记为 $\tilde{\boldsymbol{u}}_1^{(1)},\tilde{\boldsymbol{u}}_2^{(1)},\tilde{\boldsymbol{u}}_3^{(1)}$,摄像机作平移运动后关于3个控制点的图像坐标记为 $\tilde{\boldsymbol{u}}_1^{(2)},\tilde{\boldsymbol{u}}_2^{(2)},\tilde{\boldsymbol{u}}_3^{(2)}$.记 $\boldsymbol{R}=(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\ \ \boldsymbol{r}_3)$,则 $\boldsymbol{r}_3=\boldsymbol{r}_1\times\boldsymbol{r}_2$.令

$$\boldsymbol{H}^{(1)}=\boldsymbol{K}\left(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\ \ \boldsymbol{t}^{(1)}\right),\boldsymbol{H}^{(2)}=\boldsymbol{K}\left(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\ \ \boldsymbol{t}^{(2)}\right),$$

$$\boldsymbol{X}=\begin{pmatrix}x_1&x_2&x_3\\y_1&y_2&y_3\\1&1&1\end{pmatrix},\boldsymbol{M}^{(1)}=\begin{pmatrix}u_1^{(1)}&u_2^{(1)}&u_3^{(1)}\\v_1^{(1)}&v_2^{(1)}&v_3^{(1)}\\1&1&1\end{pmatrix},\boldsymbol{M}^{(2)}=\begin{pmatrix}u_1^{(2)}&u_2^{(2)}&u_3^{(2)}\\v_1^{(2)}&v_2^{(2)}&v_3^{(2)}\\1&1&1\end{pmatrix},$$

则有

$$\boldsymbol{H}^{(1)}\boldsymbol{X}=\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)}\right),\tag{2}$$

$$\boldsymbol{H}^{(2)}\boldsymbol{X}=\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{s}^{(2)}\right),\tag{3}$$

其中 $\boldsymbol{s}^{(i)}=\left(s_1^{(i)}\ \ s_2^{(i)}\ \ s_3^{(i)}\right)^T$ 是每个分量均大于 0 的未知三维向量,每个分量表示相应图像点的尺度因子,如式(1)所示.由假设(1), $\boldsymbol{X},\boldsymbol{M}^{(1)},\boldsymbol{M}^{(2)}$ 均为可逆矩阵,因此 $\boldsymbol{H}^{(i)}$ 也是可逆矩阵,所以

$$\boldsymbol{K}\left(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\ \ \boldsymbol{t}^{(1)}\right)=\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)}\right)\boldsymbol{P},\quad \boldsymbol{K}\left(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\ \ \boldsymbol{t}^{(2)}\right)=\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{s}^{(2)}\right)\boldsymbol{P},$$

其中 $\boldsymbol{P}=\boldsymbol{X}^{-1}=\left(\boldsymbol{p}_1\ \ \boldsymbol{p}_2\ \ \boldsymbol{p}_3\right)$.这样就有

$$\boldsymbol{K}\left(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\right)=\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)}\right)\left(\boldsymbol{p}_1\ \ \boldsymbol{p}_2\right)=\left(\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{p}_1\right)\boldsymbol{s}^{(1)}\ \ \boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{p}_2\right)\boldsymbol{s}^{(1)}\right),\tag{4}$$

$$\boldsymbol{K}\left(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\right)=\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{s}^{(2)}\right)\left(\boldsymbol{p}_1\ \ \boldsymbol{p}_2\right)=\left(\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{p}_1\right)\boldsymbol{s}^{(2)}\ \ \boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{p}_2\right)\boldsymbol{s}^{(2)}\right).\tag{5}$$

因此,得到关于 $\boldsymbol{s}=\begin{pmatrix}\boldsymbol{s}^{(1)}\\\boldsymbol{s}^{(2)}\end{pmatrix}$ 的线性方程组

$$\boldsymbol{G}\boldsymbol{s}=0,\tag{6}$$

其中

$$\boldsymbol{G}=\begin{pmatrix}\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{p}_1\right)&-\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{p}_1\right)\\\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{p}_2\right)&-\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{p}_2\right)\end{pmatrix},\tag{7}$$

可证明 $rank\left(\boldsymbol{G}\right)=5$ (见附录),所以

$$\boldsymbol{s}=\eta\boldsymbol{s}^*,\tag{8}$$

其中 $\boldsymbol{s}^*=\begin{pmatrix}\boldsymbol{s}^{(1)*}\\\boldsymbol{s}^{(2)*}\end{pmatrix}$ 是 $\boldsymbol{G}^T\boldsymbol{G}$ 最小特征值的每个分量均大于 0 的单位特征向量,η为大于 0 的未知常数.于是从式(4)和式(5)有

$$\tilde{\eta}\left(\boldsymbol{r}_1\ \ \boldsymbol{r}_2\right)=\boldsymbol{K}^{-1}\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)*}\right)\left(\boldsymbol{p}_1\ \ \boldsymbol{p}_2\right)\left(=\boldsymbol{K}^{-1}\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{s}^{(2)*}\right)\left(\boldsymbol{p}_1\ \ \boldsymbol{p}_2\right)\right),\tag{9}$$

其中 $\tilde{\eta}=1/\eta>0$.令 $\boldsymbol{K}^{-T}\boldsymbol{K}^{-1}=\mathrm{diag}(x,y,1)$,从式(9)可推知

$$\boldsymbol{B}^T\mathrm{diag}(x,y,1)\boldsymbol{B}=\tilde{\eta}^2\boldsymbol{I}_{2\times2},\tag{10}$$

其中 $\boldsymbol{B}=\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)*}\right)\left(\boldsymbol{p}_1\ \ \boldsymbol{p}_2\right)=\begin{pmatrix}b_{11}&b_{21}\\b_{21}&b_{22}\\b_{31}&b_{32}\end{pmatrix}\left(=\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{s}^{(2)*}\right)\left(\boldsymbol{p}_1\ \ \boldsymbol{p}_2\right)\right)$.方程(10)关于 $(x,y,\tilde{\eta}^2)$ 有惟一解的充要条件是,3 个控制点所确定的平面与摄像机坐标系的 3 个坐标轴均不平行,其证明请参考文献[12].当方程(10)有惟一解时,它的惟一解是

$$x=\frac{b_{31}b_{32}(b_{22}^2-b_{21}^2)+b_{21}b_{22}(b_{31}^2-b_{32}^2)}{b_{21}b_{22}(b_{11}^2-b_{12}^2)+b_{11}b_{12}(b_{22}^2-b_{21}^2)},\tag{11}$$

$$y=\frac{b_{31}b_{32}(b_{11}^2-b_{12}^2)+b_{11}b_{12}(b_{32}^2-b_{31}^2)}{b_{21}b_{22}(b_{11}^2-b_{12}^2)+b_{11}b_{12}(b_{22}^2-b_{21}^2)},\tag{12}$$

© 中国科学院软件研究所 http://www.jos.org.cn

$$\tilde{\eta}^2=\frac{(b_{11}b_{32}-b_{12}b_{31})(b_{21}b_{32}-b_{22}b_{31})}{b_{11}b_{21}+b_{12}b_{22}}. \tag{13}$$

因 $\tilde{\eta}=1/\eta>0$,所以

$$\tilde{\eta}=\sqrt{\frac{(b_{11}b_{32}-b_{12}b_{31})(b_{21}b_{32}-b_{22}b_{31})}{b_{11}b_{21}+b_{12}b_{22}}}. \tag{13a}$$

这样,我们得到摄像机内参数矩阵和方位的惟一解:

$$\boldsymbol{K}=\mathrm{diag}\left(\frac{1}{\sqrt{x}},\frac{1}{\sqrt{y}},1\right), \tag{14}$$

$$\boldsymbol{r}_j=\mathrm{diag}\left(\frac{\sqrt{x}}{\tilde{\eta}},\frac{\sqrt{y}}{\tilde{\eta}},1\right)\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)*}\right)\boldsymbol{p}_j\left(=\mathrm{diag}\left(\frac{\sqrt{x}}{\tilde{\eta}},\frac{\sqrt{y}}{\tilde{\eta}},1\right)\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{s}^{(2)*}\right)\boldsymbol{p}_j\right), j=1,2, \tag{15}$$

$$\boldsymbol{r}_3=\boldsymbol{r}_1\times\boldsymbol{r}_2, \tag{16}$$

$$\boldsymbol{t}^{(j)}=\mathrm{diag}\left(\frac{\sqrt{x}}{\tilde{\eta}},\frac{\sqrt{y}}{\tilde{\eta}},1\right)\boldsymbol{M}^{(j)}\mathrm{diag}\left(\boldsymbol{s}^{(j)*}\right)\boldsymbol{p}_3, j=1,2. \tag{17}$$

命题 1 成立. □

3 n=4 的情况

在本节中,我们作下述基本假设:

(1) 4 个控制点不共面;

(2) 每一个控制点都不在光心的连线上.

命题 2. 若上述基本假设成立*,则从 4 个控制点可线性惟一求解摄像机内参数矩阵 $\boldsymbol{K}$ 以及平移运动前、后的方位 $\left(\boldsymbol{R}\quad \boldsymbol{t}^{(1)}\right)$ 和 $\left(\boldsymbol{R}\quad \boldsymbol{t}^{(2)}\right)$.

命题 2 的证明方法与命题 1 类似,这里仅给出证明要点而略去其细节.令

$$\boldsymbol{H}^{(1)}=\boldsymbol{K}\left(\boldsymbol{R}\quad \boldsymbol{t}^{(1)}\right), \boldsymbol{H}^{(2)}=\boldsymbol{K}\left(\boldsymbol{R}\quad \boldsymbol{t}^{(2)}\right),$$

$$\boldsymbol{X}=\left(\tilde{\boldsymbol{x}}_1\quad \tilde{\boldsymbol{x}}_2\quad \tilde{\boldsymbol{x}}_3\quad \tilde{\boldsymbol{x}}_4\right), \boldsymbol{P}=\left(\boldsymbol{p}_1\quad \boldsymbol{p}_2\quad \boldsymbol{p}_3\quad \boldsymbol{p}_4\right)=\boldsymbol{X}^{-1},$$

$$\boldsymbol{M}^{(1)}=\left(\tilde{\boldsymbol{u}}_1^{(1)}\quad \tilde{\boldsymbol{u}}_2^{(1)}\quad \tilde{\boldsymbol{u}}_3^{(1)}\quad \tilde{\boldsymbol{u}}_4^{(1)}\right), \boldsymbol{M}^{(2)}=\left(\tilde{\boldsymbol{u}}_1^{(2)}\quad \tilde{\boldsymbol{u}}_2^{(2)}\quad \tilde{\boldsymbol{u}}_3^{(2)}\quad \tilde{\boldsymbol{u}}_4^{(2)}\right)$$

则有

$$\boldsymbol{KR}=\left(\boldsymbol{M}^{(1)}\mathrm{diag}(\boldsymbol{p}_1)\boldsymbol{s}^{(1)}\quad \boldsymbol{M}^{(1)}\mathrm{diag}(\boldsymbol{p}_2)\boldsymbol{s}^{(1)}\quad \boldsymbol{M}^{(1)}\mathrm{diag}(\boldsymbol{p}_3)\boldsymbol{s}^{(1)}\right), \tag{18}$$

$$\boldsymbol{KR}=\left(\boldsymbol{M}^{(2)}\mathrm{diag}(\boldsymbol{p}_1)\boldsymbol{s}^{(2)}\quad \boldsymbol{M}^{(2)}\mathrm{diag}(\boldsymbol{p}_2)\boldsymbol{s}^{(2)}\quad \boldsymbol{M}^{(2)}\mathrm{diag}(\boldsymbol{p}_3)\boldsymbol{s}^{(2)}\right), \tag{19}$$

其中 $\boldsymbol{s}^{(j)}=(s_1^{(j)},s_2^{(j)},s_3^{(j)},s_4^{(j)})^T$.由式(18)和式(19)得到关于 $\boldsymbol{s}=\begin{pmatrix}\boldsymbol{s}^{(1)}\\ \boldsymbol{s}^{(2)}\end{pmatrix}$ 的线性方程组:

$$\hat{\boldsymbol{G}}\boldsymbol{s}=0, \tag{20}$$

其中

$$\hat{\boldsymbol{G}}=\begin{pmatrix}\boldsymbol{M}^{(1)}\mathrm{diag}(\boldsymbol{p}_1) & -\boldsymbol{M}^{(2)}\mathrm{diag}(\boldsymbol{p}_1)\\ \boldsymbol{M}^{(1)}\mathrm{diag}(\boldsymbol{p}_2) & -\boldsymbol{M}^{(2)}\mathrm{diag}(\boldsymbol{p}_2)\\ \boldsymbol{M}^{(1)}\mathrm{diag}(\boldsymbol{p}_3) & -\boldsymbol{M}^{(2)}\mathrm{diag}(\boldsymbol{p}_3)\end{pmatrix}. \tag{21}$$

可证明 $rank\left(\hat{\boldsymbol{G}}\right)=7$, 所以

$$\boldsymbol{s}=\eta\boldsymbol{s}^*, \tag{22}$$

* 如果基本假设中任何一条不成立,则问题的解不惟一.换句话说,基本假设是线性惟一求解的充要条件.

© 中国科学院软件研究所 http://www.jos.org.cn

其中 $\boldsymbol{s}^*=\begin{pmatrix}\boldsymbol{s}^{(1)*}\\ \boldsymbol{s}^{(2)*}\end{pmatrix}$ 是 $\hat{\boldsymbol{G}}^T\hat{\boldsymbol{G}}$ 最小特征值的单位特征向量,η为未知的非 0 常数.于是从式(18)和式(19)有

$$\boldsymbol{KR}=\eta\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)*}\right)\left(\boldsymbol{p}_1\quad \boldsymbol{p}_2\quad \boldsymbol{p}_3\right)\left(=\eta\boldsymbol{M}^{(2)}\mathrm{diag}\left(\boldsymbol{s}^{(2)*}\right)\left(\boldsymbol{p}_1\quad \boldsymbol{p}_2\quad \boldsymbol{p}_3\right)\right).$$

对矩阵 $\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)*}\right)\left(\boldsymbol{p}_1\quad \boldsymbol{p}_2\quad \boldsymbol{p}_3\right)$ 进行 QR 分解,得到

$$\boldsymbol{M}^{(1)}\mathrm{diag}\left(\boldsymbol{s}^{(1)*}\right)\left(\boldsymbol{p}_1\quad \boldsymbol{p}_2\quad \boldsymbol{p}_3\right)=\boldsymbol{K}^*\boldsymbol{R}^*.$$

于是有

$$\boldsymbol{K}=\frac{1}{(\boldsymbol{K}^*)_{33}}\boldsymbol{K}^*,\ \boldsymbol{R}=\boldsymbol{R}^*,\ \boldsymbol{t}^{(1)}=\boldsymbol{K}^{*-1}\boldsymbol{M}^{(1)}\mathrm{diag}(\boldsymbol{s}^{(1)*})\boldsymbol{p}_4,\ \boldsymbol{t}^{(1)}=\boldsymbol{K}^{*-1}\boldsymbol{M}^{(2)}\mathrm{diag}(\boldsymbol{s}^{(2)*})\boldsymbol{p}_4. \tag{23}$$

这样,我们就证明了命题 2. □

下面我们来介绍一种基于分层重建的求解算法,它具有更好的鲁棒性.基于分层重建的求解算法,主要分为以下 3 个步骤:

(1) 求图像点的仿射重建.

因为 $\{\tilde{\boldsymbol{u}}_1^{(1)},\tilde{\boldsymbol{u}}_2^{(1)},\tilde{\boldsymbol{u}}_3^{(1)},\tilde{\boldsymbol{u}}_4^{(1)}\}$ 与 $\{\tilde{\boldsymbol{u}}_1^{(2)},\tilde{\boldsymbol{u}}_2^{(2)},\tilde{\boldsymbol{u}}_3^{(2)},\tilde{\boldsymbol{u}}_4^{(2)}\}$ 是一对平移视点下两幅图像的一组对应点,而且平移视点间的基本矩阵是极点所确定的反对称矩阵,所以,根据 4 对对应点可以求出图像的极点 $\boldsymbol{e}$.由于平移视点间的无穷远平面的单应矩阵为单位矩阵,所以仿射意义下的投影矩阵为[16]

$$\boldsymbol{P}_A=\begin{pmatrix}\boldsymbol{I} & 0\\ \boldsymbol{I} & \boldsymbol{e}\end{pmatrix}. \tag{24}$$

于是,由下述方程

$$\begin{pmatrix}\boldsymbol{I} & 0\\ \boldsymbol{I} & \boldsymbol{e}\end{pmatrix}\tilde{\boldsymbol{x}}_j^{(A)}=\begin{pmatrix}s_j^{(1)}\boldsymbol{u}_j^{(1)}\\ s_j^{(2)}\boldsymbol{u}_j^{(1)}\end{pmatrix} \tag{25}$$

所确定的 $\{\tilde{\boldsymbol{x}}_1^{(A)},\tilde{\boldsymbol{x}}_2^{(A)},\tilde{\boldsymbol{x}}_3^{(A)},\tilde{\boldsymbol{x}}_4^{(A)}\}$ 与 4 个控制点 $\{\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_2,\tilde{\boldsymbol{x}}_3,\tilde{\boldsymbol{x}}_4\}$ 之间相差一个 3 维仿射变换 $\boldsymbol{A}$,其中 $\tilde{\boldsymbol{x}}_j^{(A)}$ 称为 $\tilde{\boldsymbol{x}}_j$ 的仿射重建.

(2) 由 $\{\tilde{\boldsymbol{x}}_1^{(A)},\tilde{\boldsymbol{x}}_2^{(A)},\tilde{\boldsymbol{x}}_3^{(A)},\tilde{\boldsymbol{x}}_4^{(A)}\}$ 与 $\{\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_2,\tilde{\boldsymbol{x}}_3,\tilde{\boldsymbol{x}}_4\}$ 求仿射变换 $\boldsymbol{A}$.

令 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{b}\\ 0 & c\end{pmatrix}$,则有

$$\begin{pmatrix}\boldsymbol{B} & \boldsymbol{b}\\ 0 & c\end{pmatrix}\tilde{\boldsymbol{x}}_j=\tilde{\boldsymbol{x}}_j^{(A)},\ j=1,2,3,4, \tag{26}$$

因 $\{\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_2,\tilde{\boldsymbol{x}}_3,\tilde{\boldsymbol{x}}_4\}$ 不共面,故方程(26)有惟一解:

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{b}\\ 0 & c\end{pmatrix}=\left(\tilde{\boldsymbol{x}}_1^{(A)}\quad \tilde{\boldsymbol{x}}_2^{(A)}\quad \tilde{\boldsymbol{x}}_3^{(A)}\quad \tilde{\boldsymbol{x}}_4^{(A)}\right)\left(\tilde{\boldsymbol{x}}_1\quad \tilde{\boldsymbol{x}}_2\quad \tilde{\boldsymbol{x}}_3\quad \tilde{\boldsymbol{x}}_4\right)^{-1}.$$

(3) 求 $\boldsymbol{K},\left(\boldsymbol{R}\quad \boldsymbol{t}^{(1)}\right),\left(\boldsymbol{R}\quad \boldsymbol{t}^{(2)}\right)$.

由于 $\boldsymbol{P}_A\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I} & 0\\ \boldsymbol{I} & \boldsymbol{e}\end{pmatrix}\begin{pmatrix}\boldsymbol{B} & \boldsymbol{b}\\ 0 & c\end{pmatrix}=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{b}\\ \boldsymbol{B} & \boldsymbol{b}+c\boldsymbol{e}\end{pmatrix}$ 是欧氏意义下的投影矩阵[16],所以有

$$\eta\begin{pmatrix}\boldsymbol{B} & \boldsymbol{b}\\ \boldsymbol{B} & \boldsymbol{b}+c\boldsymbol{e}\end{pmatrix}=\begin{pmatrix}\boldsymbol{KR} & \boldsymbol{Kt}^{(1)}\\ \boldsymbol{KR} & \boldsymbol{Kt}^{(2)}\end{pmatrix}. \tag{27}$$

对 $\boldsymbol{B}$ 进行 QR 分解: $\boldsymbol{B}=\boldsymbol{K}^*\boldsymbol{R}^*$,于是有

$$\boldsymbol{K}=\frac{1}{(\boldsymbol{K}^*)_{33}}\boldsymbol{K}^*,\ \boldsymbol{R}=\boldsymbol{R}^*,\ \boldsymbol{t}^{(1)}=\boldsymbol{K}^{*-1}\boldsymbol{b},\ \boldsymbol{t}^{(1)}=\boldsymbol{K}^{*-1}(\boldsymbol{b}+c\boldsymbol{e}). \tag{28}$$

© 中国科学院软件研究所 http://www.jos.org.cn

4 结束语

本文针对实际应用需求,将经典 PnP 问题扩展到了摄像机作平移运动下的两幅图像的情况.提出了这类 PnP 问题的线性求解方法,并给出了解的惟一性条件.本文所提出的方法不仅能够线性求解摄像机的方位,同时还能够线性求解摄像机的内参数,其结果无论是在理论上还是在实际应用中,都具有一定的意义.

References:

[1] Fishler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 1981,24(6):381~395.

[2] Haralick RM, Lee C-N, Ottenberg K, Noelle M. Analysis and solution of the three point perspective pose estimation problem. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hawaii, 1991. 592~598.

[3] Horaud R, Conio B, Leboulleux O. An analytic solution for the perspective 4-point problem. Computer Vision Graphics and Image Processing, 1989,47:33~44.

[4] Wolfe WJ, Mathis D. The perspective view of three points. IEEE Transactions on Pattern Analysis and machine Intelligence, 1991, 13(1):66~73.

[5] Su C, Xu YQ, Li H, Liu SQ, Li DG. Necessary and sufficient condition of positive root number of perspective-tree-point problem. Chinese Journal of Computers, 1998,21(12):1084~1095 (in Chinese with English Abstract).

[6] Haralick RM. Determining camera parameters from the perspective projection of a rectangle. Pattern Recognition, 1989,22(3):225~230.

[7] Penna MA. Determining camera parameters from the perspective projection of a quadrilateral. Pattern Recognition, 1991,24(6): 533~541.

[8] Quan L, Lan ZD, Linear *N*>= 4 point pose determination. In: Proceedings of the International Conference on Computer Vision. Bombay, 1998. 778~783.

[9] Liu ML, Wong KH. Pose estimation using four corresponding points. Pattern Recognition Letters, 1999,20(1):69~74.

[10] Dhome M, Richetin M, Lapreste JT, Rives G. Determination of the attitude of 3-D objects from a single perspective view. IEEE Transactions on Pattern Analysis and machine Intelligence, 1989,11(12):1265~1278.

[11] Abidi MA, Chandra T. A new efficient and direct solution for pose estimation using quadrangular targets: algorithm and evaluation. IEEE Transactions on PAMI, 1995,17(5):534~538.

[12] Hu ZY, Lei C, Wu FC. A short note on P4P problem. Chinese Journal of Automation, 2001,27(6):770~776 (in Chinese with English Abstract).

[13] Hu ZY, Wu FC. A note on the number of solutions of the non-coplanar P4P problem. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002,24(4):550~555.

[14] Wu FC, Hu ZY. A study on the P5P problem. Journal of Software, 2001,12(5):768~775 (in Chinese with English Abstract).

[15] Wu FC, Hu ZY. A note on the P5P problem with an uncalibrated camera. Chinese Journal of Computers, 2002,25(3):1221~1226 (in Chinese with English Abstract).

[16] Hartley RI. Multi Viewpoint Geometry in Computer Vision. Cambridge: Cambridge University Press, 2000.

附中文参考文献:

[5] 苏成,徐迎庆,李华,刘慎权,李冬果.判定 P3P 问题正解数目的充要条件.计算机学报,1998,21(12):1084~1095.

[12] 胡占义,雷成,吴福朝.关于共面 P4P 问题的研究.自动化学报,2001,27(6):770~776.

[13] 吴福朝,胡占义.关于 P5P 问题的研究.软件学报,2001,12(5):768~775.

[14] 吴福朝,胡占义.关于摄像机未标定的 P5P 问题研究.计算机学报,2002,25(3):1221~1226.

附录: $rank(\boldsymbol{G})=5$ 的证明.

不妨假定 $\boldsymbol{p}_i=(p_{1i}\ ,\ p_{2i}, p_{3i})^T\,(i=1,2)$ 的每一个分量均不为 0,否则通过适当的物体坐标变换就可以保证这一点(例如:取新的物体坐标系的原点在 3 个控制点所构成的三角形的内部,x,y 轴与三角形的 3 条边均不平行即可).于是,不难推知:

© 中国科学院软件研究所 http://www.jos.org.cn

$$G \approx \begin{pmatrix} I & -(\mathrm{diag}(p_1))^{-1}(M^{(1)})^{-1}M^{(2)}\mathrm{diag}(p_1) \\ -(\mathrm{diag}(p_2))^{-1}(M^{(2)})^{-1}M^{(1)}\mathrm{diag}(p_2) & I \end{pmatrix}$$

$$\approx \begin{pmatrix} I-(\mathrm{diag}(p_1))^{-1}(M^{(1)})^{-1}M^{(2)}\mathrm{diag}(p_1)(\mathrm{diag}(p_2))^{-1}(M^{(2)})^{-1}M^{(1)}\mathrm{diag}(p_2) & 0 \\ -(\mathrm{diag}(p_2))^{-1}(M^{(2)})^{-1}M^{(1)}\mathrm{diag}(p_2) & I \end{pmatrix}$$

$$\approx \begin{pmatrix} CD-CD & 0 \\ -(\mathrm{diag}(p_2))^{-1}C\mathrm{diag}(p_2) & I \end{pmatrix},$$

其中"≈"表示矩阵的等价关系, $C=(M^{(2)})^{-1}M^{(1)}$, $D=(\mathrm{diag}(p_1))^{-1}C\mathrm{diag}(p_2)$.这样必有

$$rank(G)=rank(I)+rank(CD-DC)=3+rank(CD-DC).$$

因此,我们只须证明 $rank(CD-DC)=2$.

令 $H=I+ea^T$ 是 3 个控制点所确定的平面π的单应矩阵,其中 e 是第 2 幅图像上的极点, $a^T=n^TK^{-1}$,这里,n 为关于摄像机初始坐标系的法向量,则必有

$$(I+ea^T)M^{(1)}=M^{(2)}\mathrm{diag}(w_1,w_2,w_3).$$

于是, $C=\mathrm{diag}(w_1,w_2,w_3)-(M^{(2)})^{-1}ea^TM^{(1)}$,因此可推知

$$CD-DC=\mathrm{diag}(e')[d]_\times\mathrm{diag}(a'),$$

其中 $e'=(M^{(2)})^{-1}e\neq 0, a'=a^TM^{(1)}\neq 0$, $[d]_\times$是由向量 $d=(d_1,d_2,d_3)^T$ 所确定的反对称矩阵,这里, $d_1=(p_{32}/p_{31})-(p_{22}/p_{21}), d_2=(p_{21}/p_{11})-(p_{32}/p_{31}), d_2=(p_{22}/p_{21})-(p_{21}/p_{11})$.

(I) $rank(\mathrm{diag}(e'))\geq 2$.否则 e' 有两个分量等于 0,不妨设 $e'_1=e'_2=0$,则必有 $e'_3u_3^{(3)}=e$,这说明控制点 x_3 在两光心的连线上,与基本假设(2)矛盾.故 $rank(\mathrm{diag}(e'))\geq 2$.

(II) $rank(\mathrm{diag}(a'))=3$.因 $a^T=n^TK^{-1}M^{(1)}$,若 a' 的某个分量等于 0,不妨设 $a'_1=0$,则必有 $n^TK^{-1}u_1^{(1)}=0$,即 $n^T\perp K^{-1}u_1^{(1)}$.而以 $K^{-1}u_1^{(1)}$ 为方向的射线必通过控制点 $\tilde{x}_1$,并令 $\hat{x}$ 是自光心出发,以 n 为方向的射线与平面的交点,则在三角形 $o_cx_1x^*$ 中有两个直角,这是不可能的.因此 a' 的分量均不等于 0,故 $rank(\mathrm{diag}(a'))=3$.

因 d 至多有一个分量等于 0(否则 p_1,p_2 线性相关),因此由(I)不难看出, $rank(\mathrm{diag}(e')[d]_\times)=2$.故由(II)可知, $rank(CD-DC)=2$. □

© 中国科学院软件研究所 http://www.jos.org.cn