

# IEP $n$ P: 一种基于 EP $n$ P 的相机位姿迭代估计算法

陈鹏, 王晨骁

河北工业大学控制科学与工程学院, 天津 300130

**摘要** 近年来, EP $n$ P 算法作为一种相机位姿估计的解析算法, 因其较低的计算复杂度而得到广泛的关注, 但该算法对图像噪声的稳健性不强。提出了一种基于 EP $n$ P 算法的迭代算法, 即 IEP $n$ P 算法。IEP $n$ P 算法保留了 EP $n$ P 算法的主要思想, 构造了 4 个虚拟控制点, 利用弱透视投影模型获得相机的初始位姿, 计算出虚拟控制点在相机坐标系下的坐标, 然后通过高斯-牛顿法对虚拟控制点在相机坐标系下的坐标进行优化求解, 最终通过解决绝对定向问题来获得对相机位姿的估计。IEP $n$ P 算法简化了 EP $n$ P 算法的计算过程。在不同的图像噪声水平下进行仿真实验, 结果表明, 相比于 EP $n$ P 算法, IEP $n$ P 算法不仅保持了较高的计算效率, 而且对图像噪声具有更强的稳健性。

**关键词** 成像系统; 测量; 位姿估计; P $n$ P 问题; 单目视觉

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201838.0411001

## IEP $n$ P: An Iterative Camera Pose Estimation Algorithm Based on EP $n$ P

Chen Peng, Wang Chenxiao

College of Control Science and Engineering, Hebei University of Technology, Tianjin 300130, China

**Abstract** As an analytical camera pose estimation algorithm, EP $n$ P algorithm has attracted much attention in recent years for its low computational complexity. However, it is not robust to image noise. Hence, an iterative version of EP $n$ P algorithm, called IEP $n$ P, is proposed. The primary ideas of EP $n$ P are preserved in IEP $n$ P, 4 virtual control points are introduced, and their coordinates in the camera coordinate system are obtained through an initialization process based on weak perspective projection model. Gaussian-Newton algorithm is applied to optimize the coordinates of the virtual control points in the camera coordinate system. Finally, the pose estimation result is acquired by solving an absolute orientation problem. Meanwhile, the computational process is simplified in IEP $n$ P. Simulations under different image noise levels are implemented, and the results show that IEP $n$ P is more robust than EP $n$ P to image noise while maintaining a high computational efficiency.

**Key words** imaging systems; measurement; pose estimation; P $n$ P problem; monocular vision

**OCIS codes** 110.3925; 120.0280; 100.4995; 150.113

## 1 引 言

基于视觉的位姿估计问题是假定已知世界坐标系中  $n$  个控制点及其在图像坐标系中对应的二维投影后, 求解相机坐标系相对于世界坐标系的位置和姿态关系的问题。其是模式识别、视觉导航、机器人定位, 以及增强现实等领域中的关键问题之一<sup>[1-2]</sup>。由于三维物体点和二维图像点间的对应关系已知, 因此位姿估计问题也经常被称为  $n$  点透视 (P $n$ P) 问题。P $n$ P 问题自 1981 年提出以来<sup>[3]</sup>, 受到了学者们的广泛关注。现有的算法大致可以分为两类: 解析算法和迭代算法。当  $n > 5$  时, P $n$ P 问题可以归结

为直接线性变换 (DLT) 问题, 从而能够在线性时间内解出<sup>[4]</sup>, 因此对解析算法的研究被聚焦到 P3P、P4P 和 P5P 问题<sup>[5]</sup>。在文献 [6-7] 中提出了任意  $n$  个控制点时 P $n$ P 问题的解法, 但这类解法的计算复杂度高, 在  $O(n^5)$  到  $O(n^8)$  之间。解析算法只有在控制点像素坐标能够被精确获取时, 才会得到满意的位姿估计结果, 而控制点像素坐标在提取过程中受到噪声干扰时, 位姿估计结果将不再可靠。

相较于解析算法, 迭代算法不仅在计算过程上更加简便、直接, 而且对图像噪声的稳健性也更强。典型的迭代算法是将位姿估计问题归结为非线性最小二乘问题, 并设定约束条件为旋转矩阵正交<sup>[8-9]</sup>, 从

收稿日期: 2017-09-26; 收到修改稿日期: 2017-11-05

作者简介: 陈鹏(1983—), 男, 博士, 讲师, 主要从事机器视觉、图像处理方面的研究。E-mail: holmes83@163.com

而可以用非线性优化算法解决,如 Levenberg-Marquadt 算法或高斯-牛顿(G-N)算法等<sup>[10-11]</sup>。但这些算法设计时仅从数学角度出发,未充分利用成像模型的结构特点,最终影响了算法的性能。基于弱透视投影模型的 POSIT(Pose from Orthography and Scaling with Iterations) 算法<sup>[12]</sup> 已经作为 OpenCV 的内置函数<sup>[13]</sup>而被广泛应用,尽管 POSIT 算法的计算效率非常高,但未明确定义用于优化的目标函数,故算法的收敛性不能被严格地证明。正交迭代(OI)算法是另一种解决 PnP 问题的迭代算法<sup>[14]</sup>,该算法最重要的贡献是提出了投影空间误差的概念。基于投影空间误差,OI 算法充分利用了透视投影模型的结构特点,因此在迭代过程中保证了旋转矩阵的正交性,不需要额外约束,并且能够被证明是全局收敛的。但正交迭代算法不能确保最终的收敛结果就是真实的相机位姿,尤其是当控制点共面时<sup>[15]</sup>。

近年来,EPnP 算法在各种文献中得到了广泛关注<sup>[16]</sup>,其计算复杂度为  $O(n)$ ,该算法被认为是最高效的解析算法<sup>[17]</sup>。当图像点可以被精确提取时,EPnP 算法的估计结果可以与 POSIT 算法、OI 算法相媲美。然而作为一种解析算法,EPnP 算法对图像噪声的稳健性不强,且在参数求解过程中还需要使用复杂的方程求解技术。故本文对 EPnP 算法进行改进,提出了迭代形式的 EPnP 算法,即 IEPnP 算法,以增强算法的稳健性,同时简化算法的求解过程,使算法更简洁、直接。在不同噪声水平下进行仿真实验,结果表明,相比于 EPnP 算法,IEPnP 算法在保持较高计算效率的同时,对图像噪声具有更高的抗干扰能力。

## 2 问题描述

在透视投影模型下,成像过程中包含的坐标系如图 1 所示。经过标定后,相机焦距记作  $f$ ,像素坐标系中图像主点的坐标记作  $\mathbf{u}_0 = [u_0, v_0]^T$ ,单位像素大小记作  $d_u \times d_v$ 。将  $n$  个控制点记作  $\mathbf{p}_i, i=1, 2, \dots, n$ , 其在世界坐标系中的坐标记作  $\mathbf{p}_i^w = [x_i^w, y_i^w, z_i^w]^T$ , 在相机坐标系中的坐标记作  $\mathbf{p}_i^c = [x_i^c, y_i^c, z_i^c]^T$ ,  $\mathbf{p}_i$  在图像坐标系中的二维投影记作  $\mathbf{p}_i = [x_i, y_i]^T$ , 在像素坐标系中的坐标记作  $\mathbf{u}_i = [u_i, v_i]^T$ 。

在齐次坐标系下,图 1 的成像过程可以表述为

$$\lambda_i \bar{\mathbf{u}}_i = \mathbf{K}[\mathbf{R} | \mathbf{t}] \bar{\mathbf{P}}_i^w, \quad (1)$$

式中  $\lambda_i$  为缩放因子,  $\bar{\mathbf{u}}_i$  和  $\bar{\mathbf{P}}_i^w$  分别为  $\mathbf{u}_i$  和  $\mathbf{p}_i^w$  的齐次

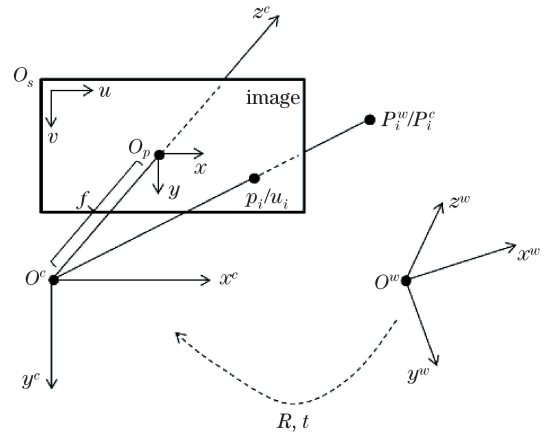


图 1 成像模型中包含的坐标系

Fig. 1 Coordinate systems involved in the imaging model

坐标。矩阵  $\mathbf{K}$  由下列相机内参数组成:

$$\mathbf{K} = \begin{bmatrix} f_u & \gamma & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

式中  $\gamma=0$ 。(1)式中  $\mathbf{R}$  和  $\mathbf{t}$  为相机的外参数,  $\mathbf{R}$  为旋转矩阵,且  $\mathbf{R} \in SO(3)$ , 表示相机的姿态,  $\mathbf{t}$  为平移向量,且  $\mathbf{t} \in \mathbf{R}^{3 \times 1}$ , 表示相机的位置。相机位姿估计的主要任务就是计算  $\mathbf{R}$  和  $\mathbf{t}$ 。

根据 EPnP 算法,在三维空间中设置 4 个虚拟控制点,记为  $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4$ 。理论上,世界坐标系中这 4 个虚拟控制点的坐标可以任意选取,但是为了简便起见,通常选取  $\mathbf{C}_1^w = [1, 0, 0]^T, \mathbf{C}_2^w = [0, 1, 0]^T, \mathbf{C}_3^w = [0, 0, 1]^T, \mathbf{C}_4^w = [0, 0, 0]^T$ 。任意一个真实控制点,如  $\mathbf{p}_i^w$ , 可以表示为 4 个虚拟控制点的线性组合:

$$\mathbf{p}_i^w = \sum_{j=1}^4 \alpha_{ij} \mathbf{C}_j^w, \quad \sum_{j=1}^4 \alpha_{ij} = 1. \quad (3)$$

由于世界坐标系和相机坐标系之间是相似变换的关系,因此(3)式中的组合系数也可以用在相机坐标系中,表示为

$$\mathbf{p}_i^c = \sum_{j=1}^4 \alpha_{ij} \mathbf{C}_j^c. \quad (4)$$

将(3)式和(4)式代入(1)式中,可得到:

$$\forall i, \lambda_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{K} \mathbf{p}_i^c = \mathbf{K} \sum_{j=1}^4 \alpha_{ij} \mathbf{C}_j^c. \quad (5)$$

记  $\mathbf{C}_j^c = [u_j^c, v_j^c, \omega_j^c]^T$ , 则(5)式可表示为

$$\forall i, \lambda_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \sum_{j=1}^4 \alpha_{ij} \begin{bmatrix} u_j^c \\ v_j^c \\ \omega_j^c \end{bmatrix}, \quad (6)$$

式中  $f_u = f/d_u, f_v = f/d_v, \lambda_i = \sum_{j=1}^4 \alpha_{ij} \omega_j^c$ 。组合系

数可以从(3)式中获得, 仅剩的未知变量为  $u_j^c$ 、 $v_j^c$ 、 $w_j^c$ , 即虚拟控制点在相机坐标系下的坐标。对(6)式进行整理可以得到关于  $u_j^c$ 、 $v_j^c$ 、 $w_j^c$  的等式:

$$\sum_{j=1}^4 \alpha_{ij} f_u u_j^c + \alpha_{ij} (u_0 - u_i) \mathbf{W}_j^c = 0, \quad (7)$$

$$\sum_{j=1}^4 \alpha_{ij} f_u v_j^c + \alpha_{ij} (v_0 - v_i) \mathbf{W}_j^c = 0. \quad (8)$$

将相机坐标系下 4 个虚拟控制点的 12 个坐标展开成一个列向量  $\mathbf{c} = [\mathbf{C}_1^{cT}, \mathbf{C}_2^{cT}, \mathbf{C}_3^{cT}, \mathbf{C}_4^{cT}]^T$ , 由(7)式和(8)式可以得到 12 个方程:

$$\mathbf{F}\mathbf{c} = 0, \quad (9)$$

式中  $\mathbf{F}$  为大小是  $2n \times 12$  的矩阵。显然, 向量  $\mathbf{c}$  位于  $\mathbf{F}$  的零空间中, 下述关系式成立:

$$\mathbf{c} = \sum_{i=1}^{N_F} \beta_i \mathbf{w}_i, \quad (10)$$

式中  $\mathbf{w}_i, i=1, 2, \dots, N_F$  为与  $\mathbf{F}^T \mathbf{F}$  的零特征值对应的特征向量。若系数  $\beta_i, i=1, 2, \dots, N_F$  被确定, 则可以根据(10)式计算出向量  $\mathbf{c}$  的值, 从而根据(4)式获得相机坐标系下真实控制点的坐标。根据  $\{\mathbf{p}_i^w\}$  和  $\{\mathbf{p}_i^c\}$  之间的对应关系, 依据文献[14]提出的绝对定向问题的解决方法即可计算出相机的位姿。

### 3 算法描述

为了获取系数  $\beta_i, i=1, 2, \dots, N_F$ , 首先要确定  $N_F$  的取值, 即  $\mathbf{F}$  零空间的维数。考虑到成像过程中唯一丢失的信息是场景点和相机之间的深度信息, 故  $N_F$  在透视投影模型下的取值应当为 1。随着焦距的增加, 投影模型从透视投影逐渐转变为正交投影, 如图 2 所示。在正交投影模型下, 场景点的投影位置与距离无关,  $\mathbf{F}$  零空间的维数  $N_F$  应为 4。 $N_F$  为 1~4 之间的整数。

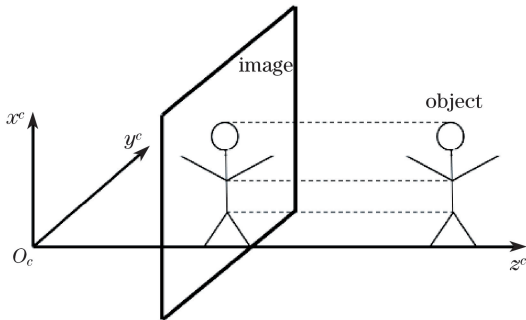


图 2 正交投影模型

Fig. 2 Orthogonal projection model

焦距  $f = 4, 35, 64, 100, 200$  mm 情况下, 进行 100 次  $\mathbf{F}^T \mathbf{F}$  的特征值计算, 然后取平均值, 得到的结果如图 3 所示。当相机焦距较小时,  $\mathbf{F}^T \mathbf{F}$  的零特

征值个数为 1, 随着相机焦距的增加, 零特征值的个数增加至 4。在 IEPnP 算法中, 将  $N_F$  的取值固定为 4, 即  $N_F = 4$ , 则(10)式可以写成:

$$\mathbf{c} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \mathbf{w}_3 \quad \mathbf{w}_4] \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \mathbf{W}\boldsymbol{\beta}. \quad (11)$$

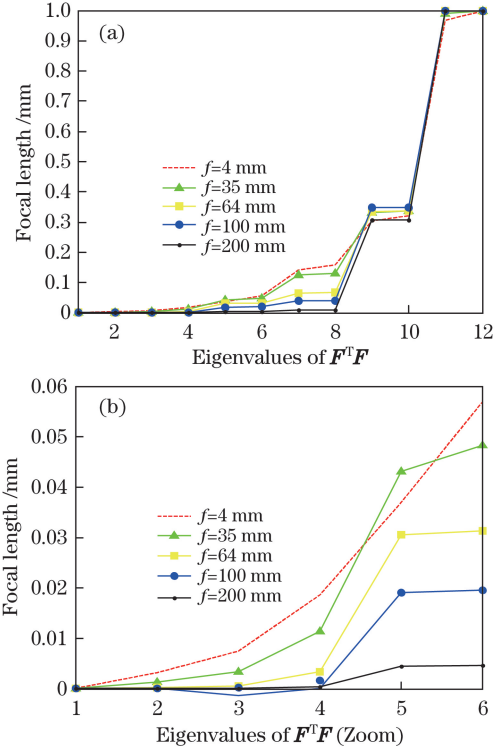


图 3  $\mathbf{F}^T \mathbf{F}$  在不同焦距下的特征值。(a)  $\mathbf{F}^T \mathbf{F}$  特征值的变化曲线;(b)  $\mathbf{F}^T \mathbf{F}$  前 6 个特征值的变化情况

Fig. 3  $\mathbf{F}^T \mathbf{F}$  eigenvalues under different focal lengths.

(a) Varying curve of eigenvalues of  $\mathbf{F}^T \mathbf{F}$ ;

(b) varying curve of the first 6 eigenvalues of  $\mathbf{F}^T \mathbf{F}$

由于  $\mathbf{R}$  和  $\mathbf{t}$  的初始值可以从弱透视投影模型中获取<sup>[14]</sup>, 记作  $\mathbf{R}_0$  和  $\mathbf{t}_0$ , 根据关系式  $\mathbf{C}_i^c = \mathbf{R}_0 \mathbf{C}_i^w + \mathbf{t}_0$  可获得向量  $\mathbf{c}$  的初值, 从而根据(11)式可得系数  $\boldsymbol{\beta}$  的计算公式为

$$\boldsymbol{\beta} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{c}. \quad (12)$$

刚性变换不改变点与点之间的距离, 因此列出如下的目标函数来对  $\boldsymbol{\beta}$  进行优化:

$$\boldsymbol{\beta} = \arg \min_{\boldsymbol{\beta}} \sum_{i,j=1, i < j}^4 \left| \|\mathbf{C}_i^c - \mathbf{C}_j^c\|^2 - \|\mathbf{C}_i^w - \mathbf{C}_j^w\|^2 \right|. \quad (13)$$

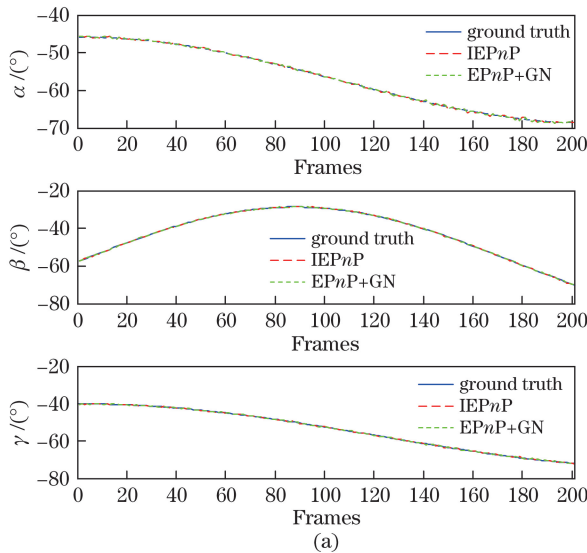
非线性最小二乘方法, 如高斯-牛顿法, 可以用来求解(13)式。获得  $\boldsymbol{\beta}$  的最优解后, 根据(10)式可得相机坐标系下虚拟控制点的坐标集合  $\{\mathbf{C}_i^c\}$ , 再由(4)式即可得到真实控制点在相机坐标系下的坐标

集合 $\{P_i^c\}$ 。通过解 $\{P_i^c\}$ 和 $\{P_i^w\}$ 之间的绝对定向问题<sup>[18]</sup>,最终求得相机的位姿 $R$ 和 $t$ 。

综上所述,将 IEP $n$ P 算法的具体步骤归结如下:1)在物体坐标系中选取 4 个虚拟控制点 $C_j^w, j=1,2,3,4$ ,根据(3)式计算组合系数 $a_{ij}$ ;2)由弱透视投影模型求得 $R_0$ 和 $t_0$ 的初始值,并计算 4 个虚拟控制点的相机坐标系坐标,从而得到向量 $c$ 的初始值;3)根据(9)式构造矩阵 $F$ ,求出 $F^T F$ 的零特征向量 $w_1, w_2, w_3, w_4$ ;4)根据(12)式解出 $\beta$ 的初始值,根据(13)式,利用高斯-牛顿法求最优解 $\beta^*$ ;5)根据(10)式求解 $C_j^{c*}, j=1,2,3,4$ ,并根据(4)式计算 $P_i^{c*}, i=1,2, \dots, n$ ;6)求解关于 $\{P_i^{c*}\}$ 和 $\{P_i^w\}$ 的绝对定向问题,最终求得相机位姿 $R$ 和 $t$ 。

#### 4 实验仿真

考虑到 IEP $n$ P 是 EP $n$ P 算法的迭代型改进,因此在仿真实验中主要将本文所提 IEP $n$ P 算法与文献[16]中提出的 EP $n$ P 算法进行比较,并按照文献[16]中的论述使用高斯-牛顿算法对 EP $n$ P 算法的结果进行优化,以增强 EP $n$ P 算法对图像噪声的稳健性,在后面的论述中将其标记为 EP $n$ P+GN。实验中,设虚拟相机的内参数如下:焦距 $f=8$  mm,图像大小为 1200 pixel $\times$ 1600 pixel,单位像元尺寸为 12  $\mu$ m $\times$ 12  $\mu$ m,主点位于图像中心。5 个虚拟控制点在世界坐标系下的坐标分别为:(0.10,0.00,0.00),(0.00,0.15,0.00),(-0.18,0.00,0.00),(0.00,-0.08,0.00),(0.00,0.00,0.16),单位为 m。仿真过程中,相机姿态和位移的变化轨迹遵循下列方程:



$$\begin{cases} \alpha = -1 + 0.2\cos(1.5t_s) \\ \beta = -1 + 0.5\sin(1.8t_s) \\ \gamma = -1 + 0.3\cos(1.3t_s) \end{cases} \quad (14)$$

$$\begin{cases} t_x = 1 + 0.1t_s \\ t_y = 1 + 0.15t_s \\ t_z = 1 + 0.2t_s \end{cases} \quad (15)$$

式中 $t_s$ 为采样间隔, $t=[t_x, t_y, t_z]^T$ ,旋转矩阵由 3 个欧拉角 $\alpha, \beta, \gamma$ 按照如下的公式构成:

$$R = \begin{bmatrix} c\beta \cdot c\gamma & c\beta \cdot s\gamma & -s\beta \\ s\alpha \cdot s\beta \cdot c\gamma - c\alpha \cdot s\gamma & c\alpha \cdot c\gamma + s\alpha \cdot s\beta \cdot s\gamma & c\beta \cdot s\alpha \\ s\alpha \cdot s\gamma + c\alpha \cdot s\beta \cdot c\gamma & c\alpha \cdot s\beta \cdot s\gamma - s\alpha \cdot c\gamma & c\alpha \cdot c\beta \end{bmatrix} \quad (16)$$

式中 $c\theta$ 为 $\cos\theta, s\theta$ 为 $\sin\theta, \theta$ 为 $\alpha, \beta, \gamma$ 中任意一个。

在仿真过程中,为控制点的图像像素坐标附加零均值、方差可变的高斯噪声,以模拟图像中无法准确提取控制点像素坐标的情况。首先令噪声方差为 0,两种算法对相机姿态变化的估计曲线如图 4 所示。由仿真结果可知,IEP $n$ P 算法和 EP $n$ P+GN 算法的位姿估计结果十分相似,这也反映在图 5 的误差比较曲线当中,图中的姿态误差 $E_{rot}$ 和位移误差 $E_{trans}$ 按照下述公式进行计算:

$$E_{rot} = \|\mathbf{q}_{true} - \hat{\mathbf{q}}\| / \|\mathbf{q}_{true}\|, \quad (17)$$

$$E_{trans} = \|\mathbf{t}_{true} - \hat{\mathbf{t}}\| / \|\mathbf{t}_{true}\|, \quad (18)$$

式中 $\mathbf{q}_{true}$ 和 $\hat{\mathbf{q}}$ 分别为相机的真实姿态与估计值的 4 元数表示, $\mathbf{t}_{true}$ 和 $\hat{\mathbf{t}}$ 分别为相机位移的真实值与估计值。

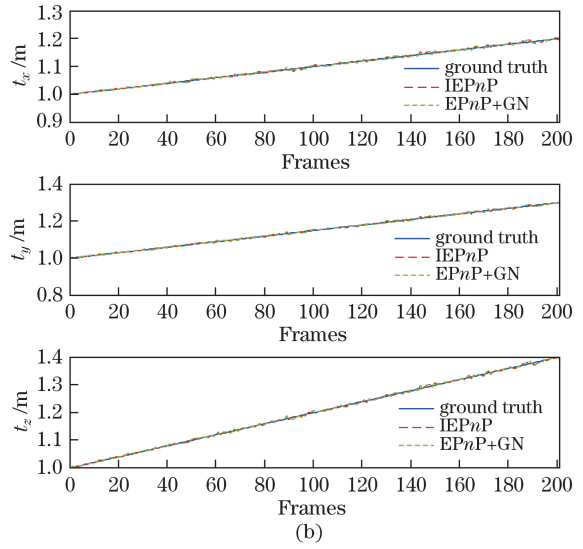


图 4 图像噪声方差为 0 时的曲线。(a)姿态估计曲线;(b)位移估计曲线

Fig. 4 Curves of the image noise when the variance is 0. (a) Attitude estimation curve; (b) translation estimation curve

为使误差比较的结果更加明晰,对图 5 中的各个数据点进一步统计,表 1 中  $\text{mean}(\cdot)$ ,  $\text{std}(\cdot)$ ,  $\text{median}(\cdot)$  分别表示对括号中的统计量求均值、标准差及中值。可以看出,IEP<sub>n</sub>P 算法与 EP<sub>n</sub>P+GN 算法在这几项数值上几乎没有差别,在图像噪声很小时,IEP<sub>n</sub>P 算法能够获得与 EP<sub>n</sub>P+GN 算法同样精确的位姿估计结果,从而证明了 IEP<sub>n</sub>P 算法的有效性。

增大图像的噪声水平,使噪声方差为 4,使用 IEP<sub>n</sub>P 算法与 EP<sub>n</sub>P+GN 算法获得的相机位姿变化曲线如图 6 所示。由图 6 可知,当图像噪声水平增大以后,EP<sub>n</sub>P+GN 算法的位姿估计结果变得不再稳定,而本文所提 IEP<sub>n</sub>P 算法仍然能够准确估计相机位姿的变化。这一结果可以进一步体现在图 7 所示的误差比较曲线中。

为了更加明晰地展示误差比较结果,对图 7 中的数据点进行统计,分别计算误差均值、标准差及中值,列在表 2 中。

表 2 中,误差均值和标准差反映了图 7 中误差比较的结果。值得注意的是误差中值,因为对误差中值的统计可以忽略不稳定的位姿估计结果的影响,反映单次位姿估计中算法能够达到的精度。

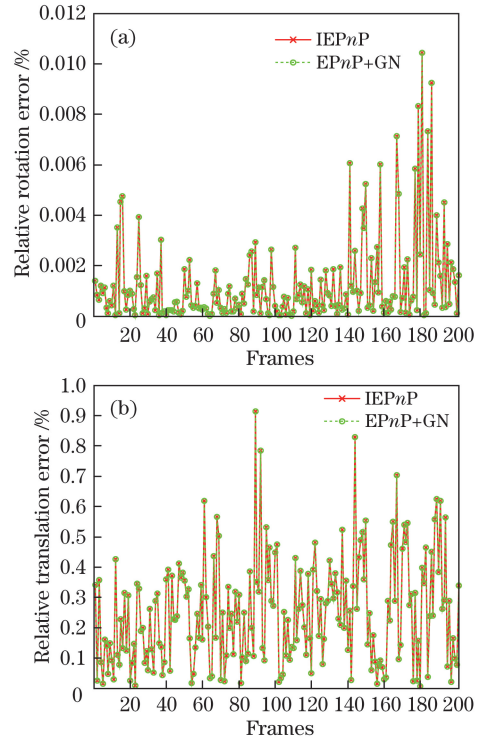


图 5 图像噪声方差为 0 时的曲线。(a)姿态估计误差；(b)位移估计误差

Fig. 5 Curves of the image noise when the variance is 0. (a) Attitude estimation error; (b) translation estimation error

表 1 图像噪声方差为 0 时的位姿误差比较

Table 1 Pose estimation error comparison when the image noise variance is 0

Algorithm	Mean( $E_{rot}$ ) / $10^{-5}$	Std( $E_{rot}$ ) / $10^{-5}$	Median( $E_{rot}$ ) / $10^{-6}$	Mean( $E_{trans}$ )	Std( $E_{trans}$ )	Median( $E_{trans}$ )
EP <sub>n</sub> P+GN	1.2121	1.6904	6.3569	0.0024931	0.0017529	0.0023812
IEP <sub>n</sub> P	1.2121	1.6904	6.3569	0.0024931	0.0017529	0.0023812

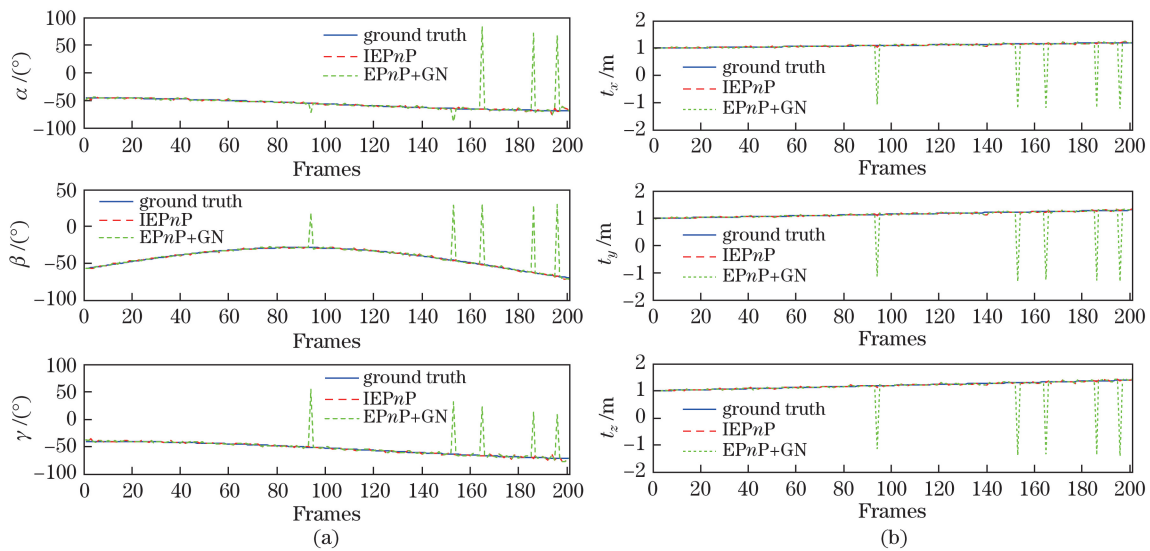


图 6 图像噪声方差为 4 时的曲线。(a)姿态估计曲线；(b)位移估计曲线

Fig. 6 Curves of the image noise when the variance is 4. (a) Attitude estimation curve; (b) translation estimation curve

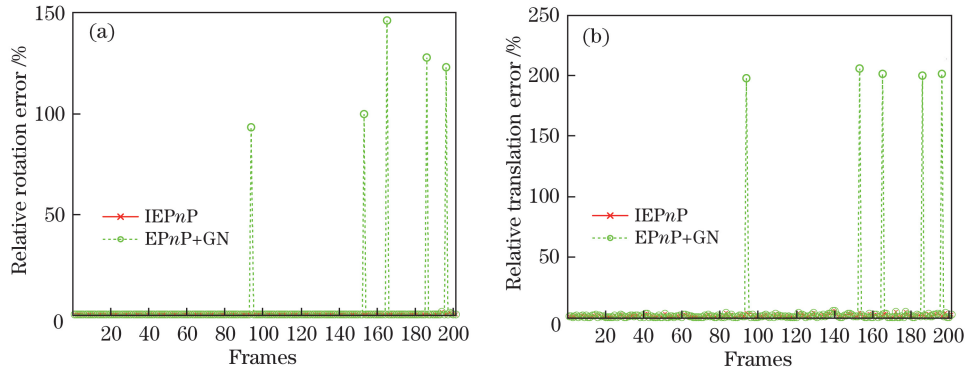


图 7 图像噪声方差为 4 时的曲线。(a)姿态估计误差;(b)位移估计误差

Fig. 7 Curves of the image noise when the variance is 4. (a) Attitude estimation error; (b) translation estimation error

表 2 图像噪声方差为 4 时的位姿误差比较

Table 2 Pose estimation error comparison when the image noise variance is 4

Algorithm	Mean( $E_{rot}$ )	Std( $E_{rot}$ )	Median( $E_{rot}$ )	Mean( $E_{trans}$ )	Std( $E_{trans}$ )	Median( $E_{trans}$ )
EPnP+GN	0.030048	0.1867100	0.00027094	0.063622	0.312440	0.012155
IEPnP	0.0006934	0.0016083	0.00025074	0.014224	0.011448	0.011982

由表可知,即使不考虑不稳定的位姿估计结果,EPnP+GN 算法在一次位姿估计中的精度仍然略低于本文所提 IEPnP 算法。类似的实验结果还可以在更大的图像噪声水平条件下获得,这说明当图像噪声水平较大时,IEPnP 算法比 EPnP+GN 算法具有更强的稳健性。

两次实验中两种算法的运行时间比较如图 8 所示。实验中对 EPnP 算法的运行结果使用高斯-牛顿法进行了优化,因此解析算法的速度优势并未体现,本文所提 IEPnP 算法反而略快于 EPnP+GN 算法。综合考虑算法的实时性与稳健性,本文所提 IEPnP 算法明显优于 EPnP+GN 算法。

## 5 结 论

基于 EPnP 算法提出了一种相机位姿估计的迭代算法,即 IEPnP 算法。算法的迭代过程借鉴了 EPnP 算法虚拟控制点的思想,选择了 4 个虚拟控制点来表示真实的控制点,通过弱透视投影模型获取相机的初始位姿,从而构建出 4 个虚拟控制点在相机坐标系下的坐标,再根据刚性变换不改变坐标点间距离的原则,建立了迭代求解的目标函数,并通过非线性最小二乘算法来更新虚拟控制点在相机坐标系下的坐标,最终通过解决绝对定向问题完成了对相机位姿的估计。

不同图像噪声水平的仿真比较结果表明,IEPnP 算法在低噪声水平时能够获得与 EPnP 算法相似的位姿估计结果,而当图像噪声水平较高时,IEPnP

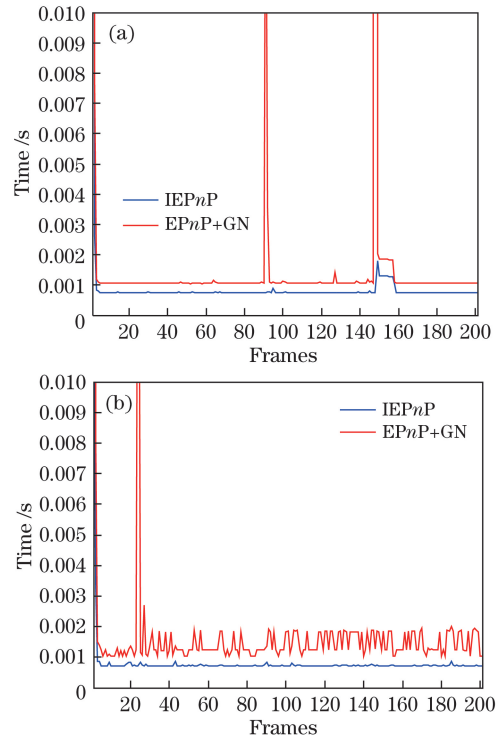


图 8 算法的运行时间比较。(a)图像噪声方差为 0;  
(b)图像噪声方差为 4

Fig. 8 Comparison of the running time of the algorithm. (a) Image noise variance is 0; (b) image noise variance is 4

算法相较于 EPnP 算法则表现出了更强的稳健性,且 IEPnP 算法的计算效率并没有因此降低。综合考虑算法的计算效率和对噪声的稳健性,IEPnP 算法优于 EPnP 算法。由于 IEPnP 算法是一种迭代算法,算法的运行结果主要依赖于初始值,当初始值

偏离真实位姿较远时,IEP<sub>n</sub>P 算法的运行结果也会偏离真实位姿。因此在今后的研究中,需要对如何减少算法对初始位姿的依赖,以及如何确保算法向真实位姿收敛等问题进行更加深入的探索。

## 参 考 文 献

- [1] Hong Y, Sun X X, Wang D, *et al.* Fast pose estimation method for unmanned aerial vehicle based on rectangular geometry feature[J]. Chinese Journal of Lasers, 2016, 43(5): 0508006.  
洪洋, 孙秀霞, 王栋, 等. 基于矩形几何特性的小型无人机快速位姿估计方法[J]. 中国激光, 2016, 43(5): 0508006.
- [2] Zhang C, Han C, Yang H M, *et al.* Large field and binocular vision calibration algorithm based on position and orientation constraints[J]. Acta Optica Sinica, 2016, 36(1): 0115003.  
张超, 韩成, 杨华民, 等. 基于位姿约束的大视场双目视觉标定算法[J]. 光学学报, 2016, 36(1): 0115003.
- [3] Fischer J, Eichler M, Bartz D, *et al.* Model-based hybrid tracking for medical augmented reality[C]// Proceedings of the 12<sup>th</sup> Eurographics Conference on Virtual Environments, 2006: 71-80.
- [4] Wu F C, Hu Z Y. A study on the P5P problem[J]. Journal of Software, 2001, 12(5): 768-775.  
吴福朝, 胡占义. 关于 P5P 问题的研究[J]. 软件学报, 2001, 12(5): 768-775.
- [5] Wu F C, Hu Z Y. A linear method for the P<sub>n</sub>P problem[J]. Journal of Software, 2003, 14(3): 682-688.  
吴福朝, 胡占义. P<sub>n</sub>P 问题的线性求解算法[J]. 软件学报, 2003, 14(3): 682-688.
- [6] Quan L, Lan Z D. Linear N-point camera pose determination[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(8): 774-780.
- [7] Ansar A, Daniilidis K. Linear pose estimation from points or lines[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 578-589.
- [8] Zhu R Z, Lin Y, Zhang L. New algorithm of solving for ranges during final approach of spacecraft rendezvous[J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(7): 764-768.  
朱仁璋, 林彦, 张磊. 航天器交会计算机视觉系统测距求解新算法[J]. 北京航空航天大学学报, 2006, 32(7): 764-768.
- [9] Li X, Zhang Y Q, Liu J B, *et al.* A direct least squares method for camera pose estimation based on straight line segment correspondences[J]. Acta Optica Sinica, 2015, 35(6): 0615003.  
李鑫, 张跃强, 刘进博, 等. 基于直线段对应的相机位姿估计直接最小二乘法[J]. 光学学报, 2015, 35(6): 0615003.
- [10] Liu M L, Wong K H. Pose estimation using four corresponding points[J]. Pattern Recognition Letters, 1999, 20(1): 69-74.
- [11] Dai D K, Wang X S, Hu C S, *et al.* Camera calibration and attitude measurement technology based on astronomical observation[J]. Acta Optica Sinica, 2012, 32(3): 0312005.  
戴东凯, 王省书, 胡春生, 等. 基于天文观测的相机标定及姿态测量技术研究[J]. 光学学报, 2012, 32(3): 0312005.
- [12] Dementhon D F, Davis L S. Model-based object pose in 25 lines of code[J]. International Journal of Computer Vision, 1995, 15(1/2): 123-141.
- [13] Bradski G, Kaehler A. Learning OpenCV[M]. Yu S Q, Liu R Z, Transl. Beijing: Tsinghua University Press, 2009.  
布拉德斯基, 克勒. 学习 OpenCV[M]. 于仕琪, 刘瑞祯, 译. 北京: 清华大学出版社, 2009.
- [14] Lu C P, Hager G D, Mjolsness E. Fast and globally convergent pose estimation from video images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(6): 610-622.
- [15] Li X, Long G C, Liu J B, *et al.* Accelerative orthogonal iteration algorithm for camera pose estimation[J]. Acta Optica Sinica, 2015, 35(1): 0115004.  
李鑫, 龙古灿, 刘进博, 等. 相机位姿估计的加速正交迭代算法[J]. 光学学报, 2015, 35(1): 0115004.
- [16] Lepetit V, Moreno-Noguer F, Fua P. EP<sub>n</sub>P: an accurate  $O(n)$  solution to the P<sub>n</sub>P problem[J]. International Journal of Computer Vision, 2009, 81: 155.
- [17] Hmam H, Kim J. Optimal non-iterative pose estimation via convex relaxation[J]. Image and Vision Computing, 2010, 28(11): 1515-1523.
- [18] Umeyama S. Least-squares estimation of transformation parameters between two point patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(4): 376-380.