

# 一种高精度的非迭代位姿估计方法

薛俊诗, 郭宁博, 刘胜恩

中国人民解放军航天工程大学研究生院, 北京 101416

**摘要** 针对增量式运动结构恢复中相机内参数未标定情况下位姿估计问题, 提出一种高精度的非迭代位姿求解方法。根据针孔成像模型, 构建了相机内参数未标定位姿估计问题描述方程, 使用单位四元数对旋转矩阵进行参数化, 并根据四重轴对称现象对旋转矩阵进行分解, 通过最小二乘原理构建目标函数后, 利用一阶最优化条件建立了关于旋转矩阵参数的二元多项式方程组, 使用 Gröbner 基求解该方程组得到旋转参数的闭合解, 进而获得相机的位姿参数和焦距。该算法的计算复杂度为  $O(n)$ , 实验验证了其有效性、高精度和高数值稳定性。

**关键词** 机器视觉; 相机位姿估计; 非迭代方法; 旋转矩阵参数化; 多项式方程组

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201838.0715001

## New Non-Iterative Method for Pose Estimation with High Precision

Xue Junshi, Guo Ningbo, Liu Sheng'en

Department of Postgraduate, Aerospace Engineering University of PLA, Beijing, 101416

**Abstract** The non-iterative method with high precision is proposed for the pose estimation of uncalibrated in-camera parameters in incremental motion structure recovery. Based on the pinhole camera model, the equations which describes uncalibrated cameras pose estimation problems are given. The rotation matrix is parameterized by unit-norm quaternion, and then decomposed according to the four-fold symmetry. To eliminate variables properly and build a bivariate objective function, the least square principle and first order optimal condition is adopted respectively. An efficient solver by using the Gröbner basis technique is developed, so as to obtain optimal solution of rotation matrix and focal length. The complexity of proposed method is  $O(n)$ . Experimental results prove the efficiency, high precision and numerical stability of the proposed solution.

**Key words** machine vision; camera pose estimation; non-iterative method; rotation matrix parameterization; polynomial equations

**OCIS codes** 150.1488; 150.6910; 120.4640

## 1 引 言

利用空间中的  $n$  个三维 (3D) 点及其在影像中的二维 (2D) 投影点, 同时估计相机位置、姿态和焦距是计算机视觉中的重要研究内容, 也是透视  $N$  点 ( $P_nP$ ) 问题在未知焦距情况下的延伸, 即  $P_nPf$  问题。其求解广泛应用于即时定位与地图构建 (SLAM)、机器人视觉导航、增量式运动结构恢复 (SFM) 和增强现实 (AR) 等方面<sup>[1-2]</sup>, 具有非常高的应用价值。

目前, 常用的  $P_nP$  解算方法主要分为迭代和非迭代两大类。迭代法主要通过构建目标函数, 然后

使用高斯-牛顿、Levenberg-Marquardt 等算法进行迭代优化。Horaud 等<sup>[3]</sup> 提出的 POSIT 算法通过弱透视变换求解初值, 然后利用迭代法逼近透视变换, 最终得到位姿解; Lu 等<sup>[4]</sup> 提出利用正交迭代算法求解目标空间共线方程得到相机位姿解, 该算法能够保证旋转矩阵的正交性, 且具有较高的精度, 但依赖弱透视投影变换求解初值, 导致 3D 空间点分布在平面或近似平面情况下收敛不稳定; Schweighofer 等<sup>[5]</sup> 将  $P_nP$  问题转化为半定规划 (SDP) 问题, 能够有效避免在迭代过程陷入局部最值, 且计算复杂度为  $O(n)$ , 但进行全局优化时需要大量的计算; 文献<sup>[6]</sup> 提出基于共面直线迭代加权最小二乘的相机位姿估计。总

收稿日期: 2017-11-13; 收到修改稿日期: 2018-02-05

基金项目: 军队探索项目 (7131145)

作者简介: 薛俊诗 (1988—), 男, 博士研究生, 主要从事机器视觉、三维实景建模等方面的研究。

E-mail: xueao2015@sina.com

体而言,迭代方法解算精度较高,但执行效率低,且易收敛于局部最优解。

最早得到广泛应用的非迭代方法为 EPnP 算法<sup>[7]</sup>,该方法将所有的 3D 点表示为 4 个控制点的加权和,并在相机坐标系构建代价函数,由绝对定向方法求解相机的位姿解。文献[8]使用迭代算法对 EPnP 进行优化,得到了更精确的求解结果;文献[9]为解决 EPnP 算法受控制点分布位置限制,将 PnP 问题转化为多个 P3P 问题后,稳定处理了控制点处于平面、非平面及近似平面的情况,并得到较高的求解精度。为进一步提高相机位姿解算精度,近几年,通常采用将旋转矩阵参数化,使 PnP 问题转化为关于旋转参数的最优化问题。直接最小二乘(DLS)算法<sup>[10]</sup>采用 CGR(Cayley-Gibbs-Rodriguez)方法对旋转矩阵进行参数化,并采用驻式法对所有 3D 点进行检索,保证旋转参数求解的稳健性;文献[11]提出的基于直线段对应的 DLS 算法,利用代数多项式直接求解目标函数;OPnP 算法<sup>[12]</sup>采用非单位四元数对旋转矩阵进行参数化,并利用 Gröbner 基对旋转参数多项式进行求解,解决了 DLS 算法旋转矩阵求解中存在的退化问题;ASPnP 算法<sup>[13]</sup>通过构建代数合适的代数误差多项式函数,增强了算法的可扩展性;UPnP 算法<sup>[14]</sup>增加了对非中心投影下相机位姿解算的适用性。

然而,上述方法均是在相机标定矩阵已知情况下进行的。当焦距未知时,共有 7 个未知量需要求解,一般而言,最少需要 4 个 2D-3D 对应点进行解算。因此,文献[15]利用 Gröbner 基求解 3D 点空间距离构建的多项式的方法,其位姿估计精度受对应点配置影响较大,且求解效率较低;文献[16]利用分枝定界算法求解 PnP 问题,求解效率依然没有较大提高;文献[17]利用 EPnP 算法中的虚拟控制点的思想和文献[15]利用 RPnP 算法中点三元组思想,均无法解决非平衡代价函数造成的噪声非稳健性问题;文献[18]利用未知焦距对旋转矩阵参数化过程中存在的四重轴对称现象,对旋转矩阵进行分解简化,最后利用 Gröbner 基对构建的多项式方程组进行求解,该方法计算精度较高,但计算复杂度为  $O(n^3)$ ,计算效率无法满足实时性要求。

本文针对相机内参数未标定情况下的位姿估计问题,根据旋转矩阵分解思想,构建不同的目标函数,提出一种计算复杂度为  $O(n)$  的非迭代位姿估计方法,旨在保证计算精度的前提下,提高计算效率,将其用于捆集调整(Bundler)中可以实现未标定

影像的高精度三维重建。该算法首先利用单位四元数对旋转矩阵进行参数化,根据四重轴对称现象进行旋转矩阵分解,之后使用最小二乘残差原理构建目标函数,并利用一阶最优化条件构建多项式方程组,使用 Gröbner 基直接求解得到旋转参数的闭合解,最终获得相机的位姿参数和焦距。该算法具有较高的实时性,且位姿估计精度较高。

## 2 PnP 问题描述

如图 1 所示,标准的相机透视成像模型可以表示为

$$\alpha_i m_i = K [R | t] P_i, \quad (1)$$

式中  $\alpha_i$  为投影深度因子,  $m_i = [u_i, v_i, 1]^T$  为像点坐标的齐次表示,  $p_i = [X_i, Y_i, Z_i]^T$  为空间点坐标,  $P_i = [X_i, Y_i, Z_i, 1]^T$  为空间点坐标的其次形式,  $K$  为相机内参数矩阵,假设唯一的未知参数为焦距  $f$ ,则  $K$  可以写成  $\text{diag}[f, f, 1]$  的形式,  $R = [r_{ij}]_{i,j=1,\dots,n}^3$  为  $3 \times 3$  的旋转矩阵,  $t = [t_x, t_y, t_z]^T$  为平移矩阵。设  $\bar{m}_n$  为  $m_i$  的归一化单位向量,则(1)式可写为

$$\alpha_i \bar{m}_i - Kt = KRp_i. \quad (2)$$

设  $I$  为  $3 \times 3$  的单位矩阵,则 PnP 问题可写为

$$\begin{bmatrix} \bar{m}_1 & & & -I \\ & \ddots & & \vdots \\ 0 & & \bar{m}_n & -I \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ Kt \end{bmatrix} = \begin{bmatrix} KR & & & \\ & \ddots & & \\ & & KR & \\ & & & KR \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}. \quad (3)$$

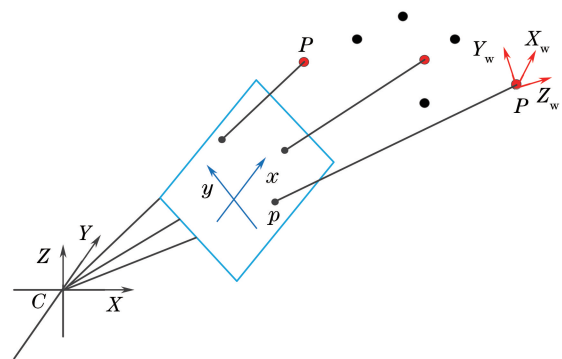


图 1 位姿估计示意图

Fig. 1 Camera pose estimation problem

## 3 算法原理

(3)式等价于  $AX = WQ$  的形式。由文献[6]可得:

$$X = (A^T A)^{-1} A^T WQ = [U \ V]^{-1} WQ, \quad (4)$$

式中矩阵  $A$  的广义逆被分解为  $U, V$  两个矩阵,其中

$U = [u_1^T \cdots u_n^T]^T$  为变换尺度的函数,  $V$  为平移矩阵的函数。应用矩阵的 Schur 补与块矩阵反演理论可得:

$$V_{3 \times 3n} = [V_1, \cdots, V_n], \quad (5)$$

$$V_i = H [\bar{m}_i \bar{m}_i^T - I] \in R^{3 \times 3}, i = 1, \cdots, n, \quad (6)$$

$$H_{3 \times 3} = (nI - \sum_{i=1}^n \bar{m}_i \bar{m}_i^T)^{-1}, \quad (7)$$

$$U_{n \times 3n} = \begin{bmatrix} \bar{m}_1^T & & \\ & \ddots & \\ & & \bar{m}_n^T \end{bmatrix} + \begin{bmatrix} \bar{m}_1^T \\ \vdots \\ \bar{m}_n^T \end{bmatrix} V = \begin{bmatrix} u_1^T \\ \vdots \\ u_n^T \end{bmatrix}, \quad (8)$$

$$u_i^T = [u_{i1}^T, \cdots, u_{in}^T]_{1 \times 3n}, i = 1, \cdots, n, \quad (9)$$

$$u_{ij}^T = \bar{m}_i^T \delta(i, j) + \bar{m}_i^T V_j \in R^{1 \times 3}, j = 1, \cdots, n, \quad (10)$$

式中  $u_i^T$  表示  $U$  的第  $i$  行,  $u_{ij}^T$  表示矩阵在第  $i$  行  $3j$  列的  $1 \times 3$  元素。则有:

$$\begin{cases} \alpha_i = u_i^T W Q \\ t = V W Q \end{cases}. \quad (11)$$

将(11)式代入(2)式可得关于旋转矩阵及相机内参数矩阵的限制表达式,即:

$$u_i^T W Q \bar{m}_i - K R p_i - V W Q = 0. \quad (12)$$

将(5)~(9)式代入(12)式,可得:

$$\bar{m}_i^T K R p_i - K R p_i + \bar{m}_i^T \left[ \sum_{j=1}^n V_j K R p_j \right] \bar{m}_i - \sum_{j=1}^n V_j K R p_j = 0. \quad (13)$$

利用单位四元数方法对旋转矩阵  $R$  进行参数化,用  $s = [s_1, s_2, s_3, s_4]^T$  表示,则有:

$$R(s) = \begin{bmatrix} s_1^2 + s_2^2 - s_3^2 - s_4^2 & 2s_2 s_3 - 2s_1 s_4 & 2s_2 s_4 + 2s_1 s_3 \\ 2s_2 s_3 + 2s_1 s_4 & s_1^2 - s_2^2 + s_3^2 - s_4^2 & 2s_3 s_4 - 2s_1 s_2 \\ 2s_2 s_4 - 2s_1 s_3 & 2s_3 s_4 + 2s_1 s_2 & s_1^2 - s_2^2 - s_3^2 + s_4^2 \end{bmatrix}. \quad (14)$$

由于  $s$  与  $-s$  表示相同的旋转情况,四元数表示本身存在二重轴对称。当焦距未知时,有  $\text{diag}\{f, f, 1\} R = \text{diag}\{-f, -f, 1\} [\text{diag}\{-1, -1, 1\} R]$ ,即存在另一种二重轴对称。根据 Hamilton 法则,  $\text{diag}\{-1, -1, 1\} R$  可用四元数  $[-s_4, -s_3, s_2, s_1]^T$  表示。

当  $s_1^2 + s_4^2 \neq 0$  时,  $R(s)$  可分解为  $R(s_\theta)$  与  $R(s_\omega)$  两个旋转分量的乘积形式。设  $s_\theta = [s_1, 0, 0, s_4]^T, \bar{x}^2 + \bar{y}^2 = 1$ , 则有

$$R(s_\theta) = \frac{1}{s_1^2 + s_4^2} \begin{bmatrix} s_1^2 - s_4^2 & -2s_1 s_4 & 0 \\ 2s_1 s_4 & s_1^2 - s_4^2 & 0 \\ 0 & 0 & s_1^2 + s_4^2 \end{bmatrix} = \begin{bmatrix} \bar{x} & -\bar{y} & 0 \\ \bar{y} & \bar{x} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (15)$$

当  $s_1 = s_4 = 0$  或者  $s_1, s_4$  的量级远小于  $s_2, s_3$  时,有  $R p_i = [R \cdot \text{diag}\{1, -1, -1\}] [\text{diag}\{1, -1, -1\} p_i] = \bar{R} \bar{p}_i$ 。其中,  $\text{diag}\{1, -1, -1\}$  表示绕  $x$  轴旋转  $180^\circ$ 。根据 Hamilton 法则,  $\bar{R}$  可由四元数  $[s_2, -s_1, -s_4, s_3]^T$  表示,此时  $\{s_1, s_4\}$  与  $\{s_2, s_3\}$  位置进行了互换,在无法对  $R$  进行分解时,可对  $\bar{R}$  进行分解,即将 3D 点绕  $x$  轴旋转  $180^\circ$  后再进行求解。

根据文献[19]可知:

$$\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{x} & -\bar{y} & 0 \\ \bar{y} & \bar{x} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} x & -y & 0 \\ y & x & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (16)$$

式中  $x = f\bar{x}, y = f\bar{y}$ 。此时,焦距与两个独立的变量融合在一起,由焦距引起的二重轴对称模糊可以消除。设  $s_\omega = [1, b, c, 0]^T$ , 则:

$$R(s_\omega) = \frac{1}{1 + b^2 + c^2} \begin{bmatrix} 1 + b^2 - c^2 & 2bc & 2c \\ 2bc & 1 - b^2 + c^2 & -2b \\ -2c & 2b & 1 - b^2 - c^2 \end{bmatrix}. \quad (17)$$

将两个旋转分量  $R(s_\theta)$  与  $R(s_\omega)$  代入,经过整理可得:

$$F(b, c) [x \ y \ 1]^T = 0. \quad (18)$$

$F(b, c)$  为  $n \times 3$  的矩阵, (18) 式展开可得:

$$G(b, c) [x \ y]^T = \eta(b, c), \quad (19)$$

式中  $G(b, c)$  为  $n \times 2$  的矩阵,  $\eta(b, c)$  为  $n$  维列向

量。则有  $(\mathbf{G}\mathbf{G}^+ - \mathbf{I})\boldsymbol{\eta} = 0$ , 其中  $\mathbf{G}^+ = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T$ , 两边分别乘以  $\det(\mathbf{G}^T\mathbf{G})$ , 可得:

$$\sqrt{\det(\mathbf{G}^T\mathbf{G})}(\mathbf{G}\mathbf{G}^+ - \mathbf{I})\boldsymbol{\eta} = 0. \quad (20)$$

由于实际解算过程中无法避免噪声的影响, 根据最小二乘残差原理可以构建目标函数

$$\varphi(b, c) = \boldsymbol{\eta}^T [\mathbf{G}(\mathbf{G}^T\mathbf{G})^* \mathbf{G}^T - \det(\mathbf{G}^T\mathbf{G})\mathbf{I}] \boldsymbol{\eta}, \quad (21)$$

式中  $(\mathbf{G}^T\mathbf{G})^*$  为  $\mathbf{G}^T\mathbf{G}$  的伴随矩阵。

## 4 基于 Gröbner 基的求解

为得到(21)式的全局最优值, 将目标函数分别对  $b, c$  求偏导, 并使之等于零, 得到方程组:

$$\begin{cases} \frac{\partial \varphi(b, c)}{\partial b} = 0 \\ \frac{\partial \varphi(b, c)}{\partial c} = 0 \end{cases}. \quad (22)$$

此时, 可得到自由度为 7, 单项式个数为 36 的多元多项式方程组。本文选用自动 Gröbner 基求解该方程组。由于 Gröbner 基的构建需要非常专业的数学知识, 直到近几年, Gröbner 基才在计算机视觉中的多项式方程组求解中得到大量应用。Kukelova 等<sup>[20]</sup>提出了一种自动 Gröbner 基生成算法, 为促进其广泛应用奠定了基础。该算法基本的求解步骤为: 1) 在分次逆序字典下生成 Gröbner 基; 2) 构建消除模板; 3) 建立作用矩阵; 4) 对作用矩阵进行 SVD 分解; 5) 获得原始多项式方程组的解。

在求解过程中, 选择完全随机的系数矩阵生成方式, 共产生 31 个方程, 消除模板矩阵的大小为  $81 \times 139$ , 作用矩阵的大小为  $31 \times 31$ 。得到  $b, c$  的解之后, 根据(18)式可以求得焦距与剩余的旋转参数, 最后根据(11)式求得平移矩阵。

## 5 算法流程

根据上述分析, 本文算法流程如下。

- 1) 根据成像点的 2D 坐标, 构建归一化成像点单位向量。
- 2) 根据针孔成像模型, 构建 PnPf 问题描述方程, 并消除平移矩阵。
- 3) 使用单位四元数法对旋转矩阵进行参数化, 并分解旋转矩阵。
- 4) 利用最小二乘法构建目标函数, 根据一阶最优优化条件构建旋转参数的多项式方程组。
- 5) 使用自动 Gröbner 基求解(4)式的多项式方

程组, 得到旋转参数、旋转矩阵、焦距和平移矩阵。

6) 若求解结果出现退化, 则将 3D 点绕  $x$  轴旋转  $180^\circ$ , 再重新进行上述步骤求解。

## 6 仿真实验及分析

为评估算法的实时性和位姿估计精度, 分别通过合成数据和实际数据进行实验验证, 并与 GPnPf 算法<sup>[15]</sup>进行对比。

### 6.1 合成数据实验

假设某合成图像的尺寸为  $640 \text{ pixel} \times 480 \text{ pixel}$ , 等效焦距为  $800 \text{ pixel}$ , 像主点位于影像中心。 $n$  个 3D 参考点的  $x, y, z$  坐标在  $[-2, 2] \times [-2, 2] \times [4, 8]$  的范围内随机生成, 设定随机生成的参考偏移向量为  $\mathbf{t}_{\text{true}}$ , 参考旋转矩阵为  $\mathbf{R}_{\text{true}}$ 。根据针孔成像模型, 生成与 3D 点相对应的 2D 影像坐标。采用文献[7]中的方法定义位姿估计误差, 以度( $^\circ$ )为单位, 定义  $\mathbf{R}_{\text{true}}$  与其估计值  $\tilde{\mathbf{R}}$  的绝对误差为  $e_{\text{rot}} = \max_{k=1}^3 \{a \cos[\text{dot}(\mathbf{r}_{\text{true}}^k, \tilde{\mathbf{r}}^k)]\} \times 180/\pi$ , 其中,  $\mathbf{r}_{\text{true}}^k, \tilde{\mathbf{r}}^k$  分别为  $\mathbf{R}_{\text{true}}, \tilde{\mathbf{R}}$  的第  $k$  列,  $\text{dot}(\cdot)$  表示两个向量的叉乘,  $\mathbf{t}_{\text{true}}$  与其估计值  $\hat{\mathbf{t}}$  的相对误差定义为  $e_{\text{trans}}(\%) = \|\mathbf{t}_{\text{true}} - \hat{\mathbf{t}}\| / \|\hat{\mathbf{t}}\| \times 100\%$ ,  $f_{\text{true}}$  与其估计值  $\hat{f}$  的相对误差定义为  $e_{\text{focal}}(\%) = \|f_{\text{true}} - \hat{f}\| / \|\hat{f}\| \times 100\%$ 。

#### 6.1.1 计算精度与稳定性

参与位姿估计的对应点数目从 4 到 20 递增, 并添加均值为 0、标准差为 2 的高斯噪声, 进行 200 次独立实验, 得到的估计结果如图 2 所示。

由图 2 可知, 随着估计点数的增加, 参数估计误差变小, 且估计值更加稳定。随着参与估计点数的增加, 旋转矩阵的估计误差均值从  $0.83^\circ$  下降到  $0.21^\circ$ ; 平移误差均值从  $1.92\%$  下降到  $0.32\%$ ; 焦距均值误差从  $2.50\%$  下降到  $0.45\%$ 。由于本文算法按照重投影误差最小化构建的代价函数, 故投影误差均值会随点数的增加而累积, 同时, 位姿估计值的稳定性也随着参与估计点数的增加而变高。一般而言, 重投影误差小于 3 pixel, 能够为捆集优化提供良好的初值, 满足三维重建的精度要求, 并减少迭代时间。与 GPnPf 算法相比, 本文算法在估计精度、数值稳定度方面有一定的提高。

选定估计点数为 10, 添加 0 均值, 标准差从 0.5 到 2.5 逐渐增加的高斯噪声, 进行 200 次独立实验, 得到的估计结果如图 3 所示。

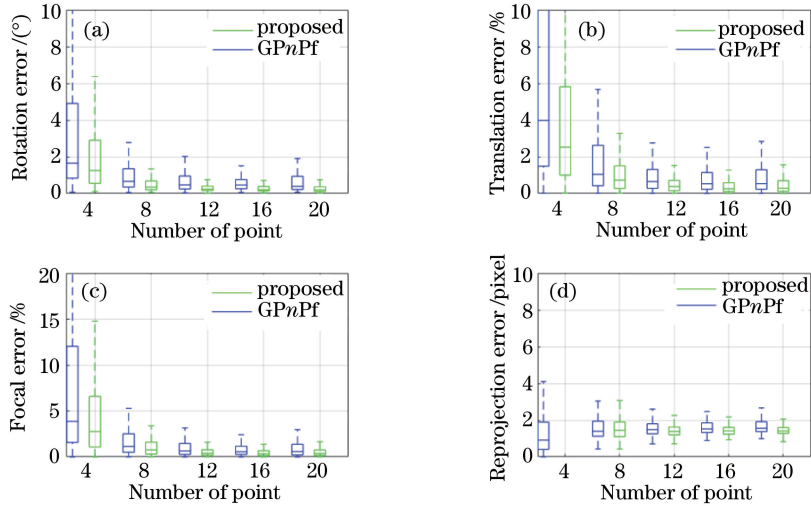


图 2 不同估计点数下的 200 次独立实验结果。(a)旋转矩阵误差;(b)平移矩阵误差;(c)焦距误差;(d)重投影误差

Fig. 2 200 times experimental results with respect to varying number of points. (a) Rotation matrix error; (b) translation matrix error; (c) focal error; (d) reprojection error

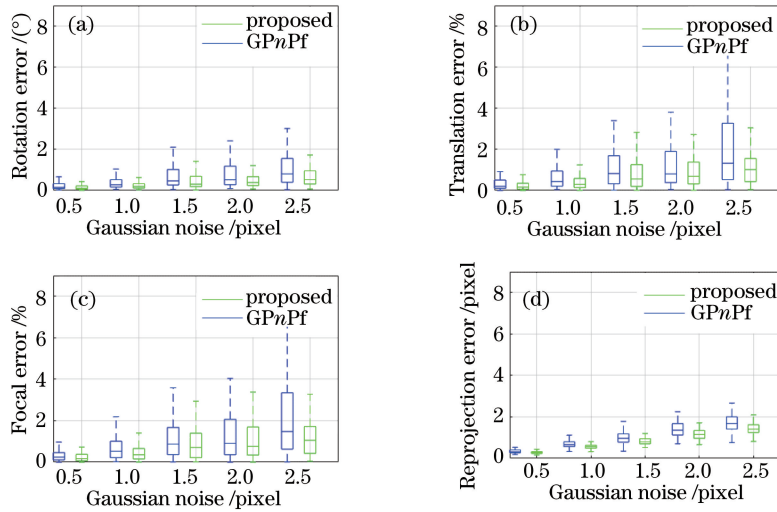


图 3 不同噪声下的 200 次独立实验结果。(a)旋转矩阵误差;(b)平移矩阵误差;(c)焦距误差;(d)重投影误差

Fig. 3 200 times experimental results with respect to varying noise level. (a) Rotation matrix error; (b) translation matrix error; (c) focal error; (d) reprojection error

随着噪声水平的不断增加,位姿估计误差逐渐变大,且稳定性有所下降。但整体而言,当噪声标准差低于 2.5 时,本文算法能够获得较高精度的位姿估计值。与 GPnPf 算法相比,本文算法具有更好的抗噪声性能。

### 6.1.2 计算时间

利用 Matlab 软件实现各算法,计算机配置为 2.8 GHz 的 Intel Core i7-4900MQ CPU。参与估计点数目变化为 10~1010,每隔 200 点计算一次执行时间,结果如图 4 所示。可见,随着参与计算点数的增加,计算时间基本呈线性增加。由于本文算法不需要进行迭代,所以计算效率较高,能够满足大部分

相机位姿估计应用的需求。

## 6.2 真实数据

分别采用无人机拍摄的某建筑影像序列和 CastleP30 影像序列进行相机位姿估计和三维重建实验。

无人机数据集共 50 张影像,影像尺寸为 4912 pixel×3264 pixel,焦距未知,该数据集的相机位置排列比较规则,且影像经过畸变校正。首先,利用 SIFTGPU 算法进行特征点提取,每张影像提取的特征点数目设定为 10000,提取特征点后使用文献[21]中的方法进行特征点匹配,得到的结果如图 5 所示,图中圆点表示特征点分布位置,直线为特征

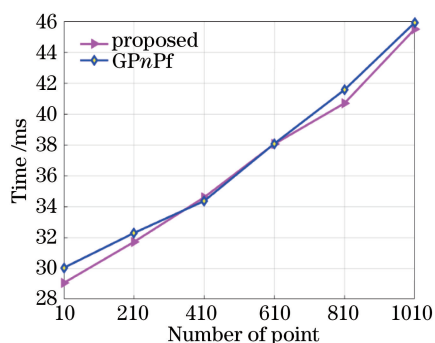


图 4 计算时间随估计点数的变化

Fig. 4 Computation time with respect to varying number of point

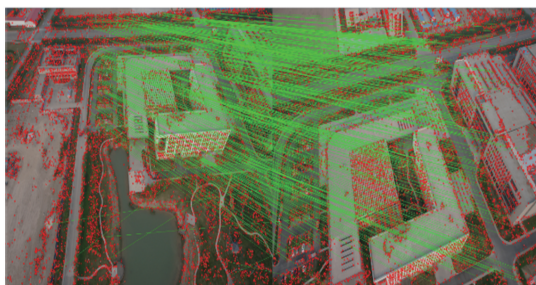


图 5 SIFT 特征点及匹配关系

Fig. 5 SIFT points and matching lines  
点匹配对应关系。

完成影像的特征匹配后,选择部分特征点与 3D 控制点建立 2D-3D 对应关系,使用本文算法进行相机位姿估计。然后,利用相机位姿估计结果重建三维场景,恢复场景的 3D 稀疏结构。

为了更清晰地表现重投影误差分布结果,部分重投影点与特征点在影像上的分布如图 6 所示,图中+表示特征点分布位置,×表示重投影点分布位置。可见,重投影点与参与重建的特征点完全重合,这意味着位置误差小于 1 pixel,进一步说明本文提出的位姿估计方法具有较高的精度。

在不提供标定参数(焦距)的情况下,无法直接使用 Bundler 算法进行三维重建,主要是因为 Bundler 算法使用 8 点法估计基础矩阵,在未知畸



图 6 SIFT 特征点与重投影点

Fig. 6 SIFT and reprojection points

变参数与焦距的情况下无法进行相机位姿估计。将本文算法应用到 Bundler 算法中,得到的三维重建结果如图 7 所示,图中蓝色锥体表示解算得到的相机位置与姿态,彩色点云表示场景的稀疏三维结构。

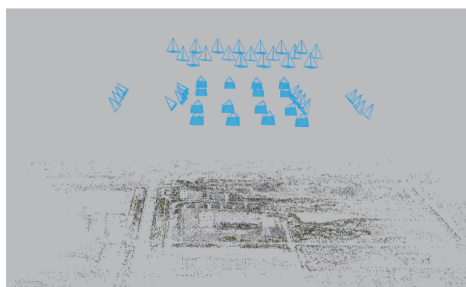


图 7 无人机影像集的三维重建结果

Fig. 7 3D reconstruction result of unmanned aerial vehicle dataset

由三维重建结果可知,本文方法重建的场景结构清晰,外点非常少,相机姿态合理。

由于无人机航拍数据没有相机的位置及姿态参考值,无法直接得到相机位姿的估计误差,只能采用重投影误差作为三维重建精度的评测标准。为了说明本文算法位姿估计的精度,采用数据集 CastleP30<sup>[22]</sup>进行位姿估计和三维重建,该数据集共 30 张影像,影像尺寸为 3072 pixel×2048 pixel,数据集提供了相机的位置和姿态参考值,可以统计位姿估计误差。使用该数据集进行三维重建的结果如图 8 所示,可以看出,三维重建结果中,建筑物立面完整,结构清晰,相机数目与参与重建影像数目一

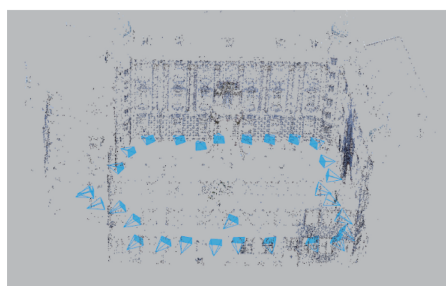
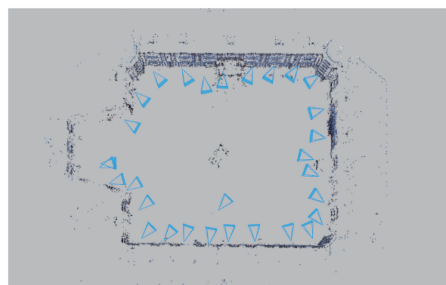


图 8 CastleP30 影像集的三维重建结果

Fig. 8 3D reconstruction result of CastleP30 dataset

致,相机位姿与场景点对应关系合理。

两组实验最终得到的标定矩阵如表 1 所示,无人机序列影像等效焦距估计值为 3811 pixel; CastleP30 序列影像的等效焦距估计值为 2764.30 pixel,其实际值为 2761.80 pixel。两组实验误差如表 2 所示,无人机序列影像平均重投影误

差为 0.43 pixel,由于没有相机位姿参考值,无法直接得出相机旋转矩阵和平移向量误差; CastleP30 的平均重投影误差为 0.27 pixel,平均旋转矩阵误差为 0.08°,平均平移向量误差为 0.11%。可见,使用本文算法进行三维重建,能够得到较高精度的相机位姿估计结果。

表 1 不同数据集标定矩阵估计结果

Table 1 Estimation results of calibration matrix with different datasets

Dataset	Number	Estimated calibration matrix	Reference calibration matrix
Unmanned aerial vehicle	50	$\begin{bmatrix} 3810.73 & 0 & 2457.50 \\ 0 & 3811.18 & 1632.00 \\ 0 & 0 & 1 \end{bmatrix}$	N/A
CastleP30	30	$\begin{bmatrix} 2761.99 & 0 & 1526.50 \\ 0 & 2766.51 & 1009.50 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 2759.48 & 0 & 1520.69 \\ 0 & 2764.16 & 1006.81 \\ 0 & 0 & 1 \end{bmatrix}$

表 2 不同数据集的估计误差

Table 2 Estimation error of different datasets

Dataset	Number	Average reprojection error/pixel	Average rotation error / (°)	Average translation error/%
Unmanned aerial vehicle	50	0.43	N/A	N/A
CastleP30	30	0.27	0.08	0.11

合成数据及真实影像实验结果表明,将本文提出的位姿估计方法应用到 Bundler 中,可以有效提高未知焦距情况下三维重建的精度和稳健性。

## 7 结 论

提出了一种高精度的非迭代位姿估计方法。该方法能够在 0 均值、标准差为 2 的噪声下,得到精度较高的位姿估计结果,并获得相机的焦距。该算法估计结果的数值稳定性好,计算效率高。将其应用于增量式 SFM 中可以大幅提高相机的位姿估计精度,实现未标定序列影像的稀疏重建。本文算法目前尚不具有非中心投影下的位姿估计能力,因此,需进一步研究。

## 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] Horaud R, Dornaika F, Lamiroy B. Object pose: the link between weak perspective, paraperspective, and full perspective [J]. International Journal of Computer Vision, 1997, 22(2): 173-189.
- [4] 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.
- [5] Schweighofer G, Pinz A. Globally optimal  $O(n)$  solution to the PnP problem for general camera models [C]. British Machine Vision Conference, 2008: 1-10.
- [6] Zhang Z J, Hao X Y, Cheng C Q, *et al.* Iteratively reweighted least squares method for camera pose estimation based on coplanar line correspondences [J]. Optics and Precision Engineering, 2016, 24(5): 1168-1175.

- 张振杰, 郝向阳, 程传奇, 等. 基于共面直线迭代加权最小二乘的相机位姿估计[J]. 光学精密工程, 2016, 24(5): 1168-1175.
- [7] Lepetit V, Moreno-Noguer F, Fua P. EPnP: an accurate  $O(n)$  solution to the PnP problem [J]. International Journal of Computer Vision, 2009, 81(2): 155-166.
- [8] Chen P, Wang C. IEPnP: an iterative camera pose estimation algorithm based on EPnP[J]. Acta Optica Sinica, 2018, 38(4): 0411001.  
陈鹏, 王晨. IEPnP: 一种基于 EPnP 的相机位姿迭代估计算法[J]. 光学学报, 2018, 38(4): 0411001.
- [9] Li S Q, Xu C, Xie M. A robust  $O(n)$  solution to the perspective- $n$ -point problem [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1444-1450.
- [10] Hesch J A, Roumeliotis S I. A direct least-squares (DLS) method for PnP [C]. IEEE International Conference on Computer Vision (ICCV), 2011: 383-390.
- [11] 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. 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.
- [12] Zheng Y Q, Kuang Y B, Sugimoto S, *et al.* Revisiting the PnP problem: a fast, general and optimal solution[C]. IEEE International Conference on Computer Vision (ICCV), 2013: 2344-2351.
- [13] Zheng Y Q, Sugimoto S, Okutomi M. ASPnP: an accurate and scalable solution to the perspective- $n$ -point problem[J]. IEEE Transactions on Information and Systems, 2013, 96(7): 1525-1535.
- [14] Kneip L, Li H D, Seo Y. UPnP: an optimal  $O(n)$  solution to the absolute pose problem with universal applicability[C]. European Conference on Computer Vision (ECCV), 2014: 127-142.
- [15] Bujnak M, Kukulova Z, Pajdla T. A general solution to the P4P problem for camera with unknown focal length[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008: 1-8.
- [16] Choi K, Lee S, Seo Y. A branch-and-bound algorithm for globally optimal camera pose and focal length[J]. Image and Vision Computing, 2010, 28(9): 1369-1376.
- [17] Penate-Sanchez A, Andrade-Cetto J, Moreno-Noguer F Exhaustive linearization for robust camera pose and focal length estimation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(10): 2387-2400.
- [18] Zheng Y Q, Sugimoto S, Sato I, *et al.* A general and simple method for camera pose and focal length determination [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014: 430-437.
- [19] Wu C C. P3.5P: pose estimation with unknown focal length[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 2440-2448.
- [20] Kukulova Z, Bujnak M, Pajdla T. Automatic generator of minimal problem solvers[C]. European Conference on Computer Vision (ECCV), 2008: 302-315.
- [21] Li X C, Larson M, Hanjalic A. Pairwise geometric matching for large-scale object retrieval [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 5153-5161.
- [22] Strecha C, von Hansen W, von Gool L, *et al.* On benchmarking camera calibration and multi-view stereo for high resolution imagery [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008: 1-8.