

参考点共面条件下的稳健相机位姿估计方法

陈津平, 吴晓亮*

天津大学精密仪器与光电子工程学院精密测试技术及仪器国家重点实验室, 天津 300072

摘要 单目视觉中的参考点共面问题是计算机视觉领域内一个基础而重要的问题。建立了针对共面参考点的 PnP (Perspective-n-Point) 问题模型, 并提出了一种稳健的解决方法, 具体包括直接法和迭代法两个部分。在直接法环节, 针对位姿恢复过程中存在的尺度不统一问题, 利用奇异值分解方法对旋转矩阵的恢复进行改进, 获得相机位姿估计结果, 并将此结果用作迭代法的初始值。在迭代法环节, 介绍了以物方空间共线性误差为目标函数的正交迭代算法, 为提升该算法的稳健性, 研究了加权正交迭代算法, 确立了重投影误差中外点的判定阈值并以此为依据引入了权重信息。实验结果表明, 在参考点数量较少或少量异常值存在的条件下, 所提方法具有良好的计算精度及稳健性, 有较高实用价值。

关键词 单目视觉; 共面参考点; 位姿估计; 共线性误差

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP230945

Robust Camera Pose Estimation Method Under Coplanar Reference Points Condition

Chen Jinping, Wu Xiaoliang*

State Key Laboratory of Precision Measurement Technology and Instruments, School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China

Abstract The problem of coplanar reference points in monocular vision is a fundamental and critical aspect within the field of computer vision. We establish a model for the coplanar Perspective-n-Point problem and present a robust method, which comprises two main components: a direct algorithm and an iterative algorithm. In direct algorithm part, we address the problem of scale inconsistency during the pose estimation process, employ the singular value decomposition method to improve the recovery of the rotation matrix, leading to an estimation of the camera pose. This estimated pose is then used as the initial value for the iterative algorithm. In iterative algorithm part, we introduce an orthogonal iterative algorithm with the object-space collinearity error as the objective function. To enhance the robustness of this algorithm, a weighted orthogonal iterative algorithm is studied. We establish a threshold for determining outliers in the reprojection error and introduce weight information for this algorithm. Experimental results demonstrate that under conditions of limited reference points or minimal outliers, the proposed method exhibits excellent computational accuracy and robustness. It may provide significant practical value.

Key words monocular vision; coplanar reference points; pose estimation; collinearity error

1 引言

相机位姿估计是图像处理和计算机视觉领域中的经典问题, 该问题描述了如何从二维图像中恢复相机在三维空间中的位置和姿态, 现已广泛应用于机器人导航^[1]、相机标定^[2]和虚拟现实^[3]等多个领域。位姿估计问题根据参考特征分为 PnP (Perspective-n-Point) 问题和 PnL (Perspective-n-Line) 问题^[4], PnL 问题以图像

中的直线信息作为参考依据, 而 PnP 问题则以图像中的参考点信息为依据。在实际应用中, 由于直线信息的获取相对困难, 因此相机位姿估计通常指的是 PnP 问题模型的求解。对于帧间相机位姿估计问题, 应使用特征点法^[5]获取参考点在不同相机视图中的匹配关系; 而对于单视图相机位姿估计问题, 通常只需要匹配参考点的空间坐标与像素坐标即可。

PnP 问题的解决方法根据有无目标函数分为直接

收稿日期: 2023-03-24; 修回日期: 2023-05-15; 录用日期: 2023-06-01; 网络首发日期: 2023-06-11

通信作者: *xwu423@tju.edu.cn

法和迭代法,直接法求解速度较快,而迭代法的计算结果精度较高。在实际应用中,迭代法通常依赖直接法或其他方法提供的初始值以避免陷入局部最优解,通过对初始值迭代优化可以得到位姿最优估计。常见的直接法包括基于几何的 P3P(Perspective-3-Point)算法^[6]、基于四元数的 OPnP 算法^[7]、基于最小二乘问题的 DLT(direct linear transform)算法以及基于刚体变换的 EPnP 算法^[8]等;而常见的迭代法包括以重投影误差为目标函数的光束平差(bundle adjustment)算法^[9]、以正交投影近似透视投影的 POSIT(pose from orthography and scaling with iterations)算法^[10]和以物方空间共线误差为目标函数的 LHM(Lu, Hager and Mjolsness)算法^[11]。其中,LHM 算法又名正交迭代算法,其全局收敛性表现出色,然而在参考点共面的条件下并不稳定^[12]。Schweighofer 等^[13]分析并推导了正交迭代算法不稳定的原因,提出了在参考点共面的情况下消除歧义的方法。此外,周润等^[14]提出了利用重投影误差评估测量误差可靠程度的方法,引入了权重信息。李鑫等^[15]通过引入 Kronecker 积提升了算法的速度,将每次运算的时间复杂度从 $O(n)$ 降至 $O(1)$ 。这些改进不但拓展了正交迭代算法的适用领域,还在效率上取得了巨大提升,使得该算法成为最先进的位姿估计迭代算法之一^[16]。

此外,参考点共面问题是 PnP 问题的一种特殊情况,该问题描述了从单张相机视图中恢复参考点平面与相机之间相对位姿的过程,可以通过两种方法解决:一种方法将参考点共面情况视为 PnP 问题的一个特例,利用 PPnP(planar-PnP)方法如 IPPE(infinitesimal plane-based pose estimation)算法^[17]来解决;另一种方法计算物体平面到图像平面之间的单应性映射关系,并通过分解单应矩阵来恢复相机与参考点平面之间的相对位姿关系^[18]。实际应用中,共面参考点问题可应用于基于二维平面图案(棋盘格、二维码和同心圆)的相机标定^[19]和无人机着陆。

本文研究了参考点共面条件下的 PnP 问题,使用奇异值分解(SVD)方法改进了相机位姿初始值估计方法,解决了旋转矩阵与平移向量恢复尺度不统一的问题。同时,由于位姿估计方法的测量精度、稳健性、稳定性依赖于参考点数量^[20],因此研究了加权正交迭代算法,解释了引入权重项的目的并确定了异常值的判断标准。实验表明所提方法具有精度高、稳健性好等特

点,适用于参考点数量较少且存在噪声的应用场景。

2 基本原理

2.1 针孔相机模型

选用针孔相机模型,设相机内参 \mathbf{K} 为

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

式中: f_x 和 f_y 对应各轴焦距; c_x 和 c_y 对应主点的相机坐标。设相机坐标系下存在某一空间点,其坐标为 $\mathbf{P}_c = [x_c, y_c, z_c]^T$,对应图像中的像素齐次坐标为 $\tilde{\mathbf{u}} = [u, v, 1]^T$,根据针孔相机模型的成像原理:

$$s\tilde{\mathbf{u}} = \mathbf{K}\mathbf{P}_c, \quad (2)$$

式中, $s = z_c$ 为深度信息,式(2)两边同时消去 s 可得

$$\tilde{\mathbf{u}} = \mathbf{K}\tilde{\mathbf{P}}_c, \quad (3)$$

式中, $\tilde{\mathbf{P}}_c = [x_c/z_c, y_c/z_c, 1]^T$ 为该点投影到归一化像平面的相机齐次坐标。设该空间点在世界坐标系下的坐标为 $\mathbf{P}_w = [x_w, y_w, z_w]^T$,使用旋转矩阵 $\mathbf{R} \in \text{SO}(3)$ (三维旋转群)和平移向量 $\mathbf{t} \in \mathbb{R}^3$ 描述世界坐标系到相机坐标系的转换关系:

$$\mathbf{P}_c = \mathbf{R}\mathbf{P}_w + \mathbf{t}, \quad (4)$$

代入式(2)得到:

$$s\tilde{\mathbf{u}} = \mathbf{K}\mathbf{P}_c = \mathbf{K}(\mathbf{R}\mathbf{P}_w + \mathbf{t}), \quad (5)$$

描述了空间点世界坐标与像素坐标之间的关系。

2.2 参考点共面问题

假设空间中有 n 个共面参考点,将某一参考点的世界坐标定义为 $\mathbf{P}_{iw} = [x_{iw}, y_{iw}, 0]^T$,该点对应的像素齐次坐标定义为 $\tilde{\mathbf{u}}_i = [u_i, v_i, 1]^T$,其中 $i = 1, 2, \dots, n$ 。定义增广矩阵为 $[\mathbf{R}|\mathbf{t}]$,将式(5)展开并改写为

$$s_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}}_{[\mathbf{R}|\mathbf{t}]} \begin{bmatrix} x_{iw} \\ y_{iw} \\ 0 \\ 1 \end{bmatrix}, \quad (6)$$

整理式(6)得到:

$$s_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \begin{bmatrix} x_{iw} \\ y_{iw} \\ 1 \end{bmatrix}. \quad (7)$$

此问题可通过 DLT 算法求解,将式(7)展开为方程组得到:

$$\begin{cases} s_i u_i = x_{iw} f_x r_{11} + x_{iw} c_x r_{31} + y_{iw} f_x r_{12} + y_{iw} c_x r_{32} + f_x t_x + c_x t_z \\ s_i v_i = x_{iw} f_y r_{21} + x_{iw} c_y r_{31} + y_{iw} f_y r_{22} + y_{iw} c_y r_{32} + f_y t_y + c_y t_z, \\ s_i = x_{iw} r_{31} + y_{iw} r_{32} + t_z \end{cases} \quad (8)$$

消掉式(8)中深度信息 s_i 可以得到两个约束。令 $t_z = 1$,其余参数可以通过4组参考点的匹配信息计算得到,此时得到的结果均失去尺度意义。

对于尺度恢复问题,常见的方法是暂时忽略旋转矩阵的正交性,引入3个独立的尺度因子来恢复旋转矩阵的前两列和平移向量,并通过外积计算直接恢复

旋转矩阵信息^[21]。这种方法削弱了旋转矩阵列向量之间的约束条件,并且在平移向量恢复与旋转矩阵寻优过程中独立了尺度的概念,因此对尺度的定义较为模糊。李红卫等^[22]利用 Gram-Schmidt 正交化方法对旋转矩阵进行处理,该方法提高了旋转矩阵的正交性但并非最佳旋转矩阵估计。

2.3 改进位姿初始值估计方法

针对上述问题,对旋转矩阵与平移向量的初始值估计作出了改进。考虑将平移向量恢复与旋转矩阵寻优这两个过程所使用的尺度因子统一为 k , 设通过外积计算得到旋转矩阵的估计值为 \mathbf{Q} , 使用 Frobenius 范数表示矩阵之间的距离, 将问题定义为

$$\arg \min_{k, \mathbf{X}} \|\mathbf{X} - k\mathbf{Y}\|_{\text{F}}^2, \text{ subject to } \mathbf{X}^T \mathbf{X} = \mathbf{I}_{2 \times 2}, \quad (9)$$

式中: $\mathbf{X} = [\mathbf{r}_1 \quad \mathbf{r}_2]$, \mathbf{r}_1 和 \mathbf{r}_2 为待恢复旋转矩阵 \mathbf{R} 的前两列; $\mathbf{Y} = [\mathbf{q}_1 \quad \mathbf{q}_2]$, \mathbf{q}_1 和 \mathbf{q}_2 为旋转矩阵估计值 \mathbf{Q} 的前两列; $\mathbf{I}_{2 \times 2}$ 是大小为 2×2 的单位矩阵。利用矩阵迹 (trace) 的性质, 展开式 (9) 中的目标函数得到:

$$\begin{aligned} \|\mathbf{X} - k\mathbf{Y}\|_{\text{F}}^2 &= \text{tr}[(\mathbf{X} - k\mathbf{Y})^T (\mathbf{X} - k\mathbf{Y})] = \\ &2 + k^2 \text{tr}(\mathbf{Y}^T \mathbf{Y}) - 2k \text{tr}(\mathbf{X}^T \mathbf{Y}), \end{aligned} \quad (10)$$

将问题转化为:

$$\begin{aligned} \arg \min_{k, \mathbf{X}} \|\mathbf{X} - k\mathbf{Y}\|_{\text{F}}^2 &= \\ \arg \min_{k, \mathbf{X}} [k^2 \text{tr}(\mathbf{Y}^T \mathbf{Y}) - 2k \text{tr}(\mathbf{X}^T \mathbf{Y})]. \end{aligned} \quad (11)$$

对于尺度因子 k , 式 (11) 目标函数取到最小值的条件为:

$$k = \text{tr}(\mathbf{X}^T \mathbf{Y}) / \text{tr}(\mathbf{Y}^T \mathbf{Y}), \quad (12)$$

代入式 (11) 得到:

$$\begin{aligned} \arg \min_{k, \mathbf{X}} \|\mathbf{X} - k\mathbf{Y}\|_{\text{F}}^2 &= \arg \min_{\mathbf{X}} \left[-\frac{\text{tr}(\mathbf{X}^T \mathbf{Y})^2}{\text{tr}(\mathbf{Y}^T \mathbf{Y})} \right] = \\ \arg \max_{\mathbf{X}} [\text{tr}(\mathbf{X}^T \mathbf{Y})]. \end{aligned} \quad (13)$$

通过 SVD 方法将矩阵 \mathbf{Y} 分解得到 $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$,

其中:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}; \quad (14)$$

\mathbf{U} 为 3×3 大小的正交矩阵; \mathbf{V} 为 2×2 大小的正交矩阵。将式 (13) 中目标函数展开得到:

$$\text{tr}(\mathbf{X}^T \mathbf{Y}) = \text{tr}[(\mathbf{X}^T \mathbf{U}\mathbf{\Sigma})\mathbf{V}^T] = \text{tr}[\mathbf{V}^T (\mathbf{X}^T \mathbf{U}\mathbf{\Sigma})], \quad (15)$$

令 $\mathbf{Z} = \mathbf{V}^T \mathbf{X}^T \mathbf{U}$, 根据其性质得到不等式:

$$\text{tr}(\mathbf{X}^T \mathbf{Y}) = \text{tr}(\mathbf{Z}\mathbf{\Sigma}) = \sum_{i=1}^2 z_{ii} \sigma_i \leq \sum_{i=1}^2 \sigma_i, \quad (16)$$

当 $\mathbf{Z} = [\mathbf{I}_{2 \times 2} | \mathbf{0}_{2 \times 1}]$ 时, 式 (16) 取最大值:

$$\max_{\mathbf{X}} [\text{tr}(\mathbf{X}^T \mathbf{Y})] = \sum_{i=1}^2 \sigma_i, \quad (17)$$

此时得到矩阵 \mathbf{X} 的解为 $\mathbf{X} = \mathbf{U}[\mathbf{V} | \mathbf{0}_{2 \times 1}]^T$ 。将式 (17) 代入式 (12) 得到:

$$k = \frac{\sigma_1 + \sigma_2}{\text{tr}(\mathbf{Y}^T \mathbf{Y})}. \quad (18)$$

最后, 利用外积计算旋转矩阵 \mathbf{R} 的第 3 列 $\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$, 利用比例因子 k 恢复平移向量。在改进后的单应矩阵分解方法中, 旋转矩阵和平移向量均可以通过同一个尺度因子恢复。

2.4 正交迭代算法

上述方法由于缺乏目标函数, 难以通过迭代的方式得到全局最优解, 只适合在噪声较低或精度要求不高的场景下使用。不同于光束平差法利用像方空间中重投影误差作为目标函数, 正交迭代算法利用物方空间中的共线性误差来定义目标函数^[11], 其中, 物方空间误差 (共线性误差) 与像方空间误差 (重投影误差) 的几何意义如图 1 所示。

设第 i 个参考点对应的像素坐标观测值为 $\hat{\mathbf{p}}_i$, 由式 (3) 得到该点的像平面投影坐标:

$$\hat{\mathbf{v}}_{ic} = \mathbf{K}^{-1} \hat{\mathbf{p}}_i, \quad (19)$$

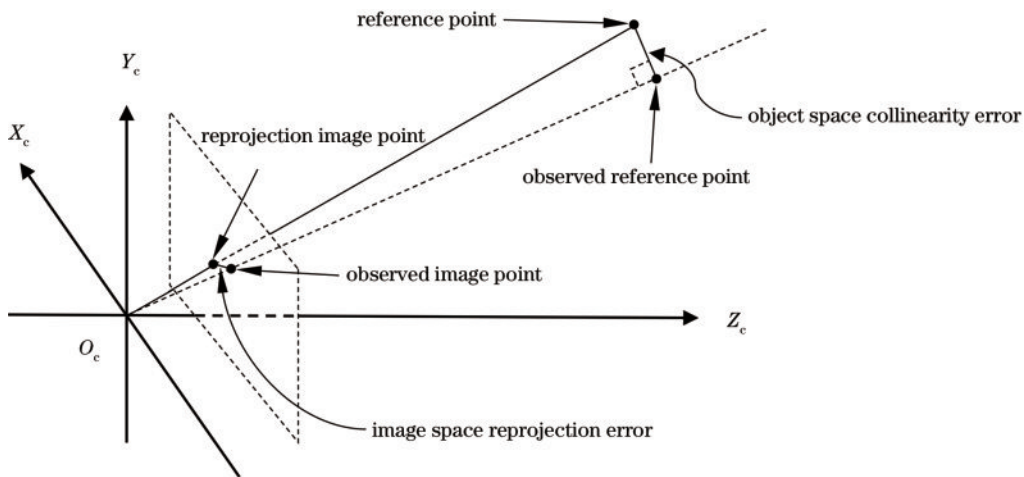


图 1 物方空间与像方空间中误差的几何意义

Fig. 1 Geometric meanings of the object-space error and image-space error

计算正交投影矩阵 \hat{V}_i :

$$\hat{V}_i = \frac{\hat{\mathbf{v}}_{ic} \hat{\mathbf{v}}_{ic}^T}{\hat{\mathbf{v}}_{ic}^T \hat{\mathbf{v}}_{ic}}, \quad (20)$$

定义物方空间共线性误差向量:

$$\mathbf{e}_i = \hat{\mathbf{P}}_{ic} - (\mathbf{R}\mathbf{P}_{iw} + \mathbf{t}), \quad (21)$$

由正交投影关系得到:

$$\hat{\mathbf{P}}_{ic} = \hat{V}_i(\mathbf{R}\mathbf{P}_{iw} + \mathbf{t}), \quad (22)$$

将式(22)代入式(21),并定义目标函数的表示:

$$E(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^n \|\mathbf{e}_i\|^2 = \sum_{i=1}^n \left\| (\mathbf{I} - \hat{V}_i)(\mathbf{R}\mathbf{P}_{iw} + \mathbf{t}) \right\|^2. \quad (23)$$

此时假设已知旋转矩阵 \mathbf{R} , 由于平移向量 \mathbf{t} 在式(23)中的最高项为二次形式, 因此平移向量 \mathbf{t} 的当前最优值有闭合解:

$$\mathbf{t}(\mathbf{R}) = \arg \min_{\mathbf{t}} \sum_{i=1}^n \left\| (\mathbf{I} - \hat{V}_i)[\mathbf{R}\mathbf{P}_{iw} + \mathbf{t}(\mathbf{R})] \right\|^2 = \left[\sum_{i=1}^n (\mathbf{I} - \hat{V}_i) \right]^{-1} \sum_{i=1}^n (\hat{V}_i - \mathbf{I}) \mathbf{R}\mathbf{P}_{iw}. \quad (24)$$

旋转矩阵 \mathbf{R} 的求解可以视为 3D-3D 绝对定向问题, 使用 Umeyama 算法^[23]求得最小二乘解。为避免绝对定向问题中协方差矩阵计算的相关性^[11], 设在第 k 次迭代后, 已知 $\mathbf{t}^{(k)}$ 与 $\mathbf{R}^{(k)}$, 并通过计算得到 $\mathbf{t}^{(k+1)}$, 令 $\mathbf{M}_i^{(k)} = \hat{V}_i[\mathbf{R}^{(k)}\mathbf{P}_{iw} + \mathbf{t}^{(k)}]$, 则 $\mathbf{R}^{(k+1)}$ 可通过公式求解, 表示为

$$\mathbf{R}^{(k+1)} = \arg \min_{\mathbf{R}^{(k+1)}} \sum_{i=1}^n \left\| \mathbf{R}^{(k+1)}\mathbf{P}_{iw} + \mathbf{t}^{(k+1)} - \mathbf{M}_i^{(k)} \right\|^2. \quad (25)$$

迭代式(24)和式(25)以更新平移向量 \mathbf{t} 和旋转矩阵 \mathbf{R} , 预设迭代次数上限、最小误差阈值、迭代灵敏度阈值均可作为该方法的停止条件。

2.5 加权正交迭代算法

理想条件下, 上述方法往往能在较少的迭代次数内达到最优值, 然而在实际应用中, 当参考点较少时, 上述算法的性能表现往往不能达到预期标准, 这并非算法的缺陷而是误差的存在使然。

考虑以正交迭代算法为基础, 引入权重项来描述每个空间参考点共线性误差的置信度。对于测量误差较大的参考点, 其计算得到的共线误差置信度较低, 应当赋予较小的权重; 反之, 对于测量误差较小的参考点, 其计算得到的共线误差置信度较高, 应当赋予较大的权重。通过对每个空间参考点样本引入权重, 可以调整不同样本对总体误差函数的贡献率。在样本充足的条件下, 合适的权重选择可以减少甚至避免异常值对迭代结果的影响; 而在样本不充足的条件下, 合适的权重选择可以避免迭代结果的过拟合。

设空间参考点的第 i 个样本对应的权重项为 w_i (注意与世界坐标系下标区分), 设目标函数为

$$E(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^n \|\omega_i \mathbf{e}_i\|^2 = \sum_{i=1}^n \omega_i \|\mathbf{e}_i\|^2 = \sum_{i=1}^n \omega_i \left\| (\mathbf{I} - \hat{V}_i)(\mathbf{R}\mathbf{P}_{iw} + \mathbf{t}) \right\|^2, \quad (26)$$

由式(24)得到第 k 次迭代后的平移向量的表达式为

$$\mathbf{t}^{(k)}(\mathbf{R}) = \left[\sum_{i=1}^n \omega_i (\mathbf{I} - \hat{V}_i) \right]^{-1} \left[\sum_{i=1}^n \omega_i (\hat{V}_i - \mathbf{I}) \mathbf{R}\mathbf{P}_{iw} \right], \quad (27)$$

类似地, 第 $k+1$ 次迭代后的旋转矩阵为:

$$\mathbf{R}^{(k+1)} = \arg \min_{\mathbf{R}^{(k+1)}} \sum_{i=1}^n \omega_i \left\| \mathbf{R}^{(k+1)}\mathbf{P}_{iw} + \mathbf{t}^{(k+1)} - \mathbf{M}_i^{(k)} \right\|^2. \quad (28)$$

与式(25)相比, 对式(28)的求解可由几个步骤完成, 首先分别计算 \mathbf{P}_{iw} 和 $\mathbf{M}_i^{(k)}$ 的质心:

$$\begin{cases} \bar{\mathbf{P}}_{iw} = \sum_{i=1}^n \omega_i \mathbf{P}_{iw} / \sum_{i=1}^n \omega_i \\ \bar{\mathbf{M}}_i^{(k)} = \sum_{i=1}^n \omega_i \mathbf{M}_i^{(k)} / \sum_{i=1}^n \omega_i \end{cases}, \quad (29)$$

令 $\mathbf{P}'_{iw} = \mathbf{P}_{iw} - \bar{\mathbf{P}}_{iw}$, $\mathbf{M}'_i^{(k)} = \mathbf{M}_i^{(k)} - \bar{\mathbf{M}}_i^{(k)}$, 并定义矩阵 $\mathbf{S}_{3 \times n}$ 和 $\mathbf{T}_{3 \times n}$, 二者的列向量分别由每个样本对应的 \mathbf{P}'_{iw} 和 $\mathbf{M}'_i^{(k)}$ 组成, 设 $\mathbf{W} = \text{diag}(\omega_1, \omega_2, \dots, \omega_n)$ 为对角矩阵, 得到协方差矩阵:

$$\mathbf{N} = \mathbf{S}\mathbf{W}\mathbf{T}^T, \quad (30)$$

使用 SVD 方法分解矩阵 \mathbf{N} 得到 $\mathbf{N} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, 当 $\mathbf{R}^{(k+1)} = \mathbf{V} \cdot \text{diag}[1, 1, \det(\mathbf{V}\mathbf{U}^T)] \cdot \mathbf{U}^T$ 时, 式(28)中目标函数取最小值, 证明加权后的正交迭代算法同样满足正交迭代算法的全局收敛性^[11-14]。

2.6 权重项的确定

在实际应用中, 由于观测所引入的误差分布与大小未知, 无法直接得到每个样本的置信度, 对异常值的正确判断是加权正交迭代算法性能提升的关键。汪佳宝等^[24]尝试使用参考点深度的倒数作为权重项, 该方法对粗差的敏感性较差。周润等^[14]提出了利用重投影误差评估测量误差, 结合对 Huber 函数^[25]的改进来确定权重项, 给重投影误差大于一定阈值的参考点分配较小的权重, 该方法通过引入二次项对外点进行惩罚。陈紫强等^[26]提出了将重投影误差四分位数中的最大值和另一个与焦距和重投影误差相关的参数作为阈值, 取二者最大值对空间参考点对应的共线误差进行二值加权, 该方法对粗差的处理效果明显。此外, 刘晨阳等^[27]在使用改进 Huber 函数引入权重项的基础上, 利用观测点到图像中心点的距离引入了新的权重项, 该方法有效提高了结果的精度。

在参考点 3D-2D 匹配信息中, 若少量异常值存在, 如某个空间参考点位置信息存在较大测量误差, 二次项引入的权重衰减会使正交迭代算法更准确地恢复相机的姿态。若不存在异常值, 如在噪声水平不高的情况下, 二次项的引入会削弱算法中正确的共线性约

束。在实际应用中,当样本量充足时,样本分布可以近似代表总体分布,而对于样本较少的场景,标准差作为异常值判别方法会产生较大误差。假设目标观测值横纵坐标误差分别为 Δu 与 Δv 且服从标准正态分布,则重投影误差服从自由度为 2 的卡方分布而非正态分布。此外,由于工程实践中样本较少,因此无法准确估计目标像素观测值横纵坐标误差的标准差。

将重投影误差样本空间均值 μ 、第一四分位数 q_1 、中位数 q_2 与第三四分位数 q_3 作为依据,并基于统计学家 Huber 提出的理念,为重投影误差值不同的参考点样本分别赋予平方项或一次项衰减系数。权重项的具体定义为

$$w_i = \begin{cases} \mu^2/r_i^2, & |r_i| > \delta_1 \\ \mu/r_i, & \delta_2 < |r_i| \leq \delta_1, \\ 1, & |r_i| \leq \delta_2 \end{cases} \quad (31)$$

其中:

$$\delta_1 = \max\left(\mu, q_2, \frac{q_1 + q_3}{2}\right); \quad \delta_2 = \min\left(\mu, q_2, \frac{q_1 + q_3}{2}\right). \quad (32)$$

使用改进初始值估计方法计算得到位姿初始值初始权重项 w_i , 利用加权正交迭代算法计算新的位姿,并生成新的权重项以便下一轮迭代,多次重复以上流程直到满足停止条件。

3 实验与分析

实验部分分为数据仿真实验和真实实验,数据仿真实验使用 Visual Studio 2019 和 MATLAB R2020b 软件,真实实验使用大疆 Phantom 4 Pro+ V2.0 型号无人机平台。通过旋转误差、平移误差以及物方空间定位误差对算法的可行性、稳健性及精度进行验证。

设 \mathbf{R}_{true} 为旋转信息真值,旋转误差的计算方式为

$$e_r = 2\arccos\left[0.5\sqrt{1 + \text{tr}(\mathbf{R}\mathbf{R}_{\text{true}}^T)}\right] \cdot 180/\pi, \quad (33)$$

设 \mathbf{t}_{true} 为平移信息真值,平移误差的计算方式为

$$e_t = |\mathbf{t}_{\text{true}} - \mathbf{t}|/|\mathbf{t}_{\text{true}}| \cdot 100, \quad (34)$$

设 \mathbf{P}_{true} 为空间参考点真实坐标、 \mathbf{P}_{obs} 为观测坐标,物方空间定位误差计算方式为

$$e_{\text{obj}} = \|\mathbf{P}_{\text{true}} - \mathbf{P}_{\text{obs}}\|_0. \quad (35)$$

3.1 仿真实验

设相机分辨率长为 1280 pixel、宽为 960 pixel,等效焦距为 800,在平面上限定区域内随机生成 n 个 Z 轴分量为 0 的共面参考点、旋转矩阵、平移向量,约束条件为相机光心对准世界坐标系原点、相机坐标系下 Z 轴分量大于 0 且像素坐标在图像分辨率范围内。在仿真过程中,给像素坐标添加高斯噪声以模拟坐标观测值。

选用以下 7 种算法进行对比实验。

1) HOMO 算法,即常见单应矩阵分解方法:通过

DLT 算法恢复 2D 到 3D 之间的变换关系,直接通过外积恢复位姿,计算速度较快。

2) 改进 SVD 正交化算法,即初始值估计算法:通过对 DLT 算法计算得到的旋转矩阵进行正交化,并引入统一的尺度因子恢复平移向量,提升了 HOMO 算法的精度。

3) LHM 算法,即正交迭代算法:计算结果精度较高,具有全局收敛性。

4) WLHM 算法,即加权正交迭代算法:提升了正交迭代算法的稳健性,同样具有全局收敛性。

5) 改进 SVD 正交化+LHM 算法:以算法 2) 的结果作为算法 3) 的初始值,进一步提升了 LHM 算法的性能。

6) 改进 SVD 正交化+WLHM 算法:以算法 2) 的结果作为算法 4) 的初始值,进一步提升了 WLHM 算法的性能。

7) 改进 SVD 正交化+高斯牛顿(GN)算法:以算法 2) 的结果作为光束平差算法的初始值,并选用 GN 算法求解其中的非线性最小二乘问题,避免了单独使用 GN 算法陷入局部最小值的问题。

与上述 7 种算法相比,单独使用 GN 算法得到的实验结果相对较差,故未列入其中。

3.1.1 参考点数目实验

该实验研究了各算法精度随参考点数目的变化关系,设异常值数目占总参考点数的 1/10(向右取整),其分布满足均值为 0、标准差为 5 pixel 的正态分布,进行 500 次独立实验,参考点数目从 4 个增加到 20 个。选用旋转误差与平移误差的平均值与中位数作为位姿恢复结果的评估方法,结果如图 2 所示,其中:横坐标为参考点数目;纵坐标为误差评估值。

由图 2 可知:

1) WLHM 算法对异常值的处理能力非常强,在引入改进 SVD 正交化算法的计算结果作为其初始值后,改进 SVD 正交化+WLHM 算法的性能在参考点较少的情况下得到了进一步的提升;

2) 相比于 WLHM 算法,初始值并没有给 LHM 算法带来巨大提升,验证了 LHM 算法的全局收敛性,并说明 LHM 算法的收敛速度较快,而 WLHM 算法以收敛速度的降低为代价,引入了衰减权重系数以提升其稳健性;

3) 由于引入了旋转矩阵内在约束,改进 SVD 正交化算法提升了 HOMO 算法的旋转精度;

4) 在有良好初始值的条件下,改进 SVD 正交化+GN 算法略优于改进 SVD 正交化+LHM 算法,但是前者十分依赖平移向量作为初始值,不适用于已知旋转信息却未知平移信息的场合。

3.1.2 噪声水平实验

噪声水平仿真实验旨在得到各算法精度随高斯噪声标准差增加的变化曲线,设参考点数目为 5,噪声服

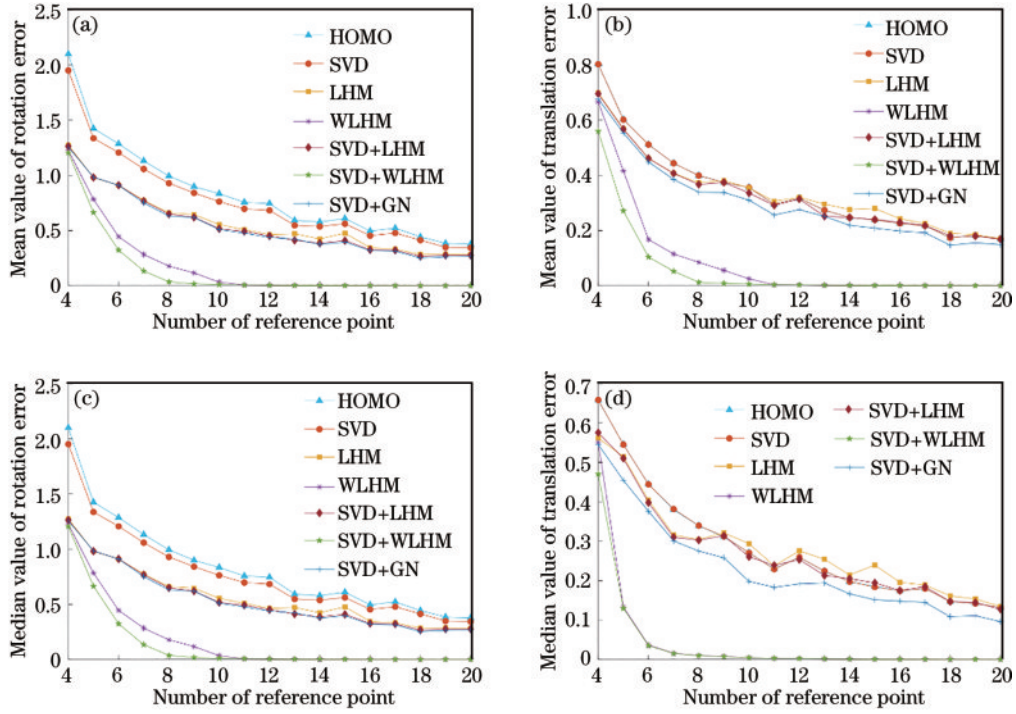


图 2 参考点数目仿真实验。(a) 旋转误差平均值;(b) 平移误差平均值;(c) 旋转误差中位数;(d) 平移误差中位数
 Fig. 2 Simulation of the number of reference points. (a) Mean value of rotation error; (b) mean value of translation error; (c) median value of rotation error; (d) median value of translation error

从高斯分布且均值为 0, 标准差取值范围为 0~10 pixel, 步长为 0.5 pixel, 对每个标准差的取值分别进行 500 次独立试验, 结果如图 3 所示, 其中: 横坐标为噪声标准差; 纵坐标为误差评估值。

从图 3 中能够看出:

1) 即使是较小的噪声, WLHM 算法对其的抑制作用也非常明显, 在改进 SVD 正交化算法提供初值后, 算法的性能提升明显, 结合参考点数目实验,

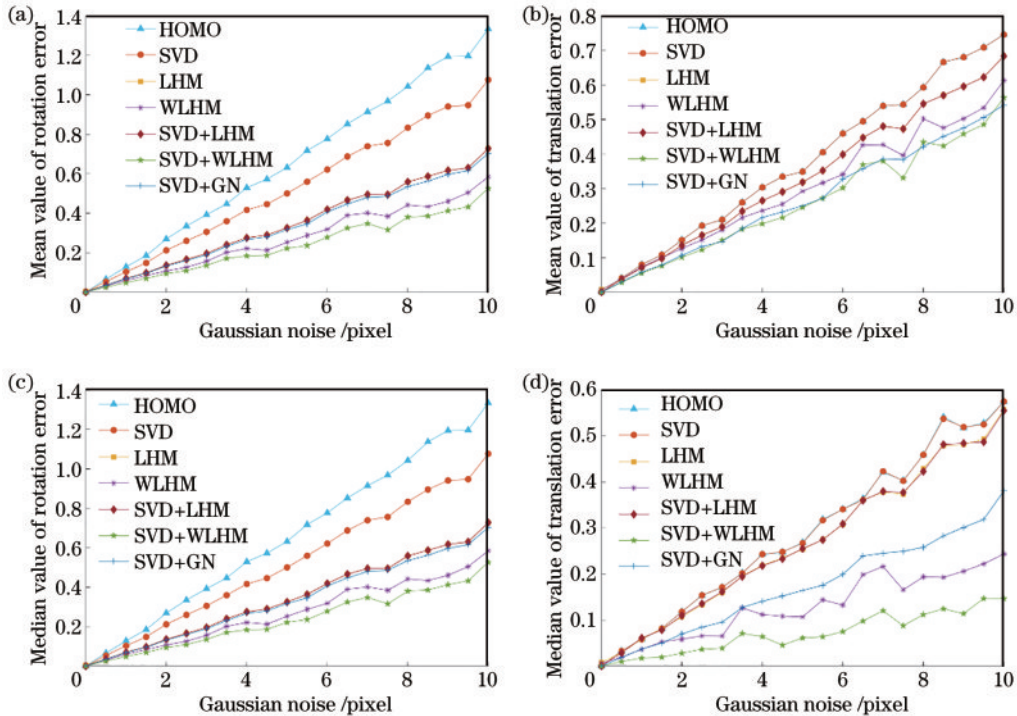


图 3 噪声水平仿真实验。(a) 旋转误差平均值;(b) 平移误差平均值;(c) 旋转误差中位数;(d) 平移误差中位数
 Fig. 3 Simulation of noise level. (a) Mean value of rotation error; (b) mean value of translation error; (c) median value of rotation error; (d) median value of translation error

WLHM算法对异常值处理的性能提升是全面的。

2) WLHM算法在标准差范围为 3~8 pixel 时,对应的曲线出现了抖动的情况,对比 LHM 算法,发现该抖动较为吻合,在 GN 算法中也出现了类似的情况,而 WLHM 算法在多数情况下得到了误差更小的结果,少数情况下精度提升相对一般,但仍优于其他算法,因此其抖动更为明显。

3) 各算法对位姿的估计与噪声的大小呈近似线性关系,可以建立结果误差模型。该结论需要得到进一步验证。

3.1.3 定位误差分布实验

该仿真实验对像面中每个像素对应的物方平面定

表 1 仿真日志
Table 1 Simulation log

Index	U-V Coordinate	Camera Coordinate	World Coordinate
1	(403, 295)	(-41.357, -32.283, 139.601)	(-44.886, -32.571, 0)
2	(999, 282)	(46.447, -25.617, 103.503)	(50.006, -25.327, 0)
3	(407, 588)	(-40.005, 18.543, 137.356)	(-43.094, 18.291, 0)
4	(1073, 673)	(53.557, 23.872, 98.950)	(58.010, 24.235, 0)
5	(461, 581)	(-29.815, 16.823, 133.250)	(-32.097, 16.637, 0)

图 4 为此次实验对应的误差分布图,其中:横坐标为像素坐标轴 U ;纵坐标为像素坐标轴 V ;色度条为定位误差,单位为 m(对应表 1 中出现的尺度)。

从图 4 中可以得到如下信息:

1) 改进 SVD 正交化算法对图像中间区域目标的定位精度很高,而对图像其余部分的定位结果较差,其中中心区域也是参考点分布密集的区域。

2) 相比前一种算法,改进 SVD 正交化+WLHM

位误差进行了研究,仅利用改进 SVD 正交化算法(算法 2)与改进 SVD 正交化+WLHM 算法(算法 6)进行位姿估计。假设目标像素点的观测值未受到噪声影响,利用空间平面与相机直线交会的方式,计算得到世界坐标系下目标点坐标的真值与估计值,再通过式(35)计算得到该目标像素点对应的定位误差。设参考点数目为 5,全部添加均值为 0、标准差为 2 pixel 的高斯噪声,由于需要对定位误差的分布进行研究,该仿真采用单次实验模式,仿真生成的像素、相机以及世界坐标如表 1 所示。

算法可看作在区域内进行了均匀化处理,以图像中心区域的精度为代价提升了其余区域的定位精度,该信息也可以从图 4(c)中更直观地获取。

3) 在实践中,应根据参考点的位置情况,考虑采用直接法与迭代法相结合的方式定位。以图 4 为例:若目标点在图像中心区域,采用改进 SVD 正交化算法即可;若目标点出现在其他区域,则应采用改进 SVD 正交化+WLHM 算法。此外,应避免目标点出

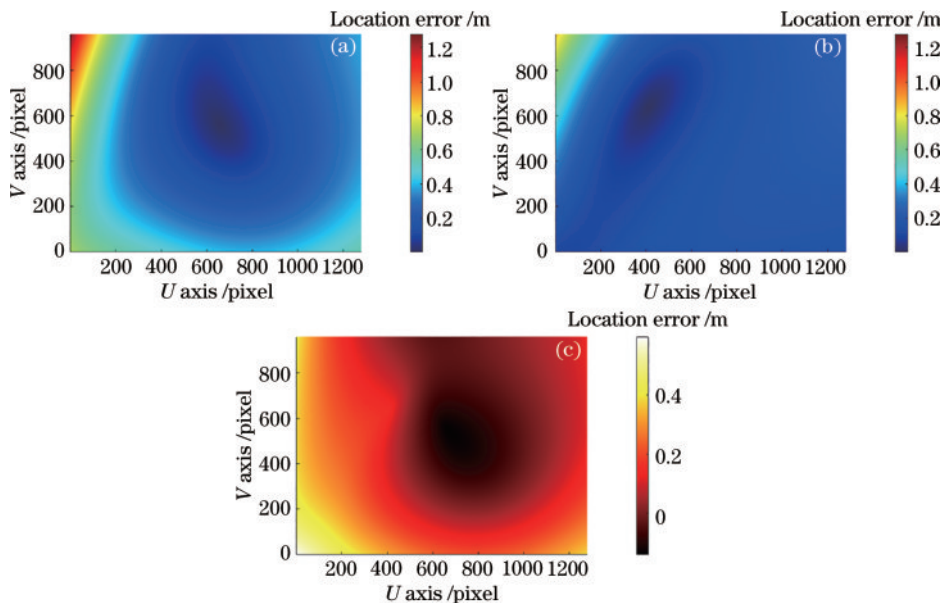


图 4 定位误差分布实验结果。(a) 改进 SVD 正交化算法定位误差;(b) 改进 SVD 正交化+WLHM 算法定位误差;(c) 图 4(a)与图 4(b)之间差值

Fig. 4 Localization error distribution experiment results. (a) Localization error using improved SVD orthogonalization method; (b) localization error using improved SVD orthogonalization + WLHM method; (c) difference between Fig. 4(a) and Fig. 4(b)

现在图像中靠近角落的位置,否则会出现定位精度极低的情况。

3.2 真实实验

真实实验选用操场作为实验场地,利用无人机相机对多个地面球形靶标拍摄图像,飞行高度在 50 m 左右。通过差分 GNSS (global navigation satellite system) 模块测量每个靶标所在位置的 GPS (global positioning system) 信息 (精度为 cm), 利用 WGS-84 大地坐标系到地心地固坐标系^[28]和地心地固坐标系到东北天坐标系的转换关系建立世界坐标系,得到各靶标的世界坐标。选用其中 5 个靶标作为参考点 (reference point 1~5), 其余作为目标点,如图 5 所示。

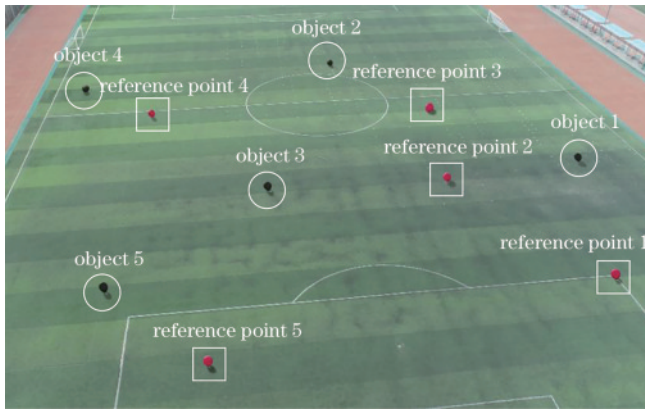


图 5 真实实验环境示意图

Fig. 5 Schematic image of real experimental environment

利用基于色域的方法确定 ROI (region of interest), 结合 Canny 算子对靶标边缘轮廓进行提取, 并对其进行圆拟合得到靶标像素坐标观测值。无人机相机内参数通过图像中 Exif 格式的元数据获取, 其中照片分辨率尺寸为 5472 pixel × 3078 pixel, 等效焦距为 24 mm。控制无人机在 10 处不同位置进行拍摄, 选用改进 SVD 正交化算法 (算法 2) 与改进 SVD 正交化 + WLHM 算法 (算法 6) 估算出机载相机的位姿数据, 然后利用该数据对每张相机视图中 5 个目标点靶标进行定位, 并计算目标点样本空间定位误差的 RMSE (root mean square error) 值, 对所提算法的精度进行实验验证。结果如表 2 所示, 其中: 第一列为实验序号; 第二列为算法 2 实验结果; 第三列为算法 6 实验结果。

从表 2 中得到如下信息:

1) 利用改进 SVD 正交化算法得到的 RMSE 值在 0.7 m 范围内, 而使用改进 SVD 正交化 + WLHM 算法得到的 RMSE 值在 0.4 m 范围内, 精度有较大提升。

2) 通过比较两种算法结果, 发现后者在任何位姿下的结果均优于前者。这是由于空间参考点的世界坐标与图像中的像素坐标均存在噪声, 在这种条件下, WLHM 算法的引入有效地提高了相机位姿估计的精度, 降低了当前视图下目标点的定位误差。

表 2 真实实验结果

Table 2 Results of real experiment

Index of experiment	RMSE of SVD /m	RMSE of SVD+WLHM /m
1	0.693	0.360
2	0.550	0.302
3	0.416	0.149
4	0.603	0.312
5	0.337	0.097
6	0.574	0.306
7	0.427	0.238
8	0.596	0.361
9	0.461	0.225
10	0.433	0.210

4 结 论

在空间参考点共面情况下研究了相机位姿问题的解决方法。所提改进初始值计算方法利用统一的尺度因子恢复了共面参考点提供的旋转和平移信息, 并以此作为迭代法的初始值。同时, 针对参考点数量较少及存在噪声的问题, 提出了一种改进正交迭代算法, 分析了加权项对异常值的抑制作用, 并对所提算法与几种常见算法的性能进行了对比。前两个仿真实验证明了改进初始值估计方法和加权正交迭代算法的有效性, 同时展示了所提方法在参考点较少时的稳健性; 第 3 个仿真实验和真实实验在噪声分布均匀的场景中进行, 通过定位精度反映了加权正交迭代算法对直接法的改进并展望了二者的适用场合。实验结果表明, 所提改进位姿初始值计算方法精度良好, 在此基础上使用加权正交迭代算法可以满足更高精度及稳健性的要求。

参 考 文 献

- [1] 张磊, 徐孝彬, 曹晨飞, 等. 基于动态特征剔除的图像与点云融合的机器人位姿估计方法[J]. 中国激光, 2022, 49(6): 0610001.
Zhang L, Xu X B, Cao C F, et al. Robot pose estimation method based on image and point cloud fusion with dynamic feature elimination[J]. Chinese Journal of Lasers, 2022, 49(6): 0610001.
- [2] 娄群, 吕钧瀚, 文立华, 等. 基于亚像素边缘检测的高精度相机标定方法[J]. 光学学报, 2022, 42(20): 2012002.
Lou Q, Lü J H, Wen L H, et al. High-precision camera calibration method based on sub-pixel edge detection[J]. Acta Optica Sinica, 2022, 42(20): 2012002.
- [3] 郭中远, 徐锋, 王贵洋, 等. 基于虚拟现实的移动机器人真实环境三维建模系统[J]. 激光与光电子学进展, 2023, 60(2): 0233001.
Guo Z Y, Xu F, Wang G Y, et al. 3D modeling system of mobile robot based on virtual reality in real environment

- [J]. *Laser & Optoelectronics Progress*, 2023, 60(2): 0233001.
- [4] Zhou L P, Koppel D, Kaess M. A complete, accurate and efficient solution for the perspective-N-line problem [J]. *IEEE Robotics and Automation Letters*, 2021, 6(2): 699-706.
- [5] 马志艳, 邵长松, 杨光友, 等. 同步定位与建图技术研究进展[J]. *电光与控制*, 2023, 30(3): 78-85, 106.
Ma Z Y, Shao C S, Yang G Y, et al. Research progress of SLAM technology[J]. *Electronics Optics & Control*, 2023, 30(3): 78-85, 106.
- [6] Wang P, Sun C K, Zhang Z M. Linear pose estimation with a monocular vision system[J]. *Chinese Journal of Scientific Instrument*, 2011, 32(5): 1126-1131.
- [7] Zheng Y Q, Kuang Y B, Sugimoto S, et al. Revisiting the PnP problem: a fast, general and optimal solution [C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2014: 2344-2351.
- [8] 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.
- [9] Triggs B, McLauchlan P F, Hartley R I, et al. Bundle adjustment: a modern synthesis[M]//Triggs B, Zisserman A, Szeliski R. *Vision algorithms: theory and practice*. Lecture notes in computer science. Heidelberg: Springer Berlin Heidelberg, 2000, 1883: 298-372.
- [10] Oberkampf D, DeMenthon D F, Davis L S. Iterative pose estimation using coplanar feature points[J]. *Computer Vision and Image Understanding*, 1996, 63(3): 495-511.
- [11] 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.
- [12] Zhou K, Wang X J, Wang Z, et al. Complete initial solutions for iterative pose estimation from planar objects [J]. *IEEE Access*, 2018, 6: 22257-22266.
- [13] Schweighofer G, Pinz A. Robust pose estimation from a planar target[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(12): 2024-2030.
- [14] 周润, 张征宇, 黄叙辉. 相机位姿估计的加权正交迭代算法[J]. *光学学报*, 2018, 38(5): 0515002.
Zhou R, Zhang Z Y, Huang X H. Weighted orthogonal iteration algorithm for camera pose estimation[J]. *Acta Optica Sinica*, 2018, 38(5): 0515002.
- [15] 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.
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.
- [16] 熊芝, 许航, 张刘港, 等. 基于加权加速正交迭代算法的相机位姿估计[J]. *红外与激光工程*, 2022, 51(10): 20220030.
Xiong Z, Xu H, Zhang L G, et al. Pose estimation of camera based on weighted accelerated orthogonal iterative algorithm[J]. *Infrared and Laser Engineering*, 2022, 51(10): 20220030.
- [17] Collins T, Bartoli A. Infinitesimal plane-based pose estimation[J]. *International Journal of Computer Vision*, 2014, 109(3): 252-286.
- [18] Malis E, Vargas M. Deeper understanding of the homography decomposition for vision-based control[D]. Paris: The National Institute for Research in Digital Science and Technology, 2007.
- [19] 黄臻臻, 黄晓梅, 郑莲玉, 等. 同心圆偏心误差补偿迭代的相机标定方法[J]. *光学学报*, 2022, 42(19): 1912003.
Huang Z Z, Huang X M, Zheng L Y, et al. Camera calibration method of concentric circle eccentric error compensation iteration[J]. *Acta Optica Sinica*, 2022, 42(19): 1912003.
- [20] 李安虎, 邓兆军, 刘兴盛, 等. 基于虚拟相机的位姿估计研究进展[J]. *激光与光电子学进展*, 2022, 59(14): 1415002.
Li A H, Deng Z J, Liu X S, et al. Research progress of pose estimation based on virtual camera[J]. *Laser & Optoelectronics Progress*, 2022, 59(14): 1415002.
- [21] Zhang Z Y. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330-1334.
- [22] 李红卫, 熊韬. 基于二维码识别的P4P算法研究[J]. *激光与光电子学进展*, 2020, 57(12): 121501.
Li H W, Xiong T. Research on P4P algorithm based on quick response code recognition[J]. *Laser & Optoelectronics Progress*, 2020, 57(12): 121501.
- [23] 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.
- [24] 汪佳宝, 张世荣, 周清雅. 基于视觉EPnP加权迭代算法的三维位移实时测量[J]. *仪器仪表学报*, 2020, 41(2): 166-175.
Wang J B, Zhang S R, Zhou Q Y. Vision based real-time 3D displacement measurement using weighted iterative EPnP algorithm[J]. *Chinese Journal of Scientific Instrument*, 2020, 41(2): 166-175.
- [25] Wang L X, Cao F Y. Huber loss based non-negative matrix factorization algorithm[J]. *Computer Science*, 2020, 47(11): 80-87.
- [26] 陈紫强, 周秉毅, 刘庆华, 等. 二值加权正交迭代相机位姿估计算法[J]. *半导体光电*, 2020, 41(5): 743-748.
Chen Z Q, Zhou B Y, Liu Q H, et al. A binary weighted orthogonal iterative camera pose estimation algorithm[J]. *Semiconductor Optoelectronics*, 2020, 41(5): 743-748.
- [27] 刘晨阳, 郑龙江, 侯培国. 一种快速加权迭代位姿估计法[J]. *激光与光电子学进展*, 2022, 59(18): 1815019.
Liu C Y, Zheng L J, Hou P G. A fast weighted iterative pose estimation method[J]. *Laser & Optoelectronics Progress*, 2022, 59(18): 1815019.
- [28] Bowring B R. Transformation from spatial to geographical coordinates[J]. *Survey Review*, 1976, 23(181): 323-327.