

视觉引导机器人的手眼标定算法

韩婧祎, 高占春

(北京邮电大学计算机学院(国家示范性软件学院), 北京 100876)

摘要: 手眼标定算法是计算机视觉领域的经典问题, 也是视觉引导机器人实现的关键技术。随着机器人视觉应用技术的不断发展, 视觉引导机器人在精度、灵活度等方面的需求逐步提升, 而传统手动的手眼标定算法在标定过程中存在人工操作造成的不可避免误差, 标定流程复杂、费时且容错性差。为应对视觉引导机器人复杂的应用环境, 本文提出了一种适应性更高、更具鲁棒性的手眼标定算法, 使用双四元数进行手眼问题建模, 同时对数据集、相机位姿估计等标定的关键环节进行算法优化, 并以标定后的重投影误差作为算法评估指标。实验结果证明, 该算法在实地应用中展现出了良好的性能, 可以完全适用于复杂光场环境。算法已在机器人 3D 视觉引导自动加油项目中落地并投入使用。

关键词: 计算机视觉; 手眼标定; 双四元数; 视觉引导机器人

中图分类号: TP391

Hand-Eye Calibration Algorithm for Visually Guided Robots

HAN Jingyi, GAO Zhanchun

(School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract: Hand-eye calibration algorithm is a classic problem in the field of computer vision, and it is also the key technology in the realization of vision guided robot. With the continuous development of robot vision application technology, the demand for visual guidance robots in terms of accuracy and flexibility has gradually increased. However, the traditional manual hand-eye calibration algorithm has inevitable errors caused by manual operation in the calibration process, and the calibration process is complicated, time-consuming and has poor fault tolerance. In order to cope with the complex application environment of vision-guided robots, this paper proposes a more adaptable and robust hand-eye calibration algorithm, which uses double quaternion to model the hand-eye problem, optimizes the algorithm for key aspects of calibration such as data set and camera pose estimation, and takes the reprojection error after calibration as the algorithm evaluation index. The experimental results show that the algorithm has good performance in field application and can be fully applied to complex light field environment. The algorithm has been put into use in the robot 3D vision guided automatic refueling project.

Keywords: Computer vision; Hand-eye calibration; Double quaternion; Vision guided robot

0 引言

手眼标定问题是计算机视觉领域的一个经典问题。在人体生理结构运作中, 要实现手抓取人眼看到的东

作者简介: 韩婧祎(2000-), 女, 主要研究计算机视觉与深度学习

通信联系人: 高占春(1967-), 男, 副教授、硕导, 主要研究计算机网络及应用. E-mail: gaozc@bupt.edu.cn

关节运动使机械手处于能抓取目标物体的位姿，从而完成抓取操作。

目前工业上机械臂的手眼协作应用场景非常丰富，随着机器人视觉应用技术的不断发展，视觉引导的机器人领域在精度、灵活度等方面的需求逐步提升，手眼标定作为关键的支撑技术，用于解决机器人在视觉与运动控制的协同问题，即相机与末端的坐标转换问题。此外，传统手动的手眼标定在标定过程中存在人工操作造成的不可避免误差，标定流程复杂、费时且容错性差，需要研究准确度更高、更易操作且封装完善、标定逻辑更清晰的高效标定算法。在机器人应用的现实环境中，由于现场的光照条件、地理位置等都可能发生变化，一次标定结果的有效期较短，需要在线自适应标定来应对实操环境的变化，从而保证标定结果长期的有效性。在机器人视觉应用领域持续发展的技术市场背景下，对于机器人的精度需求不断提高，需要持续对标定算法增进标定的精度与速度，同时保证结果的稳定有效。基于上述背景，本文提出了一种适用于视觉引导机器人的手眼标定算法。

1 相关理论与工作

1.1 手眼标定问题建模

1.1.1 刚体变换的矩阵表示与群定义

手眼系统中，机械臂与相机的位姿变换都属于三维空间中的刚体变换，实际上表示的是目标坐标系相对于世界坐标系的位置变化。使用齐次变换矩阵描述的三维空间内刚体位姿，包含旋转与平移两个部分^[1]。

对于刚体的旋转，旋转矩阵是行列式为 1 的 3*3 正交矩阵，三维旋转矩阵的集合定义为 $SO(3) = \{R \in R^{3 \times 3} \mid RR^T = I, \det(R) = 1\}$ ，即特殊正交群 $SO(3)$ 。

对于刚体的平移，使用平移向量 t 来表示，因此向量 a 到向量 a' 的变换可以表达为 $\vec{a}' = R\vec{a} + \vec{t}$ ，在向量末尾增加 1，可以转化为四维齐次坐标，因此刚体变换可以使用齐次矩阵 H 表达为

$$\begin{bmatrix} a' \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ 1 \end{bmatrix} = H \begin{bmatrix} a \\ 1 \end{bmatrix} \quad (1)$$

四维齐次矩阵的集合定义为 $SE(3) = \{H = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \in R^{4 \times 4} \mid R \in SO(3), t \in R^3\}$ ，即特殊欧氏群 $SE(3)$ 。

1.1.2 手眼问题的方程表示

手眼系统包含机械臂（Hand）、视觉相机（Eye）和标定场景中通常需要用到的标定物（棋盘格等），整个三维空间涉及四个坐标系（即机械臂基座坐标系（base）、机械臂抓手坐标系（gripper）、相机坐标系（camera）和标定物世界坐标系（cal.object）），且坐标系之间存在完整的空间变换闭环。根据手眼之间的相对位置关系可以将手眼标定场景分为眼在手上和眼在手外两种^[2]：

（1）眼在手上

该场景下，相机固定在机械臂抓手处，标定物放置于相机视野内一固定点位上。如图所示。以机械臂基座到相机两个坐标系间的变换为目标描述量可以得到如下变换关系：

$${}^{\text{base}}H_{\text{gripper}} \bullet {}^{\text{gripper}}H_{\text{cam}} = {}^{\text{base}}H_{\text{cal}} \bullet {}^{\text{cal}}H_{\text{cam}} \quad (2)$$

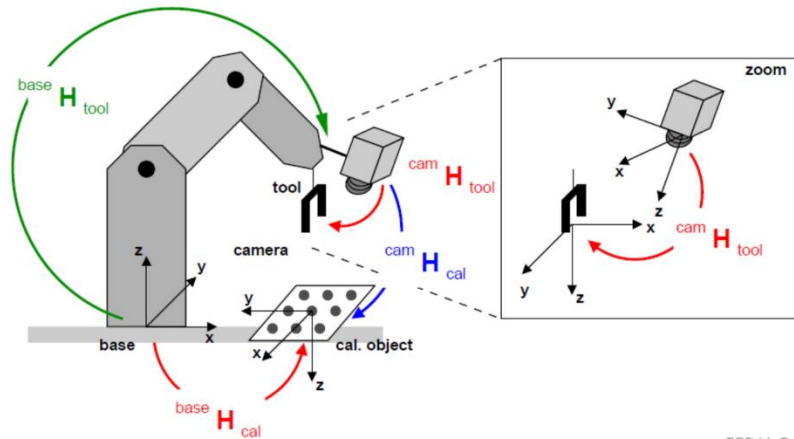
式中，以 ${}^{\text{base}}H_{\text{gripper}}$ 为例，表示机械臂基座坐标系到机械臂抓手坐标系的齐次变换矩阵。该场景下，手眼标定问题求解的量为固定量——机械臂抓手到相机坐标系的变换矩阵 ${}^{\text{gripper}}H_{\text{cam}}$ 。由于场景中标定物与机械臂基座的相对位置不变，机械臂基座到标定物世界坐标系的变换矩阵 ${}^{\text{base}}H_{\text{cal}}$ 为固定量，因此可以由两组位姿关系消掉固定量 ${}^{\text{base}}H_{\text{cal}}$ 建立方程：

$${}^{\text{base}}H_{\text{gripper}}^0 \bullet {}^{\text{gripper}}H_{\text{cam}} \bullet ({}^{\text{cal}}H_{\text{cam}}^0)^{-1} = {}^{\text{base}}H_{\text{gripper}}^1 \bullet {}^{\text{gripper}}H_{\text{cam}} \bullet ({}^{\text{cal}}H_{\text{cam}}^1)^{-1} \quad (3)$$

整理可得

$$({}^{\text{base}}H_{\text{gripper}}^1)^{-1} \bullet {}^{\text{base}}H_{\text{gripper}}^0 \bullet {}^{\text{gripper}}H_{\text{cam}} = {}^{\text{gripper}}H_{\text{cam}} \bullet ({}^{\text{cal}}H_{\text{cam}}^1)^{-1} \bullet {}^{\text{cal}}H_{\text{cam}}^0 \quad (4)$$

公式(4)形如 $A \cdot X = X \cdot B$ ，其中 A 为机械臂位姿， B 为相机位姿， X 为眼在手上问题的求解量。



注：图来自网络

图1 视觉引导的机器人手眼系统（眼在手上）

Fig. 1 Visually guided robotic hand-eye system(Eye-in-hand)

(2) 眼在手外

该场景下，标定物固定在机械臂抓手处，相机放置于空间内视野可涵盖标定物移动的一固定点位上。以机械臂基座到标定物两个坐标系间的变换为目标描述量可以得到如下变换关系：

$${}^{\text{base}}H_{\text{gripper}} \bullet {}^{\text{gripper}}H_{\text{cal}} = {}^{\text{base}}H_{\text{cam}} \bullet {}^{\text{cam}}H_{\text{cal}} \quad (5)$$

该场景下，手眼标定问题求解的量为固定量——机械臂基座到相机坐标系的变换矩阵 ${}^{\text{base}}H_{\text{cam}}$ 。由于场景中标定物与机械臂抓手的相对位置不变，同理可以通过两组位姿关系联立方程并整理得到方程：

$${}^{\text{base}}H_{\text{gripper}}^1 \bullet ({}^{\text{base}}H_{\text{gripper}}^0)^{-1} \bullet {}^{\text{gripper}}H_{\text{cal}} = {}^{\text{base}}H_{\text{cam}} \bullet {}^{\text{cam}}H_{\text{cal}}^1 \bullet ({}^{\text{cam}}H_{\text{cal}}^0)^{-1} \quad (6)$$

符合 $A \cdot X = X \cdot B$ ，其中 A 为机械臂位姿， B 为相机位姿， X 为眼在手外问题的求解量。

因此手眼标定问题可以转为矩阵方程 $A \cdot X = X \cdot B$ 中手眼矩阵 X 的求解, 其中 A 为机械臂位姿, B 为相机位姿。

1.2 手眼方程求解

传统手眼标定算法的实现主要基于空间几何关系建模、手眼方程求解与优化算法。最常用的数据采集方法为基于手眼标定板(如棋盘格、标定块等), 通过在不同的位置和机械手姿态下拍摄标定板图像, 得到标定数据集, 并以标定板真实尺寸与像素长度对齐标定结果的现实绝对尺度。

对于建模后的手眼标定方程 $A \cdot X = X \cdot B$ 、 $A \cdot X = Y \cdot B$, 从平移和旋转矩阵的计算顺序上可以分为两类: 一种采用分离的方法, 先计算旋转并利用第一步求解的旋转计算平移; 另一种算法则是同时计算旋转和平移; 从求解的方法可以分为解析解和数值解两大类。较为经典的算法代表有: TASI 两步法^[3], 引入旋转轴-角系统基于旋转对数表示旋转运动, 重写手眼方程并使用最小均方法求解方程; 使用四元数表示, 推导 X 与 Y 的显式线性解, 并使用线性最小二乘法求解平移分量^[4], 该方法在推导过程中涉及四元数代数知识, 求解方法为简单的线性最小二乘, 具有普适性; 使用 Kronecker 乘积和奇异值分解求得封闭形式的解

^[5]; Daniilidis 提出基于双四元数的方法^[6], Andreff 等人基于 Sylvester 方程的线性求解方法^[7], 以及 Condurache 和 Burlacu 基于双张量的方法^[8]; 使用拉格朗日乘子约束单位四元数最小化求解旋转的封闭解, 再用线性方法求解平移, 使用 Levenberg-Marquardt 方法和置信域方法同时求解旋转与平移^[9]; Gwak 等人的梯度/牛顿法^[10], 以及 Zhang 和 Zhang^[11]等人基于伪逆的方法。除此之外, Nguyen 和 Pham^[12]首次提出通用迭代解决方案实现对手眼问题 $AX = XB$ 中 X 的不确定性描述。两种求解方式各有利弊, 分步求解的明显缺点在于几何参数是在两个阶段分别估计的, 因此第一阶段产生的估计误差会传播至第二阶段, 导致误差遗留不能产生精度足够好的结果; 同时求解的方法在优化方面可以展现更好的性能, 但在数值迭代的过程中或耗费大量时间, 计算效率不够理想。近些年传统手眼标定算法在经典方法的基础上进行改进, 影响力较大的有: 求解特殊欧几里得群 $SE(3)$ 中的通用解析解, 依赖于四维空间的普氏分析, 提出了单位八元数表示得到一个最优的闭式特征分解解^[13];

GD- $SE(3)$ 方法^[14], 将手眼问题表述为点集匹配问题, 并提出了使用特殊欧几里得群 $SE(3)$ 上的梯度下降 (GD) 技术求解。结果表明该方法精度与基于其他点集匹配算法的精度相当; 对于形如 $A \cdot X = X \cdot B$ 、 $A \cdot X = Y \cdot B$ 的手眼标定问题, 使用符号方法导出全局最优解, 相比基于优化的求解方法, 结果更准确且 100% 可重复^[15]。传统基于标定板和最小二乘、非线性优化、重投影误差最小化等优化方法相结合的手眼标定算法已经被广泛研究与应用, 并已具有一定的准确性和鲁棒性。

此外, Kenji 和 Emanuele 提供了一种基于重投影误差最小化的位姿图优化的思路^[16], 这种方法直接使用机械手位姿和标定板图像, 通过优化手眼变换和图像位姿, 使得图像的重投影误差最小。重投影误差最小化技术在运动结构^[17]和物体姿态估计^[18]等工作上展示出了优越的准确性和鲁棒性, 但只有少数研究将该方法应用在手眼标定问题上^[19], 很多实现证

实了该方法能够利用图结构的稀疏性高效、鲁棒地解决大型非线性优化问题。

2 问题建模与算法流程

2.1 算法问题建模

手眼标定(Hand-eye Calibration)目的是求解机器人手部(Hand)和相机(Eye)之间的空间坐标系转换关系,在机器人视觉应用中至关重要要求解手眼关系并应用到机器人视觉任务中,以实现视觉与运动的协同。通常包含以下流程:(1)确定手眼系统的标定模型:常用的模型有 $A \cdot X = X \cdot B$ 、 $A \cdot X = Y \cdot B$, 其中 X 表示手坐标系和眼坐标系之间的转换关系;(2)采集标定数据:采集不同机械臂位姿数据与对应相机位姿下的图像数据,需要保证每组图像与位姿数据的时间同步性,并且手眼位姿要尽可能覆盖整个工作空间,组与组之间的位姿变换不能太接近。采集的数据量要足够多;(3)构建标定算子方程:根据采集的数据,构建手眼关系的线性方程组或非线性优化问题;(4)求解标定关系:使用闭式解法或迭代算法来求解方程组,获得手眼变换矩阵;(5)优化标定关系:采用全局优化、最小二乘等的方式进一步提高标定估计的精度;(6)验证标定结果:通过目标检测、定位或转为二维图像重投影来验证标定结果的准确性与精度。传统的手眼标定方法通常基于几何模型和解析方法,需要精确测量机器人臂和相机之间的关键点,并使用数学求解技术进行标定。然而,这些方法对于复杂的机器人系统和场景可能存在一些挑战,如非刚性变形、噪声、遮挡等。

2.1.1 四元数的鲁棒性优势

根据 1.1 节对手眼标定问题的数学建模可知,手眼标定问题可以根据三维空间上多个坐标系相对于世界坐标系(基坐标系)的刚体运动描述为 $SE(3)$ 群上的矩阵方程 $A \cdot X = X \cdot B$ 求解。对于求解手眼矩阵 X , 包含旋转与平移两部分。基于对现有算法的学习和研究,同时求解旋转和平移两个分量可以避免误差遗留问题,在鲁棒性表现上更优。

使用旋转矩阵表示旋转时,矩阵有 9 个未知量,并且存在矩阵奇异性的问题,数值不稳定;而使用动态欧拉角表示旋转时存在万向锁问题(Gimbal lock), 即当选择 $\pm 90^\circ$ 作为 Pitch 角时,会导致第一次旋转与第三次旋转等价,整个旋转表示系统被限制在只能绕竖直轴旋转,丢失了一个表示维度。相比之下,使用李群 $SE(3)$ 与四元数表示,可以实现同时求解旋转与平移,并且有效地避免了旋转矩阵奇异性问题,同时也可以规避万向锁问题。而相比李群,四元数的几何意义表达更直观,且数值计算更简单。使用四元数表示旋转,内存和运算速度更优。内存上,一个四元数值占四个浮点数,四元数相乘时可以直接在四个浮点数上进行加减乘的基本运算,相比使用旋转向量转换旋转矩阵相乘,再转回旋转向量计算次数更少,未知数个数更少,并且不涉及 Rodrigues 变换中的三角函数等高级运算,避免了矩阵运算的复杂度。由于浮点数运算中误差是不可避免的,运算越多,误差积累更多,因此理论上使用四元数相乘在精度上也更有优势。此外使用四元数可以实现平滑插值,对 Roll-Pitch-Yaw 欧拉角进行插值可能会导致结果又跳跃而不平滑,造成插值过程中的误差,而四元数更易于插值以获得平滑解,有更广泛的适用性。

2.1.2 基于双四元数的手眼问题建模

四元数的定义为 $q = a + b\vec{i} + c\vec{j} + d\vec{k}, \vec{i}^2 = \vec{j}^2 = \vec{k}^2 = \vec{i}\vec{j}\vec{k} = -1$, 结构类似于虚数,包含

标量和向量两部分，因此四元数可以表示为标量-向量有序对的形式，即 $q = [s, \vec{v}]$, $\vec{v} = (x, y, z)^T$ 。这种形式与轴-角表示旋转的形式相似，刚体运动变换矩阵中旋转部分基于轴-角变换表示为 $R = (\hat{k} \sin \frac{\theta}{2}, \cos \frac{\theta}{2})$ ，其中 \hat{k} 表示旋转轴， θ 表示绕轴旋转的角度，类似于轴角表达，使用四元数的标量向量有序对的形式可以将旋转表示为 $q = [\sin \frac{\theta}{2}, \cos \frac{\theta}{2} \vec{u}]$, $\|q\| = 1$ ，其中 \vec{u} 表示旋转轴， θ 表示绕轴旋转的角度，几何意义为以向量 \vec{u} 为轴旋转 θ 角。

单个四元数可以表示旋转矩阵的 $SO(3)$ 群，为了统一旋转与平移的表示方式，实现旋转和平移的同时求解，引入双四元数的概念，使用双四元数代替平移向量表示平移。双四元数 Q 的定义为 $Q = q + \varepsilon q'$, $\varepsilon^2 = 0$ ，其中 q 、 q' 为四元数， ε 为对偶因子，四元数 q 称为双四元数的实部， q' 为对偶部。当使用双四元数表示旋转时，只需令对偶部为 0，退化为纯四元数；对于平移向量 $\vec{t} = (t_x, t_y, t_z)^T$ ，只需令双四元数实部为单位 1，平移可以使用双四元数表示为 $Q_t = [1, \vec{0}] + \varepsilon [0, \frac{1}{2} \vec{t}]$ 。因此，刚体变换 (R, T) 可以使用双四元数表示为

$$Q = T \otimes R = R + \varepsilon [0, \frac{1}{2} \vec{t}] \otimes R \equiv q + \varepsilon q' \quad (7)$$

双四元数可以实现同时表示旋转与平移，在描述功能上与 $SE(3)$ 群等价。使用双四元数表示的 $SE(3)$ 群，群元素由单位双四元数 Q 组成， $Q = q + \varepsilon q'$ 且满足单位双四元数条件 $Q \otimes Q^* = 1$ ，其中 q 、 q' 均为单位四元数，满足约束 $\|q\| = 1$ ，且 $q \cdot q' = 0$ 。

手眼标定求解方程 $A \cdot X = X \cdot B$ 可以表示为 $\Delta C_i X = X \Delta H_i$, $i = 1, 2, \dots, N$ ，其中 ΔC_i 表示相机位姿的刚体变换， ΔH_i 表示机械手的相对位姿变换。使用双四元数表示，方程可以改写为 $Q_{C_i} \otimes Q_X = Q_X \otimes Q_{H_i}$, $i = 1, 2, \dots, N$ ，其中 $Q = q + \varepsilon q'$ ，因此方程可等价表示为方程组

$$\begin{aligned} q_{C_i} \otimes q_X &= q_X \otimes q_{H_i}, \quad i = 1, 2, \dots, N, \\ q_{C_i} \otimes q'_X + q'_{C_i} \otimes q_X &= q_X \otimes q'_{H_i} + q'_X \otimes q'_{H_i} \end{aligned} \quad (8)$$

四元数的乘法可以转化为矩阵形式，对于一个四元数 $q = a + b\vec{i} + c\vec{j} + d\vec{k}$ ，定义左乘矩阵 $L(q)$ 与右乘矩阵 $R(q)$:

$$L(q) = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}, \quad R(q) = \begin{bmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{bmatrix}$$

两个四元数相乘等价于四维空间向量左乘 L 矩阵、右乘 R 矩阵，即

$$q \otimes p \equiv L(q)p \equiv R(p)q \quad (9)$$

将 (8) (9) 两组公式联立可以得到

200

$$(L(q_{C_i}) - R(q_{H_i})q_X = 0, \quad i=1,2,\dots,N$$

$$(L(q'_{C_i}) - R(q'_{H_i}))q_X + (L(q_{C_i}) - R(q_{H_i}))q'_X = 0 \quad (10)$$

为了简化符号, 定义 $A_i \equiv L(q_{C_i}) - R(q_{H_i})$, $B_i \equiv L(q'_{C_i}) - R(q'_{H_i})$, 其中, L 表示四元数的左乘矩阵, R 表示四元数的右乘矩阵, 公式 10) 最终可以简写作

$$A_i q = 0, \quad i=1,2,\dots,N$$

205

$$B_i q + A_i q' = 0, \quad i=1,2,\dots,N \quad (11)$$

基于公式 (11) 双四元数表示的手眼标定问题模型, 可以建立最小化问题:

$$\arg \min_{q, q'} \sum_{i=1}^N (|A_i q|^2 + \alpha^2 |B_i q + A_i q'|^2) \quad (12)$$

210

需满足约束 $\|q\|=1$, 且 $q \cdot q' = 0$ 。其中引入表示固定维数的参数 $\alpha \in \mathbf{R}$, 由于旋转和平移具有不同的单位, α 可描述为权重因子, 单位为长度的倒数, 从而统一平移和旋转的单位, 在计算成本函数时认为 α 为固定可调参数。

由于二次成本函数受非线性约束, 一个经典的解决方法^[20]就是使用两个拉格朗日乘子 λ 、 μ 对应于两个约束来表示成本函数 L :

$$L = \sum_i (|A_i q|^2 + \alpha^2 |B_i q + A_i q'|^2) + \lambda(1 - q^2) - 2\mu q \cdot q' \quad (13)$$

对于成本函数的求解是在拉格朗日乘子空间中找到一个特殊点, 使得最优解对应于最小

215

λ 曲线的鞍点满足 $\frac{d\lambda}{d\theta} = 0$ 。对于使用拉格朗日乘子实现优化算法, Amit Deke 等人^[21]提出了最小化问题的几种求解方法: 最优一维线性搜索 (DQOpt)、凸松弛 (DQCovRlx)、二阶近似求解 (DQ2ndOrd)、迭代求解 (DQItr) 等。

2.2 算法流程

220

手眼标定算法的流程如图 2 所示。首先数据采集与预处理, 拍摄一定数量的标定板多视角图像, 并记录拍摄位置下的机械手位姿数据。由于相机成像与使用损耗等问题, 相机参数可能与出厂参数存在微小差异, 需要对采集后的图像使用张氏标定法^[22]进行相机标定, 得到相机的内部参数与畸变系数。同时, 在场景光照条件不同的情况下, 机械臂大幅摆动调整拍摄角度可能导致采集的图像出现亮度不一致、过曝、边缘明显畸变等问题, 因此对于采集到的二维图像需要进行一步去畸变处理, 保留矫正后的图像。随后对去畸变后的图像进行标定板角点检测, 按固定序列保存每张图像中角点的二维坐标点集, 并以标定板平面为世界坐标系的 xoy 平面, 通常以标定板左上角为坐标系原点, 根据标定板尺寸, 得到具有绝对尺度且单位统一的角点三维空间坐标。根据同一个角点在二维成像面与三维空间中的坐标对应关系, 通过透射投影模型估计出相机位姿, 即相机坐标系相对于世界坐标系的刚体运动关系。最后与机械臂输出的 6D 位姿数据 (多数机械臂的旋转表达为动态欧拉角) 代入手眼标定问题方程求解手眼矩阵。

230

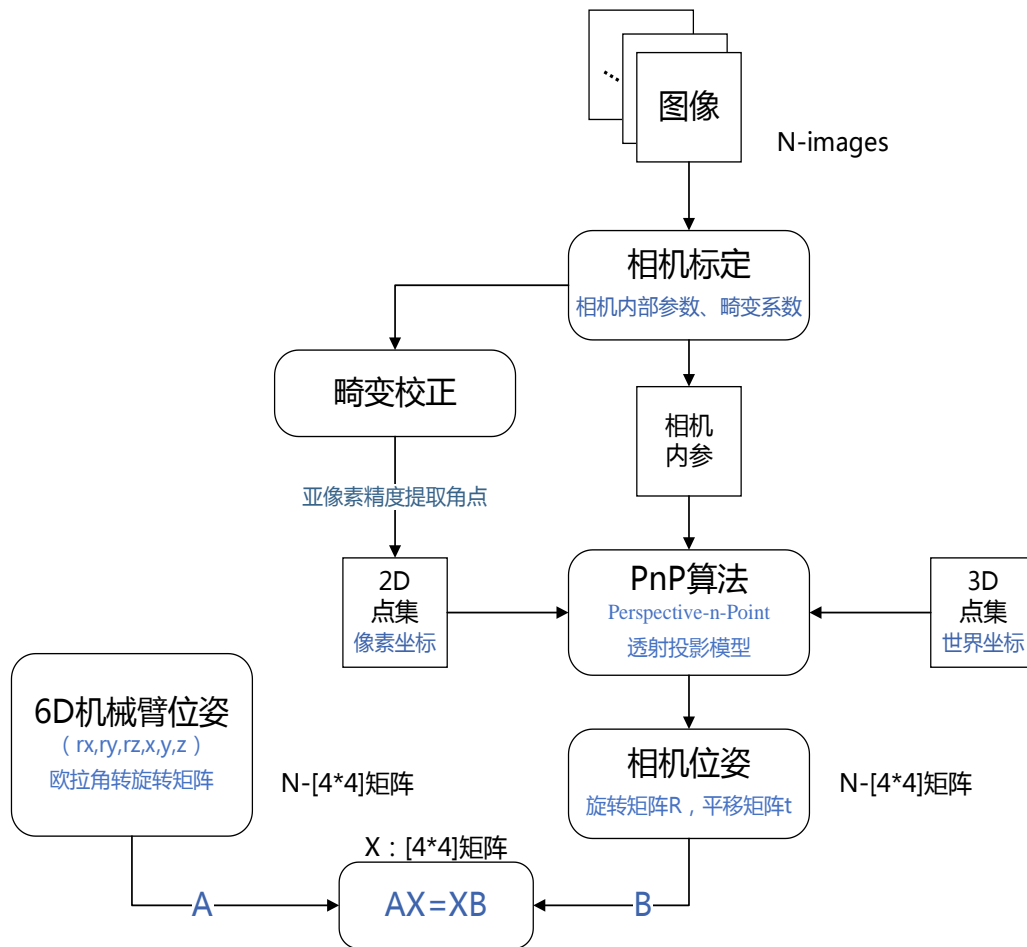


图 2 传统手眼标定算法流程

Fig. 2 Traditional hand-eye calibration algorithm flow

本课题基于企业合作项目，手眼标定的应用场景主要为室外环境。由于自然光照条件下，光强受时间、环境等因素影响较大，且采集标定数据需要机械臂尽可能大幅度移动来获取多视角标定板图像，固定曝光参数、帧频等相机工作参数后，同组采集的标定数据图像可能出现平面边缘畸变、中心虚焦、过曝光、标定板格线边界模糊等问题。图像质量直接影响后续角点提取精度，因此在图像采集环节，应注意：（1）调整拍摄距离，使标定板位于成像最佳的景深距离范围内，避免图像中标定板出现虚焦模糊的情况；（2）选择合适尺寸大小的标定板，根据实际拍摄环境使用的机械手臂展范围和相机成像景深范围，选择大小适宜的标定板，使得标定板在多视角拍摄的情况下，尽可能位于相机视野中心，占图像比例过半，避免标定板出现在图像的边角或部分出画；（3）适当干预光强，对于光照过强或过弱的情况，应对拍摄环境进行适当的遮光或补光处理，尽可能保证相机工作参数与环境情况适配；（4）实际采集的过程中实时关注图像成像细节，及时调整相机工作参数与标定物摆放位置，对于质量不佳的图像应舍弃剔除，不作为后续标定数据。

得到采集的标定图像后，对标定数据集做预处理。目前部分光结构工业相机会集成图像去畸变组件，可以实现对拍摄的图像根据相机出厂内在参数进行畸变矫正，输出矫正后的图像。但相机实际使用的真实内在参数与出厂参数会存在差异，且差异程度无法主观估计，因此对于采集的标定图像首先使用张氏标定算法完成对相机内在参数与畸变系数的估计，之后使用 OpenCV 提供的方法函数 `cv::initUndistortRectifyMap()` 和 `cv::remap()` 对标定图片进行图

像畸变矫正。首次利用相机参数计算矫正映射矩阵，再根据矫正映射函数定位畸变图像中每一个像素在矫正后图像中的对应位置，使用线性插值（INTER_LINEAR）算法对非整数像素点进行线性平滑插值，得到矫正后的标定图像。

完成数据集预处理后，对矫正后的图像进行角点检测，主要使用的检测算法为 FAST（Features from Accelerated Segment Test）算法、MSER（Maximally Stable Extremal Regions）算法，实现对角点的快速稳定检测，并细化至亚像素精度，得到角点的 2D 像素坐标。世界坐标系原点定义在标定板上，并且选定板平面 z 值为 0，很容易得到具有绝对尺度的角点 3D 空间坐标。已知每张标定图像中所有角点的像平面坐标与空间坐标，结合相机的内在参数，可以实现通过针孔相机的透射投影模型估计出相机位姿，即 R 、 t 。传统手眼标定方法中一般使用 PnP（Perspective-n-Point）算法^[23]计算相机位姿。算法通过最小化重投影误差来优化相机的位姿，使得在三维空间中的三维点经过相机投影后与对应的二维图像特征点的投影位置尽可能接近。根据已知的三维点和对应的二维特征点，通过透视变换的原理，估计相机的透视投影矩阵。利用透视投影矩阵，通过主轴的估计来求解相机的旋转矩阵和平移向量，得到相机的位姿估计。

由于课题中所有三维点均在同一个平面上，因此使用解析求解的方法，基于无限小平面的姿态估计 IPPE(Infinitesimal Plane-Based Pose Estimation)算法^[24]，通过数学推导直接求解相机的位姿，不需要迭代优化的过程。但使用 PnP 算法估计相机位姿时，若点对出现误匹配或初始化偏差，会导致结果不稳定，可能求解得到局部最优值。因此在使用 IPPE 方法时，为提高方法鲁棒性，课题结合使用了随机抽样一致性算法 RANSAC(Random Sample Consensus)^[25]进行采样和估计，以剔除误匹配的特征点对，通过随机采样，从匹配的特征点中随机选择一个子集构建假设模型，IPPE 方法估计出位姿后进行重投影误差计算，根据预定义的阈值，判断特征点对是否满足当前的假设模型，重复一定次数，最终选择具有最大内点数量的假设模型作为结果。

3 现场实验与结果分析

本文使用双四元数对手眼系统的刚体变换关系建模，并在数据集采集与相机位姿估计环节做鲁棒性优化，最终实现 IPPE 方法与 RANSAC 算法结合的相机位姿估计，并在拉格朗日乘子空间求解手眼矩阵方程的手眼标定算法，根据上述技术路线开展实验。

实验使用到了 3D 光结构相机与 DKamViewer d130 相机两款工业 RGB-D 摄像机，相机信息如表 1 所示。实验分别在室内实验室与室外现场两个场景进行数据采集，并且对眼在手上、眼在手外两种手眼系统结构进行算法验证。由于两个实验环境中使用的机械臂与视觉相机型号均不同，因此从标定的手眼矩阵结果数值上无法直接判断算法的精度与鲁棒性，需要利用二维图像的重投影像素误差或点云数据匹配的三维空间距离误差评估标定结果质量。实验中根据实验环境、机械手臂展、相机视野宽度与景深的不同，选用了三种尺寸的棋盘格标定板，如表 2 所示。

表 1 实验相机信息

Tab. 1 Camera Information

相机名称	红外摄像头	RGB 摄像头	畸变矫正	点云坐标系基准
TripleNano3D 光结构	1920*1200	1920*1080	无	与红外相机一致
DKamViewer d130s	1280*1024	1296*972	有	

注：单位为像素（pixel）

表 2 棋盘格标定板尺寸
Tab. 2Checkerboard calibration board size

编号	使用场景	角点数	格子尺寸（单位：mm）
1	室内	8*9	18
2	室内	26*27	10
3	室外	8*11	20

实验使用重投影误差作为标定结果评估指标。重投影误差（Reprojection Error）指真实三维空间点在图像平面上的投影（即初始图像上的像素点）和利用估计出的相机外参[R,t]重新投影到像平面上的虚拟的像素点之间的差值。重投影误差是一个表示投影点与对应目标二维点之间距离的度量。误差值越小表示相机标定结果与实际数据越接近。

根据公式（4）可以整理得到，

$${}^{cal}H_{cam}^1 = {}^{cal}H_{cam}^0 \bullet \left({}^{gripper}H_{cam}\right)^{-1} \bullet \left({}^{base}H_{gripper}^0\right)^{-1} \bullet {}^{base}H_{gripper}^1 \bullet {}^{gripper}H_{cam}$$

(14)

通过手眼矩阵并不能直接估计出某一图像对应的相机位姿，需要一组位姿数据作为基和目标图像对应的机械臂位姿作为已知量，与手眼矩阵代入公式（14）中，估计出目标图像的相机位姿。估计出的相机位姿与三维空间的点坐标通过针孔相机模型的投影矩阵，重投影到像平面中，得到虚拟投影点的像素坐标(u',v')。投影点与原始图像中目标点的像素坐标(u,v)计算 MSE 误差（即均方误差），得到以像素为单位的重投影误差，从而实现以数值结果评估手眼矩阵的精度。

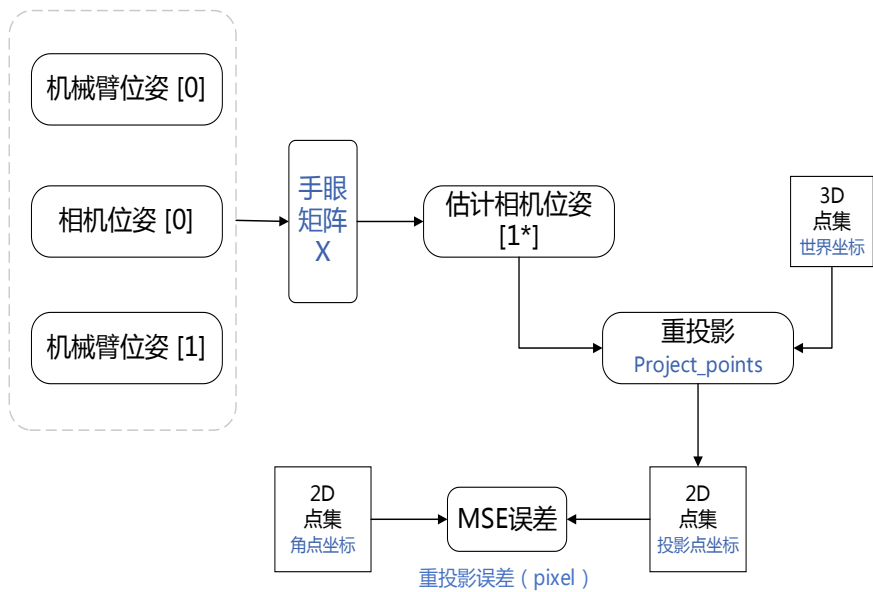


图 3 手眼矩阵参与计算重投影误差

Fig. 3Reprojection error involving hand-eye matrix

实验使用红外相机与 RGB 相机分别在室外与室内进行眼在手上、眼在手外两种手眼系统的手眼标定，并对比了使用未优化的传统算法与优化后算法的标定结果精度，实验结果如表 3 所示。

表 3 实验结果-重投影误差（单位：像素）
Tab. 3Experimental Results-Reprojection Error (Unit:Pixel)

编号	相机类型	拍摄场景	是否使用优	手眼系统	标定图片	平均单张重投
----	------	------	-------	------	------	--------

			化算法	类型	数量	影误差
1	RGB1	室内	已优化	眼在手外	28	1.15785
2			未优化			3.03107
3		室外	已优化		35	2.07408
4			未优化			10.8093
5	RGB2	室内	已优化	眼在手上	20	2.15249
6			未优化			7.0549
7		室外	已优化		35	3.16368
8			未优化			12.3996
9	红外	室外	已优化		40	1.64239
10					35	1.92869
11					25	2.58899
12					20	3.03671

注：除表中所列出的变量外，标定中使用的棋盘格尺寸也不同

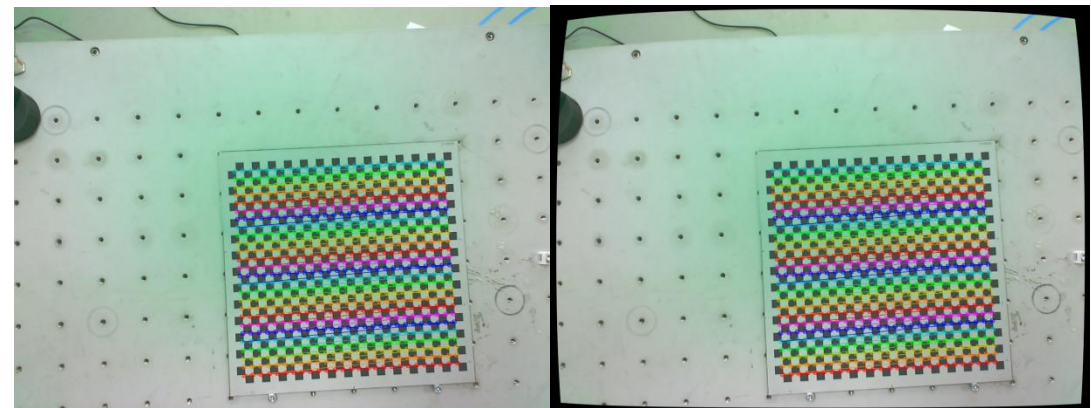
根据表中的实验结果数据可知：

无论是在眼在手上还是眼在手外的手眼系统中，使用优化后的手眼标定算法单张重投影误差均明显小于未优化的传统算法结果；

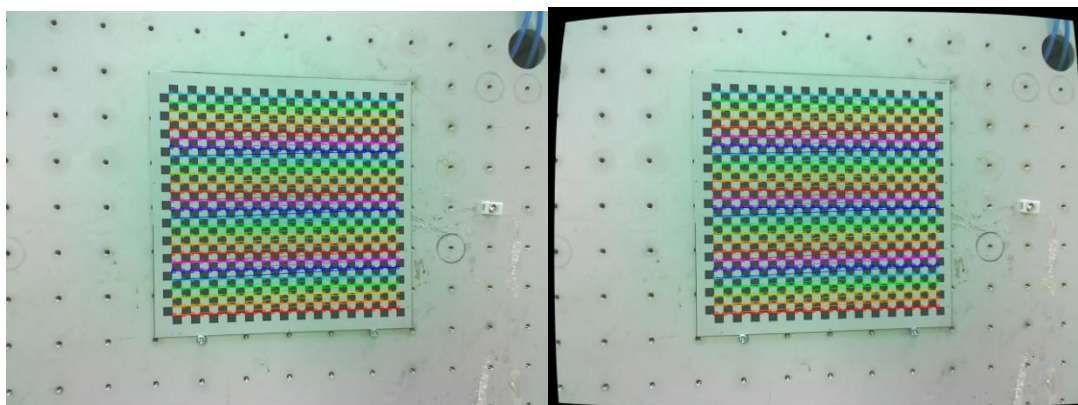
1) 对比 1-4 组、5-8 组数据，无论对于眼在手上还是眼在手外的手眼系统，使用 RGB 相机采集标定图片数据，室内稳定光照条件下的标定结果精度都更高，并且算法的优化效果在室外环境中更显著；

2) 对比 9-11 组数据，使用同一红外相机在室外环境下拍摄多组标定数据集，标定图片数量越多，标定结果精度越高，但精度结果与图片数量增势不成正比，标定图片数量以 30 张以上为宜。

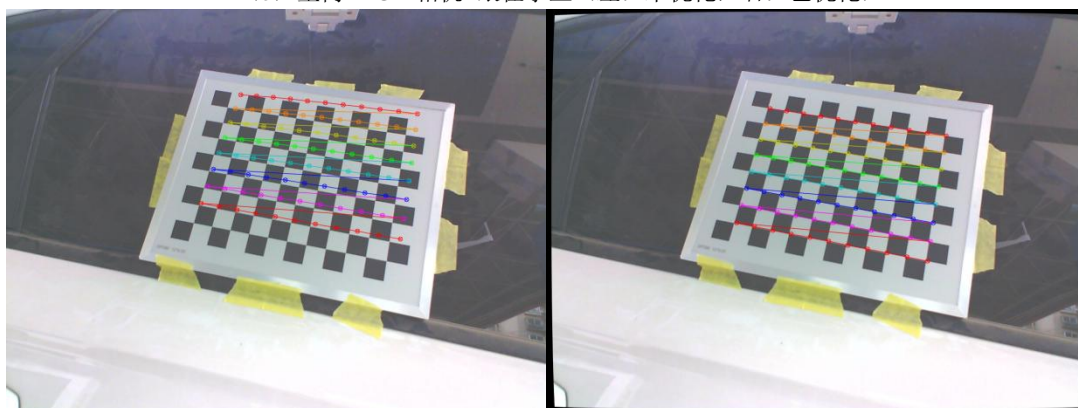
3) 对比第 7 组与第 10 组数据，在室外光照条件下，使用红外相机标定的结果精度优于使用 RGB 相机，对于复杂光场的应用场景，红外相机更适合用于做手眼标定。



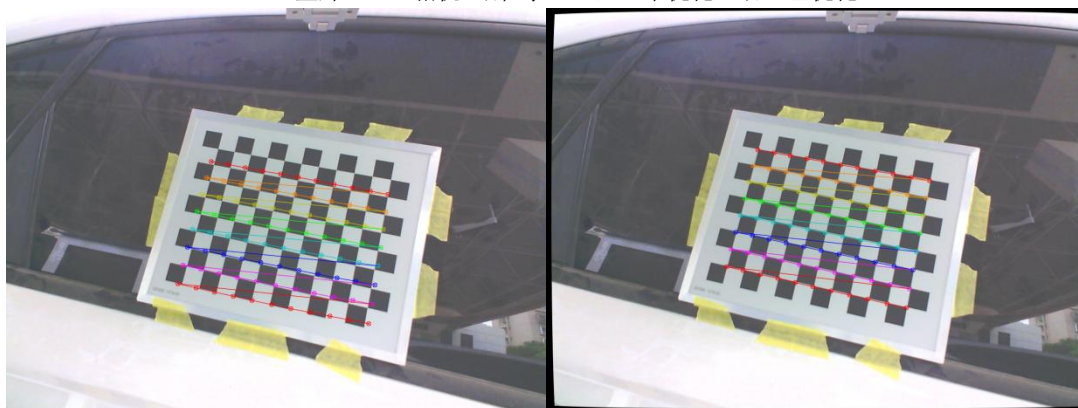
(a) 室内+RGB 相机+眼在手上（左：未优化，右：已优化）



(b) 室内+RGB 相机+眼在手上 (左: 未优化, 右: 已优化)



(c) 室外+RGB 相机+眼在手上 (左: 未优化, 右: 已优化)



(d) 室外+RGB 相机+眼在手上 (左: 未优化, 右: 已优化)

图 4 真实实验数据

Fig. 4 Real experimental data

图 4 展示了部分实验数据, 图中绘制的均为重投影后的角点位置, 可以看到使用优化后的算法的重投影点与角点基本重合, 精确度明显优于传统算法。

4 结论

本文给出了一种用于视觉引导机器人的手眼标定算法, 介绍了基于双四元数建模的传统手眼标定算法的鲁棒性优化方案, 最终通过对数据集采集质量与预处理、相机位姿估计算法使用 IPPE 算法与 RANSAC 算法结合, 实现在手眼标定的关键环节上的鲁棒性提升。同时使用拉格朗日乘子求解手眼矩阵代价函数最小化问题, 确保结果的精确度与抗噪能力。使用

重投影误差作为标定结果评估指标,实验结果表明本文提出的手眼标定算法在室内稳定光照与室外自然光照的条件下表现均优于传统算法,并且在室外复杂光场环境下仍有精确且稳定的标定结果。本算法已在室外加油站实现视觉引导机器人自动加油项目投入使用。

致谢

感谢中石油集团与中科行智(苏州)为本论文提供实验条件支持,并将本文所述算法投入实地使用。

[参考文献] (References)

- [1] 高翔,张涛. 视觉 SLAM 十四讲[M]. 北京: 电子工业出版社, 2017.
- [2] Horaud R, Dornaika F. Hand-eye calibration[J]. The international journal of robotics research, 1995, 14(3): 195-210.
- [3] Tsai R Y, Lenz R K. A new technique for fully autonomous and efficient 3 d robotics hand/eye calibration[J]. IEEE Transactions on robotics and automation, 1989, 5(3): 345-358.
- [4] Zhuang H, Roth Z S, Sudhakar R. Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form $AX = YB$ [J]. IEEE Transactions on Robotics and Automation, 1994, 10(4): 549-554.
- [5] Shah M. Solving the robot-world/hand-eye calibration problem using the Kronecker product[J]. Journal of Mechanisms and Robotics, 2013, 5(3): 031007.
- [6] Daniilidis K. Hand-eye calibration using dual quaternions[J]. The International Journal of Robotics Research, 1999, 18(3): 286-298.
- [7] Andreff N, Horaud R, Espiau B. On-line hand-eye calibration[C]//Second International Conference on 3-D Digital Imaging and Modeling (Cat. No. PR00062). IEEE, 1999: 430-436.
- [8] Condurache D, Burlacu A. Orthogonal dual tensor method for solving the $AX = XB$ sensor calibration problem[J]. Mechanism and Machine Theory, 2016, 104: 382-404.
- [9] Horaud R, Dornaika F. Hand-eye calibration[J]. The international journal of robotics research, 1995, 14(3): 195-210.
- [10] Scherer C, Weiland S. Linear matrix inequalities in control[J]. Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands, 2000, 3(2).
- [11] Zhang Z, Zhang L, Yang G Z. A computationally efficient method for hand-eye calibration[J]. International journal of computer assisted radiology and surgery, 2017, 12: 1775-1787.
- [12] Nguyen H, Pham Q C. On the Covariance of \mathbf{X} in $\mathbf{A}\mathbf{X} = \mathbf{B}\mathbf{X}$ [J]. IEEE Transactions on Robotics, 2018, 34(6): 1651-1658.
- [13] Wu J, Sun Y, Wang M, et al. Hand-eye calibration: 4-D procrustes analysis approach[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 69(6): 2966-2981.
- [14] Qiu S, Wang M, Kermani M R. A new formulation for hand-eye calibrations as point-set matching[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(9): 6490-6498.
- [15] Wu J, Liu M, Zhu Y, et al. Globally optimal symbolic hand-eye calibration[J]. IEEE/ASME Transactions on Mechatronics, 2020, 26(3): 1369-1379.
- [16] Edlund O. A software package for sparse orthogonal factorization and updating[J]. ACM Transactions on Mathematical Software (TOMS), 2002, 28(4): 448-482.
- [17] Koide K, Menegatti E. General hand-eye calibration based on reprojection error minimization[J]. IEEE Robotics and Automation Letters, 2019, 4(2): 1021-1028.
- [18] Munaro M, Antonello M, Moro M, et al. Fibremap: Automatic mapping of fibre orientation for draping of carbon fibre parts[C]//IAS-13 Workshop on ROS-Industrial in European Research Projects, Padova, Italy. 2014: 272-275.
- [19] Hesch J A, Roumeliotis S I. A direct least-squares (DLS) method for PnP[C]//2011 International Conference on Computer Vision. IEEE, 2011: 383-390.
- [20] Esquivel S. Eye-to-Eye Calibration: Extrinsic Calibration of Multi-Camera Systems Using Hand-Eye Calibration Methods[D]. , 2015.
- [21] Dekel A, Harenstam-Nielsen L, Caccamo S. Optimal least-squares solution to the hand-eye calibration problem[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 13598-13606.
- [22] Zhang Z. Flexible camera calibration by viewing a plane from unknown orientations[C]//Proceedings of the seventh IEEE international conference on computer vision. Ieee, 1999, 1: 666-673.
- [23] Marchand E, Uchiyama H, Spindler F. Pose estimation for augmented reality: a hands-on survey[J]. IEEE transactions on visualization and computer graphics, 2015, 22(12): 2633-2651.
- [24] Collins T, Bartoli A. Infinitesimal plane-based pose estimation[J]. International journal of computer vision, 2014, 109(3): 252-286.

[25] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6): 381-395.