



仪器仪表学报
Chinese Journal of Scientific Instrument
ISSN 0254-3087, CN 11-2179/TH

《仪器仪表学报》网络首发论文

题目：一种鲁棒且线性的 PnP 问题求解方法
作者：王平，周雪峰，安爱民，何倩玉，张爱华
网络首发日期：2020-10-15
引用格式：王平，周雪峰，安爱民，何倩玉，张爱华. 一种鲁棒且线性的 PnP 问题求解方法. 仪器仪表学报.
<https://kns.cnki.net/kcms/detail/11.2179.th.20201015.1107.008.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

一种鲁棒且线性的 PnP 问题求解方法

王平¹, 周雪峰², 安爱民¹, 何倩玉¹, 张爱华¹

(1. 兰州理工大学电气工程与信息工程学院 兰州 730050; 2. 广东省智能制造研究所 广州 510070)

摘要：传统DLT (Direct Liner Transformation) 方法利用3D点的投影成像模型求解PnP (Perspective-n-Point) 问题，具有效率高的特点。但由于其在求解过程中缺少附加的非线性约束条件，因此求解精度和稳定性不高。为提升DLT方法的精度和稳定性，本文推导出了一种全新的点与点之间的投影成像模型，并以此模型为基础结合传统的3D点投影成像模型，提出了具有鲁棒性的直接线性求解 (Robust Direct Linear Transformation, RDLT) PnP问题方法。仿真实验和实际实验表明：RDLT方法作为PnP问题的一种线性解，由于利用了点与点之间的约束条件，因此其求解精度和稳定性比得上现有非线性最优化的方法，而计算效率则远高于这些方法。

关键词：计算机视觉；PnP问题；位姿测量；线性求解

中图分类号：TP391 TH701 文献标识码：A 国家标准学科分类代码：460.4010

A robust and linear solving method for Perspective-n-Point problem

Wang Ping¹, Zhou Xuefeng², An Aimin¹, He Qianyu¹, Zhang Aihua¹

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China; 2. Guangdong Institute of Intelligent Manufacturing, Guangdong 510070, China)

Abstract: Traditional direct liner transformation (DLT) method uses the projective imaging model of 3D points to solve the PnP (Perspective-n-Point) problem, which has high efficiency characteristic. However, the DLT method has low accuracy and stability because of lacking of additional nonlinear constraint conditions in the solving process. To improve the accuracy and stability of DLT method, a new projective imaging model of point-to-point is derived in this paper. Based on this model and combining traditional projective imaging model of 3D points, a robust direct linear transformation (RDLT) method is proposed to solve the PnP problem. The simulation and real experiments show that the RDLT method, as a linear solution to the PnP problem, offers the accuracy and stability comparable the existing non-linear optimization methods due to the use of constraint conditions of point-to-point, however its computational efficiency is much higher than that of those methods.

Keywords: computer vision; PnP problem; pose measurement; linear solving

0 引言

摄像机的绝对位姿估计是计算机视觉中需要解决的一个关键核心问题，该问题描述为：在相机内参数已知的情况下，通过 n 个空间特征点与其所对应的图像点来估算摄像机的空间位置和姿态信息，即 PnP 问题^[1]。PnP 问题的解决可广泛应用于设备的姿态监测、无人飞行器的视觉着陆/舰、航天器的交会对接、虚拟现实和视觉 SLAM (Simultaneous Localization and Mapping) 等^[2-4]任务中。

解决PnP问题的方法最初来源于相机标定方法，如直接线性变换的 (Direct Liner Transformation, DLT) 方法，利用共面图像之间单映性的 (Homography,

HOMO) 方法，Tsai的方法等^[5,6]。这些方法直接利用空间参考点和图像之间的映射关系建立线性方程组来求解，因此效率非常高。但是，直接线性变换的方法无法对4个参考点和5个参考点的情况 (即P4P和P5P问题) 求解。Tsai无法对P4P问题求解，单映性的方法只能对参考点共面的问题求解。除此之外，这些方法仅仅利用了已知参考点和图像坐标之间的线性约束关系进行求解，并没有考虑成像过程中的非线性约束关系。因此，这些方法求解过程中涉及到的变量个数较多，且求解过程容易受到噪声的干扰导致求解误差较大。针对这些问题，后续有人提出了一些解决P4P和P5P问题的方法^[7,8]，取得了较好的效果。但这些方法仅针

* 基金项目：国家自然科学基金 (62001198, 61963025, 61866021) 项目资助

对4个或5个空间参考点的特殊情况，不具有通用性。

由于线性求解PnP问题的方法精度、抗噪能力并不好，因此有研究者提出基于迭代方式求解PnP问题的方法^[9]。迭代法相比于线性求解的方法，求解精度和稳定性都要高很多。但迭代法对初值的选取比较敏感，如果初值选择不合理，则整个迭代过程需要重复多次，将消耗大量的时间。除此之外，由于迭代法不具有闭合解，有陷入局部最优的危险，这将导致最终得到的结果不是真实所求的位姿值。在空间参考点比较少的环境下，迭代法的精度和稳定性都不令人满意。

2009年，Lepetit提出了EPnP方法^[10]，该方法是首个采用非迭代思路优化求解PnP问题的方法，计算量相比于之前的迭代方法要小很多。但EPnP在求解过程中采用了线性化的策略，因此当空间参考点比较少的时候，其求解精度较差，且抗噪能力不强。2011年，Hesch提出了DLS方法^[11]来求解PnP问题，该方法通过矩阵分解和变换的方式将平移向量用旋转矩阵参数化，然后使用CGR（Cayley-Gibbs-Rodriguez）参数表示旋转，构建出仅含有3个未知变量的最优化问题。最后通过矩阵合成的方法求解这个最优化问题得到最终的位姿参数。但是，DLS方法的整个参数化过程较为复杂，需要进行大量的矩阵分解和变换运算。除此之外，CGR参数表示旋转容易在计算过程中带来奇异性问题，影响位姿求解的精度和稳定性。2012年，Li提出了RPnP方法^[12]，该方法首先将空间参考点三个为一组进行分组，然后利用P3P方法构建约束方程，最后通过求解多项式的方式得到最终的位姿值。RPnP方法虽然性能优异，但从原理上而言是一种次优化方法，没有从全局优化的角度求解，因此精度还有进一步提高的空间。2013年，Zhang提出了两种基于全局最优化思路求解的方法（ASPNP方法^[13]和OPnP方法^[14]）。这两种方法都是采用四元数来表示旋转，进而通过旋转参数化平移向量，得到可以全局优化求解的多项式方程组。最后通过Gröbner基的方法来求解多项式方程组，得到最终的位姿值。由于采用了全局最优化的方式求解PnP问题，因此ASPNP方法和OPnP方法具有很高的求解精度。2014年，Kneip提出了UPnP方法^[15]，该方法不仅能求解基于透视成像模型的PnP问题，也可以求解基于通用相机模型^[16]的PnP问题。2015年，Nakano提出了optDLS的方法^[17]，该方法通过对OPnP方法中的优化目标函数添加新的约束条件，减少了OPnP算法求解过程中解的个数。2018年，Wang等人通过改进RPnP算法，提出了具有鲁棒性的SRPNP算法^[18]，该算法相比于

RPnP算法精度和稳定性更优。2020年，汪佳宝等通过对空间参考点设置不同的权重系数，提出了迭代求解的WIEPNP方法^[19]，该方法的求解精度类似于OPnP方法^[14]和SRPNP^[18]方法。

到目前为止，各种解决PnP问题的方法都是以空间单一点特征的成像模型为基础求解的。事实上，在PnP问题中，除了空间单一的点可以提供成像约束信息外，点与点之间也可以提供成像约束信息。然而，现有的PnP算法研究中都未曾考虑点与点之间提供的成像约束信息。为了提升线性求解PnP问题方法的精度和稳定性，本文在单一点特征成像模型的基础上，推导得到了点与点之间所提供的成像约束数学模型。并以此模型为基础，基于传统线性求解PnP问题的思路，提出了具有鲁棒性且能线性求解PnP问题的（RDLT, Robust Direct Linear Transformation）方法。仿真和实际验证表明：RDLT虽然是一种线性求解PnP问题的方法，但是由于考虑了点与点之间的成像约束关系，因此求解精度和稳定性接近于（远高于传统线性求解方法）现有非线性最优化的方法，而求解效率则远高于这些非线性最优化的方法。

1 PnP问题的表述

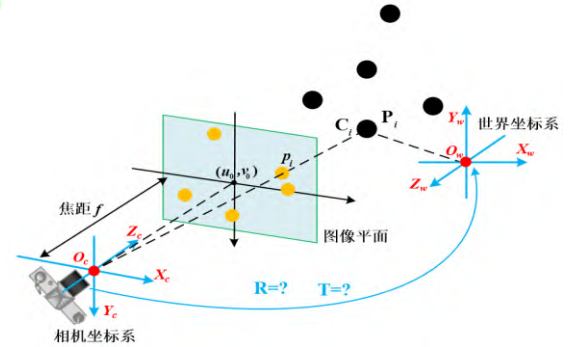


图1 PnP问题的示意图

Fig.1 Schematic diagram of the PnP problem

相机的成像模型如图1所示，其中 O_c 为相机的光心， f 为光心距离成像平面的距离（称为焦距），相机光轴与成像平面相交于主点 (u_0, v_0) 。假定空间中的某一特征点在世界坐标系 $O_w - X_w Y_w Z_w$ 下的坐标表示为 P_i ，在相机坐标系 $O_c - X_c Y_c Z_c$ 下的表示为 C_i ，则 P_i 和 C_i 之间满足关系：

$$C_i = R P_i + t \quad i = 1, 2, \dots, n \quad (1)$$

其中 $R \in SO(3)$ 为一个 3×3 的矩阵，表示相机坐标系和世界坐标系之间的旋转； $t \in \mathbf{R}^3$ 为 3×1 的向量，表示相

机坐标系和世界坐标系之间的平移。依据小孔成像模型， C_i 在图像平面上的投影为 p_i ，并且满足关系：

$$\lambda_i p_i = KC_i \quad i = 1, 2, \dots, n \quad (2)$$

其中 λ_i 为深度因子， K 为相机内参数(通过标定获得)。

PnP的目标就是在相机内参数 K 已知的情况下，利用空间参考点 $\{P_i, i = 1, 2, \dots, n\}$ 和其所对应的图像投影点 $\{p_i, i = 1, 2, \dots, n\}$ 来计算相机坐标系和世界坐标系之间的旋转 R 和平移 t 。式(2)两端左乘 K^{-1} ，将图像坐标 p_i 归一化为 \tilde{p}_i ，并将式(1)代入式(2)最终得到用于求解PnP问题的透视成像模型：

$$\lambda_i \tilde{p}_i = RP_i + t \quad i = 1, 2, \dots, n \quad (3)$$

为了后续算法的推导，令 \tilde{p}_i ， P_i ， R 和 t 的具体形式分别为： $\tilde{p}_i = [u_i, v_i, 1]^T$ ， $P_i = [X_i, Y_i, Z_i]^T$ ，

$$R = \begin{bmatrix} r_1 & r_4 & r_7 \\ r_2 & r_5 & r_8 \\ r_3 & r_6 & r_9 \end{bmatrix}, \quad t = [t_x, t_y, t_z]^T。$$

2 具有鲁棒性的PnP线性求解方法

本部分首先简要介绍基于空间点特征成像模型，即式(3)，线性求解PnP问题的方法(传统DLT方法)。然后推导出空间点和点之间提供的成像约束模型，并基于此模型提出具有鲁棒性的PnP问题线性求解新方法，即RDLT方法。

2.1 直接线性求解(DLT)的方法

传统的DLT方法仅利用了空间特征点和其图像投影点之间的关系，即基于式(3)求解 R 和 t 。该方法主要包括以下步骤：

(a) 将 \tilde{p}_i ， P_i ， R 和 t 的具体表示代入式(3)，并利用式(3)的第三行消去前两行的深度因子 λ_i ，得到等式：

$$\begin{cases} r_1 X_i - r_3 X_i u_i + r_4 Y_i - r_6 Y_i u_i + r_7 Z_i - r_9 Z_i u_i + t_x = t_z u_i \\ r_2 X_i - r_3 X_i v_i + r_5 Y_i - r_6 Y_i v_i + r_8 Z_i - r_9 Z_i v_i + t_y = t_z v_i \end{cases} \quad (4)$$

实际应用中相机所拍摄的物体处必须处在镜头的前方，因此式(4)中的 $t_z > 0$ 。式(4)两端同时除以 t_z ，

并令 $s_i = r_i/t_z, i = 1, 2, \dots, 9$ ， $t_1 = t_x/t_z$ ， $t_2 = t_y/t_z$ 。则可

以得到关于 $s = [s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, t_1, t_2]^T$ 的线性方程组：

$$A_i s = B_i \quad (5)$$

其中：

$$A_i = \begin{bmatrix} X_i, 0, -X_i u_i, Y_i, 0, -Y_i u_i, Z_i, 0, -Z_i u_i, 1, 0 \\ 0, X_i, -X_i v_i, 0, Y_i, -Y_i v_i, 0, Z_i, -Z_i v_i, 0, 1 \end{bmatrix}, \quad B_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix}。$$

(b) 空间每一个特征点都满足式(5)，对于空间中的 n 个特征点，可以得到：

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} s = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} \Leftrightarrow As = B \quad (6)$$

s 中含有11个未知数，当空间特征点数目 $n \geq 6$ 时，可采用最小二乘的方式求解得到：

$$s = (A^T A)^{-1} A^T B \quad (7)$$

(c) 由于旋转矩阵 R 单位正交，因此有：

$$r_1^2 + r_2^2 + r_3^2 = 1 \quad (8)$$

将 $s_1 = r_1/t_z$ ， $s_2 = r_2/t_z$ 和 $s_3 = r_3/t_z$ 代入式(8)，并结合 $t_z > 0$ 的条件可以得到：

$$t_z = \frac{1}{\sqrt{s_1^2 + s_2^2 + s_3^2}} \quad (9)$$

将 t_z 代回 $s_i = r_i/t_z, i = 1, 2, \dots, 9$ ， $t_1 = t_x/t_z$ 和 $t_2 = t_y/t_z$ ，

得到最终所需要的 R 和 t 。

2.2 具有鲁棒性的线性求解(RDLT)方法

本部分将具体推导出PnP问题中点和点之间的成像约束模型，然后以此模型为基础并结合传统的点成像模型，提出具有鲁棒性的PnP问题线性求解新方法(RDLT方法)。

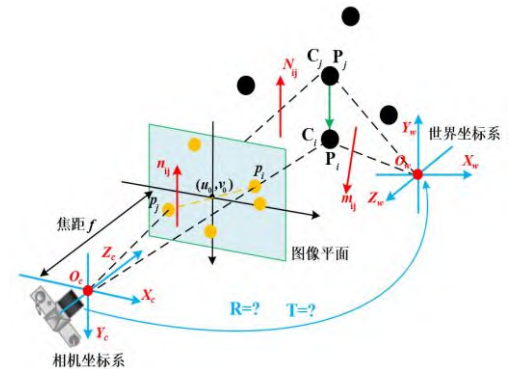


图2 PnP问题中点对点的约束

Fig. 2 The constraint of point-to-point in the PnP problem

2.2.1 点和点之间的成像约束模型

如图2中所示, P_i, P_j 为空间的两特征点, 其在相机坐标系下的表示分别为 C_i 和 C_j 。相机光心 O_c , C_i 和 C_j 共同形成平面 $C_i O_c C_j$, 该平面在相机坐标系下的法向量为 $N_{ij} = C_i \times C_j$ 。同理, C_i 和 C_j 在图像上的投影 p_i, p_j 和 O_c 之间也形成一个平面 $p_i O_c p_j$, 并且该平面和 $C_i O_c C_j$ 是同一个平面。记平面 $p_i O_c p_j$ 的法向量为 $n_{ij} = p_i \times p_j$, 则 n_{ij} 和 N_{ij} 之间满足平行关系:

$$\alpha_{ij} n_{ij} = N_{ij} \quad (10)$$

其中 α_{ij} 为比例因子。将 $N_{ij} = C_i \times C_j$ 代入式 (10) 有:

$$\alpha_{ij} n_{ij} = C_i \times C_j \quad (11)$$

其中 C_i 和 C_j 可由 P_i 和 P_j 按照式 (1) 计算得到。将式 (1) 代入式 (11) 可得:

$$\alpha_{ij} n_{ij} = (R P_i + t) \times (R P_j + t) \quad (12)$$

按照叉乘的运算规律, 将式 (12) 展开, 合并同类项得到:

$$\alpha_{ij} n_{ij} = R(P_i \times P_j) - t \times R(P_i - P_j) + t \times t \quad (13)$$

在式 (13) 中, $P_i \times P_j$ 表示在世界坐标系下形成平面 $P_i O_w P_j$ 的法向量, 记为 m_{ij} ; $P_i - P_j$ 表示世界坐标系下的向量, 记为 q_{ij} ; $t \times t$ 的值为0。将以上条件代入, 式 (13) 最终变为:

$$\alpha_{ij} n_{ij} = R m_{ij} - t \times R q_{ij} \quad (14)$$

设 $q_{ij} = [q_x, q_y, q_z]^T$, 将 t 和 R 的具体表示形式代入 $t \times R q_{ij}$ 得到:

$$\begin{aligned} t \times R q_{ij} &= \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \times \begin{bmatrix} r_1 & r_4 & r_7 \\ r_2 & r_5 & r_8 \\ r_3 & r_6 & r_9 \end{bmatrix} \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix} \\ &= \begin{bmatrix} (r_3 t_y - r_2 t_z) q_x & (r_6 t_y - r_5 t_z) q_y & (r_9 t_y - r_8 t_z) q_z \\ (r_1 t_z - r_3 t_x) q_x & (r_4 t_z - r_6 t_x) q_y & (r_7 t_z - r_9 t_x) q_z \\ (r_2 t_x - r_1 t_y) q_x & (r_5 t_x - r_4 t_y) q_y & (r_8 t_x - r_7 t_y) q_z \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} r_3 t_y - r_2 t_z & r_6 t_y - r_5 t_z & r_9 t_y - r_8 t_z \\ r_1 t_z - r_3 t_x & r_4 t_z - r_6 t_x & r_7 t_z - r_9 t_x \\ r_2 t_x - r_1 t_y & r_5 t_x - r_4 t_y & r_8 t_x - r_7 t_y \end{bmatrix}}_H \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix} = H q_{ij} \end{aligned}$$

将 H 中的每一个分量都当作新的未知数, 即令

$$H = \begin{bmatrix} h_1 & h_4 & h_7 \\ h_2 & h_5 & h_8 \\ h_3 & h_6 & h_9 \end{bmatrix}, \text{ 其中 } h_i, i=1,2,\dots,9 \text{ 为未知数, 则式}$$

(14) 最终变为:

$$\alpha_{ij} n_{ij} = R m_{ij} - H q_{ij} \quad (15)$$

式 (15) 就是本文推导得到的点和点之间的成像约束模型。

2.2.2 RDLT方法的推导

对于空间中的 n 个特征点, 随机两两组合, 总共可以得到 $k = n(n-1)/2$ 种可能。因此, 在PnP问题的线性求解中, 除了可以利用式 (5) 提供的 n 个约束信息以外, 还可以利用式 (15) 额外提供的 k 个约束条件。令式 (15) 中 $n_{ij} = [n_x, n_y, n_z]^T$, $m_{ij} = [m_x, m_y, m_z]^T$, 将 n_{ij} , R , m_{ij} , H 和 q_{ij} 的具体形式代入式 (15) 展开得到:

$$\begin{cases} \alpha_{ij} n_x = r_1 m_x + r_4 m_y + r_7 m_z - h_1 q_x - h_4 q_y - h_7 q_z \\ \alpha_{ij} n_y = r_2 m_x + r_5 m_y + r_8 m_z - h_2 q_x - h_5 q_y - h_8 q_z \\ \alpha_{ij} n_z = r_3 m_x + r_6 m_y + r_9 m_z - h_3 q_x - h_6 q_y - h_9 q_z \end{cases} \quad (16)$$

将式 (16) 的第三行分别代入前两行消去 α_{ij} 得到:

$$\begin{aligned} & r_1 m_x n_z - r_3 m_x n_x + r_4 m_y n_z - r_6 m_y n_x + r_7 m_z n_z \\ & - r_9 m_z n_x - h_1 q_x n_z + h_3 q_x n_x - h_4 q_y n_z + h_6 q_y n_x \\ & - h_7 q_z n_z + h_9 q_z n_x = 0 \end{aligned} \quad (17)$$

$$\begin{aligned} & r_2 m_x n_z - r_3 m_x n_y + r_5 m_y n_z - r_6 m_y n_y + r_8 m_z n_z \\ & - r_9 m_z n_y - h_2 q_x n_z + h_3 q_x n_y - h_5 q_y n_z + h_6 q_y n_y \\ & - h_8 q_z n_z + h_9 q_z n_y = 0 \end{aligned} \quad (18)$$

为了同时利用式(5)、式(17)和式(18)提供的约束信息,首先在式(17)和式(18)两端除以 t_z ,并令 $w_i = h_i/t_z, i = 1, 2, \dots, 9$ 。然后将式(5)中的变量 s 扩充为: $\tilde{s} = [s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, t_1, t_2, w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9]^T$ 。则式(5)变为:

$$\tilde{A}_i \tilde{s} = B_i \quad (19)$$

$$\text{其中 } \tilde{A}_i = \begin{bmatrix} X_i, 0, -X_i u_i, Y_i, 0, -Y_i u_i, Z_i, 0, -Z_i u_i, 1, 0, 0_{1 \times 9} \\ 0, X_i, -X_i v_i, 0, Y_i, -Y_i v_i, 0, Z_i, -Z_i v_i, 0, 1, 0_{1 \times 9} \end{bmatrix},$$

$0_{1 \times 9}$ 表示1行9列的零向量, B_i 和式(5)中的定义相同。根据式(19),式(6)可以重新写为:

$$\tilde{A} \tilde{s} = B \quad (20)$$

其中 $\tilde{A} = [\tilde{A}_1^T, \tilde{A}_2^T, \dots, \tilde{A}_n^T]^T$, B 和式(6)中的定义相同。

利用新定义的 \tilde{s} 将式(17)和式(18)写为矩阵的形式:

$$C_i \tilde{s} = D_i \quad (21)$$

其中 $D_i = [0, 0]^T$ 且有:

$$C_i = \begin{bmatrix} m_x n_x, 0, -m_x n_x, m_y n_z, 0, -m_y n_x, m_z n_z, 0, -m_z n_x, \\ 0, m_x n_z, -m_x n_y, 0, m_y n_z, -m_y n_y, 0, m_z n_z, -m_z n_y, \\ 0, 0, -q_x n_z, 0, q_x n_x, -q_y n_z, 0, q_y n_x, -q_z n_z, 0, q_z n_x \\ 0, 0, 0, -q_x n_z, q_x n_y, 0, -q_y n_z, q_y n_y, 0, -q_z n_z, q_z n_y \end{bmatrix}.$$

式(21)对于空间 k 个点与点的组合都满足,因此可以得到:

$$\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_k \end{bmatrix} \tilde{s} = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_k \end{bmatrix} \Leftrightarrow C \tilde{s} = D \quad (22)$$

整合式(20)与式(22)可以得到:

$$\begin{bmatrix} \tilde{A} \\ C \end{bmatrix} \tilde{s} = \begin{bmatrix} B \\ D \end{bmatrix} \Leftrightarrow E \tilde{s} = F \quad (23)$$

式(23)中不仅包含了单点特征提供的成像约束信息,而且还包含了点和点特征之间提供的成像约束信息。 \tilde{s} 中包含20个未知变量,当空间点数目 $n \geq 4$ 时,式(23)至少可以提供20个关于 \tilde{s} 的等式。此时,采用最小二乘法,可以求解得到 \tilde{s} :

$$\tilde{s} = (E^T E)^{-1} E^T F \quad (24)$$

从以上推导可以看出,本文的RDLT方法将线性求解PnP问题的特征点数目从6降低到了4,这有效扩充了线性可解PnP问题的范围。得到 \tilde{s} 以后,提取其前11

$$\text{个分量,可以得到含有比例因子的 } \tilde{R} = \begin{bmatrix} s_1 & s_4 & s_7 \\ s_2 & s_5 & s_8 \\ s_3 & s_6 & s_9 \end{bmatrix} \text{ 和}$$

$$\tilde{t} = \begin{bmatrix} t_1 \\ t_2 \\ 1 \end{bmatrix}, \text{ 其和真实的 } R, t \text{ 之间相差一个系数 } t_z. \text{ 要}$$

恢复真实的 R 和 t ,可以按照式(9)的方式先计算 t_z ,然后再计算 R 和 t 。本文中为了保证恢复 R 和 t 时的稳定性,采用绝对定向^[20]的方式来计算,具体的步骤如下:

(a) 将 \tilde{R} 和 \tilde{t} 代入式(1)得到 $\{P_i, i = 1, 2, \dots, n\}$ 所对应的 $\{C_i, i = 1, 2, \dots, n\}$,利用式(25)分别计算其均值 U_p 和 U_c :

$$U_p = \frac{1}{n} \sum_{i=1}^n P_i \quad U_c = \frac{1}{n} \sum_{i=1}^n C_i \quad (25)$$

(b) 计算 $\{P_i, i = 1, 2, \dots, n\}$ 和 $\{C_i, i = 1, 2, \dots, n\}$ 之间的协方差矩阵:

$$E_{UP} = \frac{1}{n} \sum_{i=1}^n (P_i - U_p)(C_i - U_c)^T \quad (26)$$

(c) 利用奇异值(SVD)分解的方式将 E_{UP} 分解为 U, S 和 V ,则 R 和 t 分别由式(27)得到:

$$R = UV^T \quad t = U_p - RU_c \quad (27)$$

3 仿真实验

本部分通过仿真实验来测试本文提出的RDLT算法。为了准确评估RDLT算法的各项性能,实验过程中将会和各主流PnP算法进行对比。选择对比的算法主要

有:

- ✧ **DLT 方法**^[5]: 传统线性求解 PnP 的方法, 至少需要空间 6 个点。
- ✧ **LHM 方法**^[9]: 也称为正交迭代法, 精度最好的迭代法之一。
- ✧ **EPnP 方法**^[10]: 效率非常高的非迭代求解 PnP 问题的方法, 通常在 $n \geq 6$ 的情况下有较高的求解精度, 是综合性能最优秀的非迭代求解方法之一。
- ✧ **DLS 方法**^[11]: 第一种采用矩阵合成技术最优化求解 PnP 问题的方法。
- ✧ **RPnP 方法**^[12]: 快速且具有鲁棒性的 PnP 求解方法, 在 $4 \leq n < 6$ 和 $n \geq 6$ 的情况都具有很高的求解精度。
- ✧ **OPnP 方法**^[14]: 将 PnP 问题转化为最优化问题, 通过 Gröbner 基方法求解。OPnP 方法具有很高的求解精度和鲁棒性, 其代表着目前求解 PnP 问题的最高水平。
- ✧ **SRPnP 方法**^[18]: RPnP 方法的改进, 其精度和 OPnP 方法相当, 但是效率远高于 OPnP 方法。

论文中所有对比方法的源代码都可以从作者网站 <https://sites.google.com/view/ping-wang-homepage> 下载。

3.1 仿真数据的产生及误差的定义

合成分辨率为 640×480 像素的虚拟相机, 该虚拟相机的焦距 $f = 800$ 像素单位。三维空间参考点 c_i 在相机坐标系下随机产生, 其分布范围为 $[-2, 2] \text{ m} \times [-2, 2] \text{ m} \times [4, 8] \text{ m}$ 。利用随机生成的旋转矩阵 R_{true} 和平移向量 t_{true} 将 c_i 转换到世界坐标系下, 基于小孔成像模型, 将 c_i 投影到图像平面, 并依据实验条件给投影点增加不同等级的高斯白噪声。通过算法计算得到的旋转矩阵和平移向量分别记为 R_{ture} 和 t_{ture} , 则计算值和真值之间的误差表示为:

$$e_{rot}(\text{degrees}) = \max_{k \in \{1, 2, 3\}} \cos^{-1}(r_{k, true}^T r_{k, estimation}) \times \frac{180}{\pi} \quad (28)$$

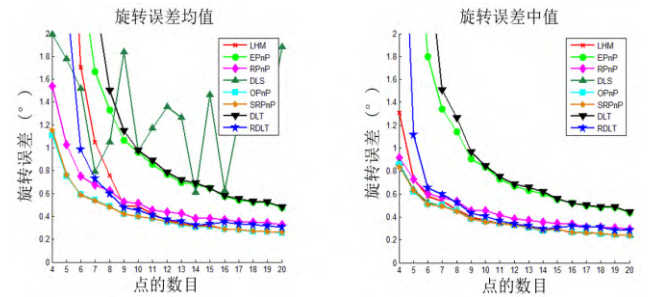
$$e_{trans}(\%) = \frac{\|t_{estimation} - t_{true}\|}{\|t_{true}\|} \times 100$$

其中 $r_{k, true}$ 和 $r_{k, estimation}$ 分别表示 R_{true} 和 $R_{estimation}$ 的第 k 列。值得注意的是, 为了误差评估的方便性和一致性, 本论文中的误差定义采用已有工作^[13,14,15,17,18]中的表示形式。

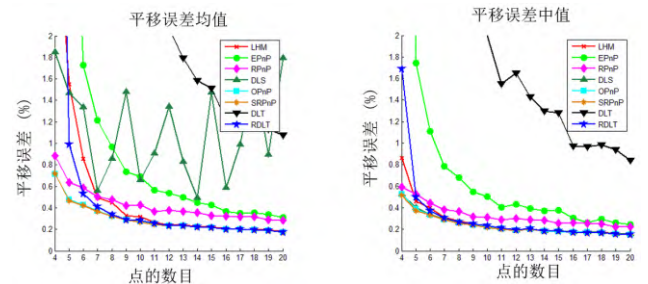
3.2 输入参考点数目对算法精度的影响

固定输入白噪声的方差为2像素, 然后通过改变参

考点数目从4到20来验证各PnP算法的求解精度, 仿真实验结果如图3所示。LHM作为一种迭代求解的方法, 在空间参考点数目较少的时候容易陷入局部最优, 求解误差较大。而在参考点较多时候, 由于有较多的冗余信息可利用, 能够计算得到全局最优值, 因此求解精度高。EPnP算法是一种使用潜在线性化策略的方法, 因此其整体求解精度不高。RPnP算法在参考点较少和较多时候都呈现出稳定的计算结果, 但由于其是一种次优化的方法, 因此精度还有提升的空间。DLS方法的中值求解精度很高, 但平均求解精度反而较低, 主要原因是DLS方法中采用了CGR参数来参数化旋转矩阵, 计算时容易出现奇异现象, 影响了算法整体的稳定性。OPnP和SRPnP算法作为非线性最优化的方法, 具有很高的求解精度。DLT作为传统的线性求解方法, 由于利用的约束信息较少, 因此精度较差。相比之下, 本文的RDLT算法不仅利用了单一点成像约束信息, 而且还利用了点与点之间的成像约束信息, 因此在同样线性求解的情况下, 取得了很高的求解精度和稳定性。随着空间点数目 n 的增加, RDLT算法的求解精度和稳定性几乎和非线性优化的方法相当。到目前为止, 在所有的PnP方法中, RDLT是首个采用了线性求解策略, 但求解精度却能够达到非线性优化方式的方法。



(a) 点数目 n 变化时旋转误差的均值和中值
(a) The mean and median of rotation errors when the number of points n varies



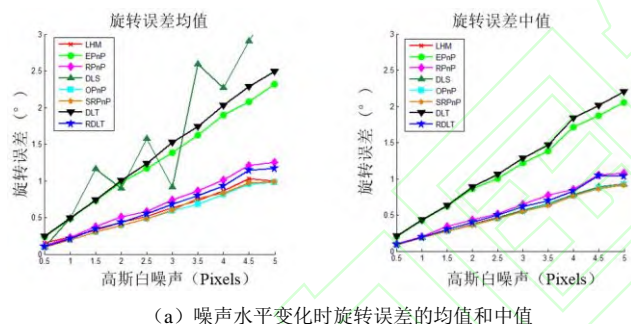
(b) 点数目 n 变化时平移误差的均值和中值
(b) The mean and median of translation errors when the number of points n varies

图 3 点数目变化时所有算法的精度

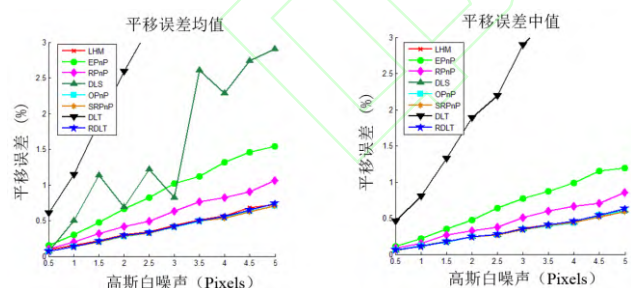
Fig. 3 The accuracies of all the methods when the number of points varies

3.3 噪声对算法精度的影响

固定空间输入参考点数目为 10，改变输入噪声的方差从 0 到 5 像素（步进为 0.5 像素）来验证各 PnP 算法的求解精度，最终的实验结果如图 4 所示。可以看出，随着输入噪声的增大，各 PnP 算法的求解精度都在下降，并且噪声增加的等级和各算法求解精度之间近似符合线性的分布关系。类似于 3.2 中的结论，RDLT 虽然是一种线性求解 PnP 问题的方法，但是其求解精度和非线性优化的方法（OPnP 和 SRPnP 方法）几乎相当，并且远高于同样采用线性求解方式的 DLT 算法。



(a) 噪声水平变化时旋转误差的均值和中值
(a) The mean and median of rotation errors when the noise level varies



(b) 噪声水平变化时平移误差的均值和中值
(b) The mean and median of translation errors when the noise level varies

图 4 噪声水平变化时所有算法的精度

Fig. 4 The accuracies of all the methods when the noise level varies

3.4 算法的运行效率

改变输入参考点的数目从 4 到 500，在每个参考点下各算法分别执行 1000 次，然后统计各算法的平均时

间，结果如图 5 中所示。LHM 因为是迭代求解的方法，效率最低。EPnP 虽然是一种最优化求解的方法，但是其求解过程中采用了线性化的策略，因此求解速度很快。DLS 和 OPnP 都是采用矩阵合成技术求解的方法，求解过程中需要花费较多的时间，求解效率较低。RPnP 算法虽然有较高的求解效率，但毕竟是一种次优化的方法，精度上还有提升的空间。SRPnP 作为 RPnP 方法的改进版本，具有很高的求解精度，但求解效率随着空间点总数的变大急剧增加。DLT 作为传统线性求解的方法，在所有对比的方法中具有最好的求解效率，但其求解精度也对应是最差的。相比之下，本文提出的 RDLT 方法，不仅求解效率高，而且求解精度接近最优化求解的方法。本文的 RDLT 算法非常适合视觉 SLAM 等实际应用项目，这些应用不仅要求视觉位姿测量算法具有较高的求解精度，而且还需要有较高的求解效率。

4 实验验证

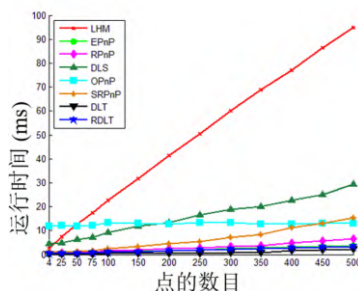
4.1 实验器材

实验中的图像采集设备为大恒 MER-132-43U3C 相机，分辨率为 1292×964 像素，像元尺寸为 $3.75 \mu\text{m}$ 。相机内参数的标定采用张正友标定法^[21]，标定中使用的棋盘格靶标由氧化铝面板和玻璃基板材料定制加工而成，其加工精度为 0.01mm ，每个棋盘格的大小为 30mm 。标定后的相机参数如表 1 中所列，其中 k_1 和 k_2 为相机镜头的径向畸变系数， p_1 和 p_2 为相机镜头的切向畸变系数。

表1 相机标定的结果

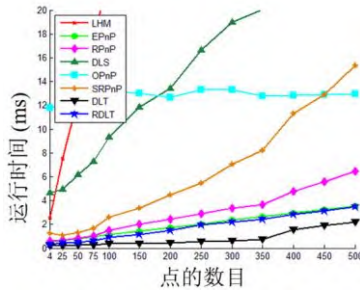
Table 1 Camera calibration result

焦距	$f_x = 2225.36002, f_y = 2225.17837$
主点	$(c_x, c_y) = (669.12275, 496.46005)$
畸变系数	$k_1 = -0.10203, k_2 = 0.18486,$ $p_1 = 0.00055, p_2 = 0.00126$



(a) 所有算法的平均运行时间分布

(a) The distribution of average running time for all the methods



(b) 图 (a) 的放大

(b) A zoomed-in version of Fig. 5 (a)

图 5 所有算法的计算效率

Fig. 5 The computational efficiency of all the methods

4.2 实验步骤

(1) 利用亚像素角点提取方法提取标定板上一定范围内（如图 6 中青色标注范围）的角点；

(2) 从提取的角点中随机选择 10 个点，利用这些点所对应的世界坐标和图像坐标求解相机和标定板之间的位置和姿态关系；

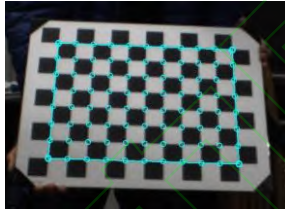


图 6 角点提取的范围

表2 所有方法的均值和中值重投影误差（单位：像素）

Table 2 The mean and median re-projection errors for all the methods

		1	2	3	4	5	6	7	8	9	10
LHM	均值	107.23	90.95	180.05	190.43	62.32	25.32	36.12	17.77	151.31	50.60
	中值	104.90	88.86	156.07	157.31	59.33	24.85	15.89	7.18	116.41	15.11
EPnP	均值	6.64	4.01	8.31	9.76	6.44	3.44	8.26	7.56	7.98	10.04
	中值	6.38	3.62	7.29	9.56	6.32	3.38	7.99	7.08	7.92	9.23
RPnP	均值	6.32	3.46	4.14	7.18	5.81	4.20	4.83	5.86	3.77	6.94
	中值	6.38	3.10	4.16	6.95	5.75	3.94	4.38	5.38	3.25	6.20
DLS	均值	248.77	250.26	258.84	277.92	253.95	241.93	211.33	277.41	269.71	255.37
	中值	6.61	6.23	5.51	8.60	5.71	7.52	5.22	4.87	6.72	5.92
OPnP	均值	16.31	134.12	9.74	7.06	5.94	440.92	4.35	5.57	3.57	6.49
	中值	6.51	93.21	6.01	6.97	5.74	7.55	4.05	5.33	2.95	6.01
SRPnP	均值	5.65	2.98	3.95	7.06	5.67	3.11	3.99	5.24	3.03	6.11
	中值	5.63	2.97	4.02	6.97	5.72	3.18	3.96	5.11	3.01	6.09

Fig. 6 The region of corner extraction

(3) 利用求解得到的位姿关系将标定板上所有角点对应的世界坐标重投影到图像上，并和亚像素提取的角点求取绝对距离。PnP 算法求解得到的位姿信息越准确，则重投影到图像平面上的角点和亚像素提取的角点距离越小，反之亦然。为了量化各 PnP 算法的求解精度，定义亚像素提取的角点位置为 $[x_i, y_i]^T$ ，通过位姿关系重投影得到的角点位置为 $[\tilde{x}_i, \tilde{y}_i]^T$ ，则图像中提取角点和其重投影角点之间的误差为：

$$\varepsilon_i = \sum_{i=1}^n (x_i - \tilde{x}_i)^2 + (y_i - \tilde{y}_i)^2 \quad (29)$$

这里 n 表示提取角点的数目。对于以上步骤，重复 50 次，将这 50 次误差的均值 $\bar{\varepsilon} = \frac{1}{50} \sum_{i=1}^{50} \varepsilon_i$ 作为最终指标来评价各 PnP 算法的求解精度；

(4) 采集 10 幅不同姿态和距离下的标定板图像，采用各 PnP 算法计算每幅图像的投影误差均值 $\bar{\varepsilon}$ ，结果如表 2 和图 7 中所示。在图 7 中，蓝色“o”表示随机测试时选择的 10 个角点，红色“+”表示通过位姿关系重投影得到的角点。

HOMO	均值	9.18	5.55	16.55	17.89	7.68	4.54	15.78	10.42	11.16	16.45
	中值	7.77	4.91	15.88	14.81	7.17	4.65	16.06	8.23	10.41	16.29
RDLT	均值	6.73	4.01	8.83	9.84	6.33	3.42	8.46	8.02	6.58	10.35
	中值	6.47	3.61	7.73	9.65	6.44	3.38	8.19	7.29	5.85	9.81

从表 2 中可以看出, LHM 算法整体投影误差较大, 这是由于 LHM 算法容易陷入局部最优所导致的。HOMO 作为 DLT 共面情况下的算法, 当选择参考点数目为 10 的时候, 含有较多的冗余信息, 因此计算精度较高。EPnP 算法虽然是一种采用最优化技术求解位姿的方法, 但由于其求解过程中使用了线性化的技术, 虽然精度高于 HOMO 方法, 但仍有提升的空间。RPnP 是一种次优化求解的方法, 求解过程相对稳定, 求解精度也相对较高。DLS、OPnP 和 SRPnP 都是从全局最优出发, 非迭代非线性求解位姿的方法, 但是由于 DLS 求解过程中使用了 Cayley 参数的方式表示旋转矩阵, 因此容易出现奇异值, 所以投影误差均值比较大。OPnP 方法仿真时的中值和均值误差都较小, 但实际测试中仅有投影误差中值符合仿真时的规律, 而投影误差均值却较大, 这是因为 OPnP 算法在求解变量的过程中需要利用较

为复杂的矩阵消去模板。在仿真测试中, 测试数据和噪声通常是比较理想的, 因此利用消去矩阵模板求解变量的过程是稳定的。但实际采集的数据中可能会存在不稳定的噪声因素, 这将导致利用消去矩阵模板求解变量时得到不稳定的解, 导致 OPnP 方法的实际投影误差均值比较大。SRPnP 算法同样作为一种最优化求解的方法, 克服了 DLS 和 OPnP 算法存在的缺点, 因此求解过程稳定, 投影误差中值和均值都比较小。RDLT 作为一种线性求解的方法, 虽然在每组测试图像上的投影误差不是最小的, 但在所有测试的方法中整体误差相对偏小, 结果分布稳定。再考虑到 RDLT 算法优异的求解效率, 可以表明本文提出的 RDLT 算法非常适合实际的位姿测量任务, 特别是同时需要保障精度和效率的位姿测量任务中。

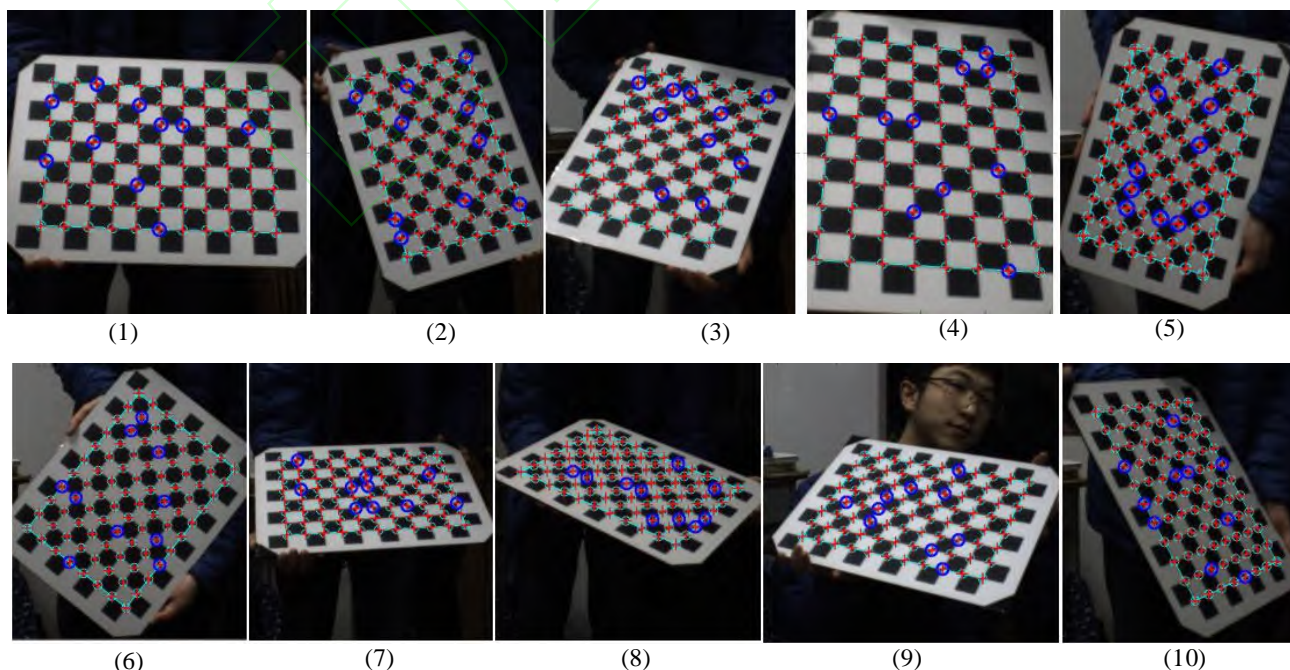


图 7 任意旋转和平移下棋盘格的角点及其投影

Fig. 7 Corners and its projections of a chessboard for arbitrary rotation and translation

5 结论

本文推导出了PnP问题中点与点之间所提供的成像约束数学模型,并以此模型为基础,结合传统单一点成像模型,提出了具有鲁棒性的PnP问题线性求解(RDLT)方法。仿真实验和实际实验表明,本文提出的RDLT算法虽然是一种线性求解方法,但是由于考虑了点与点之间的成像约束关系,因此求解精度和稳定性接近于现有非线性最优化的方法,而求解效率则远高于这些非线性最优化的方法。本文提出的RDLT方法非常适合于对求解精度和求解速度都有较高要求的视觉位姿测量场合。

参考文献

- [1] 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.
- [2] 倪得晶, 宋爱国, 李会军. 基于虚拟现实的机器人遥控操作关键技术研究 [J]. 仪器仪表学报, 2017, 38(10): 2351-2363.
NI D J, SONG A G, LI H J. Survey on robot teleoperation based on virtual reality [J]. Chines Journal of Scientific Instrument, 2017, 38(10): 2351-2363.
- [3] 于鲲, 从明煜, 段佳佳, 等. 星箭对接环抓捕点单目视觉导航方法 [J]. 仪器仪表学报, 2018, 39(12): 228-236.
YU K, CONG M Y, DUAN J J, et al. Monocular visual navigation method for capture point of docking ring [J]. Chines Journal of Scientific Instrument, 2018, 39(12): 228-236.
- [4] LIN S, GARRATT M A, LAMBERT A J. Monocular vision-based real-time target recognition and tracking for autonomously landing an UAV in a cluttered shipboard environment [J]. Autonomous Robots, 2016: 1-21.
- [5] ABDEK Y I, KARARA H M, HAUCK M. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry [J]. Photogrammetric Engineering and Remote Sensing, 2015, 81(2): 103-107.
- [6] TSAI R. An efficient and accurate camera calibration technique for 3D machine vision [C]. Proceeding of Computer Vision and Pattern Recognition, 1986: 364-374.
- [7] QUAN L, TRIGGS B, MOURRAIN B. Some results on minimal euclidean reconstruction from four points [J]. Journal of Mathematical Imaging and Vision, 2006, 24(3): 341-348.
- [8] 吴福朝, 胡占义. 关于 P5P 问题的研究 [J]. 软件学报, 2001, 12(5): 768-775.
WU F CH, HU Z Y. A study on the P5P problem [J]. Journal of Software, 2001, 12(5): 768-775.
- [9] LU C P, HAGER G D, MJOLSNES E. Fast and globally convergent pose estimation from video images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(6): 610-622.
- [10] 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.
- [11] HESCH J A, ROUMELIOTIS S I. A direct least-squares (DLS) method for PnP [C]. IEEE International Conference on Computer Vision, 2011: 383-390.
- [12] LI S, 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.
- [13] ZHENG Y Q, SUGIMOTO S, OKUTOMI M. ASPnP: An accurate and scalable solution to the perspective-n-point problem [J]. IEICE Transactions on Information and Systems, 2013, E96.D(7): 1525-1535.
- [14] ZHENG Y Q, KUANG Y, SUGIMOTO S, et al. Revisiting the PnP problem: a fast, general and optimal solution [C]. IEEE International Conference on Computer Vision, 2013: 2344-2351.
- [15] KNEIP L, LI H, SEO Y. UPnP: an optimal $O(n)$ solution to the absolute pose problem with universal applicability [C]. European Conference on Computer Vision, 2014: 127-142.
- [16] CHEN C S, CHANG W Y. On pose recovery for generalized visual sensors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(7): 848-861.
- [17] NAKANO G. Globally optimal DLS method for PnP problem with Cayley parameterization [C]. British Machine Vision Conference, 2015: 78.1-78.11.
- [18] WANG P, XU G L, CHENG Y H, et al. A simple, robust

and fast method for the perspective-n-point problem [J]. Pattern Recognition Letters, 2018, 108(6): 31-37.

- [19] 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 SH R, ZHOU Q Y. Vision based real-time 3D displacement measurement using weighted iterative EPnP algorithm [J]. Chines Journal of Scientific Instrument, 2020, 041(02): 166-175.

- [20] UMEYAMA S. Least-squares estimation of transformation parameters between 2 point patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(4): 376-380.

- [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.

作者简介



王平 (通讯作者), 2019 年于南京航空航天大学获得博士学位, 现为兰州理工大学讲师, 主要研究方向为机器视觉、机器学习和信号处理。

E-mail: pingwangsky@163.com

Wang Ping (corresponding author)

received his Ph.D. degree in 2019 from Nanjing University of Aeronautics and Astronautics; now, he is a lecturer in Lanzhou University of Technology. His main research interest includes machine vision, machine learning and signal processing.



周雪峰, 2011 年于华南理工大学获得博士学位, 现为广东省智能制造研究所副研究员, 主要研究方向为机器人技能学习、运动规划与力控制。

E-mail: xf.zhou@giim.ac.cn

Zhou Xuefeng received the Ph.D. from South China University of Technology in

2011. He is an associate research fellow in Guangdong Institute of Intelligent Manufacturing. His major research interest includes robot skill learning, motion planning and force control.



安爱民, 2010 年于兰州理工大学获得博士学位, 现为兰州理工大学教授、博士生导师, 主要研究方向为复杂系统建模与动态优化技术和人工智能技术。

E-mail: anaimin11@163.com

An Aimin received his Ph.D. degree in 2010 from Lanzhou University of Technology; now, he is a professor and PhD supervisor in Lanzhou University of Technology. His main research interest includes modeling and dynamic optimization technology of complex systems, and artificial intelligence technology.



何倩玉, 2018 年于兰州理工大学获得学士学位, 现为兰州理工大学硕士研究生, 主要研究方向为机器视觉和视觉测量。

E-mail: 1577221541@qq.com

He Qianyu received her B.Sc. degree in 2018 from Lanzhou University of Technology; now, she is an M.Sc. candidate in Lanzhou University of Technology. Her main research interest includes machine vision and vision measurement.



张爱华, 2005 年于西安交通大学获得博士学位, 现为兰州理工大学教授、博士生导师, 主要研究方向为检测技术、模式识别与智能系统。

E-mail: zhangaihua@lut.edu.cn

Zhang Aihua received her Ph.D. degree in 2005 from Xi'an Jiaotong University; now, she is a professor and PhD supervisor in Lanzhou University of Technology. Her main research interest includes detection technology, pattern recognition and intelligent system.