

# 基于单目视觉的机械臂定位与抓取算法研究

冯迎宾<sup>1</sup>, 刘 潘<sup>1</sup>, 王天龙<sup>2</sup>, 曾城锋<sup>1</sup>

(1. 沈阳理工大学 自动化与电气工程学院, 沈阳 110159;  
2. 中国科学院 机器人与智能制造创新研究院, 沈阳 110169)

**摘要:** 鉴于传统单目视觉定位方法在深度感知层面存在局限性, 本文提出一种基于单目视觉的机械臂定位与抓取算法。首先, 借助相机标定与手眼标定技术, 构建视觉系统的精确几何模型, 以此奠定视觉感知基础。其次, 对采集所得图像实施预处理操作, 并运用基于非线性优化 (Levenberg-Marquardt, L-M) 策略最小化重投影误差, 从而获取视觉基准标记 (AprilTag 码) 的位姿信息。再次, 通过 D-H 参数法搭建机械臂运动学模型, 基于几何方法完成逆运动学求解, 进而实现对目标物体的抓取。最后, 针对定位与抓取算法设计实验, 实验结果表明该算法能够对 AprilTag 码块进行精准定位与抓取, 有效提升了单目视觉系统在机械臂操作中的应用效能。

**关键词:** 单目视觉; 机械臂; 定位抓取; 相机手眼标定; AprilTag 码

**中图分类号:** TP242; TP391.4; TH166 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2025.06.003

## Positioning and Grasping Algorithm for Robotic Arms Based on Monocular Vision

FENG Yingbin<sup>1</sup>, LIU Pan<sup>1</sup>, WANG Tianlong<sup>2</sup>, ZENG Chengfeng<sup>1</sup>

(1. Shenyang Ligong University, Shenyang 110159, China; 2. Chinese Academy of Sciences, Institute of Robotics and Intelligent Manufacturing Innovation, Shenyang 110169, China)

**Abstract:** To address the limitations of traditional monocular vision positioning methods in depth perception, this paper proposes a positioning and grasping algorithm for robotic arms based on monocular vision. Firstly, by leveraging camera calibration and hand-eye calibration techniques, an accurate geometric model of the vision system is established, thereby laying the foundation for visual perception. Secondly, preprocessing operations are performed on the acquired images, and a nonlinear optimization strategy based on the Levenberg-Marquardt (L-M) algorithm is employed to minimize reprojection errors, thereby obtaining the pose information of the visual fiducial marker (AprilTag code). Third, the kinematic model of the robotic arm is established using the Denavit-Hartenberg (D-H) parameter method. The inverse kinematics is then solved based on a geometric approach, thereby enabling the robotic arm to grasp target objects. Finally, experimental validation demonstrates that the proposed algorithm achieves precise positioning and grasping of AprilTag fiducials, significantly enhancing the operational efficiency of monocular vision systems in robotic manipulation tasks.

**Key words:** monocular vision; manipulator; positioning and grasping; camera hand-eye calibration; AprilTag code

随着智能制造与自动化技术的深度发展,工业机械臂的智能化水平与控制精度成为产业升级的核心需求<sup>[1]</sup>。传统预编程轨迹控制模式难以适应动态作业环境,而机器视觉技术的引入为机械臂赋予了环境感知与自主决策能力。该技术通过实时获取目标位姿信息,显著提升了机械臂在复杂工况下的作业灵活性<sup>[2-5]</sup>。在视觉定位方案中,单目相机、双目相机和深度相机构成主流技术路线:双目视觉系统<sup>[6-7]</sup>基于视差原理实现高精度深度测量,但存在标定复杂度高与匹配算法计算量大的局限;深度相机<sup>[8]</sup>(如结构光传感器)虽可避免匹配运算直接输出高精深度图,但其硬件成本高昂,易受环境光干扰且需专用算力支持,难以满足低成本与轻量化场景需求。相比之下,单目视觉<sup>[3]</sup>凭借低成本、易部署和低算力消耗的优势,在工业机械臂领域得到广泛应用。

标志检测系统 AprilTag 因其解码速度快、抗噪声能力强等特性,成为单目视觉定位的重要技术手段<sup>[9-11]</sup>。然而,工业场景中复杂光照条件(如局部过曝、阴影遮挡)易导致图像特征退化,且传统线性测距模型在非理想视角或遮挡条件下定位误差显著增大。针对上述问题,国内外学者从图像增强与定位算法优化两方面开展研究。在光照预处理领域:苗作华等<sup>[12]</sup>提出改进单尺度 Retinex(SSR)算法,结合亮通道先验灰度化,有效提升了低光照图像的识别精度;Nusantika 等<sup>[13]</sup>针对输电线路监测高噪声、低亮度图像问题,设计了融合增强型多尺度 Retinex(MSR)的图像增强方法;吕伏等<sup>[14]</sup>提出了基于 Retinex 增强与 Gamma 校正的联合预处理方案,显著改善了图像细节表现力。在定位算法方面:张天意等<sup>[15]</sup>采用了 PnP 算法实现室内二维码定位;余泽康等<sup>[16]</sup>提出了结合二维码与目标数据库的间接定位方法。

尽管现有研究取得阶段性进展,但仍存在技术瓶颈:Retinex 类算法在提升全局光照均衡性的同时,面临计算复杂度高、噪声敏感性强及参数调节依赖场景经验等问题,难以满足工业实时性需求;传统 PnP 算法直接求解线性模型,对图像噪声与非理想视角的鲁棒性不足,易导致重投影误差累积。针对上述挑战,本文提出一种新的基于单目视觉的机械臂定位与抓取算法:首先采用对比度受限自适应直方图均衡化(CLAHE)算法消除光照干扰,然后在 PnP 算法框架下引入非线性优化策略(Levenberg-Marquardt, L-M),通过最小化重投影误差实现目标位置和姿态的精确估计。

## 1 算法框架

本文基于单目相机与标志检测系统 AprilTag,提出五自由度机械臂定位与抓取算法,算法框架如图 1 所示。在视觉感知环节,单目相机采集附有 AprilTag 码的目标图像,实时传输至上位机。上位机对图像依次进行灰度化降维、滤波降噪及 CLAHE 对比度增强等预处理。在定位阶段,算法解码 AprilTag 码获取标签的四个角点,先利用 PnP 算法得出初始位姿,再引入 L-M 算法优化重投影误差,结合相机标定参数与手眼标定关系,算出目标物体在机械臂基坐标系下的三维坐标。在运动控制阶段,算法依机械臂运动学模型规划关节角度,生成控制指令,由舵机驱动机械臂,精准抓取目标物体。

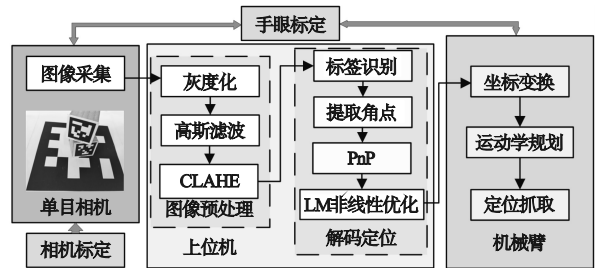


图 1 单目视觉机械臂定位与抓取算法框架

Fig. 1 Algorithmic framework for positioning and grasping of monocular vision robotic arms

## 2 视觉标定方法

### 2.1 相机标定

相机成像几何坐标系如图 2 所示,主要涉及像素坐标系  $O_{px} - uv$ 、图像坐标系  $O_i - xy$ 、相机坐标系  $O_c - X_c Y_c Z_c$  以及世界坐标系  $O_w - X_w Y_w Z_w$  之间点的坐标变换。

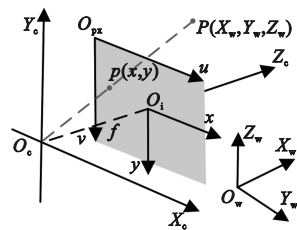


图 2 相机成像几何坐标系

Fig. 2 Camera geometry diagram

世界坐标系中的点  $P(X_w, Y_w, Z_w)$  为真实坐标,其对应的成像平面对称面的像点为图像坐标

系中的点  $p(x,y)$ ,  $Pp$  延长线上的点  $O_c$  为相机光心, 其中相机焦距为  $f$ 。

在进行相机标定时, 采用已知规格的棋盘格标定板作为标定工具。通过对棋盘格标定板拍摄多张不同角度的图像, 进而求解相关参数。相机内参矩阵涵盖了相机的焦距以及光心位置信息, 而外参则借助旋转矩阵  $R$  和平移向量  $t$  精准描述相机坐标系与世界坐标系之间的相对位姿关系。标定公式如下。

$$Z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \quad (1)$$

式中:  $dx$  和  $dy$  分别表示像素在  $x$ 、 $y$  轴方向上的物理尺寸;  $u_0$  和  $v_0$  为图像坐标系原点  $O_i$  在像素坐标系下的坐标;  $Z_c$  为光心到目标点的距离; 坐标点  $(u, v)$  是世界坐标系下的点  $P(X_w, Y_w, Z_w)$  所对应的像素坐标。

单目相机标定采用张正友标定法<sup>[17]</sup>, 重投影误差情况如图 3 所示, 结果表明重投影误差为 0.05 像素, 符合实验预期, 标定结果如表 1 所示。  $f_x$ 、 $f_y$  分别为  $x$ 、 $y$  轴方向焦距的长度。

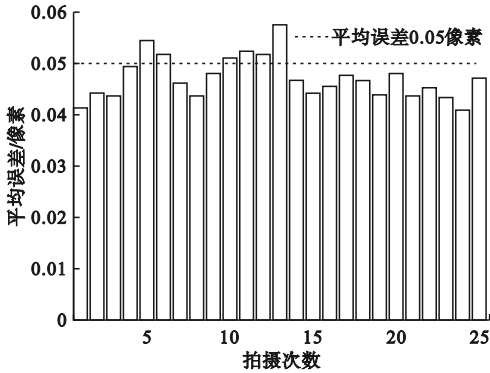


图 3 重投影误差

Fig. 3 Reprojection error

表 1 相机内参

Table 1 Camera Intrinsic Parameters

$f_x$ /像素	$f_y$ /像素	$u_0$ /像素	$v_0$ /像素
539.36	539.62	329.72	236.02

## 2.2 手眼标定

手眼标定旨在确定机器人末端执行器与相机之间的坐标映射关系, 通过标定过程获取手眼变换矩阵, 以此达成从相机坐标系到机器人坐标系的精确转换。相机与机械臂末端相对位置布局通

常存在眼在手上 (eye-in-hand) 和眼在手外 (eye-to-hand) 两种典型方式。本文将相机稳固安装于机械臂末端, 采用 eye-in-hand 构型开展标定工作, 其对应的手眼标定模型如图 4 所示。

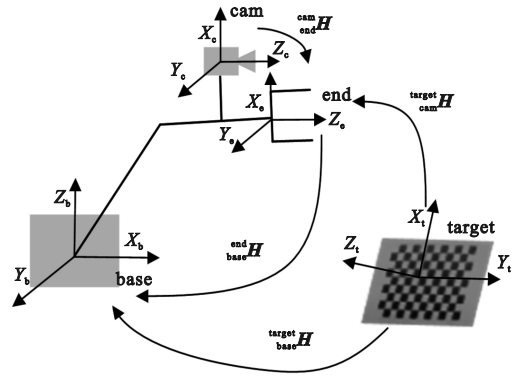


图 4 手眼标定模型

Fig. 4 Hand-eye calibration model

图 4 中,  ${}_{end}^{cam}H$  代表相机坐标系  $cam$  与末端执行器坐标系  $end$  之间的变换矩阵, 此即需通过手眼标定求解的手眼矩阵;  ${}_{target}^{cam}H$  为标定板坐标系  $target$  与相机坐标系  $cam$  的变换矩阵;  ${}_{end}^{base}H$  是基坐标系  $base$  与末端执行器坐标系  $end$  的变换矩阵;  ${}_{base}^{target}H$  则表示标定板坐标系  $target$  与基坐标系  $base$  的变换矩阵。

本文通过采集多组标定板图像, 并同步记录对应时刻机械臂末端执行器的位姿信息进行手眼标定。针对  $n$  幅标定板图像, 手眼标定对应的数学表达式为

$${}_{base}^{target}H = {}_{target}^{cam}H^{-1} \cdot {}_{end}^{cam}H \cdot {}_{end}^{base}H^{-1} \quad (2)$$

式中:  ${}_{target}^{cam}H$  可依据所拍摄的标定板图像, 进行单应性矩阵分解进行求解;  ${}_{end}^{base}H$  能够从机械臂末端执行器的姿态数据直接获取。  ${}_{base}^{target}H$  在整个标定过程中虽为未知量, 通过维持标定板与机械臂之间相对位置恒定, 可保证  ${}_{base}^{target}H$  在多组采样数据中保持不变。针对  $n$  幅标定板图像, 存在

$$\begin{cases} {}_{target}^{cam}H_2^{-1} \cdot {}_{end}^{cam}H \cdot {}_{end}^{base}H_2^{-1} = {}_{target}^{cam}H_1^{-1} \cdot {}_{end}^{cam}H \cdot {}_{end}^{base}H_1^{-1} \\ {}_{target}^{cam}H_3^{-1} \cdot {}_{end}^{cam}H \cdot {}_{end}^{base}H_3^{-1} = {}_{target}^{cam}H_2^{-1} \cdot {}_{end}^{cam}H \cdot {}_{end}^{base}H_2^{-1} \\ \dots \\ {}_{target}^{cam}H_{n-1}^{-1} \cdot {}_{end}^{cam}H \cdot {}_{end}^{base}H_{n-1}^{-1} = {}_{target}^{cam}H_n^{-1} \cdot {}_{end}^{cam}H \cdot {}_{end}^{base}H_n^{-1} \end{cases} \quad (3)$$

用基于对偶四元数<sup>[18]</sup>方法求解式 (3), 得到手眼矩阵  ${}_{end}^{cam}H$ , 手眼标定采集过程示例如图 5 所示, 手眼标定参数如表 2 所示, 其中  $x_1$ 、 $y_1$ 、 $z_1$  分别为相机坐标系相对于末端坐标系的位移,  $R_x$ 、 $R_y$ 、 $R_z$  分别为相机坐标系绕末端执行器坐标系坐标轴

旋转的角度。

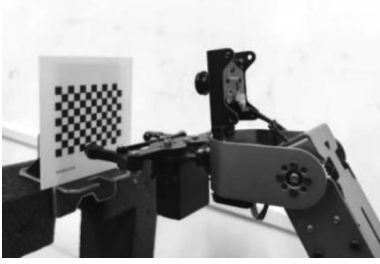


图5 采集过程示例

Fig. 5 Collection process example

表2 手眼标定参数

Table 2 Hand-eye calibration parameters

$x_1/\text{mm}$	$y_1/\text{mm}$	$z_1/\text{mm}$	$R_x/(\text{°})$	$R_y/(\text{°})$	$R_z/(\text{°})$
-112.92	-10.67	48.92	-90.82	0.31	-90.17

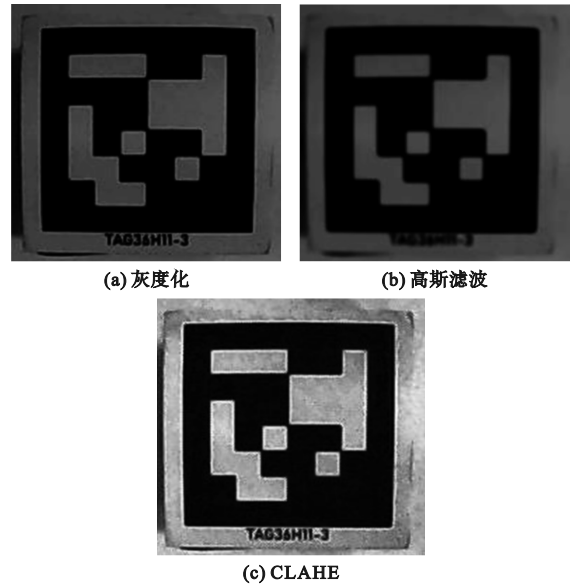


图6 图片处理结果

Fig. 6 Images processing results

### 3 位姿估计

#### 3.1 图像预处理

针对复杂光照条件下 AprilTag 目标检测难题,本文采用灰度化、滤波及光照均衡化等图像处理技术对采集的图像进行预处理,旨在提升标志检测的准确性与鲁棒性。

1) 灰度化:将彩色图像中的  $R$ 、 $G$ 、 $B$  三通道信息按照特定权重进行线性组合,转化为理想的单通道灰度图像,不仅能够有效减少数据量,降低后续计算复杂度,还能较好地保留图像中的结构与边缘信息<sup>[19]</sup>。经灰度化处理后的图像如图 6(a) 所示,该过程去除了颜色信息对定位的潜在干扰,显著提高了计算效率。图像中每个像素点的灰度值计算表达式为

$$G_{\text{ray}}(x, y) = 0.299R(x, y) + 0.587G(x, y) + 0.144B(x, y) \quad (4)$$

式中: $G_{\text{ray}}(x, y)$  表示图像在像素点  $(x, y)$  处的灰度值; $R(x, y)$ 、 $G(x, y)$  以及  $B(x, y)$  分别表示彩色图像在像素点  $(x, y)$  处红、绿、蓝通道的强度值。

2) 高斯滤波:图像经灰度化处理,仍可能存在噪声,本文运用高斯滤波技术对图像进行平滑处理<sup>[19]</sup>。处理后的图像如图 6(b) 所示,高斯滤波能够有效滤除高频噪声,同时较好地保留 AprilTag 标志的边缘信息。

3) CLAHE:完成滤波操作后,为进一步提升 AprilTag 标志区域在图像中的可视性,本文采用 CLAHE 方法<sup>[20]</sup>对图像实施对比度增强处理。该方法将图像划分为多子块,在各子块内分别独立

开展直方图均衡化。在此过程中,对直方图的分布施加限制,以避免局部区域因对比度过度增强而引入新的噪声。每个子块  $j$  内原始像素值  $H(j)$  的裁剪表达式为

$$S(j) = \min(H(j), \frac{M \times N}{P} \times L) \quad (5)$$

式中: $S(j)$  表示经 CLAHE 处理后的像素值; $L$  是预先设定的裁剪阈值; $M$  和  $N$  分别指该子块在水平与垂直方向上的像素个数; $P$  代表该子块所对应的灰度级。

完成裁剪后,子块直方图重新分配并均衡化,再通过插值算法融合各子块结果,生成完整输出图像。CLAHE 方法在复杂光照下能有效增强 AprilTag 标志区域对比度,处理后的图像中标志边缘和特征点更突出。如图 6(c) 所示。

复杂光照下,灰度化消除颜色干扰、统一亮度,高斯滤波抑制高频噪声与光照波动,CLAHE 自适应增强局部对比度,三者协同改善光照不均,提升特征信噪比、锐化边缘,大幅提高了 AprilTag 检测的精度与鲁棒性。

#### 3.2 基于非线性优化的 AprilTag 位姿估计

AprilTag 呈现为一个正方形图案,其内部由黑白正方形色块依据特定规则排列组合而成,这些色块蕴含相应的编码信息。借助 AprilTag 图像处理算法<sup>[21]</sup>,可在输入图像内迅速完成对 AprilTag 标志信息的搜寻、检测及识别,提取标志四个角点的图像坐标,为后续进行位姿估计奠定基础。

AprilTag 位姿估计主要运用 PnP 算法,该算

法基于已知的三维点及其在二维图像中的投影精确求解标志在相机坐标系中的位置和姿态。图2所示为相机几何模型(详见2.1 相机标定), PnP算法以标记四个角点的三维坐标( $X_w, Y_w, Z_w$ )及其在图像中的像素坐标( $u, v$ )为输入数据,同时结合相机内参矩阵 $K$ ,计算得出相机相对于标志的旋转矩阵 $R$ 和平移向量 $t$ 。三维坐标与二维投影坐标之间的映射关系为

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K [R \quad t] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (6)$$

式中 $s$ 为尺度因子。

通过最小化重投影误差, PnP算法能够估计出相机在世界坐标系下的位姿。鉴于PnP算法是基于线性最小二乘法求解,其输出的位姿解极易受到噪声干扰、标定误差以及视角变化等因素的影响,会降低结果的精度。为此,本文构建了三维点投影至像素坐标平面所产生的重投影误差,并采用L-M法开展非线性优化工作。优化函数定义如下。

$$E = \|\mathbf{r}\|^2 = \sum_{k=1}^N \|\mathbf{p}_k - (K[R \cdot \mathbf{P}_k + \mathbf{t}])\|^2 \quad (7)$$

式中: $E$ 表示角点重投影误差的平方和; $\mathbf{r}$ 为残差向量,由角点重投影误差构成; $\mathbf{p}_k$ 为第 $k$ 个角点在二维像素坐标系下的坐标; $\mathbf{P}_k$ 为第 $k$ 个角点在三维空间中的坐标; $N$ 代表角点的总数。通过对上述参数的调整可最小化重投影误差。

L-M算法优化的目标为位姿参数 $R$ 和 $t$ ,迭代表达式为

$$\begin{bmatrix} R_{l+1} \\ t_{l+1} \end{bmatrix} = \begin{bmatrix} R_l \\ t_l \end{bmatrix} - (J^T J + \lambda I)^{-1} J^T \mathbf{r} \quad (8)$$

式中: $l$ 代表迭代次数; $J$ 是误差函数相对于位姿参数的雅可比矩阵; $I$ 为单位矩阵; $\lambda$ 为阻尼因子,其取值对算法特性影响显著。当 $\lambda$ 取值较大时,该算法特性趋近于梯度下降法,保证算法的收敛性,确保在复杂的参数空间中稳定地向最优解逼近;当 $\lambda$ 取值较小时,算法则趋近于高斯-牛顿方法,可加速算法的收敛速度,在接近最优解时能够更快地完成参数优化。

L-M算法通过迭代方式对相机位姿参数进行动态调整,逐步实现图像中角点重投影误差的最小化。当重投影误差降低至0.05像素以下,或迭

代次数达到100次时,算法终止优化过程,进而输出满足收敛条件的高精度相机位姿估计结果。鉴于非线性优化的运算速度以及最终结果对初始值具有较强的敏感性,本文将PnP方法所得的估计结果作为初始位姿进行迭代计算,确保L-M算法能够更高效、稳定地收敛至最优解。

## 4 机械臂定位与抓取系统

为验证基于单目视觉的机械臂定位及抓取算法的可行性,本文搭建了一套以五自由度机械臂为核心的定位与抓取系统。

### 4.1 建立运动学模型

本文采用D-H参数化方法<sup>[22-23]</sup>对机械臂进行建模分析,将其简化为由连续旋转关节连接的刚性连杆,并为每个连杆 $i(i=1,2,3,4,5)$ 建立一个连杆坐标系 $O_i$ 。机械臂的模型示意图如图7所示,相应的D-H参数信息列于表3。对于每一个连杆 $i$ : $\theta_i$ 代表关节角,为绕 $Z_i$ 轴从 $X_{i-1}$ 旋转至 $X_i$ 的转角; $d_i$ 为连杆偏移量,为绕 $Z_i$ 轴从 $X_{i-1}$ 移动至 $X_i$ 的距离; $\alpha_{i-1}$ 表示连杆扭转角,为绕 $X_{i-1}$ 轴从 $Z_{i-1}$ 旋转至 $Z_i$ 的转角; $a_{i-1}$ 为连杆长度,为沿 $X_{i-1}$ 轴从 $Z_{i-1}$ 移动至 $Z_i$ 的距离。

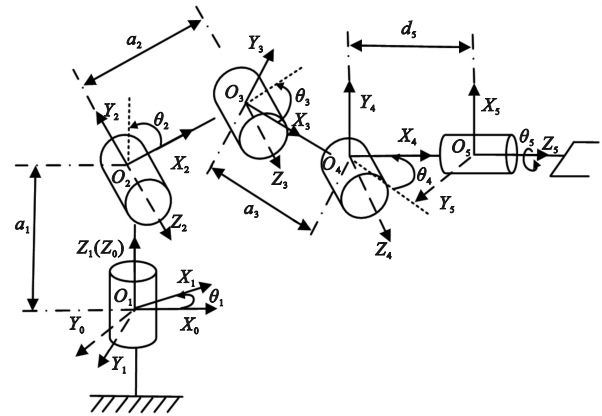


图7 机械臂结构

Fig.7 Robotic arm structure

表3 D-H参数

Table 3 D-H parameters

$i$	$d_i/\text{mm}$	$a_{i-1}/\text{mm}$	$\alpha_{i-1}/(^{\circ})$
1	0	0	0
2	0	103	-90
3	0	129	0
4	0	129	0
5	171	0	-90

正运动学借助构建连杆坐标系  $O_{i-1}$  与  $O_i$  之间的齐次变换矩阵  ${}^{i-1}T_i$ , 达成从关节角度到机械臂末端执行器位姿的映射。机械臂的正运动学表达式为

$${}^0_5T = {}^0_1T_1 {}^1_2T_2 {}^2_3T_3 {}^3_4T_4 {}^4_5T_5 T = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \quad (9)$$

式中  $\mathbf{R}$  为  $3 \times 3$  旋转矩阵。

逆运动学旨在依据机械臂末端的期望位姿来求解相应的关节角度。本文选用了计算效率较高的几何法<sup>[24]</sup>。 $\theta_1$  由世界坐标系目标点  $P(X_w, Y_w, Z_w)$  在  $X_w Y_w$  平面的投影直接确定为

$$\theta_1 = \arctan(Y_w/X_w) \quad (10)$$

鉴于各连杆的运动始终限定于由  $\theta_1$  所确定的平面内, 本文将机械臂的运动模型简化为二维平面模型, 其二维结构示意图如图 8 所示。图中,  $J_i$  为连杆  $i$  连接关节;  $\theta_i$  为第  $i$  个关节的角度;  $a_i$  为第  $i$  个连杆的长度;  $H/L$  分别为关节  $J_4$  的高度、长度;  $\varphi$  为俯仰角;  $\sigma$  为  $a_3$  和  $a_4$  组成三角形的内角。

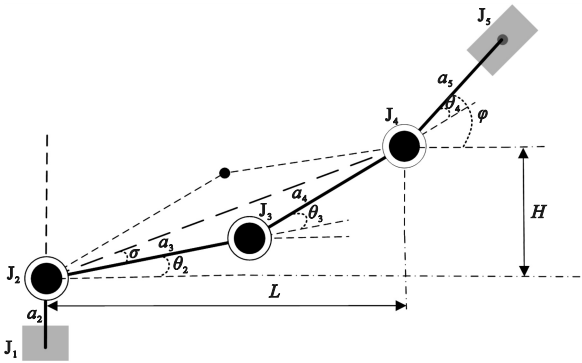


图 8 二维结构图

Fig. 8 2D structure diagram

根据图 8 几何关系可得

$$\theta_3 = \arccos\left(\frac{L^2 + H^2 - a_3^2 - a_4^2}{2a_3a_4}\right) \quad (11)$$

$$\theta_2 = \begin{cases} \arctan(H/L) + \sigma, \theta_3 \geq 0^\circ \\ \arctan(H/L) - \sigma, \theta_3 < 0^\circ \end{cases} \quad (12)$$

$$\sigma = \arccos\left(\frac{a_4^2 - (L^2 + H^2) - a_3^2}{-2a_3\sqrt{L^2 + H^2}}\right) \quad (13)$$

$$\theta_4 = \varphi - \theta_2 - \theta_3 \quad (14)$$

为确定唯一解, 选取  $\varphi$  的最小变化量作为优化条件, 同时根据横滚角  $\psi$  直接求得  $\theta_5 = \psi$ 。

## 4.2 实验及结果分析

为验证算法的有效性, 本文将分辨率为  $640 \times 480$  像素的单目相机固定在机械臂末端执行器上, 确保执行器不会对相机的视野造成遮挡。同

时, 选取表面贴有边长为 25 mm AprilTag 标志的木块, 作为测距与抓取实验的目标对象。

为验证本文提出的光照预处理结合基于 L-M 非线性优化重投影误差的 PnP 方法的有效性, 将该方法与传统 PnP 方法、仅加入 CLAHE 处理的 PnP 方法展开对比研究。在实验环节, 分别在 200、300、400 mm 三个不同距离下进行误差测距实验。实验参数详见表 4, 实验结果如图 9 ~ 11 所示。

表 4 图像预处理参数设定

Table 4 Images preprocessing parameters settings

高斯滤波		CLAHE		
$k$	$\sigma$	$L$	$M$	$N$
4	2.5	4	80	60

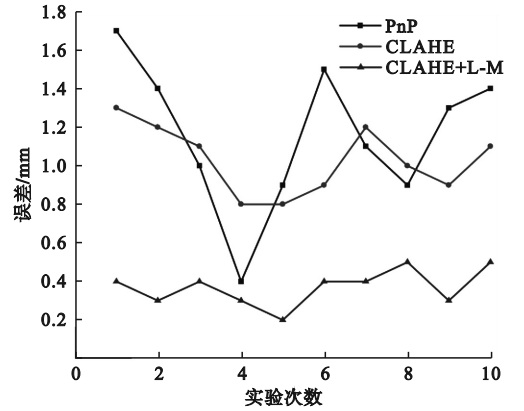


图 9 200 mm 测距误差对比图

Fig. 9 Comparison chart of 200 mm ranging error

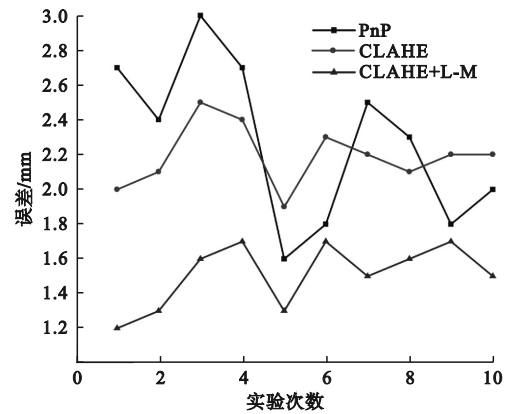


图 10 300 mm 测距误差对比图

Fig. 10 Comparison chart of 300 mm ranging error

从图 9 ~ 11 可以看出: 传统 PnP 测距方法在 200、300、400 mm 三个距离下的平均误差分别为 1.16、2.28 和 4.55 mm; 引入光照预处理后, 相应距离下的平均误差分别降至 1.03、2.19 和 4.42 mm;

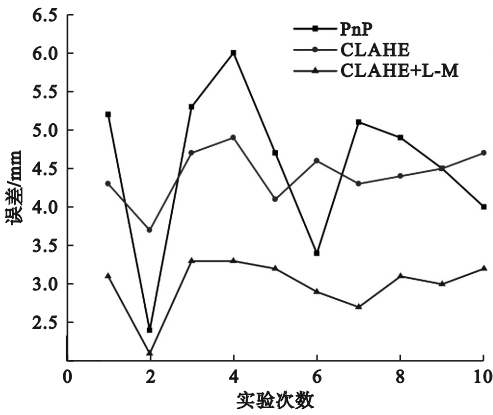


图 11 400 mm 测距误差对比图

Fig. 11 Comparison chart of 400 mm ranging error

当同时采用光照预处理与 L-M 非线性优化方法时,平均误差进一步降低至 0.37、1.51 和 2.99 mm。实验结果表明:相较于传统 PnP 方法,加入光照预处理能够有效减小系统测距的误差波动,提升系统的鲁棒性;而在此基础上进一步引入 L-M 非线性优化,可显著降低测距误差,大幅提高测距精度。

为验证本文构建的视觉标定与运动学模型在目标定位实际应用中的可行性,本研究借助单目视觉技术对 AprilTag 标志实施定位操作。通过将坐标变换计算得出的位置与末端夹爪实际位置进行对比,进而计算二者之间的相对误差,相关实验数据汇总于表 5。

表 5 坐标测量数据及误差

Table 5 D-H parameters mm

序号	实际位置	计算位置	位置偏差
1	(65.4, -12.8, 132.3)	(68.5, -11.4, 135.5)	(3.1, 1.4, 3.2)
2	(85.9, 10.5, 64.5)	(87.2, 12.4, 67.6)	(1.3, 1.9, 3.1)
3	(137.4, 22.1, 32.5)	(135.8, 23.4, 29.8)	(-1.6, 1.3, -2.7)
4	(158.7, 106.4, 87.3)	(157.4, 108.1, 90.1)	(-1.3, 1.7, 3.8)
5	(178.9, 164.1, 231.0)	(181.0, 161.0, 228.1)	(2.1, -3.1, -2.9)
6	(206.3, -39.2, 107.3)	(204.1, -40.4, 110.9)	(-2.2, -1.2, 3.6)
7	(225.3, 44.7, 192.6)	(228.6, 46.4, 195.5)	(2.3, 1.7, 2.9)

从表 5 可以看出,系统在  $X_w$  轴的误差范围处于  $-2.2 \sim 3.1$  mm 区间,  $Y_w$  轴误差范围为  $-3.1 \sim 1.9$  mm,  $Z_w$  轴误差范围为  $-2.7 \sim 3.8$  mm。实验结果表明,系统在三维空间的定位误差分布呈现均衡态势。在世界坐标系  $X_w$ 、 $Y_w$ 、 $Z_w$  三个方向上,误差绝

对值基本被有效控制在 4 mm 以内,未呈现出明显的方向性偏向,能够满足机械臂抓取作业的精度要求。

为验证本文算法于不同环境下的适用性,分别在室内自然光照、室内人工光照以及室内弱光照环境中开展机械臂抓取实验。每种环境下的抓取实验次数均超过 20 次。抓取过程的实际情况如图 12 所示,各环境的光照度以及对应的抓取实验详细数据列于表 6。

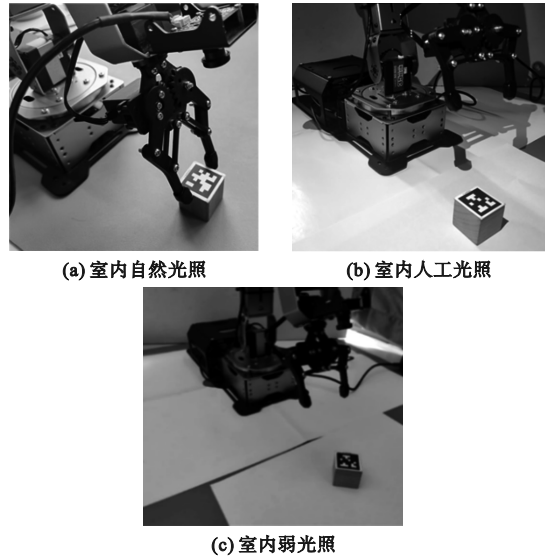


图 12 不同光照环境下抓取实验图

Fig. 12 Grasping experiments under different lights

表 6 抓取实验数据

Table 6 Grabbing experimental data

	平均光照度/Lux	总次数	成功次数	成功率/%
自然光照	651	20	20	100
人工光照	1 020	20	18	90
弱光环境	65	20	19	95

从表 6 可知,在各类光照环境下,本文所提算法均能确保机械臂的抓取成功率维持在 90% 以上,表明机械臂对于不同光照环境展现出了良好的适应性与鲁棒性。

## 5 结论

本文提出基于单目视觉的机械臂定位抓取算法,并在五自由度机械臂上进行验证。系统通过相机、手眼标定构建坐标系关系,对采集图像依次做灰度化、高斯滤波、自适应直方图均衡处理,结合基于非线性优化的 AprilTag 位姿估计方法,可精准识别、定位附有 AprilTag 标志的目标物体,实现机械臂精确控制与抓取。

实验结果表明,该系统克服了单目视觉在深度信息获取方面的固有局限。抓取系统对不同光照环境展现出良好的适应性和鲁棒性,较好地增强了机械臂在复杂环境下的工作效能。对比测距实验与定位实验中世界坐标系下  $Z_w$  轴方向的误差数据发现,定位误差相对高于测距误差。经分析,此现象可能归因于手眼标定矩阵的固有偏差以及舵机控制过程中存在的累计误差。

## 参考文献 (References):

[1] 张孟旭,高向川,尹丽楠,等. 基于机器视觉的机械臂抓取系统设计[J]. 计算机应用与软件,2024,41(8):22-27.  
ZHANG M X, GAO X C, YIN L N, et al. Design of robot arm gripping system based on machine vision[J]. Computer Applications and Software, 2024, 41(8): 22-27. (in Chinese)

[2] 杨学文,王楠. 视觉驱动机械臂自主作业技术综述[J]. 南京信息工程大学学报(自然科学版),2021,13(3):281-290.  
YANG X W, WANG N. Review on autonomous operation technology of vision-based manipulator[J]. Journal of Nanjing University of Information Science & Technology(Natural Science Edition), 2021, 13(3): 281-290. (in Chinese)

[3] SHI Y G, JIN S T, ZHAO Y M, et al. Lightweight force-sensing tomato picking robotic arm with a "global-local" visual servo[J]. Computers and Electronics in Agriculture, 2023, 204: 107549.

[4] RAKSHIT A, PRAMANICK S, BAGCHI A, et al. Autonomous grasping of 3-D objects by a vision-actuated robot arm using Brain-Computer Interface[J]. Biomedical Signal Processing and Control, 2023, 84: 104765.

[5] ABBAS M, NARAYAN J, DWIVEDY S K. A systematic review on cooperative dual-arm manipulators: modeling, planning, control, and vision strategies[J]. International Journal of Intelligent Robotics and Applications, 2023, 7(4): 683-707.

[6] 何敏军,张晓玲,张稀柳,等. 基于双目立体视觉的平面工件识别和定位研究[J]. 激光杂志,2023,44(7):199-204.  
HE M J, ZHANG X L, ZHANG X L, et al. Research on plane workpiece recognition and localization based on binocular stereo vision[J]. Laser Journal, 2023, 44(7): 199-204. (in Chinese)

[7] 江小辉,李翔翔,孙翼飞,等. 基于双目视觉距离误差测量的工业机器人运动学标定方法[J]. 计量学报,2024,45(10):1470-1479.  
JIANG X H, LI X X, SUN Y F, et al. Kinematic calibration of industrial robots based on binocular vision distance error measurement[J]. Acta Metrologica Sinica, 2024, 45(10): 1470-1479. (in Chinese)

[8] 毛成林,于瑞强,宋爱国. 一种结合 TCP 标定的深度相机手眼标定方法[J]. 仪器仪表学报,2023,44(3):280-286.  
MAO C L, YU R Q, SONG A G. A hand-eye calibration method of depth camera combined with TCP calibration[J]. Chinese Journal of Scientific Instrument, 2023, 44(3): 280-286. (in Chinese)

[9] 曹泽强,孙英惠,洪琼,等. AprilTag 图像识别应用于高差传递测量方法研究[J]. 测绘通报,2023(3):27-32.  
CAO Z Q, SUN Y H, HONG Q, et al. AprilTag image identification application to high distance transfer measurement method[J]. Bulletin of Surveying and Mapping, 2023(3):

27-32. (in Chinese)

[10] KALAITZAKIS M, CAIN B, CARROLL S, et al. Fiducial markers for pose estimation: overview, applications and experimental comparison of the artag, AprilTag, aruco and stag markers[J]. Journal of Intelligent & Robotic Systems, 2021, 101: 1-26.

[11] LIU J D, ERGAN S, ZHANG Q L. An AprilTags-based approach for progress monitoring and quality control in modular construction[J]. Buildings, 2024, 14(7): 2252.

[12] 苗作华,朱良建,赵成诚,等. 基于 Retinex 的低光照车位图像增强[J]. 汽车工程,2023,45(6):989-996.  
MIAO Z H, ZHU L J, ZHAO C C, et al. Image enhancement of low-light parking space based on retinex[J]. Automotive Engineering, 2023, 45(6): 989-996. (in Chinese)

[13] NUSANTIKA N R, XIAO J, HU X G. An enhanced multi-scale retinex, oriented FAST and rotated BRIEF(ORB), and scale-invariant feature transform(SIFT) pipeline for robust key point matching in 3D monitoring of power transmission line icing with binocular vision[J]. Electronics, 2024, 13(21): 4252.

[14] 吕伏,崔向燕,刘铁. 亮度均衡与边缘强化的多尺度低照度图像增强算法[J]. 激光与光电子学进展,2024,61(12):1237010.  
LÜ F, CUI X Y, LIU T. Multiscale low-light image enhancement algorithm with brightness equalization and edge enhancement algorithm[J]. Laser & Optoelectronics Progress, 2024, 61(12): 1237010. (in Chinese)

[15] 张天意,许一祎,冯莉芳,等. 基于二维码的室内视觉定位系统研究[J]. 光学学报,2024,44(9):157-163.  
ZHANG T Y, XU Y Y, FENG L F, et al. Research on indoor visual positioning system based on QR code[J]. Acta Optica Sinica, 2024, 44(9): 157-163. (in Chinese)

[16] 余泽康,张秋菊,吕青,等. 基于 QR 码的装配机器人视觉增强定位方法研究[J]. 现代制造工程,2022(9):27-33,8.  
YU Z K, ZHANG Q J, LÜ Q, et al. Research on the vision enhanced positioning method of assembly robot based on QR code[J]. Modern Manufacturing Engineering, 2022(9): 27-33, 8. (in Chinese)

[17] ZHANG Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.

[18] 邓仕超,梅枫,高兴宇,等. 单目机器人手眼标定研究[J]. 组合机床与自动化加工技术,2021(12):53-57.  
DENG S C, MEI F, GAO X Y, et al. Research on hand-eye calibration of monocular robot[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2021(12): 53-57. (in Chinese)

[19] 罗小青,胡荣. 一种采用改进 Canny 算子的机械臂识别与定位算法设计[J]. 现代电子技术,2024,47(13):27-31.  
LUO X Q, HU R. Design of robotic arm recognition and positioning algorithm using improved Canny operator[J]. Modern Electronics Technique, 2024, 47(13): 27-31. (in Chinese)

[20] 苗作华,张立,徐厚友,等. 基于 CLAHE-PCA 的矿井低照度图像增强研究[J]. 金属矿山,2024(6):165-172.  
MIAO Z H, ZHANG L, XU H Y, et al. Research on mine low illumination image enhancement based on CLAHE-PCA[J]. Metal Mine, 2024(6): 165-172. (in Chinese)

[21] KROGIUS M, HAGGENMILLER A, OLSON E. Flexible layouts for fiducial tags [C]//2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Macao, China: IEEE, 2019: 1898-1903.

- [22] 达悦生,郑楚悦,孙茂荣. 机器人运动学模型建立的改进DH方法及正反解计算[J]. 机械与电子,2019,37(10):72-75,80.  
DA Y S,ZHENG C Y,SUN M R. Improved DH method and forward and inverse solution calculation for robot kinematics model establishment[J]. Machinery & Electronics,2019,37(10):72-75,80. (in Chinese)
- [23] 刘拓,李艳杰,卜春光,等. 滑雪跳台修整机械臂轨迹控制方法[J]. 沈阳理工大学学报,2023,42(5):68-75.  
LIU T,LI Y J,BU C G, et al. Trajectory control method for trimming manipulator of ski jumping platform[J]. Journal of Shenyang Ligong University,2023,42(5):68-75. (in Chinese)
- [24] 张宇豪,庞辰骅,宋子韩,等. 基于OpenMV的五自由度机械臂控制方法研究[J]. 机床与液压,2024,52(5):1-7.  
ZHANG Y H,PANG C H,SONG Z H, et al. Research on control method of five degree of freedom manipulator based on OpenMV[J]. Machine Tool & Hydraulics,2024,52(5):1-7. (in Chinese)
- (责任编辑:和晓军)
- (上接第13页)
- [13] LI D H,LIU B K,GAO Q, et al. HAF-YOLO: an efficient network for the detection of safety equipment for working at heights[J]. Signal, Image and Video Processing, 2025, 19(5):401.
- [14] 付丽君,赵晨兵,杨青,等. 基于图像的集合型电机轴承故障诊断方法[J]. 沈阳理工大学学报,2020,39(5):8-12.  
FU L J,ZHAO C B,YANG Q, et al. An ensemble fault diagnosis method for motor bearings based on image[J]. Journal of Shenyang Ligong University, 2020, 39(5):8-12. (in Chinese)
- [15] PAING M P,PINTAVIROOJ C. Adenoma dysplasia grading of colorectal polyps using fast Fourier convolutional ResNet (FFC-ResNet)[J]. IEEE Access,2023,11:16644-16656.
- [16] 冯迎宾,刘文泽. 基于HPDE-YOLO的钢材表面缺陷检测算法[J]. 沈阳理工大学学报,2025,44(1):31-38.  
FENG Y B,LIU W Z. Steel surface defect detection algorithm based on HPDE-YOLO[J]. Journal of Shenyang Ligong University,2025,44(1):31-38. (in Chinese)
- [17] TAKENAGA T,HANAOKA S,NOMURA Y, et al. Development and evaluation of an integrated liver nodule diagnostic method by combining the liver segment division and lesion localization/classification models for enhanced focal liver lesion detection[J]. Radiological Physics and Technology,2024,17(1):103-111.
- [18] OUYANG D L,HE S,ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]//2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island, Greece: IEEE, 2023:1-5.
- [19] LI H L,LI J,WEI H B, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures[J]. Journal of Real-Time Image Processing,2024,21(3):62.  
(责任编辑:宋颖韬)