

本文引用格式：王欢,徐秀进,王红军,等.基于双目 eye in hand 系统的多角度树干位姿估计方法[J].自动化与信息工程, 2023,44(5):38-45.

WANG Huan, XU Xiujin, WANG Hongjun, et al. Multi-angle tree trunk pose estimation method based on binocular eye in hand system [J]. Automation & Information Engineering, 2023,44(5):38-45.

基于双目 eye in hand 系统的多角度树干位姿估计方法*

王欢¹ 徐秀进¹ 王红军¹ 温志庆²

(1.华南农业大学工程学院, 广东 广州 510642

2.季华实验室智能机器人工程研究中心, 广东 佛山 528200)

摘要: 针对采摘机器人自主行走导航过程中, 难以准确定位其与果树之间的相对位置, 难以准确估计果树树干姿态的问题, 提出基于双目 eye in hand 系统的多角度树干位姿估计方法。利用 YOLOv5 深度学习方法与半全局块匹配算法识别树干并生成局部点云; 利用半径滤波和体素滤波减少树干点云数据; 利用闭环式手眼标定方法对双目 eye in hand 系统进行标定, 并对同一树干多角度相机位置的点云数据进行拼接; 利用随机抽样一致 (RANSAC) 算法与无约束最小二乘法估计并优化树干的位置和姿态, 获取树干的圆柱体参数。通过对 30 幅标定板图像进行实验, 闭环式手眼标定方法的平均欧式误差为 3.717 7 mm; 采用半径滤波和体素滤波可减少 98.470% 的点云数据; 采用 RANSAC 算法、圆柱体估计算法拟合树干点云数据, 得到圆柱体的半径 $r = 41.277 1$ mm, $R_{MAE} = 2.571 56$ mm, $R_{RMSE} = 2.989 36$ mm; 无约束最小二乘法优化后 $r = 39.402 8$ mm, $R_{MAE} = 1.989 55$ mm, $R_{RMSE} = 2.465 88$ mm。该文通过对双目 eye in hand 系统进行标定, 建立坐标系转换关系, 多角度采集环境信息, 准确定位机器人与果树之间的相对位置, 估计果树树干的姿态。

关键词: eye in hand 手眼标定; 局部立体匹配; 树干圆柱体估计; 多角度点云拼接

中图分类号: S225; TP391.41

文献标志码: A

文章编号: 1674-2605(2023)05-0006-08

DOI: 10.3969/j.issn.1674-2605.2023.05.006

Multi-angle Tree Trunk Pose Estimation Method Based on Binocular Eye in Hand System

WANG Huan¹ XU Xiujin¹ WANG Hongjun¹ WEN Zhiqing²

(1.College of Engineering, South China Agricultural University, Guangzhou 510642, China

2.Intelligent Robot Engineering Research Center, Ji Hua Laboratory, Foshan 528200, China)

Abstract: In the process of autonomous walking and navigation for harvesting robots, it is difficult to accurately locate the relative position between them and the fruit tree, as well as to accurately estimate the posture of the fruit tree trunk. Based on the binocular eye in hand system, YOLOv5 deep learning method and classical stereo matching algorithm are applied to identify tree trunks and generate local point clouds. Radial filtering and voxel filtering are used to reduce the number of point clouds. A closed-loop hand eye calibration method is proposed to calibrate the system and concatenate point cloud data from multi angle camera positions on the same tree trunk, using the random sampling consensus (RANSAC) algorithm and unconstrained least squares method to estimate and optimize the trunk posture, and obtain the trunk cylinder parameters. Through repeated experiments on 30 sets of calibration board images, the average Euclidean error of the closed-loop hand eye calibration method is 3.717 7 mm; Reduce 98.470% of point cloud data by using radius filtering and voxel filtering; The RANSAC cylinder estimation algorithm fitting tree trunk point cloud data yields a cylinder radius of $r = 41.277 1$ mm, $R_{MAE} = 2.571 56$ mm, $R_{RMSE} = 2.989 36$ mm; After unconstrained least squares optimization, $r = 39.402 8$ mm, $R_{MAE} = 1.989 55$ mm, $R_{RMSE} = 2.465 88$ mm. The closed-loop hand eye calibration method proposed in this article calibrates the eye in hand system, establishes coordinate system conversion relationships, collects environmental information from multiple angles, effectively and accurately locates

the relative position between the robot and the fruit tree, and estimates the posture of the fruit tree trunk.

Keywords: eye in hand calibration; local stereo matching; estimation of tree trunk cylinder; multi angle point cloud stitching

0 引言

在野外环境下,通过视觉方法对树干进行精准识别和定位是采摘机器人自主行走的关键^[1-2]。精确确定采摘机器人与树干的相对位置和距离,以及准确估计树干的位置和姿态,可提升导航和采摘的定位精确度,提高采摘机器人的作业效率,降低作业成本^[3-5]。

近年来,树干模型的三维重建和姿态估计一般采用激光雷达和图像 2 种方式。基于激光点云的树干模型三维重建以计算机视觉、计算机图形学为理论基础进行激光点云去噪以及树干骨架重建,但由于激光点云数据采样量较大,给分析和计算带来巨大挑战^[6-7]。通过图像进行树干模型的三维重建常采用双目视觉、多目视觉、深度视觉等方案。如邹湘军团队^[8-9]对荔枝视觉采摘机器人进行了大量地研究,利用双目视觉定位方法估计荔枝定位误差;LIN 等^[10]提出基于深度相机与全卷积网络 (fully convolutional networks, FCN) 深度学习的番石榴识别与姿态估计方法,但深度相机受光照强度影响较大;CHEN 等^[11-12]提出基于四目系统的香蕉果轴三维重建与定位方法,但仅从 2 个角度采集环境中的图像信息,图像信息少且位置固定。OLOFSSON 等^[13]改进基于随机抽样一致 (random sample consensus, RANSAC) 模型对二维圆的拟合算法,提取树干胸径值,但在树干存在倾斜角度时,二维圆拟合误差较大。吴杭彬等^[14]结合布料模拟滤波 (cloth simulation filtering, CSF) 算法与 RANSAC 算法,有效滤除地面数据并拟合树干模型,实现树干胸径大规模提取自动化,但利用激光扫描仪采集图像数据,成本较高,无法适用于机器人行走过程中树干相对位置和姿态的估计。

本文提出基于双目 eye in hand 系统的多角度树干位姿估计方法,以机械臂和普通双目为基础,构建双目 eye in hand 系统进行双目标定、手眼标定;通过 YOLOv5 深度学习方法进行树干识别;利用经典立体匹配算法实现树干的三维重建;利用双目 eye in hand

系统进行多角度点云拼接;利用 RANSAC 算法进行树干姿态估计;利用无约束最小二乘法优化圆柱体参数,实现从多角度对同一树干进行重建,准确估计树干的相对位置和姿态。

1 方法

基于双目 eye in hand 系统的多角度树干位姿估计方法包含手眼标定、局部立体匹配、树干姿态估计等部分,如图 1 所示。

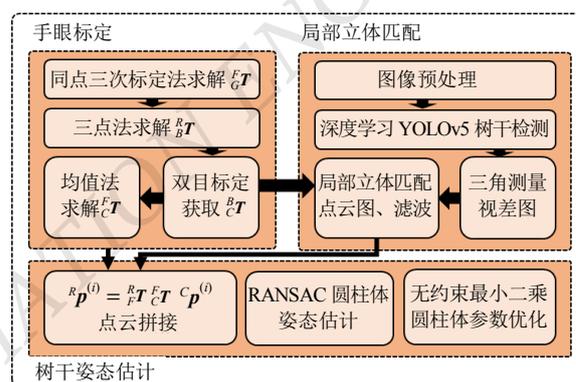


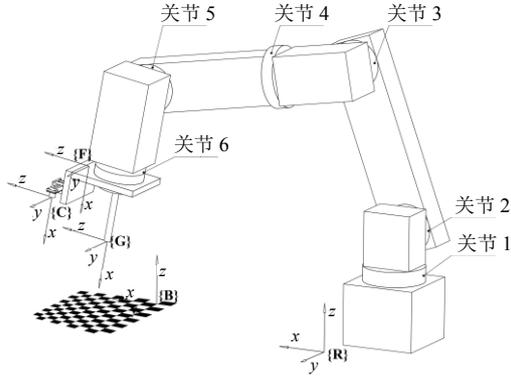
图 1 基于双目 eye in hand 系统的多角度树干位姿估计方法组成图

图 1 中,手眼标定部分建立各个坐标系之间的转换关系,并将相机坐标系下的图像信息转换到机器人坐标系;局部立体匹配部分生成深度学习检测目标的局部点云图;树干姿态估计部分获得优化后树干的近圆柱体模型。

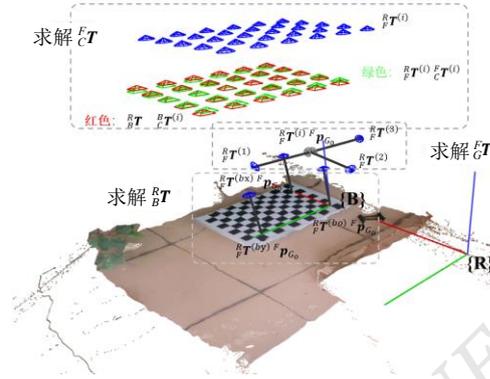
1.1 闭环式手眼标定

手眼系统主要有 2 种:一种是 eye to hand,相机固定,不随机器人运动;另一种是 eye in hand,相机随机器人运动。手眼标定通过建立各个坐标系之间的转换关系,将双目相机拍摄的图像、三维重建后的三维点云数据转换到机器人坐标系下。

在双目 eye in hand 系统中,含有机器人基坐标系 {R}、机器人末端坐标系 {F}、工具坐标系 {G}、标定板坐标系 {B} 和相机坐标系 {C} 等 5 个坐标系,如图 2 所示。



(a) eye in hand 手眼标定系统的坐标系



(b) 闭环式手眼标定步骤分布图

图2 eye in hand 手眼标定系统图

其中，机器人末端坐标系{F}可通过六轴机器人结构计算得到，仅需求解工具坐标系{G}、标定板坐标系{B}和相机坐标系{C}到机器人基坐标系{R}的转换关系。

1.1.1 同点三次标定法求解 ${}^F_G T$

工具坐标系{G}到机器人基坐标系{R}的转换关系 ${}^R_G T$ 为

$${}^R_G T = {}^R_F T {}^F_G T \quad (1)$$

其中 ${}^R_F T$ 可通过机器人结构计算得到，需求解 ${}^F_G T$ 。

${}^F_G T$ 由旋转和平移2部分组成，旋转部分为单位矩阵，平移向量 ${}^F p_{G_0}$ 的求解采用同点三次标定法。工具位于机器人的末端，让工具的尖端从3个不同的角度到达空间中的一点，则工具坐标系{G}尖端点可表示为

$${}^R p_{G_0} = {}^R_F T^{(i)} {}^F p_{G_0} \quad (2)$$

展开为

$$\begin{bmatrix} {}^R x_{G_0} \\ {}^R y_{G_0} \\ {}^R z_{G_0} \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11}^{(i)} & r_{12}^{(i)} & r_{13}^{(i)} & x^{(i)} \\ r_{21}^{(i)} & r_{22}^{(i)} & r_{23}^{(i)} & y^{(i)} \\ r_{31}^{(i)} & r_{32}^{(i)} & r_{33}^{(i)} & z^{(i)} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^F x_{G_0} \\ {}^F y_{G_0} \\ {}^F z_{G_0} \\ 1 \end{bmatrix} \quad (3)$$

建立等式

$$\begin{cases} {}^R_F T^{(1)} {}^F p_{G_0} = {}^R_F T^{(2)} {}^F p_{G_0} \\ {}^R_F T^{(1)} {}^F p_{G_0} = {}^R_F T^{(3)} {}^F p_{G_0} \end{cases} \quad (4)$$

整理得

$$\begin{bmatrix} r_{11}^{(1)} - r_{11}^{(2)} & r_{12}^{(1)} - r_{12}^{(2)} & r_{13}^{(1)} - r_{13}^{(2)} \\ r_{11}^{(1)} - r_{11}^{(3)} & r_{12}^{(1)} - r_{12}^{(3)} & r_{13}^{(1)} - r_{13}^{(3)} \\ r_{21}^{(1)} - r_{21}^{(2)} & r_{22}^{(1)} - r_{22}^{(2)} & r_{23}^{(1)} - r_{23}^{(2)} \\ r_{21}^{(1)} - r_{21}^{(3)} & r_{22}^{(1)} - r_{22}^{(3)} & r_{23}^{(1)} - r_{23}^{(3)} \\ r_{31}^{(1)} - r_{31}^{(2)} & r_{32}^{(1)} - r_{32}^{(2)} & r_{33}^{(1)} - r_{33}^{(2)} \\ r_{31}^{(1)} - r_{31}^{(3)} & r_{32}^{(1)} - r_{32}^{(3)} & r_{33}^{(1)} - r_{33}^{(3)} \end{bmatrix}$$

$$\begin{bmatrix} {}^F x_{G_0} \\ {}^F y_{G_0} \\ {}^F z_{G_0} \end{bmatrix} = \begin{bmatrix} -x^{(1)} + x^{(2)} \\ -x^{(1)} + x^{(3)} \\ -y^{(1)} + y^{(2)} \\ -y^{(1)} + y^{(3)} \\ -z^{(1)} + z^{(2)} \\ -z^{(1)} + z^{(3)} \end{bmatrix} \quad (5)$$

简记为

$$P \begin{bmatrix} {}^F x_{G_0} \\ {}^F y_{G_0} \\ {}^F z_{G_0} \end{bmatrix}^T = Q \quad (6)$$

当 $\text{rank}(P, Q) = \text{rank}(P) \geq 3$ 时，方程有唯一解。

计算得

$${}^F p_{G_0}^T = (P^T P)^{-1} P^T Q \quad (7)$$

由此得到工具坐标系{G}到机器人末端坐标系{F}的变换矩阵 ${}^F_G T$ 为

$${}^F_G T = \begin{bmatrix} E & {}^F p_{G_0} \\ 0 & 1 \end{bmatrix} \quad (8)$$

进而根据公式(1)可得工具坐标系{G}到机器人基坐标系{R}的转换关系 ${}^R_G T$ 。

1.1.2 三点法求解 ${}^R_B T$

采用三点法获取工具尖端点标定板中的3点(原点、x轴上一点、y轴上一点)在机器人基坐标系{R}下

的坐标 ${}^R p_{B_o}$ 、 ${}^R p_{B_x}$ 、 ${}^R p_{B_y}$:

$$\begin{cases} {}^R p_{B_o} = {}^R T^{(bo)} F p_{G_o} \\ {}^R p_{B_x} = {}^R T^{(bx)} F p_{G_o} \\ {}^R p_{B_y} = {}^R T^{(by)} F p_{G_o} \end{cases} \quad (9)$$

式中: ${}^R p_{B_o}$ 为工具尖端点标定板坐标原点, ${}^R T^{(bo)}$ 为机器人末端基于机器人基坐标系{R}的位姿, $F p_{G_o}$ 为工具坐标系尖端在机器人末端的平移关系。

标定板坐标系{B}下x轴方向、y轴方向、z轴方向的单位向量为

$$\begin{cases} \mathbf{v}_{B_{ox}} = \frac{{}^R p_{B_x} - {}^R p_{B_o}}{\|{}^R p_{B_x} - {}^R p_{B_o}\|_2} \\ \mathbf{v}_{B_{oy}} = \frac{{}^R p_{B_y} - {}^R p_{B_o}}{\|{}^R p_{B_y} - {}^R p_{B_o}\|_2} \\ \mathbf{v}_{B_{oz}} = \mathbf{v}_{B_{ox}} \times \mathbf{v}_{B_{oy}} \end{cases} \quad (10)$$

式中: $\mathbf{v}_{B_{ox}}$ 、 $\mathbf{v}_{B_{oy}}$ 、 $\mathbf{v}_{B_{oz}}$ 分别为标定板坐标系{B}下x轴方向、y轴方向、z轴方向的单位向量。

由此得到标定板坐标系{B}到机器人基坐标系{R}的转换关系 ${}^R_B T$ 为

$${}^R_B T = \begin{bmatrix} \mathbf{v}_{B_{ox}} & \mathbf{v}_{B_{oy}} & \mathbf{v}_{B_{oz}} & {}^R p_{B_o} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11)$$

1.1.3 求解 ${}^R_C T$

在双目 eye in hand 系统中, 双目相机相对于机器人末端固定。首先, 采用 OpenCV 中的张友正标定法进行双目相机标定, 生成标定参数; 然后, 每一个拍摄位置的双目标定板图像结合双目标定参数, 通过三角测量原理计算每一个角点在相机坐标系{C}下的三维坐标 ${}^C p^{(i)}$; 最后, 由于标定板角点在标定板坐标系{B}下的三维坐标 ${}^B p^{(i)}$ 已知, 需计算 ${}^B_C T$ 。

$${}^B_C T {}^C p^{(i)} = {}^B p^{(i)} \quad (12)$$

当机器人末端随双目相机对标定板进行拍摄标定时, 对每一个拍摄位置, 可以读取拍摄位置机器人末端基于机器人基坐标系{R}的坐标, 即 ${}^R p$; 同时该位置的双目图像对标定板中的角点进行三角测量计算, 可得到双目相机到标定板坐标系{B}的转换关系 ${}^B_C T$ 为

$$\begin{cases} {}^R_C T^{(i)} = {}^R_B T {}^B_C T^{(i)} \\ {}^R_C T^{(i)} = {}^R_F T^{(i)} {}^F_C T^{(i)} \end{cases} \quad (13)$$

式中: ${}^R_B T$ 为标定板坐标系{B}到机器人基坐标系{R}的转换关系, ${}^B_C T^{(i)}$ 为位置*i*相机坐标系{C}到标定板坐标系{B}的转换关系, ${}^R_F T^{(i)}$ 为位置*i*读取机器人末端的坐标, ${}^F_C T^{(i)}$ 为位置*i*相机坐标系{C}到机器人末端坐标系{F}的转换关系。

整理可得

$${}^F_C T^{(i)} = {}^R_F T^{(i)-1} {}^R_B T {}^B_C T^{(i)} \quad (14)$$

为了减少 ${}^F_C T$ 的计算误差, 采用均值法求平均 ${}^F_C T$ 为

$${}^F_C T = \frac{1}{N} \sum_{i=1}^N {}^R_F T^{(i)-1} {}^R_B T {}^B_C T^{(i)} \quad (15)$$

最后得到相机坐标系{C}关于机器人基坐标系{R}的转换关系为

$${}^R_C T = {}^R_F T {}^F_C T \quad (16)$$

在该双目 eye in hand 系统中, 每一个位置的双目相机, 通过公式(17), 可将相机坐标系{C}下的三维坐标点 ${}^C p^{(i)}$ 转换为机器人基坐标系{R}下的 ${}^R p^{(i)}$ 。

$${}^R p^{(i)} = {}^R_F T {}^F_C T {}^C p^{(i)} \quad (17)$$

1.2 基于深度学习的局部点云图像

固定目标双目三维重建时, 对每一个位置拍摄的双目图像, 首先, 根据双目标定参数进行图像矫正和极线对齐; 然后, 采用 YOLOv5 深度学习检测方法检测图像中的树干, 去除树干以外的背景图像; 接着, 利用三角测量原理, 对局部图像进行立体匹配, 生成视差图, 进而得到局部点云图; 最后, 先利用半径滤波法去除背景中的噪点, 得到仅含有树干的点云数据; 再进行体素下降, 减少树干点云数据。

1.2.1 YOLO 目标检测

为减小计算量, 以及无关环境信息对三维重建的影响, 采用深度学习提取指定目标。传统有监督的深度学习主要分为目标检测和语义分割 2 种; 相对于语义分割而言, YOLO 目标检测算法数据集制作较为方便, 且 YOLO 模型识别速度快、精度高、鲁

棒性强, 适用于工程领域。对于树干识别, 先利用 YOLOv5 目标检测算法, 获取双目左右图像中树干的位置矩阵, 再进行局部立体匹配。

1.2.2 局部点云图像

在局部立体匹配时, 利用半全局块匹配算法 (semi-global block matching, SGBM) 得到深度学习检测后图像的视差图。视差图中像素点包含的信息与该像素点的深度相关, 依据三角测量原理, 由视差来计算深度。视差图一般基于左图像, 对左图像中的像素点 x_u, y_v , 先根据左图像获取该像素点的 RGB 信息, 再根据视差图得到该像素点左右视图的视差 d_p 。在双目图像标定时, 得到三角测量相关的重投影矩阵 Q 。根据像素点 $(x_u^{(i)}, y_v^{(i)}, p_d^{(i)})^T$ 与重投影矩阵 Q , 可得到基于相机坐标系 $\{C\}$ 下的三维点 $c_p^{(i)}$ 为

$$c_p^{(i)} = \begin{bmatrix} X^{(i)} \\ Y^{(i)} \\ Z^{(i)} \\ 1 \end{bmatrix} = Q \begin{bmatrix} x_u^{(i)} \\ y_v^{(i)} \\ p_d^{(i)} \\ 1 \end{bmatrix} \quad (18)$$

式中: $(x_u^{(i)}, y_v^{(i)}, p_d^{(i)}, 1)^T$ 为视差图中的一点, $(X^{(i)}, Y^{(i)}, Z^{(i)}, 1)^T$ 为相机坐标系 $\{C\}$ 下以 mm 为度量单位的点, Q 为重投影矩阵。

根据公式(18), 可由双目相机拍摄的图像, 计算得到基于相机坐标系 $\{C\}$ 下的点云图像, 如图 3 所示。

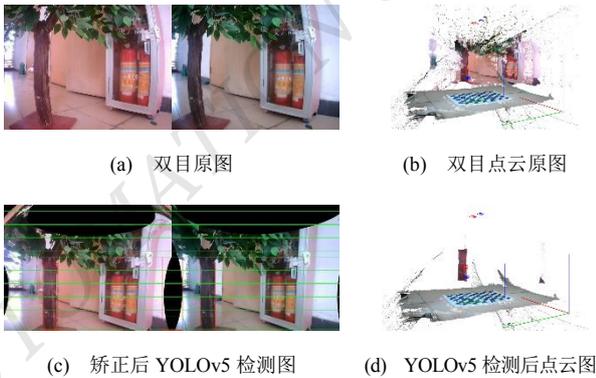


图 3 局部点云图像

1.2.3 半径滤波与体素滤波

在立体匹配时, 会生成一部分误匹配点, 影响树干模型的计算。采用滤波的方法去除立体匹配产生的

噪声点, 提取树干相关的点云数据。树干的点云数据相对噪声, 数据量多且密集, 半径滤波通过设置滤波半径与阈值点云数去除噪声点。当某一点滤波半径内的点云数大于阈值点云数, 则将该点视为数据点; 若不满足该条件, 则为噪声点。经半径滤波后只剩下树干点云数据, 由于数据密度较高, 采用体素滤波减少单位面积内的点云数据, 对点云数据进行均匀下采样。

1.3 点云拼接与树干姿态估计

1.3.1 点云拼接

对单一目标多角度的树干重建, 单一角度双目图像只能获取单一角度的树干数据, 要得到较为完整的树干点云图像, 需要对多角度点云图像进行拼接。

当机器人基坐标系 $\{R\}$ 不变时, 机器人末端带动双目相机进行图像采集和三维重建, 生成点云图像。通过公式(17)的转换关系, 得到基于机器人基坐标系 $\{R\}$ 下的多角度大规模点云图。多角度点云拼接图如图 4 所示。

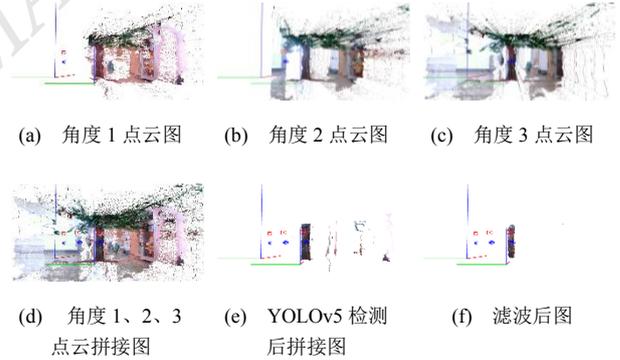


图 4 多角度点云拼接图

1.3.2 树干位置姿态估计

利用 RANSAC 算法对拼接后的树干点云数据进行圆柱体拟合, 计算拟合的圆柱体方程, 即树干的位置姿态。圆柱体的曲线方程为

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 - r^2 = \frac{[l(x - x_0) + m(y - y_0) + n(z - z_0)]^2}{l^2 + m^2 + n^2} \quad (19)$$

通过 RANSAC 算法拟合圆柱体得到圆柱体轴线的单位向量 (l, m, n) 、轴线上一点 (x_0, y_0, z_0) 和半径 r 。需要求解拟合圆柱体的最大、最小范围, 即轴线的最

大、最小值。

由于果树一般垂直于地面, 所以点云 z 轴方向的最大值、最小值可以近似为圆柱体轴线 z 轴方向上的最大值、最小值。对点云数据进行遍历, 找出 z 轴方向上的最大值 z_{\max} 和最小值 z_{\min} , 估计轴线的最大值点和最小值点。

圆柱体轴线方程为

$$q: \frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n} \quad (20)$$

圆柱体轴线的最大值点和最小值点为

$$\begin{cases} k_i = \frac{z_i - z_0}{n} \\ x_i = x_0 + k_i \times l, \quad (i = \min, \max) \\ y_i = y_0 + k_i \times m \end{cases} \quad (21)$$

通过公式(21)可计算圆柱体轴线的最小值点 $(x_{\min}, y_{\min}, z_{\min})$, 最大值点 $(x_{\max}, y_{\max}, z_{\max})$ 。

1.3.3 圆柱体参数优化

RANSAC 算法拟合得到的圆柱体参数并非最优解, 采用最小二乘法进行非线性优化, 求解最优圆柱体参数。将 RANSAC 算法拟合圆柱体得到的参数作为初始参数, RANSAC 算法拟合后筛选的点云数据作为非线性优化的输入数据, 进行非线性无约束优化。

对于有约束的最小二乘法参数优化问题, 可以描述为

$$\begin{cases} \min_x \sum_i \frac{1}{2} \rho_i(\|f_i(x_{i_1}, x_{i_2}, \dots, x_{i_k})\|^2) \\ \text{s.t. } l_j \leq x_j \leq u_j \end{cases} \quad (22)$$

式中: $\rho_i(\|f_i(x_{i_1}, x_{i_2}, \dots, x_{i_k})\|^2)$ 为残差块; $\rho_i(x)$ 为核函数, 用于去除异常值; $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ 为代价函数; x_{i_k} 为参数列表; l_j, u_j 分别为参数 x_j 的下界和上界。

在圆柱体参数优化模型中, 令

$$\begin{cases} f_i(x_{i_1}, x_{i_2}, \dots, x_{i_k}) = r_i^2 - r_0^2 \\ r_i^2 = (x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2 \\ -[l(x_i - x_0) + m(y_i - y_0) + n(z_i - z_0)]^2 \\ \rho_i(x) = x \end{cases} \quad (23)$$

则圆柱体非线性优化函数可表达为

$$\min_x \frac{1}{2} \sum_i \left(\left\| \begin{matrix} (x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2 - \\ [l(x_i - x_0) + m(y_i - y_0) + n(z_i - z_0)]^2 - r_0^2 \end{matrix} \right\|^2 \right) \quad (24)$$

式中: $x_0, y_0, z_0, l, m, n, r_0$ 为需要优化的参数, x_i, y_i, z_i 为输入的圆柱体表面的点云数据。

2 实验与结果分析

2.1 实验环境

实验环境由工业机械臂 RS007L、双目相机 WN-L2110、工具末端执行器、棋格盘标定板(棋格盘为 11×8 , 边长为 30 mm 的正方形, 精度为 0.03 mm)、荔枝模型树、笔记本电脑(CPU 为 intel i5-8250U, 内存容量为 8 GB DDR4, 固态硬盘容量为 1 TB, 运行环境为 ubuntu18.04) 等组成。棋格盘标定板位于工业机械臂的前下方, 如图 5 所示。



图 5 基于双目 eye in hand 系统的多角度树干位姿估计实验环境

2.2 手眼系统标定实验

采用同点三次标定法得到 ${}^C T$, 采用三点法得到 ${}^R T$ 。手眼标定误差主要由双目相机标定产生。在相机标定时, 于同一平面拍摄标定板图像 30 幅, 拍摄位置如图 2 所示。

为了定量分析评价手眼标定中 ${}^C T$ 的精度, 取标定板的 88 个角点为评价点, 对每一个角点, 通过计算欧式距离获取角点真实值与测量值之间的误差, 用绝对值误差计算 x_i, y_i, z_i 之间的误差。误差计算公式为

$$P_E = \frac{1}{N} \sum_{i=0}^N \sqrt{(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2} \quad (25)$$

式中： x 、 y 、 z 为真实值， N 为计算误差样本的图像数量，取30。

每个角点的平均误差计算结果如图6所示，平均欧式误差为3.7177 mm。

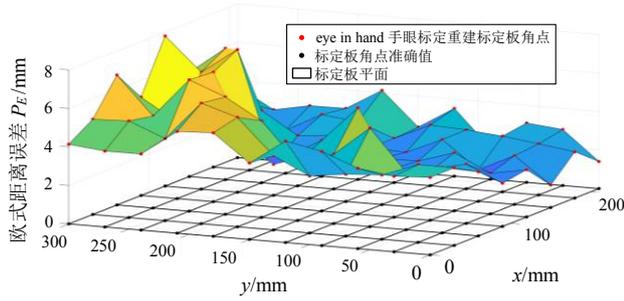


图6 手眼标定后标定板角点欧式距离误差

2.3 局部图像点云实验

在生成局部图像时，首先，通过深度学习检测得到待匹配的指定区域；然后，对指定区域匹配时，经半径滤波后，得到树干表面的局部点云图；最后，在保留树干姿态信息的前提下，尽量减少点云数量，以减少计算量。

实验时，机械臂采用圆形的轨迹规划，多角度拍摄同一棵树的双目图像，计算不同拍摄位置的点云数量和处理后的点云数量。树干点云数量变换图如图7所示。

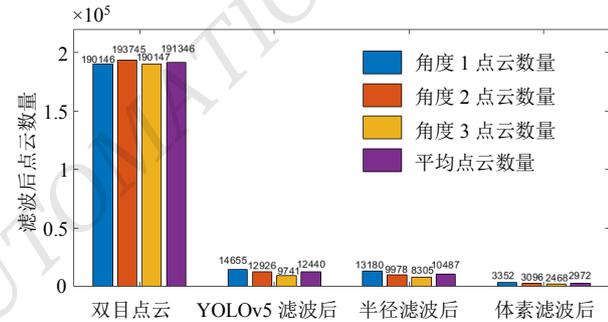


图7 树干点云数量变换图

2.4 点云拼接、位姿估计与参数优化实验

对不同角度拍摄的双目图像，通过局部立体匹配

得到该位置的点云图像。通过将点云图像转换到同一机器人坐标系下进行点云拼接。对3个位置的点云图像拼接后，进行圆柱体位置姿态估计。

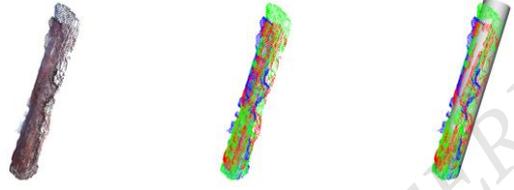


图8 角度123点云拼接图、拟合圆柱体

对拼接后的点云进行姿态估计和参数优化，计算圆柱体优化前后半径的相关误差 R_{E_i} 、平均绝对误差 R_{MAE} 、均方根误差 R_{RMSE} 。

$$r_i = \sqrt{\frac{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2 - [l(x_i - x_0) + m(y_i - y_0) + n(z_i - z_0)]^2}{l^2 + m^2 + n^2}} \quad (26)$$

$$R_{E_i} = r_i - r_0 \quad (27)$$

$$R_{MAE} = \frac{1}{N} \sum_{i=1}^N |r_i - r_0| \quad (28)$$

$$R_{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_i - r_0)^2} \quad (29)$$

式中： N 为点云数量， r_0 为圆柱体半径， r_i 为点到圆柱体轴线距离， R_{MAE} 为平均绝对误差， R_{RMSE} 为均方根误差。

对拼接后的点云进行姿态估计和参数优化，计算轴线的最小值点和最大值点。半径误差图值如图9所示，优化前后圆柱体参数如表1所示。

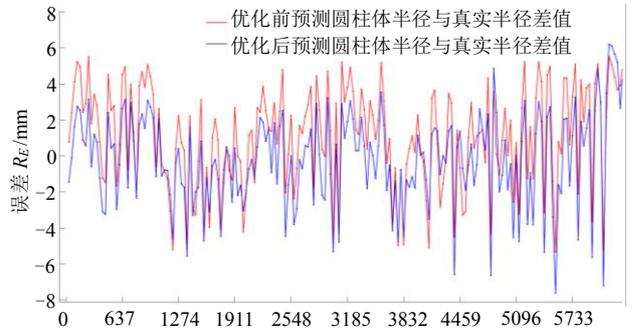


图9 优化前后预测圆柱体半径误差值

表 1 树干点云数据拟合圆柱体优化前后参数

单位: mm

	x_{\min}	y_{\min}	z_{\min}	x_{\max}	y_{\max}	z_{\max}	r_0	R_{MAE}	R_{RMSE}
优化前	1 384.13	-103.54	-112.09	1 393.38	-80.75	375.83	41.277 1	2.571 56	2.989 36
优化后	1 383.28	-101.04	-112.09	1 394.72	-82.36	375.83	39.402 8	1.989 55	2.465 88

由图 9 可知, 优化后圆柱体半径与真实半径的差值减小。

由表 1 可知, 利用最小二乘法进行非线性优化后圆柱体 $r = 39.402 8 \text{ mm}$, $R_{MAE} = 1.989 55 \text{ mm}$, $R_{RMSE} = 2.465 88 \text{ mm}$ 。

3 结论

本文提出基于双目 eye in hand 系统的多角度树干位姿估计方法, 通过计算双目 eye in hand 系统各坐标系的转换关系对标定板进行三维重建, 运用标定板角点定量计算手眼标定的误差, 平均欧式误差为 $3.717 7 \text{ mm}$; 运用半全局块匹配算法, 结合 YOLOv5 目标检测算法对树干进行局部三维重建, 采用半径滤波和体素滤波减少 98.470% 的点云数据; 采用 RANSAC 圆柱体估计算法对树干点云数据进行圆柱体拟合, 拟合后圆柱体的半径 $r = 41.277 1 \text{ mm}$, $R_{MAE} = 2.571 56 \text{ mm}$, $R_{RMSE} = 2.989 36 \text{ mm}$ 。利用最小二乘法进行非线性优化求解最优圆柱体参数, 优化后 $r = 39.402 8 \text{ mm}$, $R_{MAE} = 1.989 55 \text{ mm}$, $R_{RMSE} = 2.465 88 \text{ mm}$, 实现了树干精准位姿估计。

本文仅从多角度对单一目标进行三维重建, 在后续研究中, 通过手眼系统可以根据多角度点云图像拼接的特点, 对周围环境进行大规模的三维重建, 建立多角度多目标的三维重建系统。

参考文献

- [1] 容秀婵, 邹湘军, 李承恩, 等. 基于数据驱动的虚拟场景搭建及模型检索优化方法[J]. 中国农机化学报, 2022, 43(8): 128-135.
- [2] 李秋洁, 丁旭东, 邓贤. 基于激光雷达的果园行间路径提取与

导航[J]. 农业机械学报, 2020, 51(S2): 344-350.

- [3] 郑太雄, 江明哲, 冯明驰. 基于视觉的采摘机器人目标识别与定位方法研究综述[J]. 仪器仪表学报, 2021, 42(9): 28-51.
- [4] 王辉, 毛文华, 刘刚, 等. 基于视觉组合的苹果作业机器人识别与定位[J]. 农业机械学报, 2012, 43(12): 165-170.
- [5] 方建军. 移动式采摘机器人研究现状与进展[J]. 农业工程学报, 2004, 20(2): 273-278.
- [6] 郑似青, 熊伟, 方陆明, 等. 立木胸径近真值测量与计算模型的研究[J]. 西北林学院学报, 2021, 36(5): 125-132.
- [7] 刘鲁霞, 庞勇, 李增元. 基于地基激光雷达的亚热带森林单木胸径与树高提取[J]. 林业科学, 2016, 52(2): 26-37.
- [8] ZOU X, ZOU H, LU J. Virtual manipulator-based binocular stereo vision positioning system and errors modelling[J]. Machine Vision and Applications, 2012, 23: 43-63.
- [9] 叶敏, 邹湘军, 罗陆锋, 等. 荔枝采摘机器人双目视觉的动态定位误差分析[J]. 农业工程学报, 2016, 32(5): 50-56.
- [10] LIN G, TANG Y, ZOU X, et al. Guava detection and pose estimation using a low-cost RGB-D sensor in the field[J]. Sensors, 2019, 19(2): 428.
- [11] CHEN M, TANG Y, ZOU X, et al. Three-dimensional perception of orchard banana central stock enhanced by adaptive multi-vision technology[J]. Computers and Electronics in Agriculture, 2020, 174: 105508.
- [12] CHEN M, TANG Y, ZOU X, et al. 3D global mapping of large-scale unstructured orchard integrating eye-in-hand stereo vision and SLAM[J]. Computers and Electronics in Agriculture, 2021, 187: 106237.
- [13] OLOFSSON K, HOLMGREN J, OLSSON H. Tree stem and height measurements using terrestrial laser scanning and the RANSAC algorithm[J]. Remote Sensing, 2014, 6(5): 4323-4344.
- [14] 吴杭彬, 王旭飞, 刘春. 基于林地激光点云的树木胸径自动提取方法[J]. 同济大学学报(自然科学版), 2022, 50(7): 947-954.

作者简介:

王欢, 男, 1998 年生, 硕士研究生, 主要研究方向: 机器视觉。E-mail: 1306357145@qq.com

徐秀进, 男, 1998 年生, 硕士研究生, 主要研究方向: 机器视觉。E-mail: xxj17306690937@163.com

王红军(通信作者), 女, 1966 年生, 博士, 教授, 主要研究方向: 机器视觉、农业采摘机器人。E-mail: xtwhj@scau.edu.cn

温志庆, 男, 1964 年生, 博士, 教授级高工, 主要研究方向: 机器视觉与人工智能。E-mail: j.wen1234@outlook.com