

文章编号: 2095-2163(2024)01-0029-07

中图分类号: TP181

文献标志码: A

# 基于双目视觉 SGBM 与 WLS 融合算法的茶叶嫩芽定位研究

许宝阳, 高延峰, 张泽玺, 包俊阳

(上海工程技术大学 上海市大型构件智能制造机器人技术协同创新中心, 上海 201620)

**摘要:** 三维定位是实现采茶机器人精采名优茶的关键技术,对保证机器人采摘茶叶高品质和高产量具有重要的意义,传统的 SGBM (Semi-Global Block Matching) 算法存在匹配效果差,还原效果不高等问题。本文提出 SGBM 算法与 WLS (Weighted Least Squares) 的融合算法,使得茶叶嫩芽深视图轮廓更清晰、前后景分层更明显、还原度更高,实际定位效果更精准。实验表明:采用 SGBM 与 WLS 融合算法能够将定位误差控制在 1 mm 左右,约是同等条件下其他传统融合算法精确度的 7 倍,提高了机器人采摘茶叶时定位的工作效率,为后续实现采茶机器人智能化提供一定帮助。

**关键词:** 双目视觉; SGBM 算法; WLS 算法; 茶叶嫩芽; 采茶机器人

## Research on tea-buds location based on binocular vision with SGBM-WLS fusion algorithm

XU Baoyang, GAO Yanfeng, ZHANG Zexi, BAO Junyang

(Shanghai Collaborative Innovation Center of Intelligent Manufacturing Robot Technology for Large Components, Shanghai University of Engineering Science, Shanghai 201620, China)

**Abstract:** Three-dimensional positioning is a key technology for the successful harvesting of high-quality tea by tea-picking robots, playing a crucial role in ensuring both the high quality and high yield of tea leaves. In response to issues such as poor matching effects and low restoration effects in the traditional SGBM algorithm, this paper proposes a fusion algorithm of SGBM and WLS. This integration ultimately results in a clearer contour of the tender tea buds in deep views, a more distinct separation of foreground and background, and higher restoration accuracy, leading to a more precise practical positioning effect in tea harvesting. Experiments demonstrate that the fusion algorithm of SGBM and WLS can control positioning errors to around 1 mm, approximately seven times more accurate than traditional fusion algorithms under similar conditions. This fusion algorithm has the potential to enhance the operational efficiency of robot-assisted tea harvesting by minimizing positioning errors, thus providing significant assistance for the subsequent realization of intelligent tea-picking robots.

**Key words:** binocular vision; SGBM algorithm; WLS algorithm; tea-bud; tea picking robot

## 0 引言

随着茶叶市场需求的不断扩大,茶叶采摘技术也在不断升级。目前茶叶采摘仍以人工为主,本文自主搭建双目视觉系统实现茶叶嫩芽定位,解决现有的人工耗时耗力的采茶痛点。双目视觉技术就是为机器配备了一双“眼睛”,如何将双目中的物体特征进行有效深度还原,实现精准的三维定位一直是该领域的重难点之一。近年来,双目视觉技术随着机器视觉的进步而不断提升,因具有灵活性高,精度高等优点而被广泛应用。

在国内,文献[1]提出一种基于特征点的 ORB (Oriented FAST and Rotated Brief) 与 RANSAC (Random Sample Consensus) 融合匹配算法,有效解决了提取不充分,匹配不精确的难题,但该算法只满足点特征的提取要求,并不能提取轮廓特征;文献[2]在传统的 SIFT (Scale Invariant Feature Transform) 算法中加上相似三角形算法,将沙丘图像匹配精确度提高了 35.29%;文献[3]提出先用边缘检测粗定位,后用 SIFT 结合 RANSAC 算法精匹配的算法,减少运行时间,提高定位精度;文献[4]提出一种基于 MASK R-CNN (MASK Regions with CNN

**基金项目:** 上海市自然科学基金(21010501600)。

**作者简介:** 许宝阳(1996-),男,硕士研究生,主要研究方向:深度学习、双目视觉。

**通讯作者:** 高延峰(1974-),博士,教授,博士生导师,主要研究方向:机器人传感与智能控制。Email:gyf\_2672@163.com

**收稿日期:** 2022-12-10

features)模型的双目识别与定位算法,提高了匹配精度和速度;文献[5]在农业应用中建立了一套YOLOv3(You Only Look Once version 3)结合SIFT匹配算法的双目系统,提高棉花顶部识别与定位率,有效解决了传统机械化棉花打顶痛点;文献[6]将改进后的Harris算法与SIFT算法相融合,一定程度上解决了SIFT算法对自相似或者对称性图像匹配效率低的问题;文献[7]将改进后的Harris算法与SUFT(Speeded-Up Robust Features)算法相融合,有效去除错误匹配点,匹配效率更高,匹配速度更快;文献[8]提取目标特征,利用几何约束获取三维坐标,同时还与预先构建的模型进行ICP(Iterative Closest Point)匹配,得到三维姿态;文献[9]将双目视觉系统运用在汽车尺寸检测上,通过SGBM算法将误差控制在10 mm左右;文献[10]提出基于最小外接矩形检测方法,利用最小二乘法实现三维重建,降低误差率。

在国外,文献[11]利用双目视觉技术提高实时练习中颌部定位的精度,减小手术精度误差;文献[12]开发了一套主动双目视觉系统,能够不需要通过校正等步骤而识别、瞄准和接近物体;文献[13]采用OpenCV图像处理库,编写代码实现自动识别特征点,结合OpenGL(Open Graphics Library)技术对周围场景进行三维感知,重建得到三维模型。

茶叶生长环境复杂,叶片相似度高,嫩芽小,茶梗多藏于叶片下,因此对识别定位精度要求极高。本文自主搭建双目视觉系统,在传统的SGBM算法基础之上融合了WLS算法,一定程度上提高茶叶嫩芽定位精度,同时多维度对比算法优劣,为后续实现采茶机器人智能化奠定基础。

## 1 采茶机器人硬件平台框架

本文搭建的采茶机器人主要是由移动行走模块、视觉传感模块、机械臂运动模块和计算机控制模块组成,结构如图1所示。采茶机器人可以实现在实际复杂环境下的轻松行走、茶叶图像的采集与分析、嫩芽的识别与定位、以及机械臂的控制与运动功能。

考虑茶园地理环境受天气影响,移动行走模块采用底盘稳重,转向灵活的履带式运动结构;视觉传感模块采用双目相机,利用视差原理更好的获取茶叶嫩芽的深度信息;机械臂运动模块采用图灵STH030-500型SCARA机械臂,一共有4轴,前3轴为水平面运动,第4轴为垂直运动,该机械臂体积小,重量轻,灵活度高,易操作。机械臂前端搭载的

采摘器是在原有的基础上加入负压装置,如图2所示。采摘器利用红外线受断带动电机转动来切断嫩芽,负压装置能够借用一定的风力,使得采摘口出的嫩芽伸直,提高采摘的精确率和回收率。



图1 采茶机器人结构图

Fig. 1 Structure diagram of tea picking robot

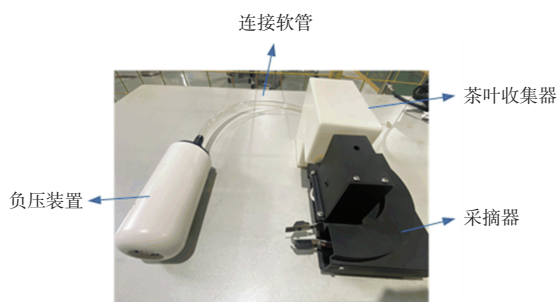


图2 采摘器

Fig. 2 Picking device

## 2 平行双目定位原理

双目视觉技术是仿照人类眼睛成像原理,利用两个不同视角的相机拍摄空间中的同一物体,产生左右两幅图像,结合相机内部参数,采用匹配算法,即可获取拍摄物体的空间三维信息<sup>[14]</sup>。平行双目成像模型如图3所示。

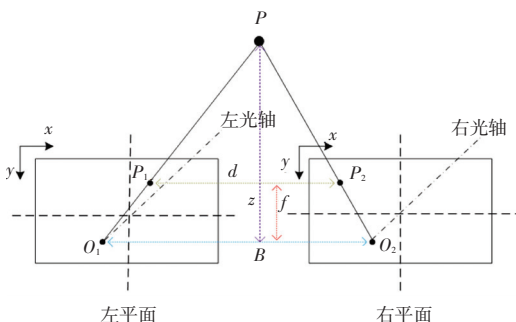


图3 平行双目成像模型

Fig. 3 Parallel binocular imaging model

图3中, $x-y$ 平面表示双目相机左右成像平面;左右相机光轴平行; $o_1$ 和 $o_2$ 表示左右相机镜头的光心,对应的横坐标为 $c_1$ 和 $c_2$ ,且 $c_1 = c_2$ ;B表示两镜头中心之间的距离,称为基线; $f$ 表示相机的

焦距;  $P(x, y, z)$  表示空间目标物体;  $P_1(x_1, y_1)$  和  $P_2(x_2, y_2)$  分别表示物体在左右像平面内映射点, 且  $y_1 = y_2$ 。根据相似三角形原理可得出式(1):

$$\frac{B - (x_1 - c_1) - (c_2 - x_2)}{B} = \frac{z - f}{z} \quad (1)$$

整理可得式(2):

$$z = \frac{f \times B}{x_1 - x_2} \quad (2)$$

其中,  $f$  和  $B$  是可知的。视差表示同一物体在左右平面映射点  $x$  坐标的差值<sup>[15]</sup>, 用  $d = x_1 - x_2$  表示视差, 代入化简得式(3):

$$z = \frac{f \times B}{d} \quad (3)$$

根据相似三角形原理也可得式(4):

$$\begin{cases} x = \frac{z \times x_1}{f} \\ y = \frac{z \times y_1}{f} \end{cases} \quad (4)$$

将上式(3)带入式(4)可得式(5):

$$\begin{cases} x = \frac{B \times x_1}{d} \\ y = \frac{B \times y_1}{d} \end{cases} \quad (5)$$

因此, 只要获取深视差  $d$ , 即可计算出空间三维坐标信息  $(\frac{B \times x_1}{d}, \frac{B \times y_1}{d}, \frac{f \times B}{d})$ 。

### 3 双目相机标定

相机标定的目的是为了获取校正与匹配过程中所需要的两种参数: 左右相机的内外参数矩阵和畸变系数, 影响着后续图像匹配与定位精度。

#### 3.1 基于张正友棋盘格标定法

张正友教授<sup>[16]</sup>提出的棋盘格标定法打破了传统标定法制作标定板难度大、操作精度高等缺点, 仅仅通过一张黑白棋盘格就能实现标定, 操作简易, 也极大提高了精度。本文使用汇博视捷科技双目相机, 具体参数见表 1。

表 1 双目相机参数

Table 1 Binocular camera parameters

指标	参数
产品型号	HBV-1780-2 S2.0
视场角/焦距	100°无畸变/2.1 mm
像素	100 万像素
分辨率	2 560 * 720
	1 280 * 720

标定手段有很多种, Matlab 自带工具箱标定具有操作简单、标定精度高等优点成为标定首选。选用 8×6 的棋盘格制作标定板, 保持相机不动, 转动标定板, 一共采集 30 组图片。图片导入 Matlab 后对图片进行角点提取, 如图 4 所示。

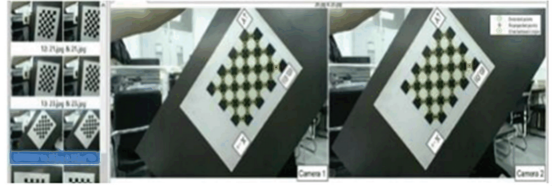


图 4 角点提取

Fig. 4 Corner extraction

标定结果同时输出左右图像标定平均误差和相机与目标之间的模拟姿态, 如图 5 和图 6 所示。删除误差较大的图片, 以提高标定精度, 本文控制误差在 0.2 左右, 最终保留 26 组有效图片。

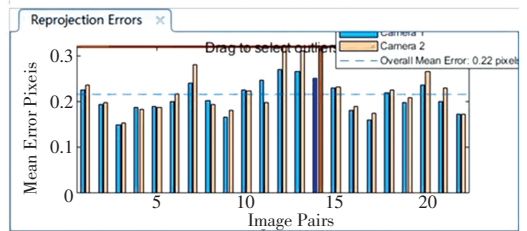


图 5 标定平均误差图

Fig. 5 Calibration average error

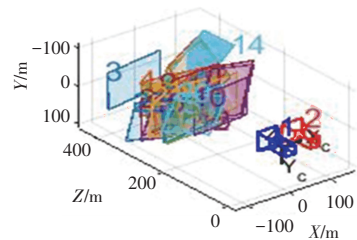


图 6 相机与目标的模拟位姿

Fig. 6 Simulated pose of camera and target

#### 3.2 内外参数获取

空间三维坐标是像素坐标通过三大坐标系利用内外参数相互转换的结果, 故标定精度决定了空间坐标精度。左右相机的内参有  $(f, \frac{1}{dx}, \frac{1}{dy}, cx, cy)$ , 外参有偏移矩阵  $T$  和旋转矩阵  $R$ , 偏移矩阵可以直接使用, 而旋转矩阵需要转置才能使用, 畸变参数包括左右相机的径向畸变  $(k_1, k_2, k_3)$  和切向畸变  $(p_1, p_2)$ 。双目相机的内外参数见表 2。

表2 相机内外参数

Table 2 Camera internal and external parameters

	左相机	右相机
内参	$\begin{bmatrix} 722.8014 & 0 & 645.7112 \\ 0 & 724.7895 & 356.3693 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 721.3669 & 0 & 627.6458 \\ 0 & 722.6804 & 345.4305 \\ 0 & 0 & 1 \end{bmatrix}$
径向畸变	$[0.0531 \quad 0.2452 \quad -0.6271]$	$[0.0769 \quad 0.0602 \quad -0.2447]$
切向畸变	$[-0.0057 \quad 0.0020]$	$[-0.0049 \quad 0.0046]$
偏移矩阵	$[-59.7906 \quad 0.1506 \quad -0.0610]$	
旋转矩阵	$\begin{bmatrix} 1 & 0.0013 & -0.0160 \\ -0.0013 & 1 & 0.0039 \\ 0.0046 & -0.0039 & 1 \end{bmatrix}$	

## 4 立体校正

由于凸透镜自身结构缺陷和光学特性,拍摄得到的图片会产生径向畸变和切向畸变,这些畸变都会使图片变形失真,故必须对左右图片进行立体校正,以提高特征提取的效率和立体匹配的精度,校正效果对比图如图7所示。



图7 立体校正对比

Fig. 7 Contrast of stereo correction

## 5 立体匹配

立体匹配是寻求左右两幅图像之间的对应关系继而生成视差图<sup>[17]</sup>。立体匹配算法目前包含局部、半全局和全局立体匹配算法。局部匹配是对像素周围进行小区域约束匹配,直观简单、速度快、精度高,但有时却不适用于缺少明显特征的环境;全局匹配则是对整幅图像进行约束匹配,可以提高匹配精度,但计算量大,复杂度高。

SGBM是介于局部和全局立体匹配之间的半全局块匹配算法,包含了预处理、代价计算、动态规划和后处理步骤,其立体匹配效果优于局部立体匹配,算法的复杂度没有全局立体匹配高,兼顾两者的效率和精度<sup>[18]</sup>。

SGBM算法都使用块匹配的方法。匹配之前会将彩色图转换成灰度图,使正常的RGB三通道像素

立体校正将原本不共面的左右两幅图片通过畸变参数和对极线校正成共面平行,使左相机平面内出现的映射点可以快速地通过平行极线在右相机平面内找到,原本二维搜索降成一维搜索,提高匹配效率同时还能够有效地降低立体匹配计算量。越是远离图像中心的部位越是易产生畸变,从图7(b)可看出边缘有明显的校正痕迹。

值转换成单通道像素值。立体匹配结果如图8所示,左相机采集的图像生成小方块A1,计算A1方块内像素点矩阵的绝对差之和SAD(Sum of absolute differences),同理右相机采集的图像同水平线上产生小方块A2、B2等,分别计算A2、B2的SAD,选取SAD差值最小块为相匹配方块。

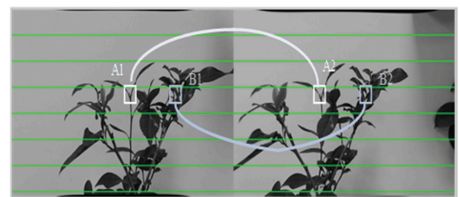


图8 立体匹配结果

Fig. 8 Stereo matching result

## 6 实验分析

光照强度和图像清晰度是影响匹配和视差精度

的关键因素,需首先对采集图像进行滤波去噪处理。无论是高斯滤波还是平均滤波都会使原始图像产生模糊现象,而中值滤波可以消除噪声的同时降低模糊影响,本文在原有的中值滤波方法基础上加以改进,因噪声点的像素值通常是 0 或者 255,故在滤波前加入一个判断条件:若该像素点是滤波窗口下的极值,则进行中值滤波;否则不进行滤波处理,保留原值。

### 6.1 运行时间与还原效果对比

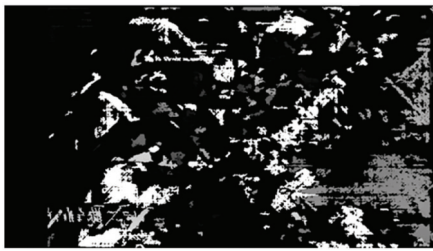
本文分别采用 BM(Boyer-Moore)算法和 SGBM 算法得到深视图,得到的深度图匹配度低,场景还原度低,误差高。因此本文在 BM 算法和 SGBM 算法基础上加入了 WLS 算法,去除掉无关信息,极大程度提高了匹配精度和视差图还原效果。各算法得到信息深度图如图 9 所示,运行时间结果见表 3。



(a) BM



(b) BM+WLS



(c) SGBM



(d) SGBM+WLS

图 9 深视图

Fig. 9 Deep view

表 3 算法运行时间

Table 3 Algorithm running time

匹配算法	运行时间/s
BM	1.578 1
SGBM	2.063 1
BM+WLS	2.156 3
SGBM+WLS	3.312 5

从运行时间看出, BM 类算法运行速度要比 SGBM 类算法快,更加适合实时性要求更高的场景。加入了 WLS 滤波器虽然会延迟运行时间,降低实时性,但融合后的算法能够去除掉无关信息,极大程度提高了匹配精度和视差图还原效果。融合前后 BM 类算法匹配效果都远远不如 SGBM 类算法。BM 类算法匹配不完整,成像效果差,边缘轮廓模糊,信息丢失严重;SGBM 类算法匹配完整,成像效果好,边缘轮廓清晰,层次分明,信息保存较好。

### 6.2 实际测试结果对比

传统的 BM 算法和 SGBM 算法成像不清,无法

准确定位,本文分别采用融合 WLS 的 BM 算法和 SGBM 算法作为对照组,从不同角度拍摄定位比较算法的精确性,即保持垂直方向距离不变,一组保持相机与茶叶平行拍摄,另一组保持 45° 角度拍摄,示意图如图 10 所示。两组分别选取相同的 5 个测试点,测试结果见表 4,误差如图 11 所示。

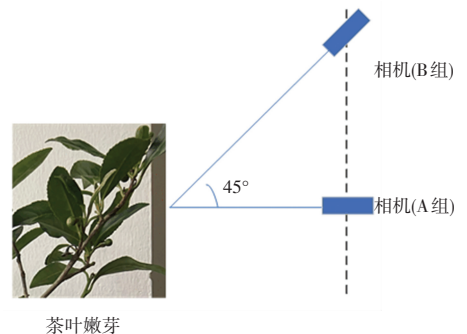


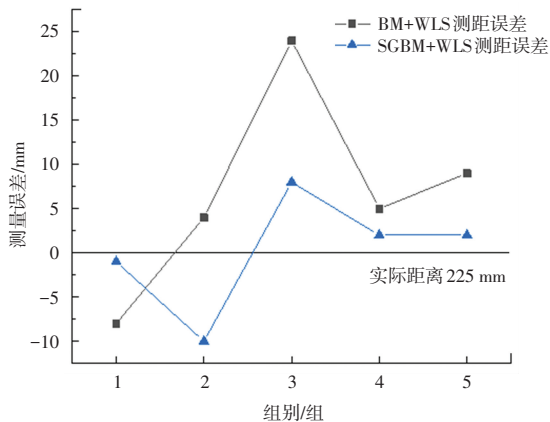
图 10 实验示意图

Fig. 10 Schematic diagram of experiment

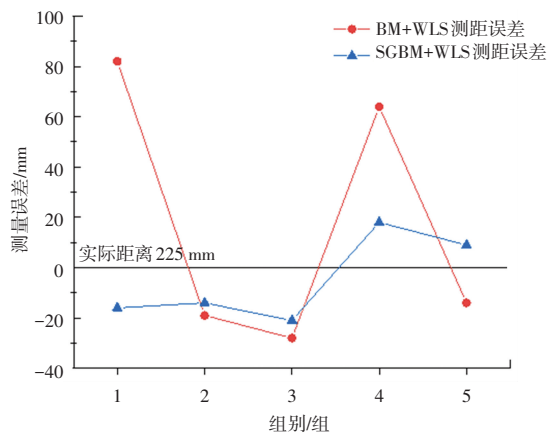
表4 算法测距结果

Table 4 Algorithm ranging results

组别	实际距离/mm	BM+WLS		SGBM+WLS	
		平行拍摄误差/ mm	45°拍摄误差/ mm	平行拍摄误差/ mm	45°拍摄误差/ mm
组1	225	-8	+82	-1	-16
组2	225	+4	-19	-10	-14
组3	225	+24	-28	+8	-21
组4	225	+5	+64	+2	+18
组5	225	+9	-14	+2	+9
平均误差 (去掉最值,取绝对值)		7.3	37	1.0	16.0



(a) 平行拍摄两种算法误差



(b) 45°拍摄两种算法误差

图11 误差图

Fig. 11 Error diagram

通过对比实验可知,改进后的 SGBM 算法定位效果比 BM 算法定位效果更加精确,误差最小可为1 mm 左右。同等角度拍摄下,融合后 BM 算法平均误差是融合后 SGBM 算法误差的 7 倍,这是因为 SGBM 算法

本身匹配精度优于 BM 算法,再加上 WLS 滤波算法能去除匹配无效点,有一定的抗干扰性,能进一步提高 SGBM 算法的精确度。而同种算法时,平行拍摄时的平均误差是俯视拍摄误差的 16 倍,主要原因在于平行拍摄得到的深度与拍摄角度同方向,可以直接有效测量茶梗,而俯视拍摄时只能拍摄茶叶表面而无法得知内部茶梗的具体位置,且俯视拍摄造成周围成像干扰多、测量误差大、定位效果差。

## 7 结束语

针对 SGBM 算法对特征提取与匹配不完整,还原度不高的缺点,本文提出了 SGBM 与 WLS 融合算法。利用双目相机依次经过数据采集、双目标定、立体校正和立体匹配过程,借助 Pycharm 开发环境和 OpenCV 视觉库,实现茶叶嫩芽的三维定位。实验结果表明,SGBM 与 WLS 融合算法能够进一步去除成像无效点,极大提高立体匹配效果和测量的精确度与可靠性;同时,平行拍摄得到的精度优于俯视拍摄,为后续智能采茶机器人的设计提供有效思路。下一步将预加入 Bbox (Bounding box) 识别程序,完善自主识别与定位系统,同时实现计算机与机械臂平台的连接,提高采茶机器人实际环境下的操作能力。

## 参考文献

- [1] 郭攀, 杜鸿. 机器人双目测距算法研究[J]. 无线互联科技, 2018, 15(5): 99-101.
- [2] 唐颖复, 王忠静, 张子雄. 基于改进 SIFT 和 SURF 算法的沙丘图像配准[J]. 清华大学学报(自然科学版), 2021, 61(2): 161-169.
- [3] 杜玉晓, 郑晓森, 陈伟杭, 等. 基于粗定位-精匹配的双目视觉目标定位方法[J]. 自动化与信息工程, 2021, 42(5): 11-17.
- [4] 彭秋辰, 宋亦旭. 基于 Mask R-CNN 的物体识别和定位[J]. 清华大学学报(自然科学版), 2019, 59(2): 135-141.

(下转第 42 页)