

同行专家业内评价意见书编号: 20240855057

## 附件1

# 浙江工程师学院（浙江大学工程师学院） 同行专家业内评价意见书

姓名: \_\_\_\_\_ 刘元浩

学号: \_\_\_\_\_ 22160042

申报工程师职称专业类别（领域）: \_\_\_\_\_ 机械

浙江工程师学院（浙江大学工程师学院）制

2024年03月27日

## 一、个人申报

**（一）基本情况【围绕《浙江工程师学院（浙江大学工程师学院）工程类专业学位研究生工程师职称评审参考指标》，结合该专业类别(领域)工程师职称评审相关标准，举例说明】**

### 1. 对本专业基础理论知识和专业技术知识掌握情况

我通过参与项目研究，系统地学习了计算机视觉、图像处理、深度学习等多个领域的知识和技术。掌握了火焰检测、定位、跟踪、识别等功能的原理和方法，了解了消防巡检机器人的设计和应用，拓宽了我的专业视野和知识面。我通过参与项目实施，锻炼了自己的动手能力和创新能力，掌握了MMDetection框架、Python语言、OpenCV库等多种工具和平台的使用，熟练了数据采集、预处理、分析等多个环节的操作。

### 2. 工程实践的经历

经过前期调研，市场上大多数消防巡检机器人停留在遥控式消防机器人阶段，国内研制的搭载的相机等视频采集设备消防机器人，仅用于将前线作业环境图像信息无线传输到火场后端，辅助消防人员遥控操作，这种被动获取视觉信息方式因为无法保证环境信息的实时性，经常导致火灾难以及时得到扑灭。目前，针对早期火焰的预警、探测、自主扑灭一体的消防巡检机器人研究仍处在初步阶段。针对消防巡检机器人的火焰自主定位的研究较少。经过充分查找早期火焰的识别与定位的文献后。我们考虑采用基于相机视觉的方法，实现火焰探测。因为基于传统点传感器的方法存在早期火焰识别难、延迟率高、服务范围有限等问题。而由于光速非常高，摄像机可以在可忽略的时间内捕捉火焰信息，能够克服基于点传感器的火焰检测系统的大部分问题。在相机选型方面由于火焰具有明显的红外热辐射特性，我们首选红外相机，但是考虑到红外相机拍摄的图像成像细节较少，易受温度影响等问题。为了实现高准确度的火焰检测与定位。我们最终采用由一个可见光相机和一个红外相机结合的双目相机来实现火焰检测与定位。具体方案为：设计一种火焰检测模型，通过可见光相机和红外相机采集火焰数据，训练火焰检测模型，使模型具有良好的火焰检测性能。我们在灭火实验室采集火焰从早期火苗到燃烧殆尽整个火焰燃烧周期中的图片，通过改变燃烧物影响火焰，构造了丰富多样的火焰形态、大小、颜色。通过添加不同的干扰源，改变环境光等方法，采集多样化的火焰数据。对如何融合异源火焰特征的问题，我与工程师们针对图像层融合、特征层融合、决策层融合进行多次讨论，考虑到硬件搭载，我们最终采用基于决策层融合的方式。在火焰定位的问题上，我们通过改进同源双目定位方案实现了基于近红外相机与可见光相机的异源双目火焰定位。

### 3. 在实际工作中综合运用所学知识解决复杂工程问题的案例

针对当前火焰公开数据集数量欠缺、多样性低、质量不高等问题，我们构建了一个多模态火焰数据集。该数据集融合了多种模态图像，包括从公开数据集中收集的可见光图像，以及在灭火实验室中采集的可见光和热红外图像。通过对数据集进行了数据增强，包括直方图均衡化、Retinex彩色恢复算法、伽马自适应亮度校正算法、图像平滑滤波和Top-hat变换等，以提高数据的质量和多样性。这为后续火焰检测技术的研究提供了坚实的数据基础。针对当前单模态图像火焰识别方法存在误报率高等问题，我们分别提出了基于可见光图像和热红外图像深度学习目标检测网络。首先，针对可见光图像火焰检测中背景干扰、小目标火焰漏检等难点，提出采用切片辅助超推理技术以提高小目标火焰的检测召回率。同时，设计了一种具有多尺度特征提取能力的级联掩码卷积神经网络ConvNeXt Cascade Mask R-CNN (CCMR)。该网络结构利用多尺度特征提取和融合技术，以及多级目标检测结构，实现对火焰目标的精确检测和分割。然后，针对热红外图像像素低、全局特征不明显等难点，改进了CCMR特征融合模块，提出了一种热红外图像火焰检测网络ConvNeXt Cascade Mask R-CNN PAFPN (CCMR-PA)。该网络在CCMR基础上利用了PAFPN路径聚合特征金字塔，该特征金字塔能够有效地提取热红外图像中火焰目标的温度特征，并增强不同尺度特征之间的信息传递和融合。针对单

模态图像火焰目标检测的不足，我们提出了一种在决策层融合单模态目标检测模型检测信息的多模态图像融合火焰检测方法。通过计算每种模态检测结果的置信度，以及不同模态目标候选框之间的Generalized Intersection over Union

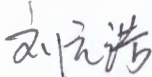
(GIoU)值，根据置信度、GIoU值对检测结果进行最终判断，以降低误报率。该方法还考虑到不同模态之间的互补性和差异性，对不同模态之间存在冲突或不一致的情况进行处理，能够有效地利用可见光和热红外两种模态的检测结果，提高火焰检测的准确性和鲁棒性。针对当前单模态双目相机火焰定位易受环境光干扰问题，我们提出了一种基于可见光相机与近红外相机的异源双目火焰定位方法。首先，使用异源双目相机标定方法对可见光相机和近红外相机进行内外参数标定；然后，使用双目立体校正方法对可见光图像和近红外图像进行几何校正，并使用异源视差匹配与计算方法对校正后的图像进行视差计算；最后，根据视差信息和相机参数计算出火焰目标在三维空间中的位置。该方法有效解决了可见光图像与红外图像之间匹配困难，视差计算复杂等问题，实现对火焰目标的精确定位。

(二) 取得的业绩(代表作)【限填3项, 须提交证明原件(包括发表的论文、出版的著作、专利证书、获奖证书、科技项目立项文件或合同、企业证明等)供核实, 并提供复印件一份】

1. 公开成果代表作【论文发表、专利成果、软件著作权、标准规范与行业工法制定、著作编写、科技成果获奖、学位论文等】

成果名称	成果类别 [含论文、授权专利(含发明专利申请)、软件著作权、标准、工法、著作、获奖、学位论文等]	发表时间/授权或申请时间等	刊物名称/专利授权或申请号等	本人排名/总人数	备注
A Flame Localization Method Combing Visible-Light and Near-Infrared Cameras	会议论文	2023年09月23日	2023 6th International Conference on Intelligent Autonomous Systems	1/6	
第四届中国研究生机器人创新设计大赛三等奖	获奖	2022年08月26日		2/3	

2. 其他代表作【主持或参与的课题研究项目、科技成果应用转化推广、企业技术难题解决方案、自主研发设计的产品或样机、技术报告、设计图纸、软课题研究报告、可行性研究报告、规划设计方案、施工或调试报告、工程实验、技术培训教材、推动行业发展中发挥的作用及取得的经济社会效益等】

<b>(三) 在校期间课程、专业实践训练及学位论文相关情况</b>	
课程成绩情况	按课程学分核算的平均成绩： 83 分
专业实践训练时间及考核情况(具有三年及以上工作经历的不作要求)	累计时间： 1.2 年 (要求1年及以上) 考核成绩： 88 分 (要求80分及以上)
<b>本人承诺</b>	
<p><b>个人声明：本人上述所填资料均为真实有效，如有虚假，愿承担一切责任，特此声明！</b></p> <p style="text-align: right;">申报人签名： </p>	



## 浙江工业大学研究生学院

## 攻读硕士学位研究生成绩表

学号: 22160042	姓名: 刘元浩	性别: 男	学院: 工程师学院	专业: 机械	学制: 2.5年						
毕业时最低应获: 24.0学分		已获得: 24.0学分		入学年月: 2021-09	毕业年月: 2024-03						
学位证书号: 1033532024602138			毕业证书号: 103351202402600364								
学习时间	课程名称	备注	学分	成绩	课程性质	学习时间	课程名称	备注	学分	成绩	课程性质
2021-2022学年冬季学期	增材制造技术		2.0	73	跨专业课	2021-2022学年夏季学期	机器人智能控制		3.0	95	专业学位课
2021-2022学年秋季学期	中国特色社会主义理论与实践研究		2.0	92	公共学位课	2021-2022学年夏季学期	自然辩证法概论		1.0	81	公共学位课
2021-2022学年冬季学期	工程中的有限元方法		2.0	83	专业选修课	2021-2022学年春季学期	工程技术发展前沿		2.0	93	专业学位课
2021-2022学年冬季学期	科技创新案例探讨与实践		2.0	88	专业选修课	2021-2022学年春季学期	工程伦理		2.0	79	公共学位课
2021-2022学年冬季学期	智能工业机器人		2.0	84	专业学位课	2021-2022学年夏季学期	研究生英语		2.0	78	公共学位课
2021-2022学年秋季学期	研究生论文写作指导		1.0	82	专业学位课	2022-2023学年秋季学期	研究生英语基础技能		1.0	66	公共学位课
2021-2022学年春季学期	人工智能制造技术		2.0	90	专业学位课						

说明: 1. 研究生课程按三种方法计分: 百分制, 两级制 (通过、不通过), 五级制 (优、良、中、及格、不及格)。

2. 备注中“\*”表示重修课程。

学院成绩核算章:(60)分

成绩核算人: 张梦依

打印日期: 2024-04-02

**6<sup>th</sup> International Conference on Intelligent Autonomous Systems**

**ICoIAS'2023**

***Acceptance Notification and Invitation Letter***

6<sup>th</sup> International Conference on Intelligent Autonomous Systems

ICoIAS' 2023

Qinhuangdao, China

Sept. 22-24, 2023

[www.ICIAS.org](http://www.ICIAS.org)

**Yuanhao Liu, Zhifan Gong, Denghu Zhu, Qiao Deng, Songyu Hu, Jianzhong Fu**

Zhejiang University, China;

<sup>3 4</sup> Cohen Think Tank Fire (Zhejiang) Company Haining, China

Email: liuyuanhao@zju.edu.cn; fancarr@zju.edu.cn; zjtt0163@163.com; dengqiao13@126.com; syhu166@zju.edu.cn; fjz@zju.edu.cn;

**Dear Yuanhao Liu, Zhifan Gong, Denghu Zhu, Qiao Deng, Songyu Hu and Jianzhong Fu,**

We are pleased to inform you that, after the double-blind peer review (please refer to the attached files), your manuscript, as identified below, has been accepted for publication and oral presentation at the 6<sup>th</sup> International Conference on Intelligent Autonomous Systems (ICoIAS' 2023), going to be held in Qinhuangdao, China during September 22-24, 2023.

**Paper ID: AS23-526**

**Paper Title: A Flame Localization Method Combing Visible-Light and Near-Infrared Cameras**

Do prepare the camera ready manuscript in IEEE format (double column), by incorporating the suggestions/modifications as recommended by the reviewers; simultaneously do complete the registration formalities within the time frame as specified. *As per the IEEE Convention, pls. do NOT add any title or Designation (Prof./Dr/Mr/Ms/Smt/Research Scholar/Director/Head-of-the-Department etc.) before/after the name of the authors in the camera ready manuscript.*

As you are aware, all the accepted papers sent on time, duly registered and presented in the conference will be published in the ICoIAS' 2023 conference Proceedings (which will be archived in the IEEE Xplore, and indexed by Ei Compendex, Scopus).

We received an overwhelming response to the CFP for ICoIAS' 2023 & were able to accept only a small fraction of the entire set due to rigorous reviewing procedure and other constraints. On



# A Flame Localization Method Combining Visible-Light and Near-Infrared Cameras

Yuanhao Liu  
Polytechnic Institute  
Zhejiang University  
Hangzhou, China  
liuyuanhao@zju.edu.cn

Zhifan Gong  
Polytechnic Institute  
Zhejiang University  
Hangzhou, China  
fancarr@zju.edu.cn

Denghu Zhu  
Cohen Think Tank Fire  
(Zhejiang) Company  
Haining, China  
zjtt0163@163.com

Qiao Deng  
Cohen Think Tank Fire  
(Zhejiang) Company  
Haining, China  
dengqiao13@126.com

Songyu Hu\*  
The State Key Laboratory of  
Fluid Power and Mechatronic  
Systems, College of Mechanical  
Engineering  
Zhejiang University  
Hangzhou, China  
syhu166@zju.edu.cn

Jianzhong Fu  
Key Laboratory of 3D Printing  
Process and Equipment of  
Zhejiang Province, College of  
Mechanical Engineering  
Zhejiang University  
Hangzhou, China  
fjz@zju.edu.cn

**Abstract**— As common natural disasters, fire incidents seriously threaten human life and property. Flame target localization is a key step to enabling timely detection and effective control of the fire incident. In this study, a novel approach that combines visible-light and near-infrared cameras in a heterogeneous binocular camera system is proposed. High-precision localization of early small-scale fire targets is achieved through calibration of the heterogeneous binocular camera, stereo rectification, heterogeneous disparity matching, and depth computation. Experimental tests were conducted within a range of 1–5 meters, validating the effectiveness of the proposed algorithm.

**Keywords**—fire incidents; flame target localization; visible-infrared; binocular stereo vision

## I. INTRODUCTION

Fire incidents pose a continuous hazard in daily life. With advanced robotics technology and firefighting techniques, autonomous robots equipped with mobile chassis and firefighting equipment have been gradually applied in practical scenarios. These robots are capable of automatically detecting flames and performing automatic firefighting. Figure 1 illustrates an autonomous firefighting robot developed by our team. Automatic flame localization is a crucial step in the operation of such robots.

Extensive research has been conducted on flame detection and localization. Li et al. [1] and Sun et al. [2] employed full-spectrum cameras to detect flame potential fields based on temperature information. Other scholars have utilized single RGB cameras and deep-learning neural network methods to detect flames using flame images [3–10]. Sousa et al. [11] detected flames using a thermal imager combined with image processing algorithms. Sun et al. [12] captured images of fire scenes using an infrared camera to detect suspected flames. These methods utilize single-camera images for flame detection without accurately obtaining the three-dimensional spatial position of the flames.



Fig. 1. Autonomous firefighting robot

Compared with monocular vision, binocular vision offers higher precision in three-dimensional position detection and stronger robustness, making it more suitable for object localization and tracking in various environments and scenarios [13]. Jia et al. [14] proposed an edge-based adaptive window algorithm that accumulates adaptive weights based on rows and columns for non-edge regions in binocular visible light camera images. The flame position is determined based on the disparity map. Chaoxia et al. [15] designed a firefighting robot that utilizes binocular visible light cameras and proposed a flame detection method designed specifically for weakly aligned images. A method based on Faster R-CNN is used to detect flames in keyframes of a localized video stream. Sun et al. [16] presented a 3D infrared imaging system based on binocular stereo vision. This system uses two visible light cameras to acquire 3D surface measurements of objects and integrates the surface temperature information obtained from the infrared camera with 3D measurement data. These binocular vision localization techniques are based on two visible light cameras. The feature point matching of the fire points are achieved through homogenous image matching.

The combination of visible light and infrared cameras has become a common setup for flame detection because of the

significant infrared radiation emitted by flames [17]. Therefore, research on integrating visible light and infrared cameras for flame localization has significant practical significance. He et al. [18] investigated the detection and localization of wildfires on transmission lines using a combination of visible light and infrared thermographic cameras. Significant variations in the grayscale field of captured photos can be observed due to the differences in imaging mechanisms between visible light and infrared thermographic cameras, making the flame depth calculation challenging. Moreover, images captured by infrared thermographic cameras have fewer imaging details, are sensitive to temperature variations, and are costly [19]. Consequently, in recent years, researchers have begun exploring the combination of visible light cameras and near-infrared cameras for flame detection [20].



Fig. 2. Thermal infrared (TIR) sensors image



Fig. 3. Near-infrared camera image



Fig. 4. Visible light camera image

## II. METHODS

The proposed method consists of four main steps: heterogeneous binocular camera calibration, stereo rectification, heterogeneous disparity matching, and depth information computation.

### A. Heterogeneous Binocular Camera Calibration

Accurately obtaining the intrinsic and extrinsic parameters of the cameras is crucial for binocular stereo ranging. In practical applications, camera calibration is typically required to acquire these parameters. Both visible light and infrared cameras can capture clear corner points on the calibration pattern. Zhang's calibration method [21] is employed in this study. Zhang's calibration method is based on a two-dimensional planar target. It involves capturing images of the planar target, such as a chessboard pattern, from different angles and then performing calculations and analysis on the corner points of the target to achieve camera calibration.

### B. Stereo Rectification

Binocular vision positioning technology is based on the ideal situation where the imaging planes of two cameras are coplanar, their optical axes are parallel, and their focal lengths are the same. However, ensuring that the camera imaging planes are coplanar can be difficult in practice. Therefore, performing stereo calibration on the imaging of the two cameras is necessary to align the imaging planes of the left and right cameras on the horizontal line and have the same viewpoint. In this way, the three-dimensional information of the scene can be restored by calculating the disparity value between the left and right images.

This paper adopts Bouguet's stereo calibration algorithm [22], as shown in Figure 5. By using the rotation and translation

Figures 2, 3, and 4 illustrate images of a typical flame target captured by an infrared thermographic camera, a near-infrared camera, and a visible light camera, respectively. By comparing the near-infrared image with the infrared thermographic image, it can be observed that the near-infrared image better preserves the fine details and textures of the flame. A comparison of the near-infrared image with the human visual system (HVS) or RGB image shows that the near-infrared camera retains the significant infrared radiation characteristics of the flame. To address the issues above, this paper proposes a heterogeneous binocular camera-based flame localization method that combines a visible light camera and a near-infrared camera. The method involves several steps, including heterogeneous binocular camera calibration, stereo rectification, heterogeneous disparity matching, and depth information computation, ultimately achieving precise flame localization.

matrices between the left and right camera coordinate systems, the two image planes are calibrated into strictly row-aligned planes. This step ensures that the number of reprojections of the two images is minimized and the observation area maximized. The specific calculation method is as follows:

First, determine the rotation matrix and translation vector between the left and right camera coordinate systems through Zhang's calibration method mentioned above; then, bring the rotation matrix and translation vector into Bouguet's stereo calibration algorithm to calculate the calibration rotation matrix. Finally, the calculated calibration rotation matrix is used to perform stereo calibration on the left and right images. Specifically, the pixel coordinate systems of the left and right images must be converted into camera coordinate systems through a common intrinsic matrix. Each of the two camera coordinate systems needs to be rotated to obtain a new camera coordinate system. Then, distortion correction and interpolation operations should be performed.

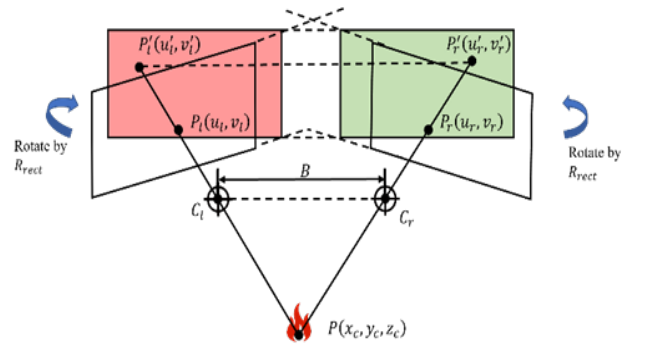


Fig. 5. Schematic diagram of stereo correction

### C. Stereo Matching

The purpose of stereo matching is to identify the corresponding point in the right image for each pixel in the left image, allowing for the calculation of disparity:

$$disparity = x_l - x_r \quad (1)$$

Where  $x_l$  and  $x_r$  represent the column coordinates of the two corresponding points in the images.

This paper adopts the Semi-Global Block Matching (SGBM) algorithm [23] based on local feature points for stereo matching. This method greatly reduces computational complexity as compared with global methods. The algorithm consists of four steps: pre-processing, cost computation, cost aggregation, and post-processing, which are implemented to achieve the matching.

1) *Pre-processing*: First, the image is pre-processed using a horizontal Sobel operator following the method shown in the following equation:

$$\begin{aligned} \text{Sobel}(x, y) = & 2[P(x+1, y) - P(x-1, y)] \\ & + P(x+1, y-1) - P(x-1, y-1) \\ & + P(x+1, y+1) - P(x-1, y+1) \end{aligned} \quad (2)$$

where the value of  $P$  represents the pixel value of the image,  $x$  and  $y$  represent the pixel coordinates in the image. The mapping function for the pixel value  $P_{NEW}$  on the image is given by the following equation:

$$P_{NEW} = \begin{cases} 0; & P < -PreFC \\ P + PreFC; & -PreFC \leq P \leq PreFC \\ 2 * PreFC; & P \geq PreFC \end{cases} \quad (3)$$

where  $PreFC$  represents the size of the mapping filter after horizontal Sobel pre-processing, with a default value of 15.

2) *Cost computation*: The cost consists of two parts: calculating the BT cost using the gradient map of the pre-processed image and calculating the BT cost using the original image without pre-processing. These two costs are added together. Finally, the obtained cost cube is summed within a rectangular window. The cost calculation formula is as follows:

$$C(p, d) = \lambda_1 * C_{\text{sobel}}^{BT}(p, d) + \lambda_2 * C_{\text{original}}^{BT}(p, d) \quad (4)$$

where  $C(p, d)$  is the cost of pixel  $p$  with disparity  $d$ ,  $\lambda_1$  and  $\lambda_2$  are weighting factors,  $C_{\text{sobel}}^{BT}(p, d)$  is the BT cost calculated from the pre-processed gradient map, and  $C_{\text{original}}^{BT}(p, d)$  is the BT cost calculated from the original image without preprocessing.

3) *Cost aggregation*: The cost aggregation is recursively calculated according to the following formula:

$$\begin{aligned} L_r(p, d) = & c(p, d) - \min_{i=d_{\min} \dots d_{\max}} L_r(p-r, i) \\ & + \min \left\{ \begin{array}{l} L_r(p-r, d) \\ L_r(p-r, d \pm 1) + P_1 \\ \min_{i=d_{\min} \dots d_{\max}} L_r(p-r, i) + P_2 \end{array} \right\} \end{aligned} \quad (5)$$

where  $L_r(p, d)$  represents the matching cost of point  $p$  along the  $r$  path with disparity  $d$ .  $p-r$  denotes the pixels located in front of  $P$  along that path.  $P_1$  and  $P_2$  represent the smooth penalties for cases with small and large disparity differences between the pixel points and their neighboring points, respectively, with  $P_1 < P_2$ . The cost  $C(p, d)$  can be either the BT cost or MI cost. The final aggregated cost is the sum of the aggregated costs along all paths, and its calculation formula is as follows:

$$S(p, d) = \sum_r L_r(p, d) \quad (6)$$

4) *Post-processing*: The aggregated cost needs to undergo post-processing, which includes uniqueness checking, subpixel interpolation, left-right consistency check, and connected component noise filtering.

### D. Heterogeneous disparity matching

The depth information is obtained by calculating the disparity using the disparity model. Figure 6 shows the image planes of the two cameras located precisely on the same plane. The optical axes are strictly parallel, at a fixed distance, and have the same focal length,  $f_l = f_r = f$ . Additionally, the left principal point,  $C_l$ , and the right principal point,  $C_r$ , have the same pixel coordinates in the left and right images, given as  $C_l(u, v) = C_r(u, v)$ .

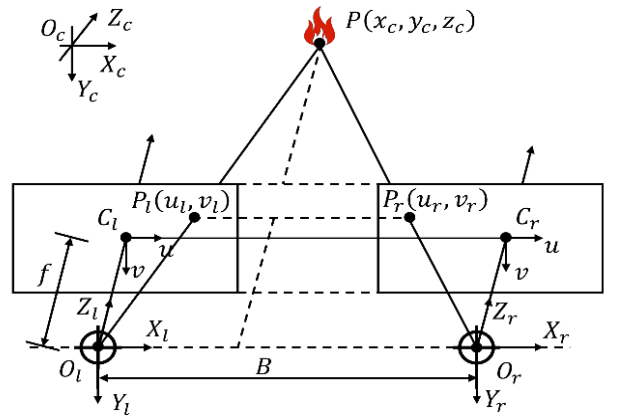


Fig. 6. Binocular vision camera model

The imaging coordinates of point  $P$  in the left and right images are denoted as  $P_l(u_l, v_l)$  and  $P_r(u_r, v_r)$ , respectively, and can be represented as follows:

$$\begin{cases} u_l = f \frac{x_C}{z_C} \\ u_r = f \frac{x_C - B}{z_C} \\ v_l = v_r = f \frac{y_C}{z_C} \end{cases} \quad (7)$$

After a simple calculation, we obtain  $P(x_C, y_C, z_C)$  as shown below:

$$\begin{cases} x_C = \frac{Bu_l}{u_r - u_l} \\ y_C = \frac{Bv_l}{u_r - u_l} \\ z_C = \frac{Bf}{u_r - u_l} \end{cases} \quad (8)$$

In the equation,  $f$ ,  $B$ ,  $u_l$ ,  $u_r$ ,  $v_l$ , and  $v_r$  represent the focal length of both cameras, the baseline distance between the two cameras, and the column and row coordinates of point  $P$  in the left and right camera images, respectively.

### III. EXPERIMENTS

In the experiments, the computer processor used was AMD Ryzen 5 5600U 2.3GHz, with 16GB DDR4 RAM, running on the Windows 11 operating system. The development environment consisted of Python 3.7 and OpenCV 4.7. The stereo camera platform used in this paper, as shown in Figure 7, consisted of a 400W pixel visible light camera and a 400W pixel near-infrared camera.



Fig. 7. Visible light and near-infrared binocular camera platform

#### A. Camera Parameter Acquisition and Stereo Rectification Experiment

According to the heterogeneous binocular vision model of visible light and near-infrared designed in Section II, 20 sets of valid calibration board images were captured simultaneously. Figure 8 shows the typical calibration board images captured by the visible-light and near-infrared cameras, respectively. The internal and external parameters of each camera were obtained after calibration. The optimal rotation and translation matrices were iteratively calculated based on the external parameters of each image. After obtaining the camera's internal and external parameters, stereo rectification was performed to obtain the reprojection matrix. The camera's internal and external parameters and the reprojection matrix are shown in Table I.



Fig. 8. Near-infrared camera images (on the left) and visible light camera (on the right)

TABLE I. STEREO CALIBRATION AND CALIBRATION PARAMETERS FOR BINOCULAR CAMERAS

Calibration Parameters	Near-infrared camera	Visible light camera
Intrinsic Parameters	$\begin{bmatrix} 512.73 & 0 & 325.13 \\ 0 & 512.21 & 248.10 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 518.19 & 0 & 323.69 \\ 0 & 518.23 & 252.67 \\ 0 & 0 & 1 \end{bmatrix}$
Baseline Length (mm)	124.543119	
Reprojection Error	0.069068	0.073972
Distortion Coefficient Matrix	$[-0.0043 \quad -0.4312 \quad -0.0004 \quad -0.0002 \quad 3.6336]$	$[-0.0043 \quad -0.4313 \quad -0.0004 \quad -0.0002 \quad 3.6336]$
Rotation Matrix	$\begin{bmatrix} 0.99999 & -0.00039 & 0.00008 \\ 0.00039 & 0.99999 & 0.00281 \\ -0.00008 & -0.00281 & 0.99999 \end{bmatrix}$	
Translation Matrix	$[-123.63824 \quad -0.43634 \quad -14.97943]^T$	
Reprojection Matrix	$\begin{bmatrix} 1 & 0 & 0 & -242.11692 \\ 0 & 1 & 0 & -242.71332 \\ 0 & 0 & 0 & 530.54718 \\ 0 & 0 & 0.00803 & 0 \end{bmatrix}$	

### B. Flame Target Localization Experiment

We used stereo matching to generate a disparity map after calibrating the images captured by the cameras. Figure 9 shows the flame targets captured by the left and right cameras, respectively. Figure 10 (a) shows the generated disparity map, and (b) shows the depth map calculated from the disparity map. The experimental results show the coordinate of the flame center is (253, 266), the center distance between the flame center and the center line of the binocular camera's optical center is depth = 1518.989mm, and the error between the calculated and measured data is 18.989mm.

The flame targets at different distances from the camera were tested to measure the ranging accuracy of this system, and the experimental results are shown in Table II. Table II shows the ranging error of the algorithm is within 7%, indicating a high measurement accuracy. Additionally, the measurement accuracy is highest at a distance of 3m because the camera in this study was calibrated at 3m, thereby obtaining higher accuracy in the camera's internal and external parameters.



Fig. 9. Flame targets captured by the left and right cameras

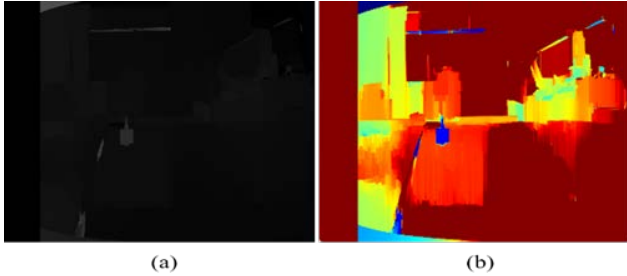


Fig. 10. Disparity map and depth map.(a) is DisparityMap. (b) is Depth map

TABLE II. TEST DATA OF FIRE POINT LOCATION

Actual distance/m	Disparity pixel coordinates of the flame centroid	Depth distance in millimeters/mm	Percentage error/%
1.0	(266,244)	993.624	6.4
1.5	(256,254)	1493.243	0.4
2.0	(248,245)	1990.991	0.4
2.5	(245,241)	2470.131	1.2
3.0	(232,240)	3003.454	0.1
3.5	(233,239)	3633.045	3.8
4.0	(239,240)	3802.935	4.9
4.5	(239,242)	4479.729	0.4

### IV. CONCLUSION

This paper proposes a flame target localization method based on near-infrared and visible light cameras. First, Zhang's calibration method is used to obtain the camera's intrinsic and extrinsic parameters. Then, the Bouguet stereo rectification algorithm is employed to obtain the camera rectification matrix. Finally, the SGBM algorithm is utilized for flame stereo matching to determine the two cameras' disparity. The relative distance from the camera to the flame is calculated using the triangulation model. Experimental results demonstrate that the proposed method can accurately locate the flame source with a depth positioning error of  $\leq 6.4\%$  for small target flames within 5m, exhibiting high measurement accuracy and meeting practical application requirements. Future research can focus on enhancing camera self-adaptive calibration based on the existing algorithm further to improve the algorithm's measurement accuracy and robustness.

### ACKNOWLEDGMENT

We appreciate the valuable opinions from various experts and researchers. This work was supported by the Key Research and Development Plan of Zhejiang Province [Grant No. 2022C01015], and the Fundamental Research Funds for the Central Universities [Grant No. 226-2023-00087].

### REFERENCES

- [1] Li T-J, Sun J, Yuan Y, et al. Simulation of calibration process in flame measurement by plenoptic camera [J]. Applied Thermal Engineering, 2018, 135: 179–187.
- [2] Sun J, Xu C, Zhang B, et al. Three-dimensional temperature field measurement of flame using a single light field camera [J]. Optics Express, 2016, 24(2): 1118.
- [3] Huang J, Zhou J, Yang H, et al. A Small-Target Forest Fire Smoke Detection Model Based on Deformable Transformer for End-to-End Object Detection [J]. Forests, 2023, 14(1): 162.
- [4] Kim B, Lee J. A Video-Based Fire Detection Using Deep Learning Models [J]. Applied Sciences, 2019, 9(14): 2862.
- [5] Abdusalomov B, Islam B M S, Nasimov R, et al. An Improved Forest Fire Detection Method Based on the Detectron2 Model and a Deep Learning Approach [J]. Sensors, 2023, 23(3): 1512.
- [6] Yang C, Pan Y, Cao Y, et al. CNN-Transformer Hybrid Architecture for Early Fire Detection [G]/PIMENIDIS E, ANGELOV P, JAYNE C, et al. Artificial Neural Networks and Machine Learning – ICANN 2022. Cham:Springer Nature Switzerland, 2022, 13532: 570–581.
- [7] Shahid M, Hua K. Fire Detection using Transformer Network [C]. Proceedings of the 2021 International Conference on Multimedia Retrieval. Taipei Taiwan:ACM, 2021: 627–630.
- [8] Zhao L, Zhi L, Zhao C, et al. Fire-YOLO: A Small Target Object Detection Method for Fire Inspection [J]. Sustainability, 2022, 14(9): 4930.
- [9] Shen D, Chen X, Nguyen M, et al. Flame detection using deep learning [C]. 2018 4th International Conference on Control, Automation and Robotics (ICCAR). Auckland:IEEE, 2018: 416–420.
- [10] Wang S, He Y, Yang H, et al. Video smoke detection using shape, color and dynamic features [J]. Journal of Intelligent & Fuzzy Systems, 2017, 33(1): 305–313.
- [11] Sousa M J, Moutinho A, Almeida M. Thermal Infrared Sensing for Near Real-Time Data-Driven Fire Detection and Monitoring Systems [J]. Sensors, 2020, 20(23): 6803.
- [12] Sun F, Yang Y, Lin C, et al. Forest Fire Compound Feature Monitoring Technology Based on Infrared and Visible Binocular Vision [J]. Journal of Physics: Conference Series, 2021, 1792(1): 012022.

- [13] McCann B C, Hayhoe M M, Geisler W S. Contributions of monocular and binocular cues to distance discrimination in natural scenes [J]. *Journal of Vision*, 2018, 18(4): 12.
- [14] Jia D, Wang Z, Liu J. Research on flame location based on adaptive window and weight stereo matching algorithm [J]. *Multimedia Tools and Applications*, 2020, 79(11–12): 7875–7887.
- [15] Chaoxia C, Shang W, Zhang F, et al. Weakly Aligned Multimodal Flame Detection for Fire-Fighting Robots [J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(3): 2866–2875.
- [16] Sun J, Ma H, Zeng D. Three-dimensional infrared imaging method based on binocular stereo vision [J]. *Optical Engineering*, 2015, 54(10): 103111.
- [17] Rossi L, Molinier T, Akhloufi M, et al. A 3D vision system for the measurement of the rate of spread and the height of fire fronts [J]. *Measurement Science and Technology*, 2010, 21(10): 105501.
- [18] He Lif, Lu Jiazheng, Liu Yu, et al. Precise positioning technology for transmission line mountain fire using visible light-infrared multi-light source [J]. *High Voltage Technology*, 2018, 44(8): 2548-2555.
- [19] Kalashnikov D A, Paterova A V, Kulik S P, et al. Infrared spectroscopy with visible light [J]. *Nature Photonics*, 2016, 10(2): 98–101.
- [20] Burnett J D, Wing M G. A low-cost near-infrared digital camera for fire detection and monitoring [J]. *International Journal of Remote Sensing*, 2018, 39(3): 741–753.
- [21] Zhang Z. A flexible new technique for camera calibration [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330–1334.
- [22] Ma Yongjun. Binocular vision target detection and positioning based on visible light and infrared thermal imager [D]. Dalian Maritime University, 2018.
- [23] Zhang Shuaishuai. Research on the visual system of fire-fighting robot fire extinguishing operation [D]. China University of Mining and Technology, 2022.



# 荣誉证书

中国研究生创新实践系列大赛  
“申昊杯”第四届中国研究生机器人创新设计大赛

龚志帆、刘元浩、洪纵横

参赛作品“实现早期火焰自主识别和扑灭的消防机器人系统”荣获

## 三等奖

中国学位与研究生教育学会



中国科协青少年科技中心



2022年8月26日