

分类号： TP391

密 级： 公开

学校代码:10363

学 号: 2210330131



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 大视场高分辨率下的高速飞行
目标速度与姿态测量

论 文 作 者 董家辉

校 内 导 师 王力超

校 外 导 师 郝刚

专业学位类别 电子信息

研究方向（领域） 检测技术与自动化装置

论文提交日期： 2024 年 5 月 31 日

类号: TP391
密类级: 公开

单位代码: 10363
学 号: 2210330131

题 目 大视场高分辨下的高速飞行目标速度及姿态测量

题 目 Velocity and attitude measurement of High-speed
Moving taigent with large field of View and high resolution

学 生 姓 名: 董家辉

校 内 导 师: 王力超

校 外 导 师: 郝刚

专业学位类别 : 电子信息

研究方向(领域): 检测技术与自动化装置

论文答辩日期: 2024 年 5 月 22 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：董家辉

日期：2024年5月31日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：董家辉 指导教师签名：王彬

日期：2024年5月31日 日期：2024年5月31日

大视场高分辨率下的高速飞行目标速度及姿态测量

摘 要

随着计算机视觉技术的快速发展,视觉测量逐渐成为测量高速飞行目标运动学参数的重要方法,通过视觉测量可以获得飞行目标的技术参数,这对日常的生产和生活具有重要的工程实践意义。为了解决相机在获取高速飞行目标图像时面临的目标分辨率高低和拍摄视场大小之间的矛盾,文中设计了一种利用振镜与高速相机结合的测量系统,通过振镜改变相机视角,让目标始终保持在视场中央,实现了在较大范围内对高速飞行目标的跟踪拍摄,文中还设计了显著性目标检测算法和立体视觉测量算法,完成了对高速飞行目标运动学参数测量的任务。

本文的研究内容和主要贡献主要有以下三点:

(1) 为了解决高速相机拍摄飞行目标过程中视场大小和目标分辨率之间的矛盾,本文设计了一个振镜加高速相机组合的测量系统,设计内容主要包括系统整体架构设计,振镜的运动模型推导,核心器件选型,光路结构布局等,通过实验表明该设计可以在较大范围内实现对高速目标的跟踪拍摄,目标细节清晰可见。扩大了传统高速相机的视场范围。

(2) 为实现对被测量目标特征点的三维重构,我们需要对目标和背景进行精确分割,为此我们设计了一种基于深度学习的显著性目标检测算法,算法基于 U 型结构,在基础结构中加入多尺度混合注意力模块,以得到更加丰富和全面的特征表达,更好地融合语义和尺度不一致的特征,提高了模型的性能和泛化能力,增强模型在复杂场

景下的性能。同时还加入了边缘注意力模块，增强网络对目标边缘的关注以得到更加精细的边缘信息，实验结果表明该网络在挑战较大的 DUT-OMRON 的数据集中表现良好，F-measure, MAE 和 S-measure 指标分别为 0.768, 0.052 和 0.839。消融实验证明本文设计的模块使基准网络在 DUTS-TEST 和 ECSSD 数据集上的 F-measure 指标分别上升了 0.058, 0.105, MAE 指标分别下降了 0.053, 0.049, 证明了本文设计的有效性。

(3) 由于单相机加振镜的拍摄方式无法直接获取视差图的缘故，本文设计了一种基于重构单应矩阵的目标特征点三维重构方法，计算由振镜产生的虚拟相机间的外参矩阵以及目标在虚拟相机上的投影点来实现对目标特征点的三维重构，并依据此方法完成对飞行目标的速度，姿态等运动学参数的测量，实验结果表明目标速度测量的平均相对误差为 2.03%，姿态角测量的平均相对误差为 2.01%。

关键词：高速飞行目标，高速振镜，视觉测量，三维重构，深度学习

VELOCITY AND ATTITUDE MEASUREMENT OF HIGH-SPEED MOVING TARGETS WITH LARGE FIELD OF VIEW AND HIGH RESOLUTION

ABSTRACT

With the rapid development of computer vision technology, visual measurement technology has gradually become an important method to measure the kinematic parameters of high-speed flight targets, and the technical parameters of flight targets can be obtained through visual measurement which has engineering practice significance for daily production and life. In order to solve the contradiction between the target clarity and the shooting field of view size that the camera faces when acquiring images of high-speed flying targets, a measurement system combining a galvanometer and a high-speed camera is designed in the paper, which realizes the tracking and shooting of high-speed flying targets in a wide range by changing the camera's angle of view through the galvanometer and keeping the target in the center of the field of view, and the saliency detection algorithm and the stereo vision measurement algorithm are designed in the paper to complete the measurement of the motion of high-speed targets and to realize the tracking and shooting of high-speed flying targets. In the paper, a salient target detection algorithm and a stereo vision measurement algorithm are also designed, which accomplish the task of measuring the kinematic parameters of high-speed flying targets.

The research and main contributions of this paper are threefold:

(1) In order to solve the contradiction between the field of view size and target resolution in the process of high-speed camera shooting the flying target, this paper designs the shooting system design scheme of the combination of galvanometer and high-speed camera, and the main contents of the scheme include the overall architecture design of the system, the derivation of the motion model of the galvanometer, the selection of the core device, the layout of the optical path structure, etc. The experiments show that the design can realize the shooting of high-speed targets in a wide range, and expand the field of view range of traditional high-speed camera. It is shown through experiments that the design can realize the shooting of high-speed targets in a larger range, which enlarges the field of view of traditional high-speed cameras.

(2) In order to realize the 3D reconstruction of the measured target feature points, we need to accurately segment the target and the background, for this reason, we design a deep learning-based saliency target detection algorithm, which is based on the U-shape structure, and adds a multi-scale hybrid attention module to the base structure to get richer and more comprehensive feature representations, and to better fuse the semantic and scale-inconsistent features, which improves the model's performance and generalization ability, and enhance the performance of the model in complex scenes. The edge attention module is also added to enhance the network's attention to the target edges to obtain more fine-grained edge information, and the experimental results show that the network performs well in the dataset of the challenging DUT-OMRON, with the F-measure, MAE and S-measure metrics of 0.768, 0.052, and 0.839, respectively. The ablation experiments demonstrate that the module designed in this paper increases the F-measure metrics of the benchmark network on the DUTS-TEST and ECSSD datasets by 0.058, 0.105, and

decreases the MAE metrics by 0.053,0.049, respectively, proving the effectiveness of the design of this paper.

(3) Due to the limitation that parallax map cannot be obtained directly by the shooting method of single camera and single galvanometer, this paper designs a three-dimensional reconstruction method of target feature points based on reconstructed single response matrix, which can calculate the external reference matrix between virtual cameras generated by galvanometer and the projection point of the target on the virtual camera to realize the three-dimensional reconstruction of the target feature points, and complete the measurement of the kinematic parameters of the target's speed, attitude and so on based on this method. The experimental results show that the average relative error of target velocity measurement is 2.03%, and the average relative error of attitude angle measurement is 2.01%.

KEY WORDS: high-speed flying target, high-speed galvanometer, vision measurement, 3D reconstruction, deep learning

目 录

第 1 章 绪论	1
1.1 研究背景和意义	1
1.2 国内外研究现状	2
1.2.1 高速飞行目标测量方法研究现状	2
1.2.2 显著性目标检测算法研究现状	4
1.2.3 立体视觉测量技术研究现状	7
1.3 本文主要研究内容及章节安排	8
1.3.1 主要研究内容	8
1.3.2 本文章节安排	9
1.4 本章小结	9
第 2 章 高速飞行目标测量系统设计	11
2.1 本章引言	11
2.2 测量系统工作原理	11
2.3 振镜旋转模型建立	13
2.4 测量系统光路结构设计	17
2.5 本章小结	21
第 3 章 显著性目标检测算法研究	22
3.1 本章引言	22
3.2 显著性目标检测模型结构设计	22
3.2.1 模型整体框架	22
3.2.2 多尺度特征融合模块	25
3.2.3 边缘注意力融合模块	26
3.3 数据集和评估指标	28
3.3.1 数据集	28
3.3.2 评估指标	28
3.4 实验设置与结果分析	30
3.4.1 训练细节	30
3.4.2 损失函数	30
3.4.3 定量分析	31
3.4.4 定性分析	32
3.4.5 消融实验	33
3.5 本章小结	34

第 4 章 飞行目标速度及姿态测量算法研究	35
4.1 本章引言	35
4.2 视觉测量原理	35
4.2.1 相机成像坐标系	35
4.2.2 立体视觉原理	36
4.3 目标运动学参数计算	38
4.3.1 特征点三维重构	38
4.3.2 目标飞行速度计算	43
4.3.3 目标姿态角计算	43
4.3.4 坐标系转换	46
4.4 本章小结	47
第 5 章 实物实验及结果分析	48
5.1 本章引言	48
5.2 实验平台搭建	48
5.3 高速飞行目标拍摄实验	51
5.4 静态目标测量实验	52
5.4.1 目标位移测量实验	52
5.4.2 目标姿态角测量实验	55
5.5 飞行目标测量实验	57
5.6 本章小结	59
第 6 章 总结与展望	60
6.1 本文总结	60
6.2 展望	61
参考文献	62
攻读学位期间发表的学术论文目录	70
致谢	71

第 1 章 绪论

1.1 研究背景和意义

测量飞行目标的运动学参数对于研究目标的运动状态特别是高速飞行目标的运动特征、运动规律和动力学性质是十分重要的，可以通过运动参数，如速度变化率、速度的方向角等，更全面地描述和分析飞行目标的运动状态，这对于深入理解目标的运动过程、实现运动控制、评估运动效果和进行运动模拟具有重要的意义^[1,2]。随着测量技术的持续进步与发展，涌现出了多种多样的测量方法，主要可分为传感器测量和视觉测量两大类，常见的传感器包括加速度计、陀螺仪^[3]、GPS、惯性测量单元（IMU）^[4,5]等，通过将传感器安装在飞行目标上，完成参数测量，这些传感器可以直接获取物体的加速度、角速度、位置等关键参数，并以数字信号形式输出，便于后续的深入分析与处理。尽管这类传感器测量法具有较高的精度，但在某些情况下，尤其针对小型或高速运动目标时，传感器安装可能会改变目标结构，影响飞行状态，从而降低测量结果的准确性。此外，传感器可能存在的采样率不足、响应速率低等问题也会导致对快速变化的加速度和速度捕捉不准确，造成数据延迟与失真。高速运动目标的测量数据通常具有大量的噪声和干扰，需要进行有效的信号处理和数据滤波，使处理和分析这些数据的复杂性增加，这就需要更复杂的算法和技术支持，使用不便^[6,7]。随着计算机视觉技术的发展，视觉测量方式逐渐受到青睐，与传统的传感器测量方法相比，视觉测量能提供更为精确详尽的运动学参数，尤其适用于捕捉高速运动目标的微小变化和快速动作，图像可以提供丰富的视觉信息，包括位置、速度、加速度、运动轨迹、形态等。这些信息可以被用于更全面和细致的运动分析，并被广泛应用于姿态估计、形状重建等领域等，视觉测量的优势在于非接触式，适应性强，不会因安装传感器而影响目标的自然运动状态，使测量结果的精准度下降，它提供的图像和视频数据便于直观展示运动过程，通过对这些信息的精细分析，有助于详细解读飞行目标的运动特征与规律。这些优点使视觉测量在飞行目标运动学参数测量领域的应用越来越广泛^[8,9,10]。

利用高速相机拍摄飞行目标的一般方法是在目标的运动轨迹旁边定点摆放

高速摄影设备来获得高速目标的图像,但受到摄影机视场的限制,获得的目标运动段较短,解决办法是在目标运动轨迹方向上架设多台高速摄影设备,通过后期拼接,得到相对完整的目标运动图像,这些图像可以观测目标运动过程中的姿态变化以及可能出现的异常情况,为人们后续分析和改进飞行目标性能提供依据,同时这些图像还可以用于视觉测量,分析目标的运动速度,目标的姿态角如俯仰角,横滚角,偏航角^[11,12]等,对于旋转目标还可以分析目标的转速等运动学参数。但这种拍摄方法需要多台高速相机同时进行拍摄,多台高速摄影设备的同步,标定以及摆放位置的选择会耗费大量的人力物力,成本较为高昂,使用起来对环境要求较高,非常不便,并且后期图像拼接也会导致视频不连贯,容易漏掉许多关键信息,不利于后期对图像的分析,成本和使用效果不匹配^[13]。针对上述问题,人们尝试了另一种方案,即使用一台安装在跟踪云台上的高速相机进行拍摄,通过旋转云台来跟踪拍摄运动目标。然而,由于传统高速摄影设备体积较大,云台电机的启动速度及转速往往难以满足高速运动目标的实时拍摄需求^[14,15]。所以在此基础上,人们又提出基于高速扫描振镜的凝视跟踪系统,在高速相机前方放置振镜,由于振镜质量小、启动速度快、旋转速率高,可通过调整振镜的角度来改变相机视角,从而实现对目标的连续跟踪拍摄,有效地捕获完整的运动目标图像。这种拍摄方式解决了目标视场和分辨率之间的矛盾,可以获得高分辨率,大视场的目标飞行图像,为后续的测量提供了基础,基于上述背景本课题设计了一个基于振镜的高速目标跟踪系统,对近一维的高速飞行目标进行跟踪拍摄,并设计了目标运动学参数的测量算法,测量目标飞行过程的速度以及姿态角。

1.2 国内外研究现状

1.2.1 高速飞行目标测量方法研究现状

随着测量技术的不断发展,对高速飞行目标的测量手段也不断涌现,主要分为接触式测量和非接触测量两种方式,对于接触式测量方法,是将传感器安装在被测目标上,通过对传感器数据的分析和计算获得目标的运动学参数。如王德田等人利用任意反射面激光干涉和全光纤干涉仪测量内弹道高速弹丸的速度^[16]。冯志勇等人使用陀螺仪和加速度计信号融合的方法更加稳定地测量旋翼飞行器的姿态和速度^[17]。马秀丽^[18]利用磁阻和陀螺组合的方式测量目标的飞行姿态,建立了 UKF 的姿态角解算方法可以根据回传信号的变化快速计算出高精度的姿态

角。傅忠云等人^[19]通过改进传感器 MPU6050 的滤波算法,使其适用于多种运动姿态的测量系统。牛春峰^[20]利用陀螺仪和 GPS 组合测量高速飞行目标的姿态角。邵艳巡等人^[21]利用惯性与磁力传感器融合的方式测量目标姿态,这些通过传感器的方式可以直接实现目标运动学参数的测量,但是测量精度取决于传感器本身,这对传感器的精度要求很高,会使成本增大,并且在一些高空高速的恶劣条件下,传感器的精度会受到很大影响。

由于接触式测量方法的一些缺点,非接触测量方法逐渐受到研究人员的欢迎,具有很强的研究意义和应用价值。非接触测量主要通过相机记录目标图像,通过分析图像,测量目标的运动学参数。在测量高速飞行目标时,相机拍摄由于其非接触性,测量精度较高,获取信息量大,便于部署,信息容易保存等优点,被广泛应用于高速飞行目标的拍摄和测量中^[22]。主要的拍摄方法有狭缝摄影法,阴影照相法,高速摄影机法,跟踪式摄影法。狭缝摄影法的原理是目标和胶片的运动方向和速度是一致的,用胶片的运动补偿目标的运动,使目标在成像曝光过程中没有因目标运动而产生位移。狭缝垂直于目标像的运动方向,是目标像从头到尾依次局部曝光来形成一幅完整的像^[23]。李金珂等人^[24]设计了一种基于线阵 CCD 的同步式狭缝摄影,可以用于测量飞行目标的运动学参数,但受到 CCD 尺寸的影响,只能用于测量低速目标。李冬冬^[25]设计了一种解决狭缝相标定和晃动的问题,对拍摄的图像进行矫正并减小了由于相机晃动带来的误差。王可东等人^[26]提出使用 CMOS 阵列式狭缝摄影对高速飞行目标进行拍摄,相比于 CCD 传感器 CMOS 具有集成度高,功耗低,成本低廉等优点。狭缝摄影法可以捕捉目标的空间信息和时间信息,为测量提供了便利,同时狭缝摄影受视角限制较小,目标可以在自由的位置,只要目标从头到尾经过相机视场即可,但是通过狭缝摄影方式得到的图像解译与测量较为困难。

阴影照相法通过在目标飞行轨迹沿线设立多个阴影照相站点,构建起一条摄像轨道,在每一个照相站点内部署两台阴影相机,分别位于轨道两侧,以正交角度对飞行目标进行拍摄,从而获取两幅飞行目标的阴影图像。接下来,在轨道区域内建立一套精密的空间基准系统,以便进一步分析和处理这些图像信息^[27,29]。如乔志旺^[30]利用阴影照相技术去测量外场的高速飞行目标的运动学参数,改进了外场测量校准系统,大大提高了测量精度。李璐^[31]改进了阴影照相图像的识别算法,实现对图像的自动判读,提高了系统的工作效率和测量精度。但阴影照相站

的缺点是操作过程复杂，调试效率低，记录的信息较少仅有目标的瞬时照片，无法拍摄目标长距离的飞行轨迹^[32]。

随着高速相机技术的不断发展，利用两台或多台高帧率相机对飞行目标进行拍摄，对不同视角拍摄的目标进行特征点匹配，利用标定好的相机内外参数与真实物理世界的距离，对目标特征点进行三维重构，得到目标的速度，加速度，角度^[33]等参数，这是目前常用的非接触式的测量方法，但是这种方法也存在着一些问题如单台相机视场过小，无法兼顾目标清晰度和拍摄范围，使用多台相机同时拍摄又面临着相机间的同步，后期图像拼接等问题。为了弥补这些问题又出现了跟踪拍摄设备，如利用光电经纬仪，跟踪拍摄飞行目标，但该设备体积大，结构复杂，位置改变较慢仅适用于远距离，慢速目标跟踪拍摄^[34]。为了适应高速飞行目标的大视场跟踪拍摄任务，出现了振镜加相机的组合，即在高速相机前摆放一面可以旋转的反射镜，反射镜以一定角速度旋转，跟踪记录高速飞行目标的运动过程，该系统可以清晰拍摄到目标在运动过程中的细节^[35]，如英国 MS 公司利用该原理研制了 Flight Follow 系统，用于对高速飞行目标的跟踪拍摄^[36]。东京大学实验室研发了一款利用二维振镜的跟踪系统，利用双轴振镜，改变相机的视角，可以在一个平面内捕获一个运动的目标^[37]。西安工业大学的张文娇，张嫚等人^[38,39]设计了一种弹道跟踪镜，利用步进电机驱动振镜，拍摄高速的弹丸，用于靶场实验。通过对比不同的高速飞行目标运动学参数的测量方法，利用图像进行的非接触式测量使用更加方便，对被测目标的影响更小，在拍摄方法上，基于反射镜的高速相机系统，可以在拍摄长距离的目标运动过程的同时，使目标一直保持在视场中央，解决了高速摄影中视场大小和目标清晰度间的矛盾，所以本课题计划设计一个基于扫描振镜的高速拍摄系统，并设计测量算法，利用拍摄的图像测量目标的运动学参数。

1.2.2 显著性目标检测算法研究现状

显著性目标检测技术在计算机视觉领域中有着非常重要的作用，它是一项对图像中最显著或最有区分度的视觉特征进行提取的技术，模拟人类的注意力，检测结果与人类感知很相似。显著性目标检测技术可以为其他的计算机视觉任务提供预处理的帮助，实现算法优化，更精准高效地实现视觉任务。例如在目标检测，图像分割，图像检索等方面。图 1-1 展示了场景中最显著的目标检测效果。



图 1-1 显著性目标检测的效果图

传统的目标显著性检测方法主要基于颜色、纹理、边缘和对比度等低级视觉特征的提取。这些特征可以使用直方图、梯度、拉普拉斯算子^[40]、中心偏差^[41]、对比度先验信息^[42]和背景先验信息^[43]和高斯金字塔等算法进行计算。随后，这些特征被整合生成图像的显著性图。由 Itti 等^[44]学者提出了一种基于人类视觉系统的显著性检测模型，包括三个处理阶段：低级特征提取、特征融合和显著性图生成。在低级特征提取阶段，模型采用高斯金字塔和差分金字塔等算法，提取图像的边缘、颜色和亮度等特征。在特征融合阶段，模型将所有特征结合起来生成显著性图，但由于复杂度较高，无法进行实时处理图像。Achanata 等^[45]学者提出了一种基于谱聚类的显著性检测模型，可快速准确地检测图像中的显著性区域。该模型首先对图像进行超像素计算，将图像划分为小区域，并将这些超像素分成不同的类别，进而计算类别之间的相似度，接下来，根据相似度矩阵，模型将超像素区域分为显著性和非显著性区域将显著性检测定义为二进制分割问题，这些作者受到早期 Ma 等人^[46]基于对比度分析的图像注意力检测模型的启发。Yuan 等人^[47]提出了一种显著性回归校正过程，通过定位并删除边界相邻的前景超像素，从而提高了基于边界先验的显著性估计的准确性和鲁棒性，还提出了正则化的随机游走排序模型，通过综合考虑区域和像素图像特征，在图像的每个像素上引入先验显著性估计，从而得到与像素细节和超像素无关的显著性图。Shan 等人^[48]提出了一种新的方法来提高现有流形排序算法的性能。首先使用背景权重图为流形排序提供种子点，接着将传统的流形排序扩展到二阶公式，并为其拟合项添加了权重掩模，最后为了进一步提高性能，建立了一个三阶平滑框架来优化显著性图。Wu 等人^[49]针对之前方法存在容易错过与背景对比度较低的一些目标区域提出了一种新的传播模型，利用变形平滑性约束来解决这个问题，通过在局部范围内对节点及其邻居进行正则化，变形平滑性约束可以防止错误的标签传播，从而使得与背景对比度较低的目标区域能够浮现出来。传统的显著性目标检测方法，往往基于低级特征提取，对于复杂场景和多目标检测任务的处理效果有限。由于

仅仅提取图像的边缘、颜色和亮度等基本特征，这种方法往往难以捕捉到目标的高级语义信息和上下文关系，很容易受到图像中的噪声干扰，这会导致显著性检测结果不够准确和稳定。尽管基于低级特征提取的方法在某些场景下检测效果较好，但其缺点限制了算法应用在更复杂、更精细的目标检测任务中。为了解决这些问题，研究者们不断探索新的方法和技术，例如利用神经网络模型以提高显著性目标检测的性能和应用范围。下面一部分就是针对基于深度学习的显著性目标检测算法的介绍。

随着卷积神经网络（CNN）的普及^[50]和全卷积神经网络的引入^[51,52,53]，显著性目标检测的研究出现了新的方向。与大多数的传统方法不同，CNN 网络通过局部连接和权值共享的机制来提取图像中的局部特征，从而更好地捕捉目标的空间局部性，这使得 CNN 网络在处理显著性目标检测任务时更具优势，能够准确地定位和识别目标。通过堆叠卷积层和池化层，可以从原始图像中学习到高级抽象的特征表示，CNN 网络在显著性目标检测任务中能够从原始图像中自动学习显著性特征，无需手动设计特征，并且 CNN 网络通过较大的感受野来融合更广泛的上下文信息，这使得网络能够考虑目标周围的背景信息，并改善目标的边界检测和分割。基于 CNN 的模型正在逐渐成为显著物体检测的主流方向^[54]。Lee 等人^[55]提出了 ELD-Net，这是一个用于准确高效的显著性检测的统一深度学习框架，证明了手工设计的特征可提供互补信息，以增强仅使用高级特征进行显著性检测的效果，高级特征使用 GoogLeNet 提取，而低级特征则通过评估局部区域与图像中其他区域之间的差异来评估其相对重要性。随着研究的持续进行中，注意力模块开始应用于各种视觉任务模型中^[56,57]，比如，Li 等人^[58]为了解决网络无法重视小区域目标，在区分前景与背景的复杂或精细结构时存在困难的问题，提出了一种新颖的卷积神经网络，具有净化机制和结构相似度损失。它基于空间和通道注意力机制，以促进对显著区域的关注，并加入了提出了矫正注意力，它是从错误预测区域学习而来的，并引导网络聚焦于易错区域，纠正错误。Cong 等人^[59]提出了一种关系推理模块，它结合了空间和通道维度，通过利用高层编码器特征来推测语义关系，从而促进生成更完整的检测结果。同时，还提出了并行多尺度注意力模块，通过利用多尺度注意力修正的底层特征，有效地恢复细节信息，并处理显著对象的尺度变化。Tang 等人^[60]提出了一种用于精确显著性目标检测的深度网络，为了充分利用底层和高层特征的不同优势，他们提出了一种新颖的

非局部跨层级注意力可以捕获长距离特征依赖关系,增强完整显著对象的区分度,同时还设计了一种新型的跨层级监督(CLS),通过像素级、区域级和对象级的学习,学习复杂结构的互补上下文信息,从而能够很好地恢复显著对象的细节和边界。综上所述,目前提高显著性目标检测算法的精确度的主要方法可以概括为:多尺度特征融合使网络适应图像中大小不一的目标;注意力机制的加入使网络建立起语义信息,细节信息等特征之间的相关性,使检测目标更加准确和完整;边缘信息的加入进一步增强了网络检测目标的精准度,提升了检测效果。本课题后续设计的显著性目标检测网络也将从这几个方面进行考虑。

1.2.3 立体视觉测量技术研究现状

为了利用视觉技术实现对飞行目标的运动学参数测量,我们对视觉测量算法的研究现状进行了调研。当前主流的视觉测量方法是基于立体视觉,立体视觉技术是通过不同角度的两台或多台摄像机对同一场景的图像进行采集,或者通过变换一台摄像机的空间位置拍摄同一场景,再基于视差原理计算出空间坐标点在摄像机图像中的像素坐标,进而获得该空间坐标点的三维空间坐标值,再结合物体真实的物理尺寸可测量飞行目标相对于相机坐标系的位移量^[61,62,63]。如 Liu 等人^[64]提出了一种基于立体视觉的喷管位移测量方法,测量误差为 0.5mm,精度较高但需要使用多台相机同时拍摄,系统复杂度较高,标定困难仅适用于目标小位移的测量。Nomura 等人^[65]使用立体视觉测量系统,对船舶上的特征点进行三维建模,从而计算船舶的靠岸距离精度在 1m 以内,有很强的实用性,但在使用中需要事先得到目标的模型数据,应用场景范围较小。Yang 等人^[66]提出了一种基于双目立体视觉系统的测速站,使用检测网络对车牌进行检测,再进行特征匹配重构车牌上三维点可以测量通过车辆的速度,测量误差为 3.8%,Peng 等人^[67]提出双目立体视觉和激光雷达的结合,获得目标较为完整的点云数据,完成目标速度和姿态的测量,由于激光雷达的加入系统有较高精度,但成本较高。Chen 等人^[68]提出了基于双目视觉的工件高精度测量方法,测量精度达到 0.018mm,但只能用于形状规则的工件,使用局限性较大。这些传统的立体视觉具有以下缺点:

- 1、使用两台或多台相机会增加测量成本,两台相机的物理参数必须保持一致,否则会在测量中引入较大误差。
- 2、多台相机必须同时对目标进行拍摄,否则可能无法找到对应的匹配点,这多台相机间的同步系统要求较高,增加使用成本。

3、多台相机间的位置必须保持不变，一旦一台相机出现移动就需要对整个系统进行重新校准，对使用环境要求较高。除了双目或多目测量系统，基于单目视觉的测量系统也有很大的发展如：Lin 等人^[69]提出将单目相机放置在旋转平台上进行拍摄，实现获得视差图的效果，可以测量目标的三维坐标，误差在 2%以内，但该系统对一个目标需要进行两次拍摄无法用于运动物体的测量。Zhong 等人^[70]提出了利用单目高速相机测量高速旋转目标转速，系统结构简单，精度较高，但事先需要在目标物上做大量标记，操作复杂。Li 等人^[71]提出了一种基于激光位移传感器和红外视觉相机的测量方法，该系统可以在强烈的光照变化下测量物体的位姿但该方法无法测量高速运动目标。除了使用单目摄像头从多角度捕捉目标之外，人们还采用了使用镜子构建虚拟摄像头的方法，以捕捉目标的多角度图像。例如，Murmu 等人^[72]建议在摄像机前安装两个凸透镜，以便从单目摄像机的一次拍摄中获得一组视差图像，提取视差图像的成本更低，而且还能测量目标的位移。误差约为 5%。然而，该系统的视场大小受到凸透镜尺寸的限制，无法测量大范围的目标位移。Liu 等人^[73]通过棱镜旋转生成了一个空间自适应虚拟相机视场，该系统允许视场中的多个虚拟相机从不同角度捕捉任何目标，进行三维测量。Han 等人^[74]使用一台相机和一个双轴反射镜组成三维测量系统，利用反射镜改变相机的视角，整个系统结构简单，测量范围较大，然而，这两种系统只能用于静态目标的三维测量，无法应用于移动目标。综上，根据现有视觉测量方法的局限性，我们提出了一种基于旋转振镜与单目相机组合进行跟踪拍摄的视觉测量方法，通过将相机与振镜相结合，扩大了相机的视场范围，实现了远距离跟踪拍摄和测量。

1.3 本文主要研究内容及章节安排

1.3.1 主要研究内容

本文是基于视觉方式对高速飞行目标运动学参数的测量方法进行研究，主要研究内容包括：

(1) 高速飞行目标测量系统的设计，利用振镜加相机的组合拍摄出大视场高分辨率的高速飞行目标图像；

(2) 显著性目标检测算法研究，设计显著性目标检测算法的网络结构，用于分割图像中的目标和背景，为后续的测量进行预处理；

(3) 特征点的三维重构, 设计算法将图像中目标的像素点重构为三维坐标, 用于三维测量;

(4) 进行实验平台的搭建, 对测量系统的拍摄效果进行观测, 验证测量算法的有效性并对误差进行计算。

1.3.2 本文章节安排

本文分为六章, 具体的章节结构安排如下:

第 1 章: 绪论。主要论述了高速飞行目标测量技术的应用背景和研究意义, 概括了目前相关技术, 如高速目标运动学测量方法、视觉测量方法、显著性目标检测算法的研究现状, 并对本文的主要研究内容和文章结构进行了介绍。

第 2 章: 高速飞行目标测量系统设计。主要介绍了高速目标测量系统的工作原理, 系统组成, 振镜的旋转曲线计算, 光路结构设计等内容。

第 3 章: 显著性目标检测算法研究。主要介绍了用于分割图像目标和背景的显著性目标检测网络, 介绍网络的整体架构, 以及其中重要的模块。说明训练细节, 测评指标, 并在测试集上进行测试, 同时进行了消融实验验证我们设计的模块的有效性。

第 4 章: 飞行目标速度及姿态的测量算法研究。主要介绍了根据上文中高速测量系统拍摄的图片, 对目标上的特征点进行三维重构, 由于拍摄的图像是基于单目相机的图像, 无法使用传统方法进行重构, 我们设计了一种基于重构单应矩阵的算法, 计算特征点的虚拟对应点, 完成特征点的重构, 实现对目标的运动速度的测量, 并采用基于区域主轴理论的方式提取中轴线测量目标的姿态角。

第 5 章: 实验及结果分析。首先在实验室搭建高速飞行目标测量系统的实验平台和高速目标发射装置, 拍摄飞行目标进行验证, 证明拍摄系统的有效性, 并对图像进行处理, 利用测量算法测量目标的速度和姿态。

第 6 章: 总结与展望。对本文的工作和研究内容进行总结, 并对研究中出现不足提出建议和展望。

1.4 本章小结

本章介绍了高速飞行目标测量的研究意义与研究现状, 通过对文献调研和实际研究, 确定了本文的研究内容和研究目标, 并对全文结构进行说明。

第2章 高速飞行目标测量系统设计

2.1 本章引言

要通过视觉方式对高速目标进行测量，首先需要获取目标的高清图像，为此本文设计了一种针对高速飞行目标的测量拍摄系统，该测量系统的拍摄方式是利用高速振镜快速改变相机视角，使目标一直保持在视场的中央，从而解决了单个高速相机在拍摄时要获得大视场图像与捕捉高分辨率目标图像之间的矛盾。这套装置将振镜放置于高速相机前，并通过控制振镜的旋转角度来实现跟踪飞行中的高速目标，确保目标影像始终经由振镜反射在相机的感光元件上，进而实现对如飞行中的子弹、炮弹等在近一维平面上高速运动小目标的清晰成像。本章主要从高速飞行目标测量系统的工作原理，振镜的旋转模型建立以及系统的光路设计等几个方面进行介绍。

2.2 测量系统工作原理

要使目标始终保持在相机视场内，必须保证振镜转速与目标的速度匹配，首先需要测量飞行目标的初始速度，系统使用光纤传感器组成的测速系统，对飞行目标进行测速，将信号传输到单片机进行处理，得到目标的初始速度后根据建立的振镜的运动曲线模型得到振镜的角度变换曲线，将计算的角度转换为电信号输出用于控制振镜旋转，同时触发高速相机开始拍摄，记录目标飞行的过程，由于振镜的存在，目标在相机的视场里保持相对静止，消除了由于高速运动带来的图像模糊，并且扩大了拍摄范围，得到了高速飞行目标的大视场高分辨率的图像。

对于振镜触发控制方案，目前主要有两种：一种是仅有一个光纤传感器作为触发装置，振镜的旋转角度曲线根据目标的预计运动速度进行确定储存在控制系统中，当目标经过光纤传感器时触发振镜按照设置好的运动曲线进行旋转，实现对目标的凝视拍摄，该方法设计简单较为容易实现，但弊端是使用麻烦，需要提前了解拍摄目标的初始速度并计算曲线，不同的目标需要重新计算振镜的控制角度，并且需要控制器有大量的储存空间存储振镜的角度信息，当预估速度不准确时，容易使目标脱离相机的视场范围。另一种方式是利用两个光纤传感器，当

目标经过第一个光纤传感器时，开始计时，当目标经过第二个开关时计时结束，根据两个对射开关间的距离以及目标运动的时间，计算目标的飞行速度传输给系统控制器，根据速度计算振镜旋转曲线，以实现针对不同速度目标的有效跟踪。相较于第一种方法，该方案更具通用性和灵活性，对各种飞行速度的目标均有出色的跟踪性能。本文选择采用第二种设计方案。

该系统的组成部分主要包括，激光测速系统，嵌入式控制系统，高速振镜系统，高速相机系统。系统的整体结构如图 2-1 所示：

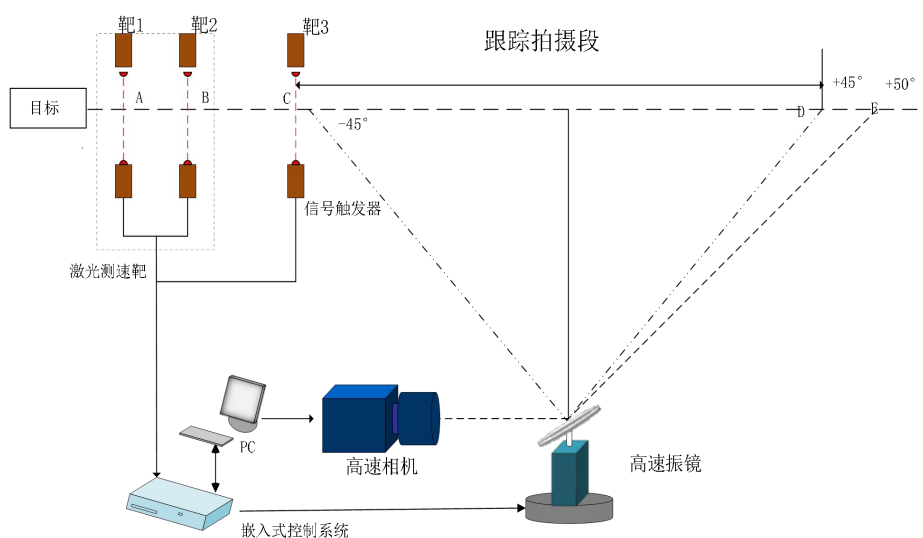


图 2-1 高速飞行目标测量系统结构图

在图 2-1 中激光测速靶 1 和靶 2 组成测速系统，当目标被发射并进入飞行状态后，首先穿越激光测速靶 1，触发控制系统启动定时器记录时间，当目标飞过测速靶 2 时，计时结束，结合激光靶间的距离和目标飞行时间，计算目标的飞行速度（由于目标为小型高速飞行目标且飞行时间较短，假设目标在飞行过程中不受外界因素干扰，在跟踪拍摄的过程中目标保持匀速直线运动）。得到目标飞行速度后根据推导出的振镜旋转曲线方程，计算振镜的位置，得到离散的振镜位置点，并转化为控制信号驱动振镜与目标保持相对静止，减少目标与相机视场的相对位移，消除高速目标的模糊现象。然而控制系统从计算目标飞行速度到依据振镜控制曲线方程计算出每个位置点所需的时间，会形成一定的延时，所以设置 BC 段为系统的延时段，确保系统有充足的时间来准备下一步操作，在 C 点设置激光靶 3，这个激光靶的主要作用是提供振镜开始旋转以及相机开始记录的触发信号，随后振镜开始按照计算的运动曲线进行运动。拍摄开始后，CD 段就是高

速相机的记录段，记录目标在这一段的飞行过程。系统选择的振镜是检流计扫描振镜，通过电压对振镜进行控制，不同的电压对应不同位置，在跟踪时，控制系统以一定时间间隔发送电压信号，控制振镜到达预定位置，使高速相机的视角可以始终凝视目标，由相机对振镜反射回来的图像进行记录，最终得到高速飞行目标清晰的运动图像，用于后续的测量分析。

2.3 振镜旋转模型建立

为了实现高速飞行目标的清晰成像，振镜系统是整个测量系统的重要组成部分，是清晰成像的关键。首先简要介绍振镜的基本原理和控制方式，系统使用扫描振镜来跟踪目标，这是一种高速扫描器件，广泛应用于激光打标设备，振镜通过通电线圈在磁场中产生力矩的原理来控制其转动。当线圈通电时，在转子上施加机械纽簧或电子方法的复位力矩，大小与转子偏离平衡位置的角度成正比，在振镜电机内部，通电线圈产生的电磁力矩会促使转子偏转至某一设定角度，此时电磁力矩与回复力矩达到动态平衡，转子因此不会连续旋转，而是精准地保持在一个可控的摆动范围内。振镜的偏转角度与所通入电流强度之间呈直接的线性关系，类似于电流计的测量原理，故也被形象地称为电流计扫描振镜。这种扫描振镜具有诸多优点，包括极高的响应速度、高效的高速扫描能力、微小的角度精确控制以及紧凑的体积和轻量化的负载特性。鉴于这些优势，扫描振镜被广泛应用于对快速移动目标进行实时跟踪的场景，尤其对于捕捉高速飞行物体的清晰影像，它是一个不可或缺的关键部件。

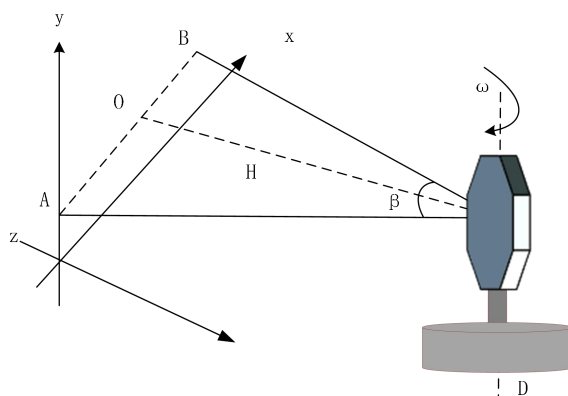


图 2-2 扫描振镜原理图

在图 2-2 中，扫描振镜通过旋转可以改变相机的视角变换到 X 轴上的指定位

置。振镜偏转的角度受电压控制其偏转角度为 ω ，两者的关系如式(2-1)：

$$\omega = KV \quad (2-1)$$

其中 V 为电压， K 为系数，通过控制电压的大小就可以控制振镜旋转的角度，要使高速运动的目标在相机上清晰成像，就要让相机的视角始终凝视目标，保证振镜与目标同步运动，在目标跟踪段通过振镜偏转改变相机视线。设振镜偏转 ω ，相机视角偏转 $\Delta\beta$ 。只有让相机视角保持对飞行目标的凝视，才能使目标始终出现在相机视场的中央，因此针对不同速度运动的目标，需要计算不同的振镜扫描曲线，以实现针对不同速度飞行目标的跟踪，由于振镜是由电压控制它的旋转角度，所以需要计算振镜旋转角度与目标速度之间的关系，再由控制系统输出相应的电压。图 2-3 是振镜跟踪段的示意图，接下来将根据示意图对振镜旋转模型进行推导。

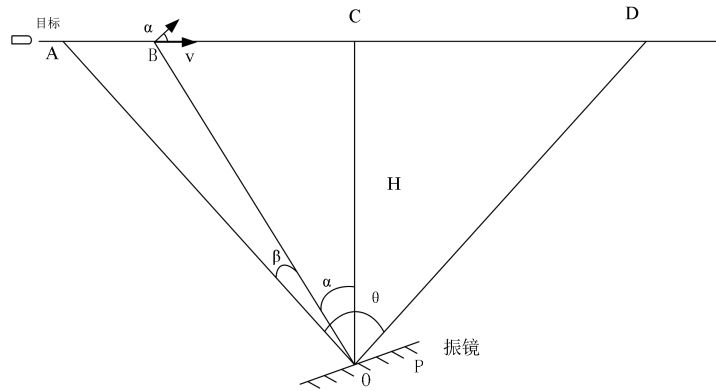


图 2-3 目标跟踪段示意图

目标进入跟踪段后，由光纤传感器负责触发信号，振镜开始旋转跟随目标，保证目标在视场中央。设目标的飞行速度为 V ，振镜跟踪的最大视场角为 θ ，振镜中心到目标轨迹的距离为 H ， C 为振镜可跟踪轨迹的中点，设目标到振镜中心的连线为相机的凝视线，需要一直跟随目标运动，凝视线与 CO 之间的夹角为 α ，此时凝视线从 AO 向 BO 旋转的角度是 $\Delta\beta$ ，目标运动时间为 t ，跟踪范围为 AC 全长为 S ，根据以上的参数设置，计算振镜的运动曲线。

跟踪段全长 S 为：

$$S = 2H \tan \frac{\theta}{2} \quad (2-2)$$

当目标在 AC 段时， BO 与 CO 的夹角 α 为：

$$\alpha = \arctan \frac{H \tan \frac{\theta}{2} - vt}{H} \quad (2-3)$$

当目标在 CD 段时夹角 α 为:

$$\alpha = \arctan \frac{vt - H \tan \frac{\theta}{2}}{H} \quad (2-4)$$

此时 AO 向 BO 转过的角度 $\Delta\beta$ 为:

$$\Delta\beta = \begin{cases} \frac{\theta}{2} - \alpha \\ \frac{\theta}{2} + \alpha \end{cases} \quad (2-5)$$

通过振镜旋转使 AO 向 BO 随时间变换的角度 $\Delta\beta$, 需要控制振镜旋转的旋转角度 ω , 使 BO 与 CO 的夹角 α 与 $\Delta\beta$ 的关系符合式(2-5)。 $\Delta\beta$ 与 ω 的关系如图 2-4 所示:

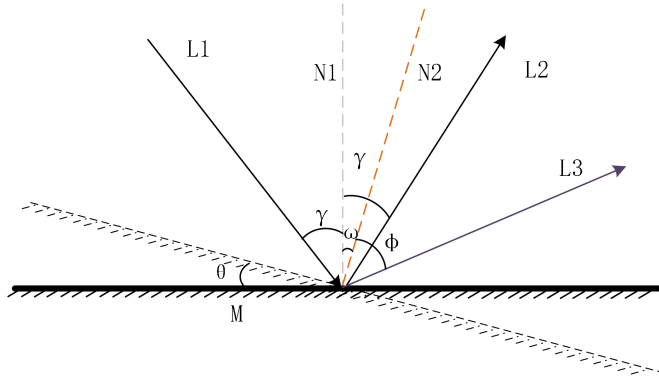


图 2-4 振镜旋转角度与相机视角间的关系

当入射光线 L1 的入射角为 γ 时, 反射角也为 γ , 当振镜 M 旋转 ω 后, 入射光线不变, 入射角为 $\theta + \gamma$, 反射角 ϕ 变为 $\theta + \gamma$, 法线由 N1 变为 N2 变化量为 θ , 所以反射光线 L2 与法线 N1 的夹角为 $2\theta + \gamma$, 所以当振镜旋转 ω 后, AO 向 BO 变换的角度 $\Delta\beta$ 为 2ω 。

所以在图 2-3 中, AO 向 BO 变换的角度 $\Delta\beta$ 时, 振镜的旋转角度为 ω , 两者关系如式(2-6):

$$\omega = \frac{\Delta\beta}{2} \quad (2-6)$$

结合式(2-5), (2-6)可得振镜的角度 ω 与目标运动时间 t 以及速度 v 的关系为

$$\omega = \frac{1}{2} \left(\frac{\theta}{2} - \arctan \frac{vt - H \tan \frac{\theta}{2}}{H} \right) \quad (2-7)$$

如果振镜拍摄的视场角最大为 90° ，目标运动速度 40m/s ， $H=2\text{m}$ ，不考虑目标飞行过程中目标所受的空气阻力和重力影响，目标为匀速直线运动，根据公式 (2-7) 可以得出振镜的角度随时间变换的曲线，如图 2-5 所示。在图中振镜角度随时间变化的曲线是一条连续的曲线，但振镜无法完全按照曲线进行转动，实际控制振镜转动的信号是离散的，由控制系统间隔 Δt ，输出电压值，刷新振镜的角度，通过这些离散的点，拟合出振镜角度随时间变化的曲线，为了提高跟踪精度，尽量使控制点更加密集，使飞行目标始终保持在相机的视场中央。 Δt 的确定是根据目标的飞行速度，相机视场大小以及振镜的阶跃响应时间来确定的，按照上文设定的目标飞行速度 v 为 40m/s ， H 为 2m ，最大视场角为 90° ，最大跟踪距离为 4m ，目标飞行时间为 100ms 。设相机的视场范围 w 为 80cm 。为保证拍摄效果，目标的中心点需要保持在视场中间部分，运动范围不超过整个视场宽度的 $3/5$ ，所以最大的间隔时间 $\Delta t_{\max}$ 为：

$$\Delta t_{\max} = \frac{3w}{5v} \quad (2-8)$$

代入数值求得 $\Delta t_{\max}$ 为 12ms 。为了拍摄视频的连贯性以及目标的成像大小，我们设定 Δt 为 3ms ，全过程由控制系统输出 33 次控制信号，即 33 个位置点，如图 2-5 中橙色点所示。

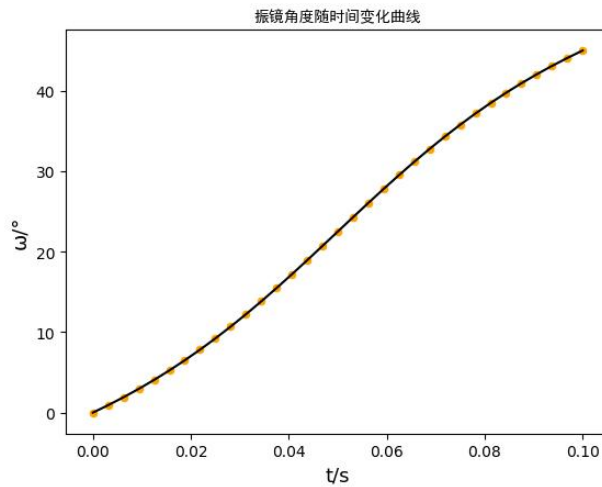


图 2-5 振镜角度随时间变化曲线及离散控制点图

2.4 测量系统光路结构设计

高速飞行目标测量系统的拍摄方式是通过振镜将目标图像反射进相机，所以为保证拍摄的视场大小和拍摄范围，需要对测量系统的光路结构进行设计。对于振镜加相机的组合，传统的光路设计一般是平行式设计，结构如图 2-6 所示，在图中相机与振镜处于同一轴线上，通过振镜的左右旋转来拍摄运动的目标。设置相机的光心为 C，相机光心和振镜中心的连线为相机的凝视线，代表相机视角的方向。根据图示，当振镜位于相机左侧并向左偏转到一定程度时，由于振镜反射面与相机位置重叠，可能会导致相机阻挡部分反射画面，这是振镜左偏的极限条件。跟踪开始后振镜向右偏转时，由于视角限制，振镜反射的目标区域在相机内所形成的图像面积过小，从而无法获取充分的有效信息，这是振镜向右偏转的极限条件。下面将针对这种平行式结构，计算其有效的跟踪范围。

如图 2-6 所示这种平行式的结构会限制拍摄系统的成像范围。当振镜向左转时目标通过振镜反射至相机的视线与相机相切时为振镜左转的极限角度，设此时振镜的宽度 AB 为 $2W$ ，相机到振镜中心的距离为 d ，以相机的光轴 CO 为基准，旋转角度为 α_1 ， $\angle COA = \alpha_1$ ，我们需要计算 α_1 的大小。以振镜中心为坐标原点，CO 方向为 x 轴，与 CO 垂直沿 AB 方向为 y 轴，则 C 点坐标为 $(-d, 0)$ ，A 的坐标为 $(-W \cos \alpha_1, W \sin \alpha_1)$ ，则直线 AC 的斜率 K_1 为 $W \sin \alpha_1 / (d - W \cos \alpha_1)$ ，直线 AO 的斜率 K_2 为 $-\tan \alpha_1$ ，AD 作为入射光线，通过振镜的反射光线为 AC，法线垂直于 AO，法线的斜率 K_3 等于 $-1/K_2$ 为 $1/\tan \alpha_1$ ，振镜左转的极状态就是振镜的左边反射光线与相机相切，切点为 D，相机本体的宽度为 $2W_v$ ，则 D 点的坐标为 $(-d, -W_v)$ ，直线 CA 与直线 AD 关于法线对称，设 AD 的斜率为 K_4 设直线 AD 的方程为 $y = K_4 x + b_4$ 则 K_1, K_3, K_4 满足公式(2-9)：

$$\frac{K_3 - K_1}{1 + K_3 K_1} = \frac{K_4 - K_3}{1 + K_4 K_3} \quad (2-9)$$

将 $K_1 = W \sin \alpha_1 / (d - W \cos \alpha_1)$ 和 $K_2 = 1/\tan \alpha_1$ 代入得到 K_4 的值。将 A 点坐标 $(-W \cos \alpha_1, W \sin \alpha_1)$ 代入求得 b_4 ，这样可以计算出 AD 的直线方程，再将 D 点代入直线方程，求解角度 α_1 ，该角度是振镜左偏的最大角度。

当振镜右偏时相机拍摄振镜的视场角 $\angle ACB$ 会不断减小，当减小到一定程

度时会影响目标的成像质量，拍摄的图像过于狭小，无法进行观测以及后续的测量工作。一般限定这个角度数为 20° ，此时振镜的偏转角为 α_2 ，为振镜右偏的极限角度。根据上文设置的参数，A 的坐标为 $(-W \cos \alpha_2, W \sin \alpha_2)$ ，B 的坐标为 $(W \cos \alpha_2, -W \sin \alpha_2)$ ，C 的坐标为 $(-d, 0)$ ，直线 AC 的斜率 K_1 ，直线 BC 的斜率为 $K_5 = -W \sin \alpha_2 / (W \cos \alpha_2 + d)$ ，直线 AC 和 BC 为 $\angle ACB$ 的两条边，所以对于 $\angle ACB$ 有：

$$\tan \angle ACB = \frac{K_1 - K_5}{1 + K_1 K_5} \quad (2-10)$$

极限条件下 $\angle ACB$ 为 20° ， $\tan 20^\circ = 0.36$ ，代入公式(2-10)可以计算出振镜的右偏极限角度 α_2 ，综上对于平行式光路结构要保证拍摄的图像清晰振镜的转动范围为 (α_1, α_2) ，由于相机凝视线变换的角度是振镜转动角度的 2 倍，所以该结构可以拍摄的视场范围为 $(2\alpha_1, 2\alpha_2)$ 。从上面的推导可以看出，对于平行式的拍摄结构，会存在图像会受到相机本体的遮挡以及随着镜子的偏转，相机的拍摄镜子的视场角过小的问题。

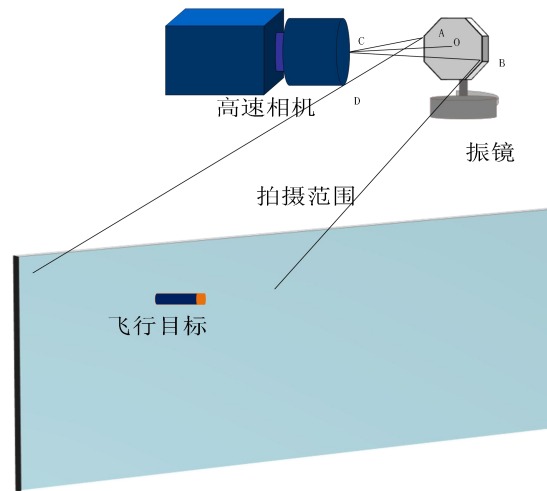


图 2-6 平行式光路结构

为了解决这一问题，扩大相机可拍摄的范围，系统采用仰角式光路系统，如图所示相机以一定角度上仰，摆放在振镜前，当振镜左右旋转时就不会出现相机遮挡图像，或者由于偏转角度过大而出现相机视场角过小的问题，大大增加了相机的拍摄范围。该结构的限制是相机仰角大小和相机距振镜的距离，下面对该结构的参数进行计算。

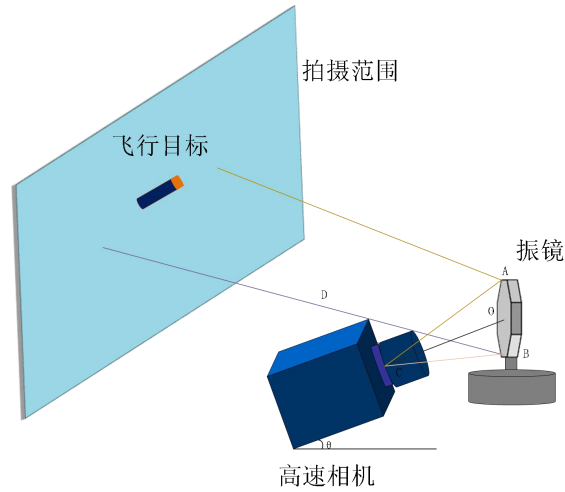


图 2-7 仰角式光路结构示意图

如图 2-7 所示，相机的仰角为 θ ，相机仰角的最小角度时振镜下沿反射的光线与相机上沿相切，设振镜的中心 O 为坐标系的原点，直线 AB 方向为坐标系的 Y 轴，过 O 点与镜面方向垂直的方向为 X 轴， A 点为振镜的上沿， B 为振镜的下沿， AB 长为 $2W$ ，相机光心到振镜中心的距离 OC 为 L ，则 A 点的坐标为 $(0, W)$ ， B 点的坐标为 $(0, -W)$ ，相机光心 C 点的坐标为 $(-L \cos \theta, -L \sin \theta)$ ，所以直线 CB 的斜率 $K_6 = L \sin \theta - W / L \cos \theta$ ， DB 通过振镜反射线为 CB ，所以 DB 与 CB 关于过 B 点的法线对称，法线方程为 $y = -W$ ，可以计算出 DB 的斜率为 $-K_6$ ，设 DB 的直线方程为式(2-11)

$$y = -k_6 x + b_6 \quad (2-11)$$

相机的高度仍然为 W_v ， D 点的坐标为 $(-L \cos \theta - W_v \sin \theta, -L \sin \theta + W_v \cos \theta)$ ，将 B 点坐标代入式(2-11)，可以计算出 b_6 ，将 D 点坐标代入式(2-11)，可以计算出相机的仰角 θ_1 ，即相机本体不遮挡振镜的最小仰角。与平行式光路结构一样，当相机仰角过大时，相机在振镜上的视场角会越来越小，当视场角过小时会影响成像质量，同样设置最小的视场角为 20° ， $\angle ACB = 20^\circ$ ， $\tan \angle ACB = 0.36$ ，相机的最大仰角为 θ_2 ， A 点的坐标为 $(0, W)$ ， B 点的坐标为 $(0, -W)$ ，相机光心 C 点的坐标为 $(-L \cos \theta_2, -L \sin \theta_2)$ ，可以求出直线 AC 的斜率 $K_7 = W + L \sin \theta_2 / L \cos \theta_2$ ，直线 BC 的斜率为 $K_8 = L \sin \theta_2 - W / L \cos \theta_2$ ， AC 和 BC 为 $\angle ACB$ 的两条边，所以：

$$\tan \angle ACB = \frac{K_7 - K_8}{1 + K_7 K_8} \quad (2-12)$$

将 K_7 , K_8 的表达式代入式(2-12), 可以计算相机的最大仰角 θ_2 , 综上相机的仰角范围为 $[\theta_1, \theta_2]$, 由于相机在左右方向上不会遮挡振镜, 当相机在合适的仰角范围内摆放时, 镜子的物理转角就是拍摄系统振镜可以使用的转角范围。

根据上述约束条件, 我们在 Matlab 中进行仿真计算, 计算不同参数下两种光路结构的拍摄范围, 主要包括当相机宽度, 振镜宽度, 相机光瞳到振镜中心的距离这些参数不同时振镜可以覆盖的范围, 结果如表 2-1 所示:

表 2-1 不同结构参数下仰角式和平行式振镜跟踪范围对比

结构参数 (单位: mm)			平行式光路结构 振镜可跟踪范围($^{\circ}$)	仰角式光路结构 振镜可跟踪范围($^{\circ}$)
W	W_v	d		
20	120	100	15.43	32.13
15	120	100	10.82	30.31
20	100	100	17.81	31.21
20	140	100	14.43	31.52
20	120	80	18.97	32.23
20	120	120	13.47	30.34

在表 2-1 中 W 表示振镜的宽度, W_v 表示相的宽度, d 表示相机光瞳到振镜中心的距离。从表中可以看出在不同的结构参数下仰角式结构都比平行式光路结构可以更大程度地利用振镜, 获得更大的拍摄视场。

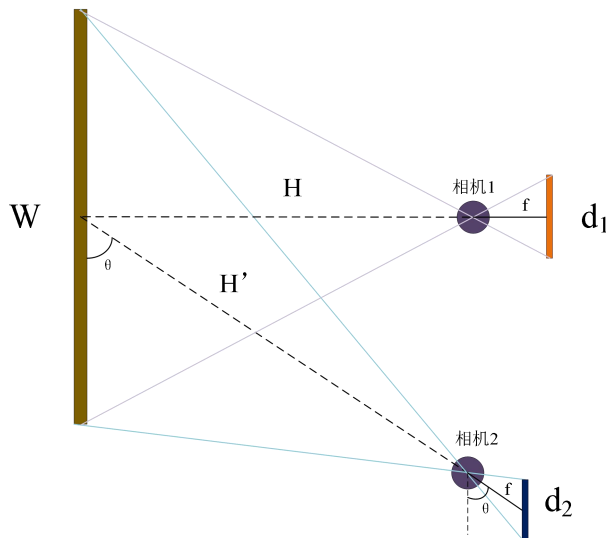


图 2-8 目标成像示意图

除了相机的倾斜角度和振镜的旋转角度影响拍摄范围外, 相机距目标的距离 d 也影响了跟踪范围和成像大小, 为保证目标在相机上成像的大小, 需要选择振镜距相机的合适距离 H。通过图 2-8 来分析物体在像中的大小与拍摄距离及拍摄角度的关系。在图中, 目标高为 W, 当目标与镜头平行放置时距镜头 H, 当镜

头与目标间有仰角为 θ ，距离为 $H'=H/\sin\theta$ ，镜头焦距为 f 。根据上图关系可以得出目标在相机 1 上的成像大小为 d_1 ，在仰角为 θ 的相机 2 上的成像大小为 d_2 。

$$d1 = \frac{f}{H} \times W \quad (2-13)$$

$$d2 = \frac{f}{H \sin \theta} \times W \quad (2-14)$$

设相机的视场范围为 L ，要实现成像清晰，目标在相机视场内占的比例越大拍摄效果越好，但目标所占相机视场的比例过大，会导致跟踪困难，目标会很容易脱离视场，为了平衡拍摄效果和跟踪效果，目标至少要占到相机视场的 $1/5$ ，但同时目标占据视场长度不能过大，所以设目标成像最大不超过视场 $3/5$ 的长度。所以对于仰角为 θ 的相机，相机到目标的距离范围为：

$$\frac{5fW}{3l \sin \theta} \leq H \leq \frac{5fW}{l \sin \theta} \quad (2-15)$$

2.5 本章小结

本章主要介绍高速飞行目标测量系统的工作原理，振镜跟踪模型的建立以及系统的光路结构选择。该系统的主要工作流程是通过激光测速靶，测速靶测量出目标的飞行速度后，由控制系统计算出振镜运动曲线，通过振镜改变相机的凝视点，跟踪飞行目标，保证飞行目标始终在相机视场中间以获取高速目标大视场高分辨率的图像，为后续的测量提供图像。在实际控制过程中我们根据目标飞行速度和振镜的阶跃响应时间确定控制信号的间隔时间 Δt ，计算每个间隔时间内振镜所要转动的角度，这些角度与时间的点组成一个离散的曲线，尽量使离散曲线和真实曲线吻合，提高拍摄画面的质量。接着为保证振镜可以拥有最大的跟踪范围，且保证拍摄过程中画面清晰不被遮挡，我们比较了平行式光路结构和仰角式光路结构，说明了仰角式光路结构的优点，并且计算振镜与相机之间的位置关系和边界参数，为后续实验平台的搭建和实际拍摄奠定了基础。

第3章 显著性目标检测算法研究

3.1 本章引言

在上一章中得到由高速飞行目标测量系统拍摄的目标图像,而想要对图像中目标的运动学参数进行测量,就需要对图像进行预处理,具体来说就是识别和匹配目标区域及其背景中的特征点,为此我们需要将目标和背景进行分割,传统的图像分割包括基于阈值,边缘,区域,图论以及能量泛函的分割方法,主要依据图片的灰度、颜色、纹理、形状等特征将图像进行划分区域,让区域外显差异性,区域内呈相似性。不过,上述传统分割技术可能无法满足对目标区域进行精细分割的要求,这直接关系到特征点匹配的准确性,进而影响测量的总体精度。所以课题采用基于深度学习的目标分割方法,在计算机视觉领域该项任务为显著性目标检测。在本章中设计了一种显著性目标算法对图像进行分割,主要内容包括检测算法的网络架构设计,各个模块的具体实现,训练和测试数据集以及评价指标和测试结果。

3.2 显著性目标检测模型结构设计

3.2.1 模型整体框架

显著性目标检测算法的核心任务是识别并精确提取图像中最为突出的物体区域。通过对图像所有像素点的显著性进行评估排序,生成显著性图,其中数值高的像素点代表其所在区域具有更高的显著性。显著性目标检测一般分为两种形式,一种是仅检测图像中最显著的部分,另一种是像素级的精细分割。第二种形式是本课题需要的方式,所以针对第二种形式进行研究,目前显著性目标检测主要面对的一些问题主要包括 1、目标尺度变化问题:大尺度目标一般在深层特征图中可以被捕捉到,而小尺度目标在图像下采样的过程中可能会被网络忽略,尤其在浅层特征中,语义信息较少,无法知道网络对目标进行定位,影响检测的准确性。2、相似背景和复杂背景的影响:当目标和背景特征差别较大时,检测效果较好。但在实际检测场景中,背景的纹理一般比较复杂,且可能和目标相似,导致部分背景被误判为目标,这极大地降低了检测的准确性。3、边缘信息模糊

的问题: 目前很多检测模型都注重区域检测的精度, 对于边缘的细化关注度不高, 由于边缘信息需要捕捉较小的感受野, 而语义信息需要较大的感受野, 这两种信息的捕获需要不同的网络通道来解决。

基于深度学习的目标检测算法发展出了多种网络架构, 早期的网络主要为单特征输出, 只输出网络最深层的预测结果, 这种网络结构简单但往往会丢失目标浅层特征, 使预测目标区域变得模糊, 接着出现了多层融合的网络结构, 通过融合不同尺度特征, 使预测结果有更加丰富的细节, 随着注意力机制在网络模型中的大量运用, 注意机制也被引入到显著性目标检测模型的特征提取部分中来, 通过对特征图中富有目标显著性信息的通道作为注意力权重, 增强模型对目标通道信息和空间信息的关注度, 这些改进都有助于获得更加进行的显著性目标检测图。本文在以上的研究基础上, 提出了一种基于多尺度特征融合和边缘注意力机制的显著性目标检测算法, 在该模型使用三个不同的分支提取特征图的注意力权重并将两个不同尺度的图片进行加权融合, 以得到更加丰富和全面的特征表达, 更好地融合语义和尺度不一致的特征, 提高模型的性能和泛化能力, 同时加入边缘注意机制模块, 增强网络对边界信息的关注度, 以提高检测目标边缘完整性方面的效果。

本文设计的基于多尺度特征融合和边缘注意力机制的显著性目标检测网络结构整体采用 U 型结构, 主要由三个部分组成, 特征提取网络 (ResNet-50)^[10]、多尺度特征融合模块 (MS-MAM 和 AFF)、边缘注意力融合模块 (EAM), 结构图如图 3-1 所示。

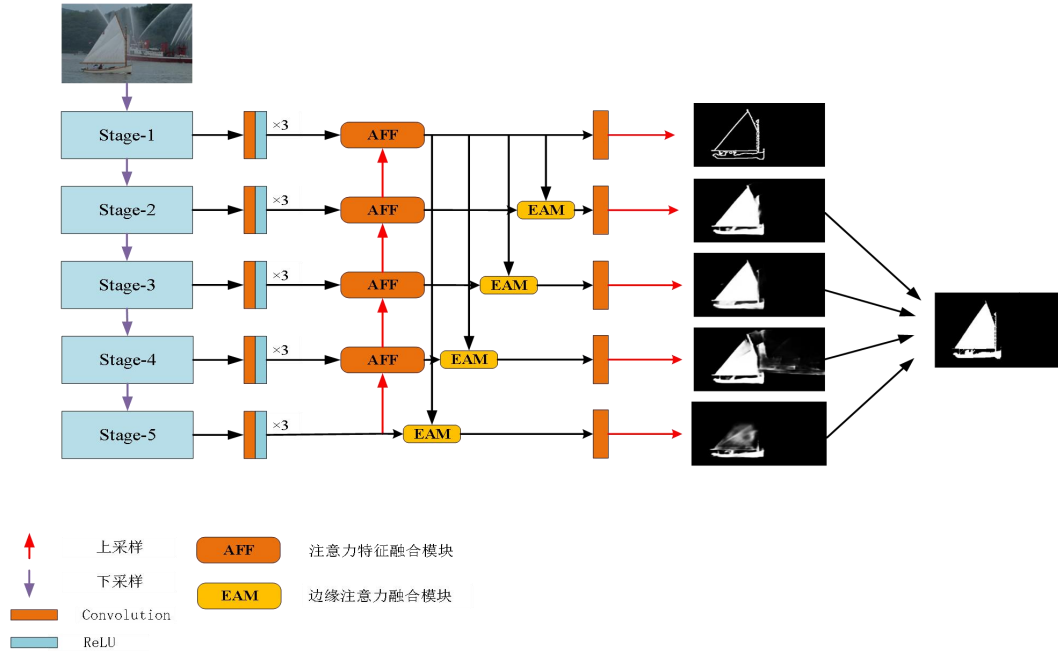


图 3-1 显著性目标检测网络框架

该网络采用经典的 U 型结构，通过编码器和解码器的结合，能够充分整合图像中的上下文信息，即在特征提取过程中能够同时保留局部和全局信息，这有助于提高模型对图像语义信息的理解，并且通过上采样和跳跃连接的设计，能够有效地融合多尺度的特征信息，这有助于模型对不同尺度的目标进行更准确的分割，同时也提高了模型对尺度变化的适应能力。一般传统的卷积神经网络只对图像单向的处理，在特征提取和压缩过程中会丢失部分细节信息，而 U 型网络结构的解码器部分可以逐步恢复特征图的大小，减少了信息丢失，有利于保留更细微的图像特征。基于以上优点，网络的整体结构采用 U 型结构。

网络的具体流程是：首先原始图像经过 ResNet-50 主干网络进行特征提取，得到 5 层不同尺度的特征后再进行 3×3 卷积和 ReLU 激活函数后，得初始的 5 层特征，为了更好地进行特征融合，我们设计了一种多尺度特征融合模块 AFF 对 5 层特征进行自下而上进行融合，为了获得预测图像更好的边缘细节，我们设计了边缘注意力机制模块，以增强对目标边缘特征的关注度，由于上层的特征图保留更多的语义信息，所以将最上层的融合特征作为边缘特征，将边缘特征与融合后显著型目标特征输入到边缘注意力机制模块，增强特征对目标边缘的关注度，增强后的 4 张特征图进行通道和尺度的对齐后进行特征融合再上采样至原图大小，输出显著性预测结果。

3.2.2 多尺度特征融合模块

为了更好地融合图像多尺度特征，丰富特征图所含信息，我们使用了一个多尺度特征提取模块。如图 3-2 所示，对特征图使用三个不同的分支进行注意力权重的提取，包括全局空间注意力，全局通道注意力，以及局部特征的通道注意力。其中空间注意力机制旨在强调图像或特征图中不同位置信息的重要性分配，因为并非所有区域对任务的贡献均等，只有那些与任务紧密相关的区域才是模型需要着重关注的焦点，通过学习并凸显图像中的关键位置权重，空间注意力机制能够让模型更专注于图像的具体部分，从而提升图像处理与理解的效能。网络使用不同尺度的两个分支关注特征图的通道注意力，它通过学习并量化每个通道的价值，进而为特征赋予相应的权重值。这一机制使得神经网络能够捕捉那些对当前任务至关重要的特征通道，同时弱化无关紧要的通道表达，如此一来，通道注意力便能够优化模型的特征提取能力，有针对性地强化有益于当前任务的特征图通道，从而增进模型的整体表征性能，我们参考 MS-CAM 模块，融合全局信息与局部信息的思想，将通道注意力分为全局和局部进行提取，并且增加了全局空间注意力模块，用于关注需要检测的目标区域。具体的结构如图 3-2(a)所示，以输入大小为 (B, C, H, W) 的特征图 X 为例，进入该模块后进入三个模块进行处理，得到三个分支为 $L(X)$ ， $S(X)$ ， $G(X)$ ，如公式(3-1)，(3-2)，(3-3)所示：

$$L(X) = B(pwconv_2(R(B(pwconv_1(Avg_1(X)))))) \quad (3-1)$$

$$S(X) = R(B(pwconv_3(Avg_2(X) + Map(X)))) \quad (3-2)$$

$$G(X) = B(pwconv_2(R(pwconv_1(X)))) \quad (3-3)$$

在上式中， B 表示 BatchNorm 层， R 表示 ReLu 激活函数， Avg_1 表示在空间维度上的平均池化， Avg_2 与 Map 分别表示通道维度上全局平均池化与全局最大池化， $pwconv_1$ 表示利用 1×1 的卷积核将输入特征 X 通道减少为原先的 $1/n$ ， $pwconv_2$ 表示利用 1×1 的卷积核将通道数目回复成与原输入通道相同的大小， $Sigmoid$ 为激活函数。特征图 X 经过多层特征提取模块后得到注意力权重 X' 经过 $Sigmoid$ 激活函数后得到特征图的注意力权重 A ，计算公式如式(3-4)所示：

$$A = M(X) = Sigmoid(L(X) \oplus S(X) \oplus G(X)) \quad (3-4)$$

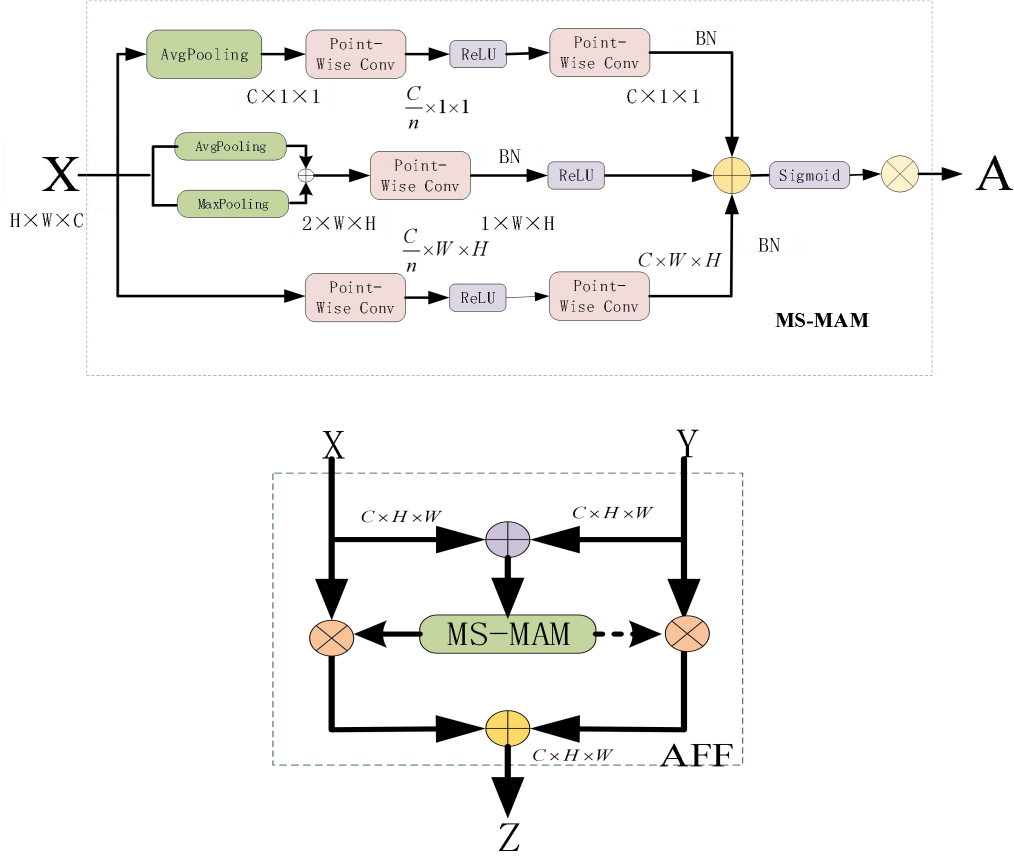


图 3-2 多尺度混合注意力模块

由于三种注意力加权模块的尺寸不一致所以三者相加需要广播操作。

利用特征融合模块(AFF)，对两个不同尺度的特征图 X ， Y 进行初始融合，融合后的特征图经过 MS-MAM 模块，得到不同分支的融合权重 A ，再对 X ， Y 做加权平均，这样通过训练让网络确定各自的权重，相比于对不同尺度的特征图的简单相加，融合模式也采用注意力融合的机制，使输入的特征图有更加完整的感知机制以得到更加丰富和全面的特征表达，更好地融合语义差别较大和不同尺度的特征，提高模型的性能和泛化能力。计算公式如式(3-5)所示，其中 $\oplus$ 表示将两个初始特征进行融合， M 表示 MS-MAM 模块， $\otimes$ 表示对应元素相乘。

$$Z = M(X \oplus Y) \otimes X + (1 - M(X \oplus Y)) \otimes Y \quad (3-5)$$

3.2.3 边缘注意力融合模块

为增强网络对目标边缘信息的关注度，我们设计了一个边缘注意力融合模块。由于第一层的浅层特征保留了更好的边缘信息，通常从 stage-1 中提取局部边缘

信息，但是想要得到的融合特征作为边缘特征，仅通过局部信息是不够的，还需要高级的语义信息和位置信息加持，所以除了 stage-1 的特征，还需要在模块中通过自下而上的传播方法，将低层特征和高层特征相融合，如图 3-3 所示得到用于提取边缘的特征图 F_e ，接着进行全局最大池化和均值池化得到 E_{Max} 和 E_{Avg} ，将生成的特征图进行维度上的拼接，在经过 3 层卷积后通过激活函数得到边缘注意力特征图 E_A ，之后需要将注意力图与特征图进行融合，一般方法是将边缘信息和显著性特征图从下到上逐步融合，但这种方法会稀释显著性边缘信息，所以我们采用一对一的融合模式，从下至上的特征图都与边缘注意力进行融合，以增强网络对于边缘的关注度使检测的细节变得更好。整个过程可以使用公式表示

$$F_e = F^{(1)} + UP(trans(F^{(2)}, F^{(3)}, F^{(4)}, F^{(5)}); F^{(1)}) \quad (3-6)$$

$$E_A = S(conv(Max(F_e) \oplus Avg(F_e))) \quad (3-7)$$

$$F^{(i)} = f(UP(F^{(i)}; F^{(1)}) \otimes E_A) \quad (3-8)$$

在式(3-6)中，trans 表示将每层的特征的通道数进行统一与 $F^{(1)}$ 相同， $UP(*; F^{(1)})$ 表示将各层特征图进行线性插值使其尺寸与 $F^{(1)}$ 大小相同。然后与 $F^{(1)}$ 相加，得到 F_e 边缘特征图。式(3-7)表示将 F_e 进行全局平均池化和最大池化，通过卷积后再进行激活函数得到边缘注意力特征图 E_A ，式(3-8)中 $F^{(i)}$ 表示第 i 层的特征图，将每层特征图线性插值到与 E_A 相同的尺寸，与 E_A 进行点乘，使每层特征图都增强对边界信息的关注度。

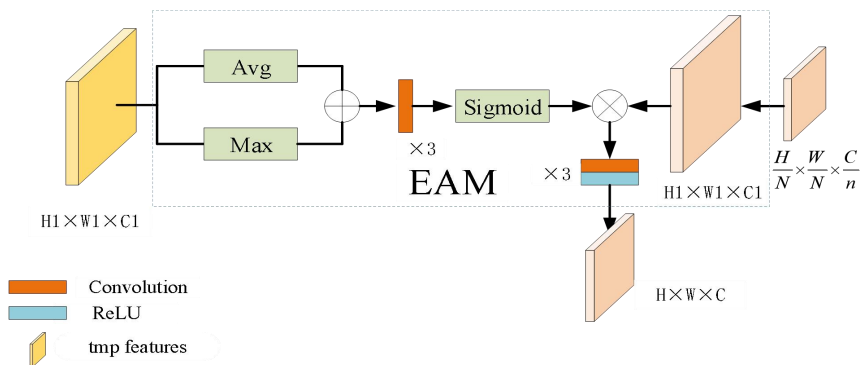


图 3-3 边缘注意力融合模块

3.3 数据集和评估指标

3.3.1 数据集

本实验共在 5 个最常用的显著性目标检测数据集上进行评估，分别是 ECSSD^[75]，PASCAL-S^[76]，DUTS^[77]，DUT-OMRON^[78]，HKU-IS^[79]。ECSSD 数据集共 1000 张图像，包含许多复杂场景下的图像，其中显著性的目标往往与复杂的背景相互交织，且具有丰富的上下文信息。这个数据集的特点在于它的多样性及包含的复杂结构，有助于评估显著性检测算法在处理复杂场景时的表现能力。DUTS 数据集是目前最大的显著性目标检测数据集之一，包括两个子集 10553 张训练图像以及 5019 张测试图像，它覆盖了广泛的场景、对象类别和规模，非常适合用来训练和评估显著性检测模型。PASCAL-S 数据集包含 850 张图像，它是从 PASCAL-VOC 数据集中扩展而来的，它包含了具有挑战性的自然场景图像，并提供了精确的人工标注显著性边界。此数据集强调的是对目标与背景边界模糊情况下的显著性检测精度。DUT-OMRON 数据集包含 5168 张图像，也专注于显著性目标检测，特点是包含了许多小尺寸且细节丰富的显著目标。由于其高难度和多样性，该数据集被广泛用于衡量显著性检测算法的性能极限。HKU-IS 数据集包含 4447 张图像，特别关注室内场景中的显著性目标检测，提供了一系列具有复杂室内环境和多对象交互的图像，该数据集对于检验显著性检测算法在特定环境下的鲁棒性和准确性非常有用。

3.3.2 评估指标

本实验使用 4 个指标对显著性结果进行评价，分别为 P-R 曲线(Precision-Recall Curve)，F-measure(mF)，平均绝对值误差(MAE)，S-measure(S)。

P-R 曲线是一种用于评估二分类模型性能的图形工具。该曲线有两个重要指标：精确度(Precision)和召回率(Recall)。精确度指模型在预测为正类别的样本中真正为正类别的比例又被称为查准率。召回率指模型成功预测为正类别的样本占实际正类别的比例又被称为查全率。P-R 曲线以不同的阈值为基础，展示了精确度和召回率之间的关系，可以衡量模型的泛化能力。在 P-R 曲线中，横轴为 Recall 召回率，纵轴为 Precision 精确率。对于 Precision 和 Recall 可以由以下公式算得：

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3-9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3-10)$$

在上式中 TP 为真正例，表示预测结果为正例，真实结果为正例，FP 为假正例，表示预测结果为正例，真实结果为反例，FN 为假反例，表示预测结果为反例，真实结果为正例。

平均绝对值误差 MAE 也叫平均绝对误差，在计算的时候，先对真实值与预测值求和，然后再取平均值。可以反映出预测显著性图 S 与真值 G 之间像素级的差异，用于评价模型的准确性和稳健性，MAE 的值越小，表示预测值和实际值之间的差异越小，模型预测性能越好，其计算公式如(3-11)所示：

$$\text{MAE} = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H \|S(x, y) - G(x, y)\| \quad (3-11)$$

在式中 W 和 H 分别为图像的宽和高， $S(x, y)$ 是网络预测显著性图中坐标为 (x, y) 处的像素值， $G(x, y)$ 是 Ground Thruth 中坐标为 (x, y) 的像素值。

查准率 P 与查全率 R 是一对矛盾的度量，一般来说，查准率高时，查全率往往偏低，查全率高时，查准率往往偏低。因此人们设计了一些综合考虑查全率、查准率的性能度量：F-measure，它是 Precision 和 Recall 加权调和平均如式(3-12)所示：

$$F_{\beta} = \frac{(1 + \beta^2) \times P \times R}{(\beta^2 \times P) + R} \quad (3-12)$$

其中 P 和 R 分别为精确率(Precision)与召回率(Recall)， β 为非负权重，当 $\beta > 1$ 时 P 有更大的影响， $\beta < 1$ 时 R 有更大的影响， β 为 1 时 F-measure 就是常见的 F1 度量，F1 综合了 P 和 R 的结果，当 F1 较高时说明所验证的方法有效性越高，本文的评价指标就选取 F1 度量。

结构度量(S-measure)作为一种新颖的显著性评价方法，它不仅考虑了像素级别的匹配程度，还关注显著图的结构一致性。相比于传统的基于精度(Precision)、召回率(Recall)或 F-measure 的评价方式，S-measure 特别强调了预测显著图在形状保持和空间分布上的准确性。

3.4 实验设置与结果分析

3.4.1 训练细节

本文所有的实验是基于 PyTorch1.12 的框架,使用 NVIDIA RTX3090 GPU 进行训练,训练集使用 DUTS 数据集共 10553 张训练图像。在输入网络前提前对图像进行预处理,将图像缩放到 320×320 像素的尺寸,并且随机进行水平、垂直翻转以及随机旋转以避免网络过拟合,以 Adam 优化器进行优化, Batch_size 设为 32, Image_size 设为 224×224 ,初始化学习率为 $5e-5$ 每 15 个 Epoch 将学习率设为原来的 $1/10$,共训练 50 轮。实验以在 ImageNet 数据集上预训练的 ResNet-50 作为 Baseline。

3.4.2 损失函数

模型分两个部分进行训练,第一部分是边缘检测部分,我们选择交叉熵损失函数,我们首先统计边缘预测结果中值为 1 的像素点(正样本)和值为 0 的像素点(负样本)的数量。接着,根据正负样本的比例,计算权重参数 α 和 β 。其中 α 表示负样本的比例, β 表示正样本的比例,但放大 1.1 倍,用于调整正负样本的权重。接下来,根据像素点的取值(1 或 0)和对应的权重参数计算权重。如果像素点的取值是 1,那么使用权重 β ;如果像素点的取值是 0,那么使用权重 α ,接着使用二元交叉熵函数进行损失计算,由于预测结果为图像边缘,像素值为 1(正样本)所占比例较低,这个更新后二元交叉熵损失函数考虑了正负样本的不平衡情况,并使用了权重参数来调整损失函数的计算,使得模型更加关注少数类。计算公式如式(3-14):

$$L_{\text{edge}} = -\frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W w(i, j) (G(i, j) \log S(i, j) + (1 - G(i, j)) \log(1 - S(i, j))) \quad (3-14)$$

其中 $S(i, j)$ 和 $G(i, j)$ 分别为预测显著性图和真值图中对应坐标为 (i, j) 处的像素值。 $w(i, j)$ 是根据预测图像素的正负样本情况动态调整的权重。

第二部分为增强后特征图的损失计算,我们每层除了使用二值交叉熵函数 L_{bce} 以外还使用结构相似性损失 L_{ssim} 作为损失函数。SSIM 损失函数类比图片的 3 个特征(亮度、对比度和结构性)来判断两张图是否相似,公式如下所示:

$$L_{ssim} = 1 - \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3-15)$$

其中 μ_x, μ_y 表示 N 个像素点的平均亮度通过对所有像素值进行平均来表示, σ_x, σ_y 表示图片明暗变化剧烈程度取所有像素值的标准差(方差的平方根)来表示, σ_{xy} 是它们的协方差, 表示图片的结构相似性。为了防止分母为 0, C_1 一般取值 0.01^2 , C_2 一般取值 0.03^2 。所以最终经过边缘特征强化后的损失函数如公式(3-16)所示:

$$L = L_{bce} + L_{ssim} + L_{edge} \quad (3-16)$$

3.4.3 定量分析

本文分别在 ECSSD, PASCAL-S, DUTS-TE, DUT-OMRON, HKU-IS 五个经典的数据集上对本文提出的模型进行了测试, 测试指标为 P-R 曲线(Precision-Recall curve), F-measure(mF), 平均绝对值误差(MAE), S-measure(S), 并与 C2SNet^[80], RAS^[81], R³Net^[82], PAGE^[83], AFNet^[84], CPD-R^[85], PoolNet^[86], RCSBNet^[87], NSS^[88]等经典网络模型进行对比, 为了公平起见, 本文引用了作者发布的预测结果, 使用相同的评估代码进行评估。

表 3-1 与当前先进算法定量比较

Methods	ECSSD			PASCAL-S			DUTS-O			HKU-IS			DUTS-TE		
	1000images			850images			5168images			4447images			5019images		
	MAE	S	mF	MAE	S	mF	MAE	S	mF	MAE	S	mF	MAE	S	mF
C2SNet	0.059	0.882	0.583	0.086	0.822	0.761	0.079	0.780	0.664	0.051	0.873	0.839	0.066	0.817	0.710
RAS	0.055	0.894	0.890	0.102	0.792	0.782	0.063	0.812	0.711	0.045	0.888	0.874	0.060	0.838	0.750
R ³ Net	0.051	0.910	0.853	0.101	0.809	0.775	0.073	0.819	0.690	0.047	0.894	0.853	0.067	0.837	0.716
PAGE	0.042	0.912	0.906	0.077	0.835	0.810	0.062	0.824	0.736	0.037	0.903	0.882	0.052	0.854	0.777
AFNet	0.042	0.913	0.908	0.070	0.844	0.821	0.057	0.826	0.738	0.036	0.905	0.888	0.046	0.867	0.792
CPD-R	0.037	0.918	0.917	0.072	0.842	0.822	0.056	0.825	0.747	0.034	0.905	0.891	0.043	0.869	0.805
PoolNet	0.039	0.921	0.915	0.074	0.845	0.822	0.055	0.836	0.747	0.032	0.916	0.899	0.040	0.883	0.809
RCSNet	0.033	0.925	0.912	0.060	0.859	0.841	0.045	0.855	0.752	0.027	0.920	0.920	0.036	0.890	0.840
NSS	0.035	0.913	0.925	0.064	0.901	0.913	0.064	0.824	0.783	0.322	0.901	0.913	0.044	0.853	0.840
Ours	0.032	0.925	0.925	0.061	0.857	0.841	0.052	0.839	0.768	0.028	0.918	0.911	0.034	0.889	0.841

从表中可以看出本文提出的算法的 MAE, F-measure, S-measure 指标在 5 个公开数据集上均有良好的表现, 优于其他算法。

如图 3-4 所示, 本文算法与其他 10 种算法在 5 数据集上的 P-R 曲线对比, 通常 P-R 曲线越靠近右上表示模型的效果越好。结果图表明本文提出的结构在 5 个公开数据集上具有良好的表现:

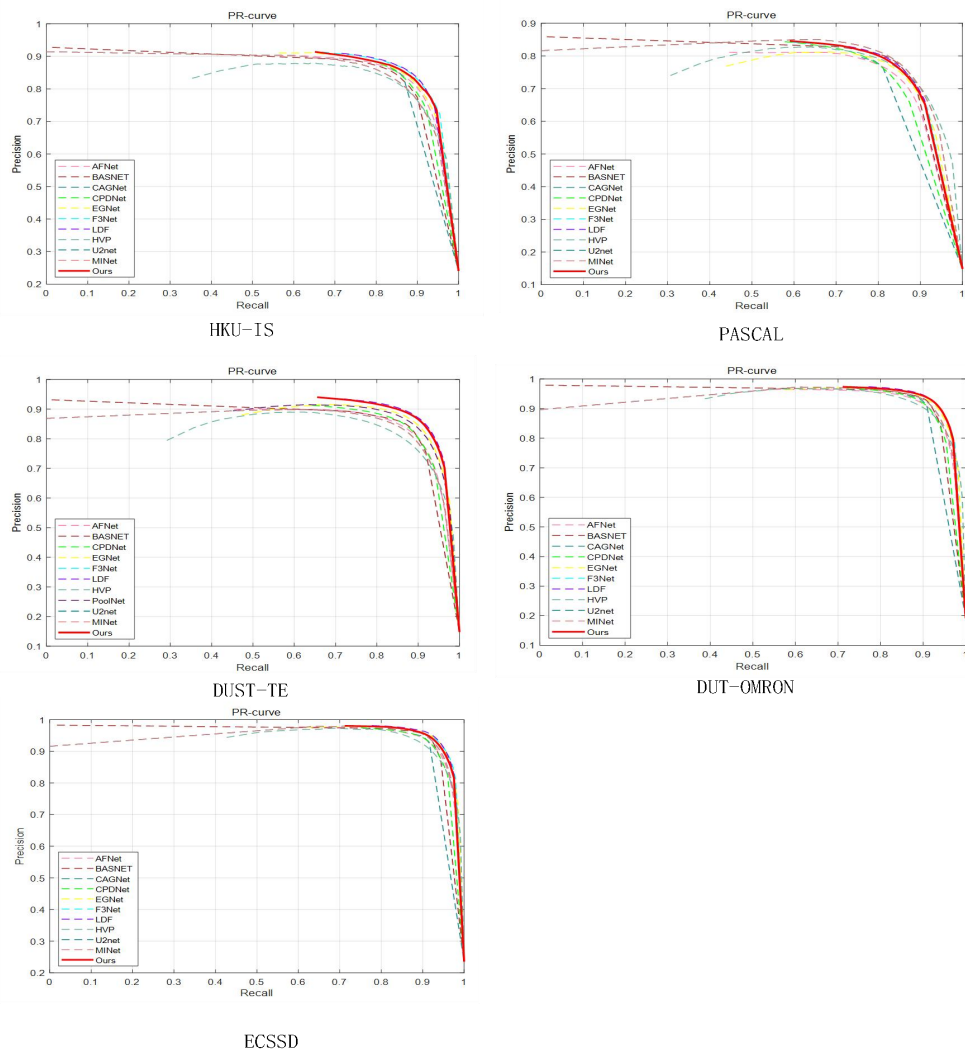


图 3-4 种算法在 5 个数据集上的 P-R 曲线

3.4.4 定性分析

为更加直观地展示本文提出网络模型的分割效果在图 3-5 中, 与 BASNet, F3Net, U2Net, C2SNet 四个模型比较了检测效果, 可以看到在复杂场景下, 与其他算法相比本文提出的算法分割的目标都更接近真实标签, 并且在边缘细节上也更加精细。

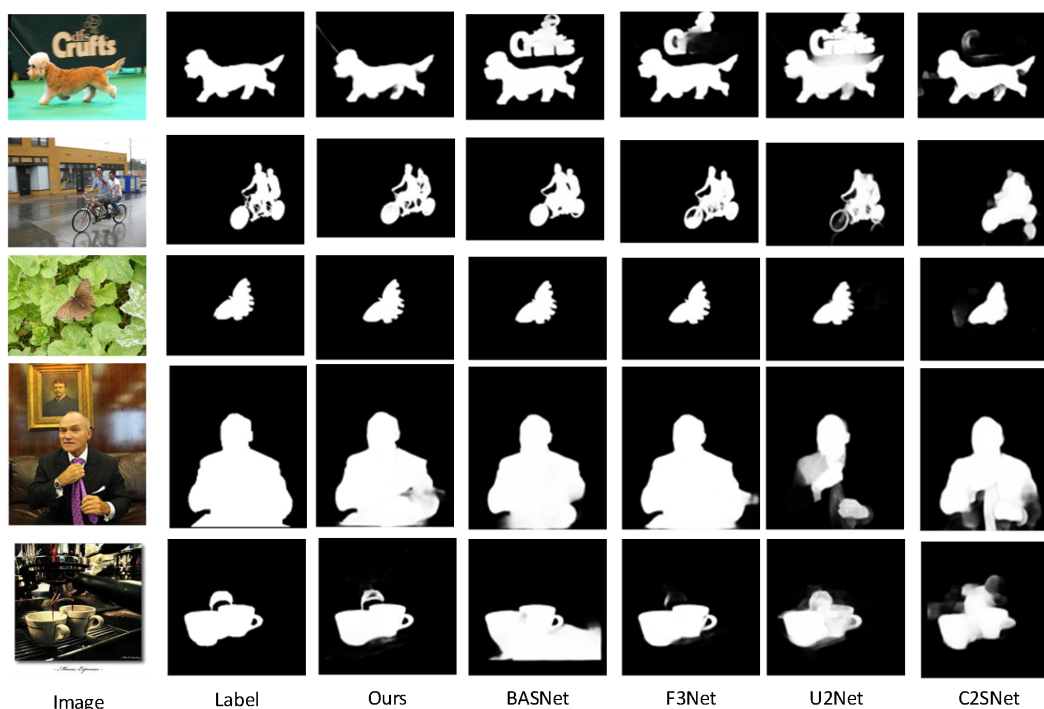


图 3-5 5 种算法的检测结果对比

3.4.5 消融实验

为了展示多尺度特征融合模块和边缘注意力融合模块的有效性,本小节进行了消融实验。所有实验均以 ResNet-50 网络为主干网络,在 DUTS-TR 数据集上进行训练,在 DUTS-TE 和 ECSSD 数据集上进行实验,实验结果如表 3-2 所示,可以看出两个模块的加入使 F-measure 指标得到了很大的提高,MAE 指标有所降低,说明预测图和真实标注图的差异减小,证明了本文设计的模块对目标检测精度具有显著效果。

表 3-2 消融实验

模型结构	ECSSD 1000 images		DUTS-TE 5019 images	
	MAE	mF	MAE	mF
Baseline	0.087	0.812	0.094	0.773
AFF	0.041	0.831	0.070	0.795
EAM	0.030	0.855	0.059	0.807
AFF+EAM	0.038	0.917	0.041	0.831

在计算机视觉测试中,热力图通过展示模型对输入数据的注意力分布,可以解释模型在产生特定预测时的决策过程。这有助于我们理解模型的行为,并提升模型的可解释性了可以直观地看出我们设计的模块对网络的性能的提升,所以我们绘制了网络输出的热力图,观察添加不同模块的网络对图像不同区域的关注度,以验证我们设计的模块的有效性。

图 3-6 展示了加入各模块后的热力图，可以看出仅通过基准网络检测的图像会出现目标定位不准的情况，在加入多尺度特征融合模块网络更加完整地关注目标整体，如图在加入边缘注意力融合模块后网络更加关注目标边缘信息，在同时加入两个模块后，网络效果达到最优。

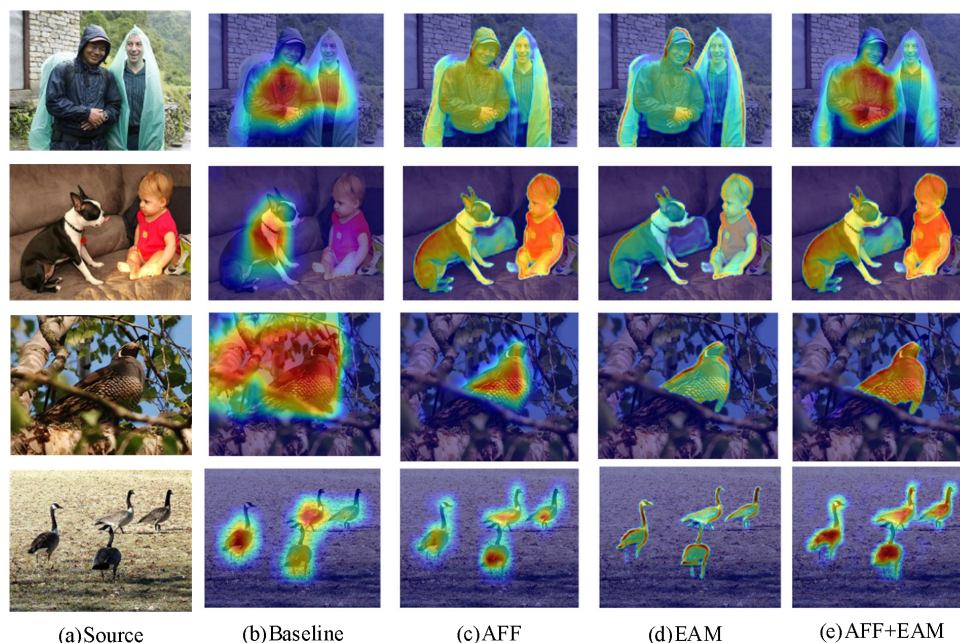


图 3-6 消融实验热力图

3.5 本章小结

本章主要介绍了一个基于深度学习的显著性目标检测网络模型用于分割图像中的目标和背景，为下文对图像的分析做准备。我们基于 U 型结构，设计了多尺度特征融合模块和边缘注意力融合模块，提升网络对不同大小目标的检测能力和对目标边缘的关注度。我们在 5 个测试集上对算法进行了验证，在 MAE、F-measure、S-measure 等指标上均有良好的表现，我们还进行了消融实验并绘制了热力图也证明了我们所设计模块的有效性。

第4章 飞行目标速度及姿态测量算法研究

4.1 本章引言

本课题的主要目的是利用计算机视觉测量高速飞行目标的运动学参数,在第二章中设计了振镜加高速相机组成的测量系统,获得大视场高分辨率的高速飞行目标图像,由于使用了振镜加单相机的组合,传统的立体视觉测量方法已经不再适用,所以我们需要设计新的算法。本章针对近一维高速飞行目标的运动特点和相机所拍摄图像的特点,设计了一种分别利用背景和目标上的特征点重构单应矩阵的方法,实现计算目标特征点的虚拟视差图,对飞行目标上特征点进行三维重构,再根据标定数据计算出目标在世界坐标系中的真实三维坐标,从而计算出目标的位移量,结合相机帧率可计算目标的运动速度。再通过提取目标的中轴线,计算目标的姿态角。

4.2 视觉测量原理

4.2.1 相机成像坐标系

相机成像过程中涉及四个关键坐标系,分别是世界坐标系、相机坐标系、图像坐标系和像素坐标系。下面我们对这四个坐标系进行说明。

像素坐标系: 像素坐标系是基于图像自身定义的,图像实质上由众多像素点组成,通常以图像左上角的像素点为原点建立坐标系,如图 4-1 用 O_0-uv 表示像素坐标系。在像素坐标系中,每个像素点都拥有唯一的坐标 (u,v) , 表示其在图像中的精确位置。在灰度图像中,每个像素点除了位置坐标外,还有一个与其亮度或灰阶相对应的数值,以此来表达该像素点的灰度信息。

图像坐标系: 图像坐标系通常以成像平面与相机光轴的交点,即图像的光学中心为原点,并且该点同时也被视为像素坐标系的中心点。参照图 4-1 所示坐标系通常表示为 O_1-xy , 图像坐标系的原点在像素坐标系中的坐标为 (u_0, v_0) 。

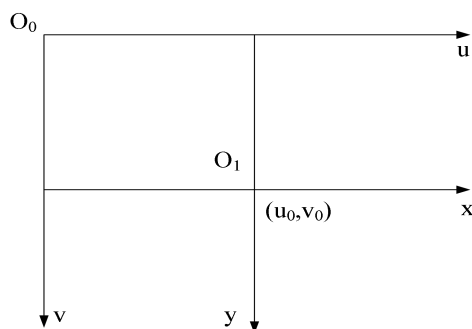


图 4-1 像素坐标系与图像坐标系

相机坐标系：相机坐标系通常以相机光心为原点，如图 4-2 所示相机坐标系通常表示为 $O_c - X_c Y_c Z_c$ ，相机的光轴被定义为 z 轴，并且垂直于图像坐标平面。相机坐标系的 X 轴和 Y 轴与图像坐标系中的相应轴平行且方向相同。相机坐标系原点到图像坐标系原点的距离 $O_1 O_c$ 为相机的焦距 f 。

世界坐标系：如图 4-2 所示世界坐标系通常表示 $O_w - X_w Y_w Z_w$ ，这是一个独立基准坐标系，用来统一标定和描述其他关联坐标系的位置和姿态。它不依赖于任何更大的外部坐标系来定义自己的位置。为了简化从世界坐标系到相机视角下的几何变换计算，通常将相机坐标系与世界坐标系对齐，即相机坐标系的原点和坐标轴方向可与世界坐标系一致，从而简化从全局视角到局部相机视点下的几何变换计算。

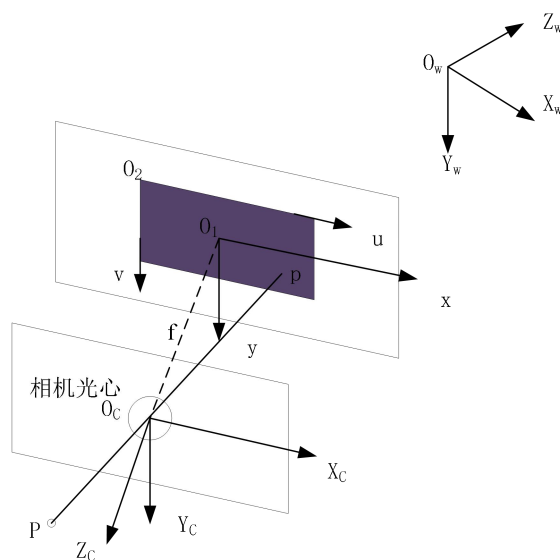


图 4-2 相机成像坐标系间的变换

4.2.2 立体视觉原理

上一小节中我们介绍了相机成像过程中的四个关键坐标系：世界坐标系、相

机坐标系、图像坐标系和像素坐标系。在拍摄过程中，相机将空间中的目标转化到二维图像上也就是图像坐标系上，而在视觉测量中我们想要通过相机拍摄的二维图像测量飞行目标的运动学参数，就需要将像素坐标系上的二维点重构为世界坐标系上的三维点。一般由目标在两张图像中的像素投影点重构出它的三维坐标，具体由双目系统完成即：两台相机对一个目标进行拍摄，得到同一空间点在不同相机投影平面上的像素坐标，再对像素点进行三角化，可得像素点在相机坐标系的三维坐标。

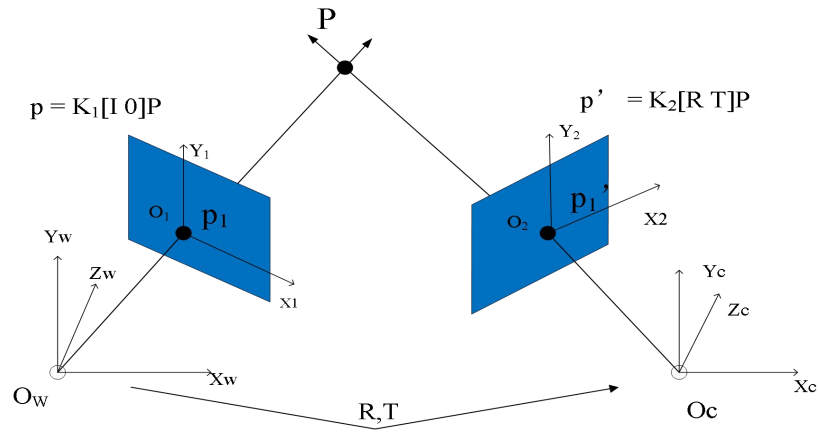


图 4-3 三角化求三维坐标

这里将像素坐标重构为世界坐标系中的三维坐标就涉及三角化问题，三角化的问题可以描述为已知一个三维点 P 在 m 张图像上投影的像素点 p_i ($i=1, 2, \dots, m$) 坐标以及 m 张图像对应的摄像机内参数矩阵 K ，外参矩阵 $[R \ T]$ ，求解三维点 P 的坐标。根据三角化原理如图 4-3，引入以第一个相机光心 O_1 作为世界坐标系的原点建立世界坐标系 $O_w X_w Y_w Z_w$ ，第二个相机的相机坐标系 $O_c X_c Y_c Z_c$ ，以及图像坐标系 $O_1 X_1 Y_1$ ， $O_2 X_2 Y_2$ ，得到图 4-4 中三维点 $P(X, Y, Z)$ 到像素点 $p_1(u, v, l)$ ， $p_1'(u', v', l)$ 的映射关系为式(4-1)所示：

$$\begin{cases} p_1 = K[I \ 0]P = MP \\ p_1' = K[R_B \ t_B]P = M'P \end{cases} \quad (4-1)$$

在式(4-1)中 $K^{[90]}$ 为相机的内参矩阵， M 与 M' 分别是虚拟相机 O_{c1} 和 O_{c2} 的投影矩阵，由相机的内参矩阵和外参矩阵共同构成。我们设置以 O_{c1} 为原点的相机坐标系与世界坐标重合，在相机模型 O_{c1} 中投影矩阵中的旋转矩阵为单位矩阵 I ，平移矩阵为零矩阵，相机模型 O_{c2} 中投影矩阵的旋转矩阵和平移矩阵分别是两个

虚拟相机之间的旋转矩阵和平移矩阵记为 $\mathbf{R}_B$, $\mathbf{t}_B$, 对于投影矩阵 $\mathbf{M}$ 与 $\mathbf{M}'$ 可以写作式(4-2):

$$\begin{cases} \mathbf{M} = \mathbf{K}[\mathbf{I} & \mathbf{0}] = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} \\ \mathbf{M}' = \mathbf{K}[\mathbf{R}_B & \mathbf{t}_B] = \begin{bmatrix} m_1' \\ m_2' \\ m_3' \end{bmatrix} \end{cases} \quad (4-2)$$

根据式(4-1)与式(4-2)可知三维点 P 在虚拟相机上的投影点 $p_1(u, v, 1)$, $p_1'(u', v', 1)$ 应满足式(4-3):

$$\begin{bmatrix} u_1 m_3 - m_1 \\ v_1 m_3 - m_2 \\ u_1' m_3' - m_1' \\ v_1' m_3' - m_2' \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = 0 \quad (4-3)$$

由(4-1), (4-2), (4-3)可知, 得到空间三维点在两张图像上的对应投影点和相机的内外参数矩阵即可求得像素点在世界坐标系下的三维坐标, 综上所述我们已知 P 在两个虚拟相机上的投影点 p_1, p_1' , 以及相机间的外参矩阵 $[\mathbf{R}_B \quad \mathbf{t}_B]$, 并标定出相机的内参矩阵 $\mathbf{K}^{[90]}$, 就可以求得图 4-3 中特征点 P 三维坐标。

4.3 目标运动学参数计算

4.3.1 特征点三维重构

由三角化的原理可知要对目标上的特征点进行三维重构, 需要目标三维点在两台相机上的投影点, 相机的内参 $\mathbf{K}$ 以及相机间的外参矩阵包括旋转矩阵 $\mathbf{R}$ 与平移矩阵 $\mathbf{T}$, 所以一般的视觉测量都是利用双目相机进行拍摄, 但本课题所使用的测量系统是基于单目相机进行拍摄。如图 4-4 所示, 本文的拍摄系统利用振镜改变机的视角产生一系列的虚拟相机对目标进行拍摄, 每帧图像是通过振镜形成的虚拟相机进行拍摄的, 实际上在同一时刻只有一个虚拟相机参与了拍摄, 缺少重构的参数, 所以针对本课题的相机和飞行目标的特点, 改进了算法, 通过将目标和背景特征点分开匹配计算出相机间的位姿关系以及目标特征点在两个虚拟相机上的投影点。如图 4-5 所示本文选取两帧图像进行分析, 分别选取由虚拟相机 O_{c1} 和 O_{c2} 拍摄的图像进行分析, 两幅图像的拍摄间隔为 n 帧。

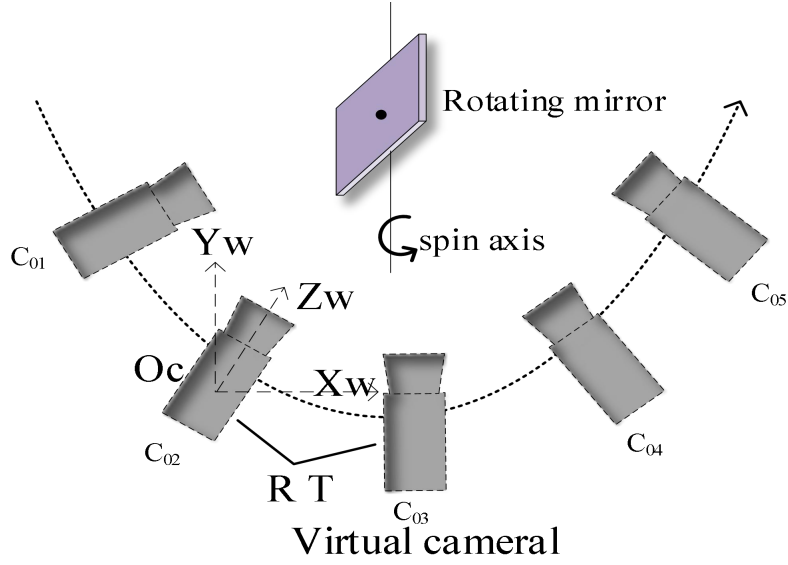


图 4-4 虚拟相机位姿变换

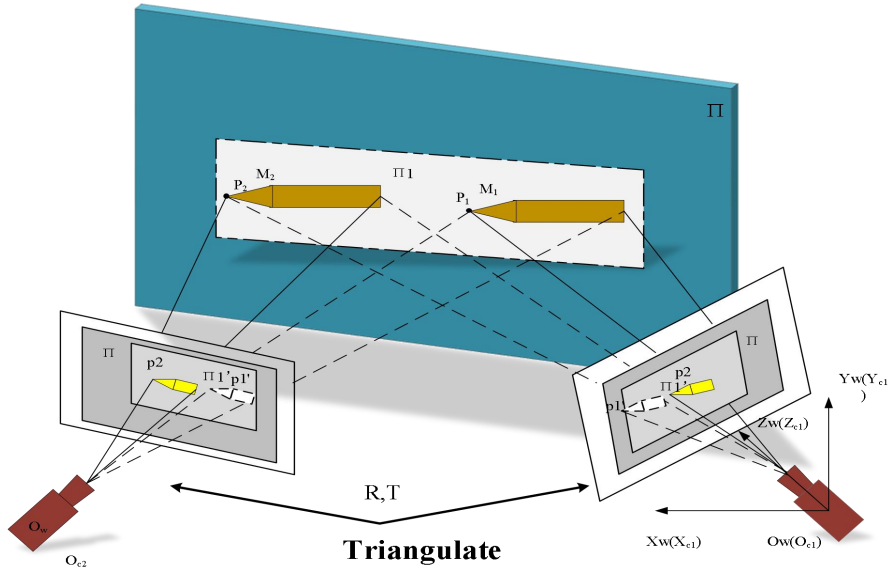


图 4-5 飞行目标投影模型

设系统的世界坐标系为 $O_w X_w Y_w Z_w$ 与虚拟相机坐标系 $O_{c1} X_{c1} Y_{c1} Z_{c1}$ 重合，如图 4-5 所示：高速飞行目标，从 M_1 运动到 M_2 处，目标上的特征点 P 由 P_1 运动至 P_2 ，在相机 C_{01} , C_{02} 上的投影分别为 p_1 , p_2 ，虚拟相机间的旋转平移矩阵分别为 R_0 , t_0 。由于目标是不断运动的，对于 P_1 , P_2 如式 4-4 所示：

$$P_1 \neq R_0 P_2 + t_0 \quad (4-4)$$

所以 p_1 , p_2 不是同一个三维点的在两台相机上的投影点，无法进行三角化求解 P 在空间中的三维坐标，所以必须求出 p_1 和 p_2 在 O_{c2} 和 O_{c1} 上对应的虚拟投影点

p_1' 和 p_2' 。

在拍摄飞行目标时，目标的表面深度远小于相机距目标的距离，也就是拍摄时相对场景深度远小于其与相机的距离时，且视场较小，可以将目标视为在一个平面上运动，令目标所在平面为 Π_1 。当物体上的三维点都位于同一个平面时，则平面三维点在两个视角下相机上的投影点 $p(u,v,1)$ ， $p(u',v',1)$ ，之间满足如下约束关系：

$$p = H p' \quad (4-5)$$

在式(4-5)中， H 称为单应矩阵该矩阵约束了同一平面上 3D 空间点在两个像素平面上投影的 2D 坐标，两个像素平面上的投影点可通过单应矩阵进行一一转换^[91]。已知空间中同一平面上的 3D 点在一张图像上的像素坐标和单应矩阵可求出该 3D 点在另一张图像唯一对应的像素坐标， H 可写为式(4-6)：

$$H = K(R - \frac{n^T t}{d})K^{-1} \quad (4-6)$$

在式(4-6)中， K 为相机的内参矩阵， R, t 为虚拟相机间的旋转平移矩阵， n^T 为目标所在平面的单位法向量， d 为平面到世界坐标系原点的距离。根据式(4-6)可知计算出目标平面在两台虚拟相机间投影对应的单应矩阵 H_T 即可求出虚拟对应点 p_1' 与 p_2' 。 p_1' 和 p_2' 计算流程如图 4-6 所示：

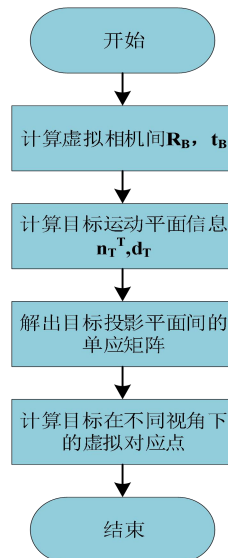


图 4-6 虚拟对应点像素坐标计算流程图

(1)计算虚拟相机的外部参数，即虚拟相机间旋转和平移矩阵。首先选取图像静态背景上的特征点进行特征匹配，由于 SIFT 特征点对于旋转和尺度均具有

不变性，并且对于噪声、视角变换和光照变化具有良好的鲁棒性，即使少数几个物体也可以产生大量的 SIFT 特征向量^[92]，综上本文选择 SIFT 特征点作为匹配点，若背景上的特征点不是来自一个平面，我们可以使用基础矩阵 $\mathbf{F}$ ，其在相机上的投影点 $\mathbf{x}$ ， $\mathbf{x}'$ 满足如下关系：

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0 \quad (4-7)$$

其中 $\mathbf{F}$ 的表达式为：

$$\mathbf{F} = \mathbf{K}^{-T} \mathbf{t}_x \mathbf{R} \mathbf{K}^{-1} \quad (4-8)$$

其中 $\mathbf{K}$ 是相机的内参矩阵， $\mathbf{t}$ 是两个虚拟相机 O_{c1} 和 O_{c2} 之间的平移矩阵， $\mathbf{R}$ 是虚拟相机 O_{c1} 和 O_{c2} 间的旋转矩阵。 $\mathbf{F}$ 是一个 3×3 秩为 2 的矩阵，一般记 $\mathbf{F}$ 为式(4-9)：

$$\mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \quad (4-9)$$

结合式(4-8)，式(4-9)，对于虚拟相机上的一对匹配点 $\mathbf{x} = [u \ v \ 1]$ ， $\mathbf{x}' = [u' \ v' \ 1]$ ，可得式(4-10)：

$$[u \ v \ 1] \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = 0 \quad (4-10)$$

整理可得式(4-11)：

$$uu'f_{11} + uv'f_{12} + uf_{13} + vu'f_{21} + vv'f_{22} + vf_{23} + u'f_{31} + v'f_{32} + f_{33} = 0 \quad (4-11)$$

由此得到每对匹配点可提供一个线性方程，利用八对点可以得到八个方程组成一个线性方程组，解方程组即可得到基础矩阵 $\mathbf{F}$ ，再对 $\mathbf{F}$ 进行奇异值分解得到虚拟相机间的旋转矩阵和平移矩阵 $\mathbf{R}_B$ ， $\mathbf{t}_B$ 。

由于单应矩阵 $\mathbf{H}$ 描述了两个平面之间的映射关系，若背景中的特征点来自同一个平面，则三维点在相机上的投影满足式(4-5)， $\mathbf{H}$ 可以记作式(4-12)：

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (4-12)$$

结合式(2-5)，式(4-12)，对于虚拟相机上的一对匹配点 $x = [u \ v \ 1]$ 和 $x' = [u' \ v' \ 1]$ ，可得式(4-13)：

$$\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (4-13)$$

式(4-13)展开后有三个等式，将第 3 个等式代入前两个式子可得式(4-14)：

$$\begin{aligned} u' &= \frac{h_1 u + h_2 v + h_3}{h_7 u + h_8 v + h_9} \\ v' &= \frac{h_4 u + h_5 v + h_6}{h_7 u + h_8 v + h_9} \end{aligned} \quad (4-14)$$

整理可得：

$$\begin{aligned} u' &= h_1 u + h_2 v + h_3 - h_7 u u' - h_8 v u' \\ v' &= h_4 u + h_5 v + h_6 - h_7 u v' - h_8 v v' \end{aligned} \quad (4-15)$$

由式(4-15)可知一对匹配点可以构造两个方程，4 组对应点就可以构建八个方程求解单应矩阵，再根据 Faugeras^[89]提出单应矩阵的分解方法进行分解得到虚拟相机间的旋转矩阵和平移矩阵 $\mathbf{R}$ ， $\mathbf{t}$ ，记 $\mathbf{R}$ ， $\mathbf{t}$ 为 $\mathbf{R}_B$ ， $\mathbf{t}_B$ 。

(2) 计算目标运动平面信息：由于目标上的点来自一个平面，由式(4-6)可知求目标平面在两台相机上投影平面间对应的单应矩阵，需要平面的单位法向量 $\mathbf{n}_T^T$ 以及平面到世界坐标系原点的距离 d_T 。在图像中选取目标区域进行特征点匹配，由式(4-15)计算出单应矩阵 $\mathbf{H}$ ，再分解单应矩阵得到目标所在平面的参数 $\mathbf{n}_T^T$ ， d_T 。

(3) 由式(4-16)组合出新的目标平面所对应单应矩阵 $\mathbf{H}_T$ ：

$$\mathbf{H}_T = \mathbf{K}(\mathbf{R}_B - \frac{\mathbf{n}_T^T \mathbf{t}_B}{d_T})\mathbf{K}^{-1} \quad (4-16)$$

由式(4-5)可求得特征点 P 在虚拟相机 O_{c1} 中的投影点 p_1 在 O_{c2} 上的虚拟对应点 p_1' 的像素坐标：

$$p_1' = \mathbf{H}_T p_1 \quad (4-17)$$

同理可求得目标上特征点 P 在相机 O_{c2} 上的投影点 p_2 在相机 O_{c1} 上的虚拟对应点

p_2' 的像素坐标:

$$p_2' = H_T p_2 \quad (4-18)$$

4.3.2 目标飞行速度计算

根据上述计算结果,我们得到了包括目标上三维点 P 不同时刻在两个虚拟相机上对应投影点的像素点分别为 p_1, p_1', p_2, p_2' , 虚拟相机间的外参矩阵以及相机的内参矩阵在内的参数, 根据 4.2.2 小节中的三角化原理我们可以重构出图像中目标特征点在不同时刻的三维坐标 $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2)$, 由距离公式 (4-19)求得两个点间的距离 d_1 , 但这些三维点坐标没有尺度因子, 尺度因子 s 需要通过目标上特征点间距离 d_1 和测量出的目标物上特征点间的实际距离 d_1' 计算得到, 如式(4-20)所示:。

$$d_1 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (4-19)$$

$$s = \frac{d_1}{d_1'} \quad (4-20)$$

目标的实际位移 d' 的计算采用 K 个特征点的平均位移表示如式(4-21)所示:

$$d' = \frac{s}{k} \sum_{i=0}^k d_i \quad (4-21)$$

我们选取的两幅图像间隔 n 帧率, 相机的帧率设为 2000FPS, 目标的飞行速度 v 可表示为式(4-22), 单位为 m/s:

$$v = \frac{2000d'}{n} \quad (4-21)$$

4.3.3 目标姿态角计算

对于对称目标的姿态测量一般采用中轴线法, 即提取图像中目标区域中轴线以获取目标姿态信息。在 4.2 节中提到我们利用单相机进行跟踪拍摄, 想要重构图像上的特征点, 需要假设目标运动轨迹在一个平面内, 所以目标的偏航角是利用上文中目标特征点的轨迹测出, 而俯仰角利用中轴线进行计算, 想要提高整个系统的姿态测量精度, 就要缩小轴线提取误差, 轴线的提取一般有基于目标边缘直线的中轴线提取法和基于区域信息的中轴线提取法两种。毋庸置疑, 区域比边界含有更多的信息量, 从区域信息出发提取目标主轴, 扩大了算法取样范围, 为

提高精度提供了可能，在目标边界检测困难或准确性受限的情形下，利用目标区域内在信息来提取中轴线不失为一种有效途径，相较于依赖边界信息的轴线提取技术，这种方式凭借对目标区域信息的运用，增强了结果的稳定性。此外，区域主轴法因其原理简洁、计算负担小、速度快的特点，尤为适用于对实时性要求较高的测量场合。在第三章我们训练了一个显著性目标检测网络，可以较为精确得到目标区域，所以我们在提取中轴线的方法上选择了基于区域主轴理论的方式进行提取，如图 4-7 所示。

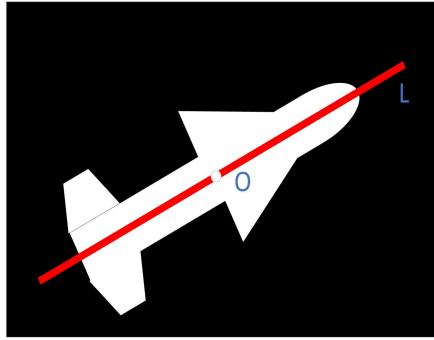


图 4-7 目标中轴线提取

基于区域主轴理论提取目标中轴线的方式就是将目标物体在二维图像平面上的投影区域视为一个由无数等质量点构成的虚拟刚体，通过计算这个虚拟刚体的最小转动惯量主轴，间接确定目标物体的实际中轴线在二维图像平面上的投影位置。具体算法实现步骤如下：首先对原始图像进行图像分割操作，目的是分离出目标物体所在的区域，排除背景和其他干扰因素的影响，接着通过对目标区域像素点坐标的分析，计算其二阶以内（即包括零阶矩、一阶矩和二阶矩）的原点矩 $m_{ij} (i, j < 2)$ 。其中，零阶矩代表质心位置，一阶矩反映质心与图像坐标系原点的偏移，二阶矩反映了目标区域在不同方向上的质量分布（即像素密度分布） $m_{ij} (i, j < 2)$ 如式(4-22)所示：

$$m_{ij} = \iint x^i y^j f(x, y) dx dy \quad (4-22)$$

考虑到目标物体的质量沿其轴线呈现对称分布特性，我们可以假设各像素点的权重均为一致，故此处取 $f(x, y)=1$ ，这对于主轴线的计算没有影响。因此，我们采用区域矩的方法来精确地确定区域质心及其主轴方向，进而推导出中轴线的表达式。具体计算过程如下。区域质心为：

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \quad (4-23)$$

区域转动惯量为：

$$I = \iint_D [(x - \bar{x}) \sin a - (y - \bar{y}) \cos a]^2 dx dy \quad (4-24)$$

I 取极小值时，过质心且倾角为 a 的直线即为区域主轴。经推导得此时 a 为：

$$\alpha = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \quad (4-35)$$

其中中心距 μ_{ij} 由原点矩计算可得：

$$\mu_{20} = m_{20} - \bar{x}m_{10}, \mu_{02} = m_{02} - \bar{y}m_{01}, \mu_{11} = m_{11} - \bar{x}\bar{y}m_{00} \quad (4-26)$$

这样我们得到了目标区域中轴线的质心 $(\bar{x}, \bar{y})$ 和倾角 α ，轴线的斜率为 k ， $k = \arctan(\alpha)$ 则轴线的直线方程为

$$y = (x - \bar{x})k + \bar{y} \quad (4-27)$$

在得到目标的质心的像素坐标 $O_1(x_1, y_1)$ 以及中轴线方程后，再找到中轴线上的另一个点的像素坐标设为 $O_2(x_2, y_2)$ ，利用 4.3 节的方法 求出由振镜产生的两个虚拟相机平面间的单应矩阵 H_T ，由式(4-15)可以计算出中轴线上的两个点 O_1 ， O_2 在另一相机视角下虚拟投影点为 $O_1'(x_1', y_1')$ 和 $O_2'(x_2', y_2')$ ：

$$O_1' = H_T O_1 \quad (4-28)$$

$$O_2' = H_T O_2 \quad (4-29)$$

再根据式(4-1)，(4-2)，(4-3)的三角化原理，由相机的外参矩阵，两组对应点，求解出中轴线上的点在世界坐标系中的三维坐标分别为 $O_a(X_1, Y_1, Z_1)$ 和 $O_b(X_2, Y_2, Z_2)$ ，目标中轴线的方向向量 n 为：

$$\bar{n} = (X_2 - X_1, Y_2 - Y_1, Z_2 - Z_1) \quad (4-30)$$

中轴线的方向向量在空间坐标系上各个平面的夹角为目标姿态角，由于我们获取的图像是单目相机拍摄的，在三维重构时假设目标在一个平面内运动，所以我们只对目标的俯仰角进行计算。目标的俯仰角为目标中轴线与平面 XOY 的夹角，

平面的法向量可以表示为 $d=(0, 0, 1)$ ，使用向量的夹角公式得到余弦值 $\cos\theta$ ：

$$\cos\theta = (d \bullet \vec{n}) / (|d| \times |\vec{n}|) \quad (4-31)$$

其中 $\bullet$ 表示向量点乘， $|d|$ 和 $|n|$ 表示向量的模，目标的俯仰角 $\theta = \arccos(\cos\theta)$ 。

4.3.4 坐标系转换

4.3.2 节中目标特征点的三维重构以及 4.3.3 节中姿态角所参考的世界坐标系都是与虚拟相机的相机坐标系重合的，计算速度时我们选取相邻两帧计算目标间的距离，通过匹配背景上的特征点得到两帧之间的外参矩阵，只需将两个特征点坐标转换到同一个坐标系中，不需要所有的特征点在同一个坐标系中。但对于姿态角来说，其参考坐标系的原点应为发射目标的发射口即将发射口坐标系定义为世界坐标，所以我们需要坐标系的转换，如图 4-8 所示：我们使用张氏标定法标定出两个坐标系间的旋转矩阵 $\mathbf{R}$ ，进行角度补偿。 $\mathbf{R}$ 可写为式(4-32)：

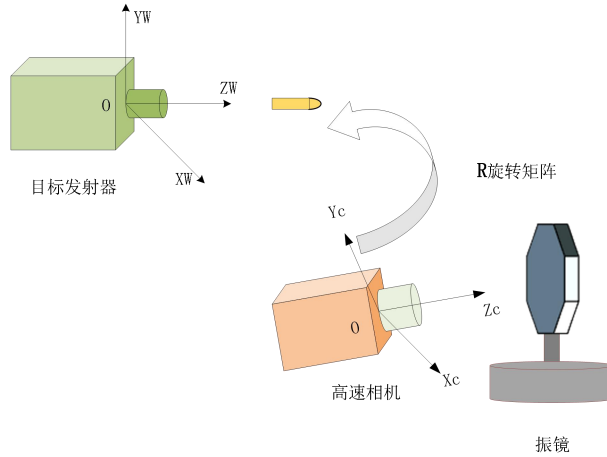


图 4-8 角度补偿示意图

$$\mathbf{R} = \begin{bmatrix} \cos(\alpha)\cos(\beta) & \sin(\gamma)\sin(\alpha)\cos(\beta) - \cos(\gamma)\sin(\beta) & \sin(\alpha)\cos(\gamma)\cos(\beta) + \sin(\gamma)\sin(\beta) \\ \cos(\alpha)\sin(\beta) & \sin(\gamma)\sin(\alpha)\cos(\beta) + \cos(\gamma)\sin(\beta) & \sin(\alpha)\sin(\beta)\cos(\gamma) - \sin(\gamma)\cos(\beta) \\ -\sin(\alpha) & \sin(\gamma)\cos(\alpha) & \cos(\gamma)\cos(\alpha) \end{bmatrix} \quad (4-32)$$

在式(4-32)中 α, β, γ 分别代表坐标系绕 x, y, z 轴旋转的角度，由于要计算俯仰角，所以需要对相机绕 z 轴旋转的角度进行补偿，在标定时使用振镜从不同角度拍摄棋盘格，由图 4-8 可知我们的振镜是绕 y 轴进行旋转，虚拟相机的与 z 轴的夹角是固定的，所以每幅图像的标定结果与 z 的夹角是相同的。 $\mathbf{R}$ 为 3×3 的矩阵可写成式(4-33)，绕 z 轴的角度可用式(4-34)计算。

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (4-33)$$

$$\theta_z = \arctan 2(-r_{31}, r_{11}) \quad (4-34)$$

所有计算出来的角度都需要经过角 θ_z 的补偿,这样的角度才具有实际的参考意义。

4.4 本章小结

本章我们主要讨论了相机坐标系的转换方法以及视觉测量算法,由于高速振镜加单相机的系统拍摄的图像丢失了深度信息,传统的双目视觉的视觉算法已不再适用,所以提出分别利用背景和目标上的特征点进行匹配,重构虚拟相机平面间的单应矩阵,计算目标特征点的虚拟视差图,实现对飞行目标上特征点的三维重构,再根据标定数据计算出目标在世界坐标系中的真实三维坐标,从而计算出目标的位移量,结合相机帧率可计算目标的运动速度。同时基于区域主轴原理,提取目标中轴线,对中轴线进行三维重构,完成目标的姿态测量。

第 5 章 实物实验及结果分析

5.1 本章引言

在前面的章节中,我们介绍了高速飞行目标测量系统的工作原理,系统结构,光路结构,又介绍了分离图像背景和目标的显著性检测算法,接着利用处理好的图像,设计测量算法,计算目标飞行的速度和姿态角。本章将进行实物实验验证,首先搭建实验平台即高速飞行目标测量系统,验证对目标的跟踪拍摄能力,接着利用静态实验平台对本文提出的测量算法进行验证,验证其有效性和精度,最后将二者结合实现对高速飞行目标的速度和姿态角进行测量。

5.2 实验平台搭建

我们按第二章设计的高速飞行目标测量系统的结构进行实验平台搭建,实验平台主要由高速相机,嵌入式控制系统,高速振镜,驱动模块以及外部光源组成。实验平台的实物图如图 5-1 所示,采用仰角式结构,相机与反射镜间的位置关系如图 5-2 所示。



图 5-1 实验装置实物图



图 5-2 相机与反射镜的位置关系图

该测量系统主要由高速相机（千眼狼 M220），振镜系统（天传奇 SC50），嵌入式控制系统（STM32F103），对射型光纤传感器组成，下面对使用到的部件进行介绍。

振镜是测量系统重要的组成部分之一，振镜的响应速度，定位精度，最大偏转角等参数可以保证对运动跟踪的准确性以及获得更大的视场范围，通过对比我们选用了天传奇公司的 SC50 激光扫描振镜，该款振镜的参数符合我们测量系统的实际需求，主要参数性能如下表 5-1 所示：

表 5-1 振镜主要参数

参数名称	参数规格
电源电压	$\pm 24\text{V}/1\text{A}$
位置电压比例系数	$0.5\text{V}/^\circ$
模拟位置电压输入范围	$\pm 10\text{V}$
1%阶跃响应时间	0.12ms
最大扫描角度	$\pm 20^\circ$
镜片反射率	$>99\% @ 45^\circ \text{入射}$
扫描速度	50Kpps
工作温度	$0\text{-}50^\circ\text{C}$

由上表可知该款振镜具有响应速度快仅 0.12ms ，扫描角度大，最大可以覆盖 40° 的范围，根据式(2-6)相机的视线可以获得近 80° 的凝视范围，满足我们测量系统的设计需求。除了跟踪的振镜以外我们还需要高速相机记录图像，由于振镜的使用，减少了目标与相机间的相对位移，所以对相机帧率降低了要求，综合考虑设计需求和相机功能后，我们选用了千眼狼 M220 高速相机，其满画幅分辨率为 1920×1080 ，在该分辨率下，相机帧率为 2000FPS ，通过万兆网口与电脑同

步，工作温度为 -20°C - 60°C ，满足我们的拍摄需求以及后续的测量需求。关于控制信号，使用电压信号控制振镜的角度变化，所以我们使用 DAC（电压输出的数字/模拟转换器）进行电压的输出，使用 STM32F103 作为系统的控制模块，该芯片有 512KBFLASH，4 个通用定时器，2 个基本定时器，3 个 12 位 ADC，主要用于接收速度测量系统的测速结果，计算振镜要到达的位置点，输出相应数字量并通过 DAC 模块转换为电压对振镜进行控制。测速模块以及触发装置采用博亿精科对射型光纤传感器 YIBO-NA11，主要参数表 5-2 所示：

表 5-2 光纤传感器参数表

参数名称	参数规格
输出方式	NPN 常开型
工作电压	DC12-24V
响应时间	$\leq 200\mu\text{s}$
检测区域	0.3-120mm
检测物体	非透明物体

在工业领域一般使用激光对射开关作为触发传感器，但是由于我们实验拍摄的飞行目标体积小，速度快，通过传感器的时间一般在 2ms 以内，而传统的激光对射开关的响应时间为 2.5ms 左右，响应速度较慢，而我们选择的对射型光纤传感器具有响应时间较短仅为 200 μs ，检测区域覆盖范围大等优点，可以检测到较小的模拟发射物，满足我们的实验需求。当目标经过第一个传感器时，控制板开始计时，在经过第二个测速装置后，计时结束，再根据两个传感器之间的距离，计算目标的飞行速度，当目标经过第三个光纤传感器时，触发振镜按照对应的控制曲线开始旋转，实现相机视线对目标的跟踪，保证目标始终在视场中央。

为了完成高速飞行目标的拍摄和测量实验，我们还在实验室搭建了一个模拟发射装置用于发射泡沫子弹，发射速度最大约为 40m/s。具体设计如图 5-3 所示：

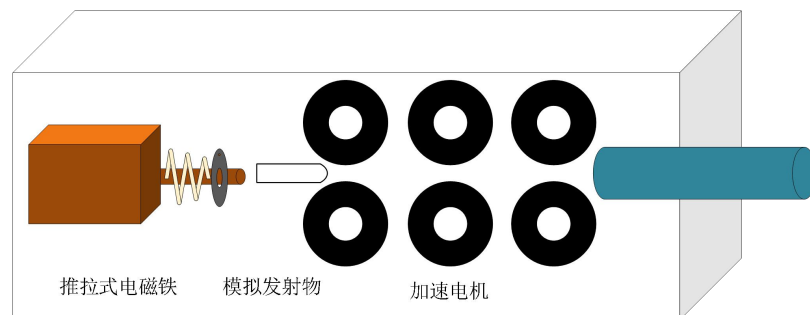


图 5-3 高速目标模拟发射装置示意图

该发射装置的基本原理为通过尾部的推拉式电磁铁推动发射物进入由电机带动转轮构成的加速部分进行加速,使目标获得较大初速度,并通过发射管发射出去,电机的速度是可以调节的,这样可以通过调节电机的转速,改变目标的发射速度,可以测试系统对不同速度目标的跟踪效果,图 5-4 为模拟发射装置的实物图:

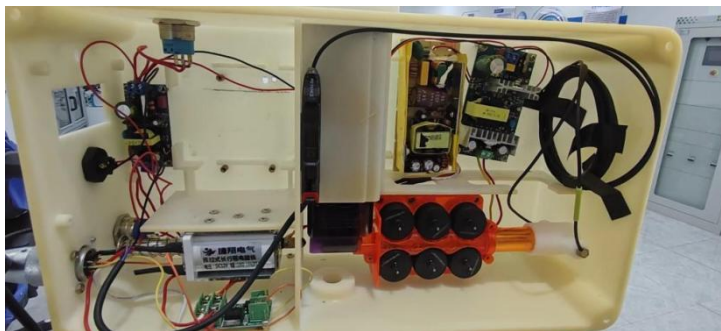


图 5-4 模拟发射装置实物图

5.3 高速飞行目标拍摄实验

在器件选型和实验平台搭建完成后,我们进行实际的拍摄实验,实验平台采用仰角式结构设计,振镜到相机的距离为 10mm,振镜到目标运动轨迹的距离为 1.5m,拍摄对象为上文设计的模拟发射装置发射模拟子弹,设置发射速度约为 20m/s,下图为拍摄的效果图。通过图 5-5 可以看到,我们设计的测量系统对高速飞行目标跟踪良好,目标始终保持在视场的中央,可以观测到目标姿态变化,并且目标的细节也非常清晰。

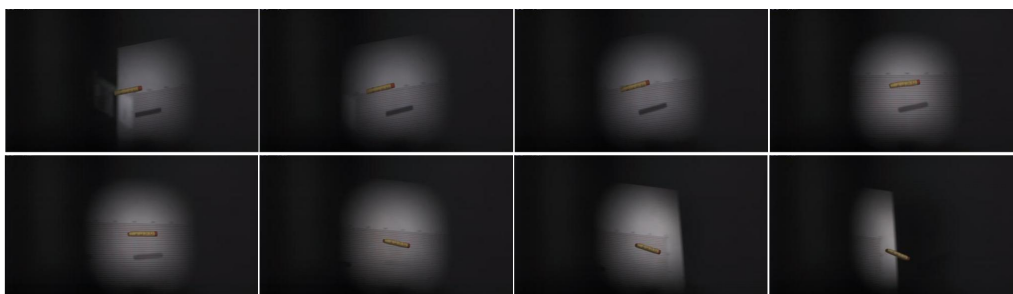


图 5-5 测量系统跟踪拍摄的图像

为了进一步展现使用振镜改变相机视角的拍摄方式在高速目标拍摄上的优势,我们拍摄了速度更快的飞行目标,速度约为 40m/s,并且与仅用单台高速相机拍摄的图像进行对比,从图 5-6(a)可以看出使用本文系统拍摄出来的目标图像依然清晰,可以观察到细节,而仅用高速相机拍摄的目标如图 5-6(b),已经非常模糊,并且观测范围也非常狭小,无法满足实际的观测和测量的需求。

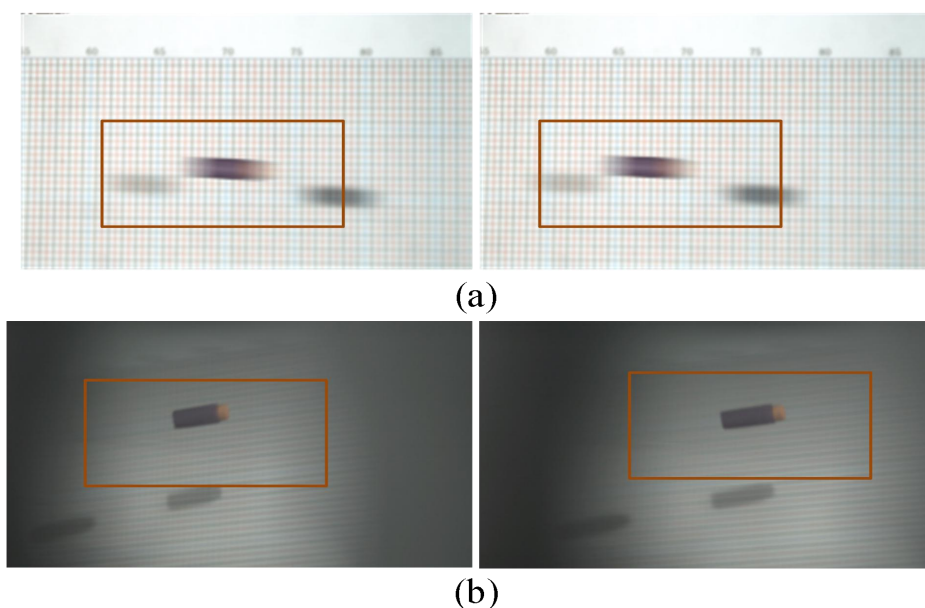


图 5-6 单台相机与测量系统拍摄对比图

5.4 静态目标测量实验

5.4.1 目标位移测量实验

这一节我们将验证第 4 章中提出的针对测量系统拍摄图像的测量算法有效性。主要包括对目标速度以及姿态角的测量，验证算法的有效性和精度。由于目标是在高速飞行中，很难得到实时的速度变化和姿态变化的准确值，没有参考值也就无法验证本文提出的测量算法的有效性，为此我们建立一个静态实验系统，使用高精度机械臂（型号：EFORT-MZ07），棋盘格标定板，数码相机，三脚架和模拟背景板组成。具体操作如下：将标定板安置在机械臂上，带有标记物的背景板放置在机械臂后面，数码相机架设在三脚架上。首先利用张正友标定法对相机进行标定得到内参矩阵 K ，对机械臂拍摄第一张图片，作为目标的起点，然后操作机械臂在平面内平移一段距离，模拟目标在平面内移动，相机变换视角再拍摄一张图片，模拟在凝视跟踪系统中，虚拟相机对目标的跟踪拍摄。实验场景如图 5-7 所示：



图 5-7 测量算法精度验证平台

机械臂的位置信息可以通过控制器精确地记录下来,本文以机械臂的坐标为基准,检验本文提出算法的有效性,实验中目标和背景均为手工提取。通过 4 组实验来验证在不同条件下算法的精度,实验包括: 1、物体单次位移量不同的测量实验(包括大位移和小位移量)。2、不同光照环境下物体位移量的测量实验。3、不同的背景环境下物体位移量的测量实验。

(1) 在室内正常光照条件下,使用 1920×1080 分辨率的相机进行拍摄,共计拍摄 10 张照片,目标位移 9 次,每次移动的距离为 5mm-10mm,使用本文提出的方法进行测量,相对误差为 2.82%,具体实验结果如表 5-3 所示:

表 5-3 自然光下目标小位移测量

Number	True displacement(mm)	Measure displacement(mm)	Absolute error(mm)	Relative error(%)
1	5.0	4.88	-0.117	2.34
2	5.8	5.64	-0.157	2.71
3	6.0	6.19	0.190	3.17
4	6.7	6.89	0.196	2.92
5	7.0	7.21	0.212	3.03
6	7.8	7.99	0.196	2.51
7	8.4	8.70	0.269	3.21
8	8.9	8.62	-0.278	3.13
9	9.5	9.27	-0.225	2.37

(2) 在正常室内照明条件下,使用 1920×1080 分辨率相机共拍摄了 10 张照片。目标移动了 9 次,每次移动的距离为 5-10 厘米。使用本文提出的测量方法,测量平均相对误差为 2.03%。具体实验结果如表 5-4 所示:

表 5-4 在自然光下目标大位移测量

Number	True displacement(cm)	Measure displacement(cm)	Absolute error(cm)	Relative error(%)
1	5.23	5.120	-0.110	2.11
2	5.78	5.675	-0.105	1.83
3	5.45	5.563	0.113	2.07

4	6.82	6.955	0.135	1.98
5	7.14	6.978	-0.162	2.27
6	7.48	7.625	0.145	1.95
7	8.27	8.440	0.170	2.06
8	8.71	8.522	-0.188	2.16
9	9.39	9.569	0.179	1.91

(3) 关闭外置光源, 模拟较暗的环境, 使用 1920×1080 分辨率相机共拍摄 10 张照片。目标位移 9 次, 目标位移 5cm-10cm。由于光线较暗, 对特征点的搜索和匹配有一定影响, 平均相对误差为 5.36%。具体实验结果如表 5-5 所示:

表 5-5 弱光条件下目标位移测量

Number	True displacement(cm)	Measure displacement(cm)	Absolute error(cm)	Relative error(%)
1	5.24	4.962	-0.278	5.32
2	5.47	5.787	0.317	5.58
3	6.13	5.810	-0.320	5.23
4	6.71	6.386	-0.324	4.83
5	7.34	7.702	0.362	4.93
6	7.93	8.384	0.454	5.73
7	8.37	7.916	-0.454	5.43
8	8.73	9.211	0.481	5.51
9	9.63	10.179	0.549	5.71

(4) 在正常室内照明条件下, 使用 1920×1080 分辨率相机共拍摄了 10 张照片。目标移动了 9 次, 每次移动的距离为 5 厘米-10 厘米。背景为实验室真实背景, 环境较为复杂。但测试结果依然良好, 平均相对误差为 2.33%。具体实验结果如表 5-6 所示:

表 5-6 复杂背景下目标位移测量

Number	True displacement(cm)	Measure displacement(cm)	Absolute error(cm)	Relative error(%)
1	5.31	5.434	0.124	2.34
2	5.45	5.332	-0.118	2.17
3	6.23	6.047	-0.156	2.51
4	6.76	6.613	0.147	2.17
5	7.12	6.945	-0.175	2.47
6	7.67	7.847	0.177	2.31
7	8.45	8.642	0.192	2.27
8	8.85	8.642	-0.208	2.35

9	9.72	9.955	0.235	2.42
---	------	-------	-------	------

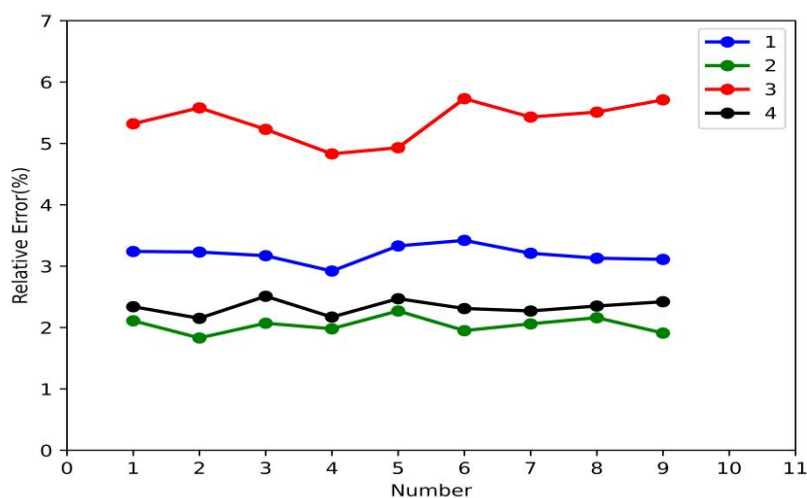


图 5-8 不同条件下相对误差统计图

由上述实验数据以及误差统计图 5-8 可知本方法在自然光条件下测量单次 5-10cm 位移量时精度较高，平均相对误差仅为 2.03%，由实验 2 和实验 4 可以看出，该方法在测量较小距离的位移时和在较为复杂的环境背景下测量时也有很高的精度，说明该方法可以用于多场景多尺度的测量工作中。根据表 3 所示，该系统在暗光环境下，由于特征点寻找较为困难，测量精度有所下降，但考虑到测量时多在正常自然光下进行，所以该算法可以满足一般情况下的测量工作。根据图 5-8 可知，在多次测量实验中，每次测量的相对误差变化幅度较小，可以证明该方法鲁棒性较高。

5.4.2 目标姿态角测量实验

除了对目标位移的测量精度进行验证，本文还对目标姿态角的测量精度进行验证，依然首先进行静态实验验证。将目标物放在姿态传感器上，一边改变目标姿态一边改变相机姿态进行拍摄，模拟使用振镜跟飞行目标的场景。实验图像如图 5-9 所示：

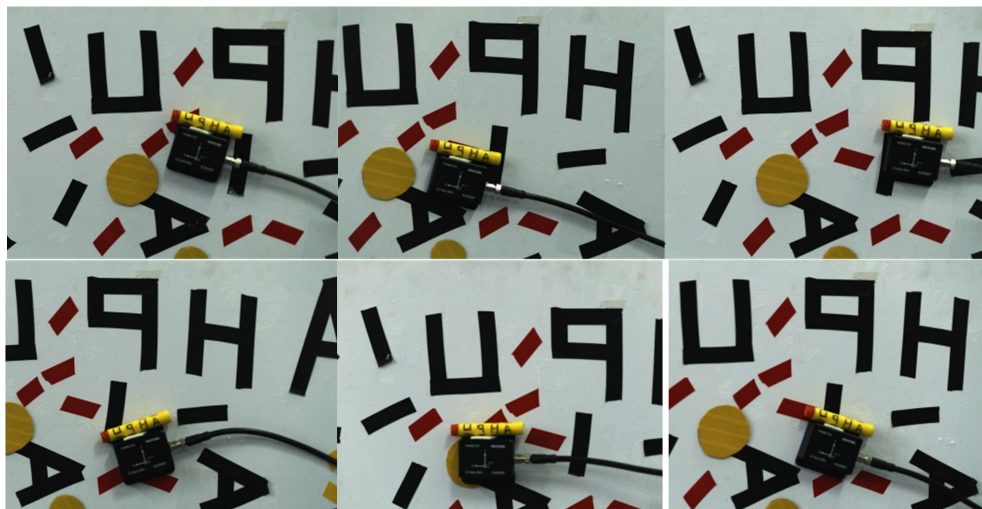


图 5-9 目标静态姿态测量拍摄图

因为拍摄方式为单目拍摄本文使用了重构单应矩阵的方式计算目标特征点的三维坐标，这就需要假设目标在一个平面上运动，所以本文只对目标的俯仰角进行测量比较。本文使用高精度的 IMU 姿态传感器作为基准如图 5-10 所示，将目标物安装在传感器上，传感器的数值就是目标姿态角的数值。



图 5-10 姿态传感器实物图

主要方法是利用分割后的图像，根据区域主轴理论提取目标的中轴线，然后由中心点坐标和中轴线倾角计算目标中轴线方程，这样可以找到在中轴线上的两个像素坐标，通过第四章的方法对两个坐标点进行三维重构，计算中轴线与 XOY 平面的夹角，中轴线的姿态角作为目标姿态角。

实验共选取 10 张图片进行测量，表 5-7 中展示了包括算法测量值，目标的实际角度以及测量的绝对误差与相对误差。与传感器的数值相比，实验测量值的平均相对误差仅为 2.01%，精度较高，并且相对误差的标准差仅为 0.1，测量结果一致性较好说明了算法的稳定性。

表 5-7 姿态角测量误差对比

Measured value(°)	Actual value(°)	Absolute error(°)	Relative error(%)
16.13	15.83	0.30	1.93
-7.65	-7.81	0.15	1.97
17.52	17.88	-0.36	2.03
8.94	8.77	0.17	1.89
-13.55	-13.84	0.29	2.12
13.50	13.79	-0.29	2.07
-17.22	-16.85	-0.37	2.18
-13.86	-14.11	0.25	1.83
21.06	20.65	0.41	1.97
-13.04	-12.78	0.26	2.05

5.5 飞行目标测量实验

完成静态算法验证后，接着对测量系统实际拍摄的目标图片进行测量。首先在目标物上做好标记点，并测量标记点间的真实距离，进行相机标定，完成准备工作后再进行拍摄。拍摄效果如图 5-11 所示，选取两组特征点匹配图进行展示分别为第 37(a)-38(b)帧，39(c)-40(d)帧，这是不同视角下虚拟相机拍摄的目标运动图像，图 5-11 中红色点和黄色点分别表示目标上的特征点在相机上的真实投影以及计算出的虚拟投影，接着对特征点进行三维建模，结合真实距离得到比例系数，计算目标上特征点间的距离。

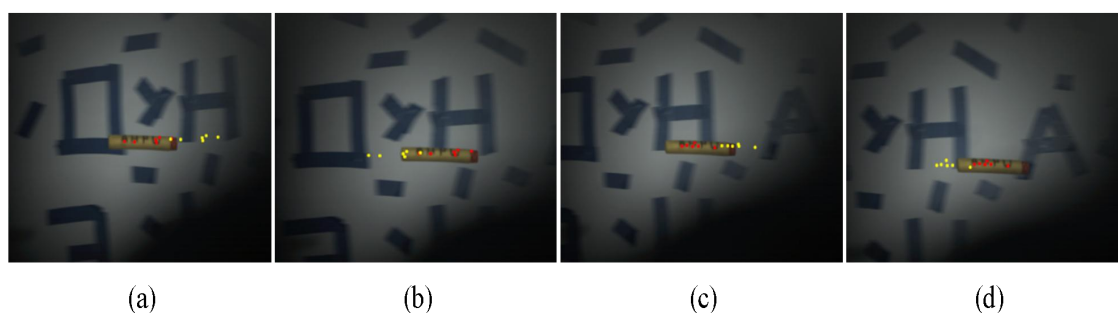


图 5-11 高速飞行目标特征点匹配图像

测速时选取相邻两帧图片为一组进行分析，选取 12 组相邻的两帧图像进行分析，并绘制跟踪范围内目标的速度变化曲线。其中目标从 10 帧出现，200 帧消失，选取的编号为 10-11，30-31，50-51，60-61，70-71，100-101，110-111，120-121，130-131，140-141，150-151，190-191。图 5-12 为 12 组图片测得的目标速度曲线图：

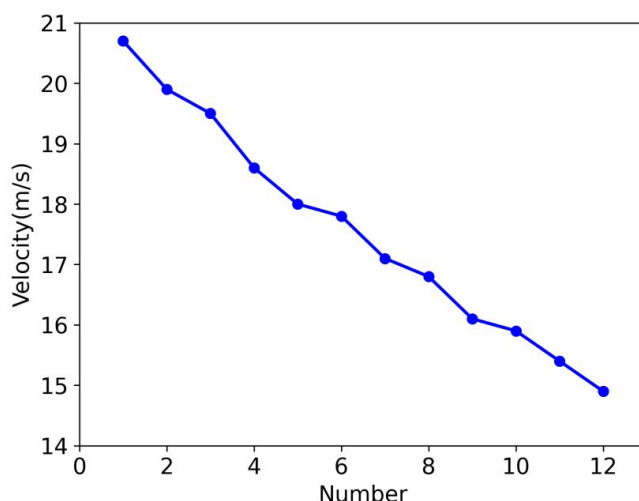


图 5-12 目标速度变化曲线

用本文提出的算法对测量系统拍摄的图片进行分析,可得到在跟踪范围内的目标速度变化曲线,从图中可以看出随着目标的运动,其速度不断降低,符合物理规律,从侧面反映出该方法的真实有效,可以完成高速飞行目标测速的任务。

在测量完速度后,又对目标的姿态角进行了测量。在本文 4.3.4 小节中提到要将坐标系进行转换,所以首先进行了标定,标定出相机坐标系对于发射口坐标的旋转矩阵进行角度补偿,标定图像如图 5-13 所示:

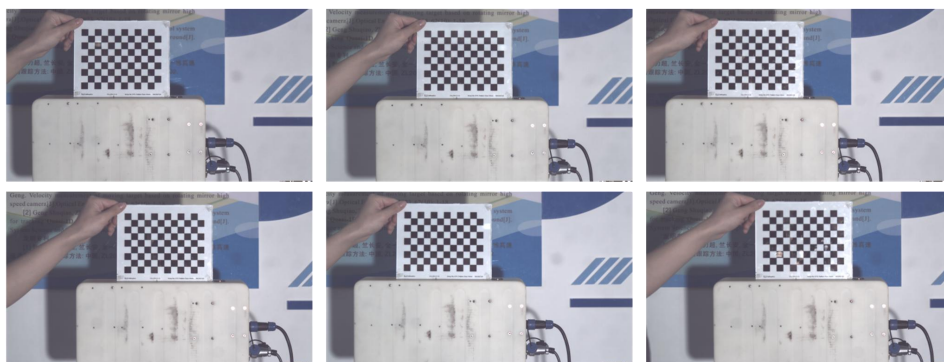


图 5-13 旋转矩阵标定

实验利用 Matlab 工具箱进行标定,计算出相机坐标系相对于发射口坐标系绕 z 轴旋转角度为 5.22° ,我们将对算法计算出的角度进行补偿。

如图 5-14 所示首先通过显著性目标检测算法,将目标区域提取出来,再根据区域主轴理论对目标中轴线进行提取,将像素平面上的中轴线重构为空间中的直线即可计算出目标姿态角。我们同样选取 10 张图片进行测量,具体结果如图 5-14 所示。

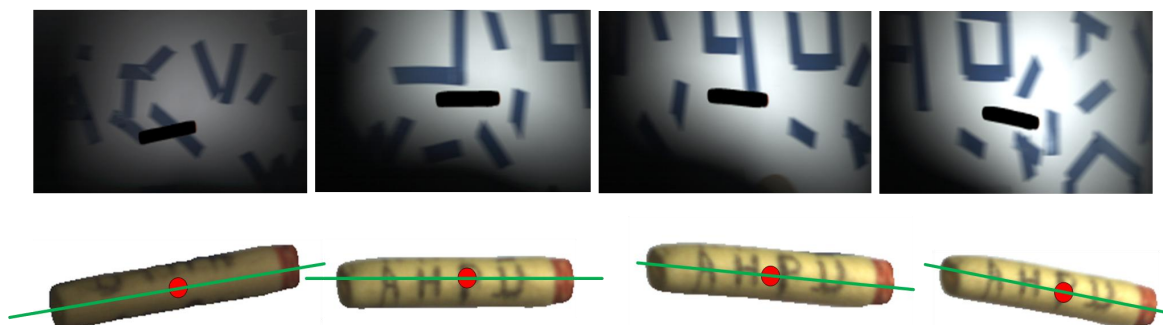


图 5-14 目标分割及中轴线提取

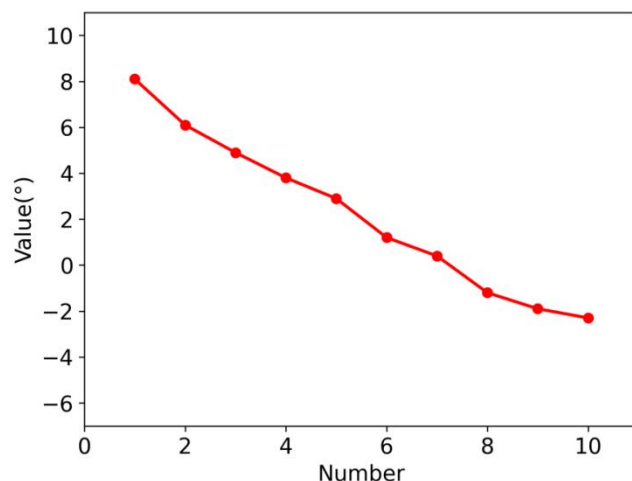


图 5-15 飞行目标姿态角变化曲线图

在曲线图 5-15 中可以看出目标发射出来后，仰角为正，目标飞行姿态为上仰，由于目标前端较重，会逐渐下降，仰角逐渐减小直至负角度，实验证明本文设计测量系统可以在一定范围内连续地测量目标俯仰角的变化。

5.6 本章小结

本章中我们完成了高速飞行目标测量系统的实验平台搭建，测试了拍摄效果，实验结果表明实验装置可对高速飞行目标进行凝视跟踪拍摄，保证飞行目标始终处于视场中央，接着又对测量算法进行验证，使用高精度机械臂和姿态传感器进行静态实验，将测量数据与算法的解算数据进行比较，通过对比我们测得在正常光照的条件下目标位移的平均相对误差为 2.03%，姿态角测量的平均相对误差为 2.01%，完成静态实验后，我们又对测量系统拍摄的图像进行分析，实验证明我们的算法依然可以正常工作，证明了本课题设计的可行性。

第 6 章 总结与展望

6.1 本文总结

本文的主要研究内容围绕高速飞行目标速度及姿态的视觉测量问题展开,传统视觉测量方式大多通过双目或多目相机系统进行拍摄,主要面临着拍摄视场和目标清晰度之间的矛盾,本文设计了一种振镜加高速相机组合的测量系统,利用振镜改变相机视角,在一定范围内保证目标始终在画面中央,以较低的成本兼顾了视场范围和目标清晰度,并针对拍摄图像的特点,设计运动学参数的测量算法,最后通过实验验证了拍摄系统的有效性和测量算法的有效性。

本文的主要研究工作如下:

(1) 为了实现对高速飞行目标的有效跟踪和测量,我们设计了振镜加相机的高速飞行目标测量系统,主要由高速相机,振镜,控制系统,光纤传感器等部件构成,通过振镜偏转改变相机的视角,确保在一定范围内对运动目标实现稳定的跟踪拍摄。我们推导了振镜的运动模型,确保相机视角能够实时跟随目标变化,并适用于拍摄不同速度飞行目标的需求。此外,还对系统的光路结构进行设计,并通过对比仰角式和平行式两种结构,确定仰角式结构在拍摄范围和减少遮挡等方面的优越性。

(2) 为了实现通过视觉方式对目标的运动学参数进行测量,我们需要对图像进行分割即将图像中的目标和背景进行分割。通过比较不同的分割方法,我们选择了基于深度学习的显著性目标检测算法。通过对传统 U 型网络架构进行改进,加入了边缘注意力模块和混合注意力模块,在 DUTS 数据集上进行训练,并与其他分割算法在 5 种数据集上进行比较,实验结果显示本文所提出的方法在分割精细度和准确性方面均有提升,消融实验证明本文设计的模块使基准网络在 DUTS-TEST 和 ECSSD 数据集上的 F-measure 指标分别上升了 0.058, 0.105, MAE 指标分别下降了 0.053, 0.049, 证明了本文设计的有效性。

(3) 在完成图像拍摄和预处理后,我们设计了视觉测量算法,由于我们使用振镜改变相机视角进行跟踪拍摄,实际上形成了类似单目相机系统的工作模式,无法直接利用已有的双目算法进行对目标的三维重构,所以我们提出了一种新方

法：在相邻帧的两幅图像中，分别对背景和目标进行特征点匹配，重构两幅图像目标所在平面对应的单应矩阵，找到空间点在两个像素平面上的投影，再利用三角化对目标特征点进行三维重构，根据标定的数据就可以得到目标特征点的空间坐标，结合相机的帧率实现对目标飞行速度的测量。对于目标姿态，我们采用基于区域主轴定律找到目标的中轴线，使用中轴线的方向代表目标的方向。

（4）实验验证环节，在实验室搭建模拟发射器发射高速飞行目标，首先测试了测量系统拍摄的效果，实验结果表明我们的系统工作正常，拍摄高速飞行目标的效果良好。接着我们验证测量算法的有效性，首先进行静态实验，利用高精度机械臂作为基准，与算法测量的结果进行比较，通过对比我们的算法的相对误差为 2.03%。我们又使用姿态传感器作为基准验证姿态测量的准确性，通过实验我们得出测量算法的相对误差为 2.01%，在完成静态实验后我们对测量系统拍摄的飞行目标运动学参数进行测量，实验证明我们的算法依然是可行有效的。

6.2 展望

本文主要解决的问题是大视场高分辨率条件下的高速飞行目标速度及姿态测量问题，在课题中设计了基于振镜的高速飞行目标测量系统以及相应的视觉测量算法，并验证了整个设计的可行性，但由于时间，实验条件以及作者的水平的限制，本研究依然有不完善的地方，未来还需要进行改进，主要改进方向如下：

（1）本文提出的算法对光照强度有较高的要求，而受限于振镜的大小导致拍摄的图像亮度不高，需要使用大功率光源设备进行补光，这就造成该系统使用不便，后期需要考虑采用尺寸更大的振镜增大镜头的进光量，并且优化测量算法使其在较低的亮度下也有较高的准确率。

（2）本文中振镜的控制算法比较简单，用离散的位置点进行控制，导致振镜在整个运动过程连续性欠佳，可能会使目标与振镜之间产生一定的相对位移，如果目标运动进一步加快这种不连贯的运动可能导致拍摄图像模糊，所以在后续工作中，有必要对振镜的控制算法进行深入优化，以确保振镜运动的连贯和平顺性。

（3）本文采用单振镜进行跟踪拍摄，该系统仅适用于对运动轨迹在平面内的目标进行测量，使用起来局限性比较大，未来我们计划使用双振镜系统进行拍摄，并重新设计测量算法，使其能测量三维空间中高速飞行目标相关运动学参数。

参考文献

- [1] 李晶. 碰撞试验中加速度传感器的应用和常见问题分析[J]. 科技创新与应用, 2020(29): 181-182.
- [2] 陈胜政, 杨波, 宋宇航, 等. 制导炮弹姿态测量技术综述[J]. 探测与控制学报, 2021, 43(04): 9-13.
- [3] 姬曼, 邵伟平, 郝永平. 基于一种利用惯性转子陀螺的弹丸姿态控制的研究[J]. 光电技术应用, 2019, 34(5): 67-71.
- [4] 周晖, 熊天武, 黄宗卫. 基于GPS信号的目标位置姿态信息解算方法[J]. 舰船电子对抗, 2020, 43(1): 105-107.
- [5] Nazarahari M, Rouhani H. 40 years of sensor fusion for orientation tracking via magnetic and inertial measurement units: Methods, lessons learned, and future challenges[J]. Information Fusion, 2021, 68: 67-84.
- [6] 陈胜政, 杨波, 宋宇航, 等. 制导炮弹姿态测量技术综述[J]. 探测与控制学报, 2021, 43(04): 9-13.
- [7] Hu Haitao, Cao Yiping, An Haihua, etc. High-precision 3D shape measurement of rigid moving objects based on the Hilbert transform.[J]. Applied Optics, 2021, 60(27): 8390-8399.
- [8] Daniel E. Kakaley, Russell E. Altieri, Gregory D. Buckner. Non-contacting measurement of torque and axial translation in high-speed rotating shafts[J]. Mechanical Systems and Signal Processing, 2020, 138(4)
- [9] Jiang Bixiao, Yang Xianqing, Ma Ping, etc. Contactless Measurement of Hypervelocity Projectile Wake Velocity Distribution.[J]. IEEE Access, 2019, 7.
- [10] 周富强, 王晔昕, 柴兴华等. 镜像双目视觉精密测量技术综述[J]. 光学学报, 2018, 38(08): 59-67.
- [11] 侯艳丽, 苏显渝, 陈文静. 旋转视觉测量系统相机光心与转轴距离的标定方法[J]. 光学学报, 2022, 42(21): 69-79.
- [12] 谭强俊, 程永生, 唐彬, 等. 俯仰/滚转耦合的大角度姿态测量算法[J]. 太赫兹科学与电子信息学报, 2022, 20(09): 946-952.
- [13] Gao G S, Partridge R. High speed digital TDC for D0 vertex reconstruction[J].

- IEEE Transactions On Nuclear Science, 1991, 38(2): 286-289.
- [14]高胡震,杜昌平,郑耀.基于扩张状态观测器的无人机云台系统控制算法[J].计算机应用,2024,44(02): 604-610.
- [15]孙训红,都海波,陈维乐,等.基于移动机器人机载视觉云台的有限时间目标跟踪控制[J].控制与决策,2023,38(10): 2875-2880.
- [16]王德田,彭其先,刘俊等.激光干涉测速技术在内弹道弹丸速度测量中的应用研究[J].高压物理学报,2011,25(02): 133-137.
- [17]冯智勇,曾瀚,张力等.基于陀螺仪及加速度计信号融合的姿态角度测量[J].西南师范大学学报(自然科学版), 2011,36(04): 137-141.
- [18]马秀丽. 基于磁阻/陀螺组合测量弹体飞行姿态的数据融合方法研究[D]. 南京: 南京理工大学, 2012.
- [19]傅忠云, 朱海霞, 孙金秋等. 基于惯性传感器MPU6050 的滤波算法研究[J]. 压电与声光, 2015, 37(05): 821-825.
- [20]牛春峰. 弹道环境下的陀螺/GPS组合姿态测量方法研究[D]. 南京: 南京理工大学, 2015.
- [21]邵艳巡,邓林红,戎海龙.基于惯性—地磁组合的多模型融合姿态测量算法[J].传感器与微系统,2023,42(01):109-112.
- [22]Wu Yanxue, Wu Gaoxu, Li Lingchi, etc. Inner shifting-phase method for high-speed high-resolution 3-d measurement[J]. IEEE Transactions On Instrumentation And Measurement, 2020, 69(9): 7233-7239.
- [23]周全,王飞行,文贡坚.基于狭缝摄影的靶场目标运动测量综述[J].现代电子技术, 2010, 33(19): 54-57.
- [24]李金珂, 陈良益. 基于线阵CCD的弹道同步式狭缝摄影系统[J]. 激光与红外,2009, 39(03): 300-303.
- [25]李冬冬. 基于交比不变量的狭缝摄影测量相机标定和晃动问题研究[D]. 湖南: 国防科学技术大学, 2017.
- [26]王可东, 孙道虎, 晏磊, 等. 低成本CMOS阵列式高速成像系统 [J] . 半导体光学. 2005, 26(2): 1074-1080.
- [27]Sun Ce, Jia Yangyu, Zhang Dongjiang. High-speed laser retroreflective shadowgraph system for large-scale transient studies on a field-test range[J].

- Optical Engineering, 2020, 59(9): 80-93.
- [28]谢爱民, 黄洁, 宋强等. 多序列激光阴影成像技术研究及应用[J]. 实验流体力学, 2014, 28(04): 84-88.
- [29]张文博. 弹丸运动参数激光阴影照相测试技术研究[D]. 西安: 西安工业大学, 2020..
- [30]乔志旺. 基于高速成像技术的外场弹丸运动姿态测量方法研究[D]. 南京: 南京理工大学, 2020.
- [31]李璐. 基于正交阴影图像处理的弹丸特征点识别与参数辨识[D]. 南京: 南京理工大学, 2021.
- [32]史元元, 冯斌, 张文博, 胥磊, 杨曼曼. 基于激光阴影照相系统的研究[J]. 应用光学, 2020, 41(02): 282-287.
- [33]Zaarane A, Slimani I, Al Okaishi W, etc. Distance measurement system for autonomous vehicles using stereo camera[J]. Array, 2020, 5: 102-124.
- [34]符宾, 吴坦. 光电经纬仪姿态高精度测量研究[J]. 测绘与空间地理信息, 2018, 41(10): 236-238.
- [35]Jiang Mingjun, Ryo Sogabe, Shimasaki Kohei, etc. 500-fps omnidirectional visual tracking using three-axis active vision system[J]. IEEE Transactions on Instrumentation and Measurement, Shanghai: 2021, 70: 1-11.
- [36]W.HB. Flight Follower追踪系统[J]. 军民两用技术与产品, 2007(8):28-30.
- [37]Okumura K, Oku H, Ishikawa M. High-speed gaze controller for millisecond-order pan/tilt camera[C]//2011 IEEE International Conference On Robotics And Automation, 2011: 6186-6191.
- [38]张文娇. 弹道跟踪镜技术研究[D]. 西安: 西安工业大学, 2010.
- [39]张嫚. 弹道跟踪转镜伺服控制与驱动技术研究[D]. 西安: 西安工业大学, 2011.
- [40]Ullah I, Jian M, Hussain S, etc. A brief survey of visual saliency detection[J]. Multimedia Tools And Applications, 2020, 79: 34605-34645.
- [41]Aksac A, Ozyer T, Alhajj R. Complex networks driven salient region detection based on superpixel segmentation[J]. Pattern recognition, 2017, 66: 268-279.
- [42]Cheng M M, Mitra N J, Huang X, etc. Global contrast based salient region detection[J]. IEEE Transactions On Pattern Analysis And Machine Intelligence, 2014, 37(3): 569-582.

- [43]Yang C, Zhang L, Lu H, etc. Saliency detection via graph-based manifold ranking[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, 2013: 3166-3173.
- [44]Itti, L.; Koch, C.; Niebur, E. A model of saliencybased visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [45]Achanta R, Estrada F, Wils P, etc. Salient region detection and segmentation[C]//International Conference On Computer Vision Systems, 2008: 66-75.
- [46]Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing[C]//Proceedings Of The Eleventh Acm International Conference On Multimedia, 2003: 374-381.
- [47]Yuan Y, Li C, Kim J, etc. Reversion correction and regularized random walk ranking for saliency detection[J]. IEEE Transactions On Image Processing, 2017, 27(3): 1311-1322.
- [48]Shan D, Zhang X, Zhang C. Visual saliency based on extended manifold ranking and third-order optimization refinement[J]. Pattern recognition letters,2018, 116: 1-7.
- [49]Wu X, Ma X, Zhang J, etc. Salient object detection via deformed smoothness constraint[C]//2018 25th IEEE International Conference On Image Processing (ICIP), 2018: 2815-2819.
- [50]He K, Zhang X, Ren S, etc. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2016: 770-778.
- [51]Pang Y, Zhao X, Zhang L, etc. Multi-scale interactive network for salient object detection[C]//Proceedings of the IEEE/CVF Conference On Computer Vision And Pattern Recognition, 2020: 9413-9422.
- [52]Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2015: 3431-3440.
- [53]Li G, Yu Y. Visual saliency based on multiscale deep features[C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2015: 5455-5463.
- [54]Borji A, Cheng M M, Hou Q, etc. Salient object detection: A survey[J].

- Computational visual media, 2019, 5: 117-150.
- [55]Lee G, Tai Y W, Kim J. ELD-net: An efficient deep learning architecture for accurate saliency detection[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 40(7): 1599-1610.
- [56]Zhang P, Wang D, Lu H, etc. Learning uncertain convolutional features for accurate saliency detection[C]//Proceedings of the IEEE International Conference on computer vision, 2017: 212-221.
- [57]Wang X, Girshick R, Gupta A, etc. Non-local neural networks[C]//Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition, 2018: 7794-7803.
- [58]Li J, Su J, Xia C, etc. Salient object detection with purificatory mechanism and structural similarity loss[J]. IEEE Transactions on Image Processing, 2021, 30: 6855-6868.
- [59]Cong R, Zhang Y, Fang L, etc. RRNet: Relational reasoning network with parallel multiscale attention for salient object detection in optical remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-11.
- [60]Tang L, Li B. CLASS: cross-level attention and supervision for salient objects detection[C]//Proceedings of the Asian Conference on Computer Vision, 2020.
- [61]Shan F Q, Wang F, Zhang H F. Research on 3D pose measurement algorithm based on binocular vision[C]//Journal of Physics: Conference Series. IOP Publishing, 2021, 1792(1): 012-043.
- [62]Hamzah R A, Ibrahim H. Literature survey on stereo vision disparity map algorithms[J]. Journal of Sensors, 2016.
- [63]Zhang G, Yang S, Hu P, etc. Advances and prospects of vision-based 3D shape measurement methods[J]. Machines, 2022, 10(2): 124.
- [64]L.Bo, G. Chen, Measurement Method Based on Stereo Vision for Nozzle Motion[J], 2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control, 2012, 1413-1416.
- [65]Nomura Y, Yamamoto S, Hashimoto T. Study of 3D measurement of ships using dense stereo vision: towards application in automatic berthing systems[J]. Journal of Marine Science and Technology, 2021, 26: 573-581..
- [66]Yang L, Li M, Song X, etc . Vehicle speed measurement based on binocular stereovision system[J]. IEEE Access, 2019, 7: 106628-106641.

- [67]Peng J, Xu W, Liang B, etc. Pose measurement and motion estimation of space non-cooperative targets based on laser radar and stereo-vision fusion[J]. IEEE Sensors Journal, 2018, 19(8): 3008-3019.
- [68]Chen L, Zhong G, Han Z, etc. Binocular visual dimension measurement method for rectangular workpiece with a precise stereoscopic matching algorithm[J]. Measurement Science and Technology, 2022, 34(3): 035010.
- [69]Lin H Y, Tsai C L. Depth measurement based on stereo vision with integrated camera rotation[J]. IEEE Transactions On Instrumentation And Measurement, 2021, 70: 1-10.
- [70]Zhong J, Zhong S, Zhang Q, etc. Vision-based measurement system for instantaneous rotational speed monitoring using linearly varying-density fringe pattern[J]. IEEE Transactions On Instrumentation And Measurement, 2018, 67(6): 1434-1445.
- [71]Li Z, Pei Y, Qu C, etc. A position and attitude measurement method based on laser displacement sensor and infrared vision camera[J]. IEEE Transactions On Instrumentation And Measurement, 2022, 71: 1-9.
- [72]N. Murmu., Relative velocity measurement using low cost single camera-based stereo vision system[J], Measurement 2019, 141: 1–11.
- [73]Liu X, Li A. Multiview three-dimensional imaging using a Risley-prism-based spatially adaptive virtual camera field[J]. Applied Optics, 2022, 61(13): 3619-3629.
- [74]Han Z, Zhang L. A virtual multi-ocular 3D reconstruction system using a galvanometer scanner and a camera[J]. Sensors, 2023, 23(7): 3499.
- [75]Yan Q, Xu L, Shi J, etc. Hierarchical saliency detection[C]//Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition, 2013: 1155-1162.
- [76]Li Y, Hou X, Koch C, etc. The secrets of salient object segmentation[C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2014: 280-287.
- [77]Wang T, Borji A, Zhang L, etc. A stagewise refinement model for detecting salient objects in images[C]//Proceedings of the IEEE International Conference On Computer Vision, 2017: 4019-4028.
- [78]Yang C, Zhang L, Lu H, etc. Saliency detection via graph-based manifold ranking[C]//Proceedings of the IEEE Conference On Computer Vision And

- Pattern Recognition, 2013: 3166-3173.
- [79]LI G, YU Y. Visual saliency detection based on multiscale deep CNN features[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5012-5024.
- [80]Li X, Yang F, Cheng H, etc. Contour knowledge transfer for salient object detection[C]//Proceedings Of The European Conference On Computer Vision (ECCV), 2018: 355-370.
- [81]Chen S, Tan X, Wang B, etc. Reverse attention for salient object detection[C]//Proceedings Of The European Conference On Computer Vision (ECCV), 2018: 234-250.
- [82]Deng Z, Hu X, Zhu L, etc. R3net: Recurrent residual refinement network for saliency detection[C]//Proceedings of the 27th international joint conference on artificial intelligence. Menlo Park, CA, USA: AAAI Press, 2018, 684-690.
- [83]Zhang X, Wang T, Qi J, etc. Progressive attention guided recurrent network for salient object detection[C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2018: 714-722.
- [84]Feng M, Lu H, Ding E. Attentive feedback network for boundary-aware salient object detection[C]//Proceedings of the IEEE/CVF Conference On Computer Vision And Pattern Recognition. 2019: 1623-1632.
- [85]Wu Z, Su L, Huang Q. Cascaded partial decoder for fast and accurate salient object detection[C]//Proceedings of the IEEE/CVF Conference On Computer Vision And Pattern Recognition, 2019: 3907-3916.
- [86]Liu N, Han J, Yang M H. Picanet: Learning pixel-wise contextual attention for saliency detection[C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2018: 3089-3098.
- [87]Ke Y Y, Tsubono T. Recursive contour-saliency blending network for accurate salient object detection[C]//Proceedings Of The Ieee/Cvf Winter Conference On Applications Of Computer Vision. 2022: 2940-2950.
- [88]Gao S, Zhang W, Wang Y, et al. Weakly-supervised salient object detection using point supervision[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022, 36(1): 670-678.
- [89]Liu J J, Hou Q, Cheng M M, etc. A simple pooling-based design for real-time salient object detection[C]//Proceedings of the IEEE/CVF Conference On Computer Vision And Pattern Recognition, 2019: 3917-3926.
- [90]Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions

On Pattern Analysis And Machine Intelligence, 2000, 22(11): 1330-1334.

[91]Faugeras O D, Lustman F. Motion and structure from motion in a piecewise planar environment[J]. International Journal Of Pattern Recognition And Artificial Intelligence, 1988, 2(03): 485-508.

[92]Li J, Wang H, Zhang L, etc. The research of random sample consensus matching algorithm in PCA-SIFT stereo matching method[C]//2019 Chinese Control And Decision Conference (CCDC), IEEE, 2019: 3338-3341.

攻读学位期间发表的学术论文目录

- [1] Wang Lichao, Dong Jiahui, Cheng Qingyang, et al. Velocity measurement of moving target based on rotating mirror high speed camera[J]. Optical Engineering, 2023, 62(10): 104102-104102.
- [2] Wang L, Shang Y, Cheng Q, et al. Multilayer feature fusion and saliency-attention object tracking[J]. Journal of Electronic Imaging, 2023, 32.
- [3] 王力超, 董家辉, 程清扬, 尚永健, 耿树巧 一种基于转镜式高速相机的运动目标速度测量方法, 2023.05. (实审)
- [4] 王力超, 尚永健, 董家辉, 程清扬, 耿树巧 一种采用显著性捕获交互的目标跟踪方法及系统, 2023.05. (实审)

致谢

在此，我怀着无比诚挚的心情写下这篇硕士学位论文的致谢部分。首先，我要衷心感谢我的导师王力超教授，您的严谨治学态度、深厚的学术造诣以及无私的教诲引导，为我在学术研究之路上照亮了前行的方向。您对我论文选题的精心指导，对研究过程中的困惑与难题给予耐心解答，使我得以顺利完成这项研究。您的言传身教，不仅让我在专业领域有了更深的理解和掌握，更教会了我如何做一个求真务实、严谨细致的研究者。这份感激之情，非言语所能尽述。

其次，我要深深感谢我的同窗好友尚永健，程清扬，师兄李子阳，师姐王义琼，还有刘俊楠，缪陆，刘飞瑶，刘祥玉，尹冰，汪冉，黄志成师弟师妹们是你们在我科研探索道路上的陪伴和支持，让这段旅程充满欢笑与温暖。我们共享知识、共同探讨，一起度过了无数个日夜。此外，实验室团队的每一位成员都以各自的方式给予了我帮助，你们的专业素养和团队精神是我宝贵的财富。

同时，我要特别向提供实验数据、技术支持以及参与讨论的所有老师和同学们表达由衷的感谢。没有你们的支持与付出，这篇论文也无法顺利的完成。

最后，我要深深地感谢我的父母家人和亲爱的花玉莲同学，你们的无条件支持与关爱是我勇往直前的动力源泉。你们默默的付出，我都铭记在心。这份学位论文背后，凝聚着你们无私的爱与期待。

在未来的人生旅途中，我会继续秉持这份感恩之心，砥砺前行，用实际行动回馈社会，回报师长，亲友不负青春韶华。

再次感谢所有人，谢谢！

尚德敏学 唯实惟新

