

基于 YOLOv5s 的精密视觉检测系统快速调焦方法

胡新宇 刘锡阳* 张骏巍 严爽 李云翔 叶旭辉

湖北工业大学机械工程学院, 武汉, 430068

摘要: 针对视觉检测系统在测量时因受生产精度、装配误差等外界因素影响, 图像存在离焦模糊的问题, 提出了一种基于 YOLOv5s 的精密视觉检测系统快速调焦方法, 该方法采用粗精结合调焦策略。首先利用训练的 YOLOv5s 模型搜索清晰成像的景深区间, 准确率达到 97.6%, 900 ms 内完成粗调焦过程; 然后利用清晰度评价函数及改进搜索算法实现精调焦, 在景深区间内快速准确地找到最佳成像平面。实验结果表明, 在 ± 4 mm 的离焦区间内, 调焦精度达到 0.04 mm, 平均用时不超过 1600 ms, 较现有基于图像处理的方法缩短了 47.6%, 具有速度快、精度高、适应性强等优点, 可应用于视觉检测系统的在线精密测量。

关键词: 自动调焦; 视觉检测; 区间搜索; 清晰度评价函数

中图分类号: TP272; TP391.4

DOI: 10.3969/j.issn.1004-132X.2025.04.024

开放科学(资源服务)标识码(OSID):



Fast Focusing Method for Precision Vision Detection System Based on YOLOv5s

HU Xinyu LIU Xiyang* ZHANG Junwei YAN Shuang LI Yunxiang YE Xuhui

School of Mechanical Engineering, Hubei University of Technology, Wuhan, 430068

Abstract: During measurement, the visual inspection system was influenced by factors such as production accuracy and assembly errors, leading to defocused and blurred images. Consequently, a rapid focusing method for the precision visual inspection system was developed based on YOLOv5s. This method employed a combination of coarse and fine focusing strategies. Initially, the trained YOLOv5s models were utilized to search for clear imaging depth ranges with an accuracy of 97.6%, completed the coarse focusing processes within 900ms. Subsequently, the clarity evaluation function and an improved search algorithm were applied for precise focusing, swiftly identified the optimal imaging plane within the depth range. Experimental results indicate that within a defocus range of ± 4 mm, the focusing accuracy reaches 0.04 mm, with an average time not exceeding 1600 ms, which reaches a 47.6% reduction compared to existing methods. This method offers rapid speed, high accuracy, and strong adaptability, making it ideally suited for online precision measurements in visual inspection systems.

Key words: autofocus; visual inspection; interval search; clarity evaluation function

0 引言

在光学系统中, 正焦位置是指物体在图像平面上形成最清晰焦点的位置^[1]。当物体与其对应的像满足共轭关系时, 成像系统就能够在正焦位置上实现最佳的成像质量。当焦点移动到物体的前侧或后侧时, 光线未能在图像平面上正确定位, 系统的成像会变得模糊^[2]。

视觉检测技术在工业制造、智能交通、医疗诊断等领域获得了广泛应用^[3]。图像处理是视觉检测的核心组成部分, 高质量的清晰图像能够降低图像处理中特征提取的难度, 从而提高系统处理速度^[4], 因此研究视觉系统的自动调焦技术有着

重要价值。

自动调焦技术的演进历程经历了多个阶段。从最开始的手动调焦发展到被动式和主动式的自动调焦, 随着数字图像处理技术的飞速发展, 基于图像处理的自动调焦技术成为研究热点^[5]。国内对基于图像处理的自动调焦技术进行了大量研究。潘宏亮等^[6]提出了改进 Sobel 算子图像清晰度评价函数与粗细调焦结合的变步长两段式快速搜索自动调焦方法, 提高了评价函数的准确度, 并加快了搜索策略的速度。罗其俊等^[7]提出了一种基于圆边响应曲线的离焦估计新算法, 实现了调焦参数的标定和望远物镜的自动调焦。上述研究内容集中在改进图像清晰度评价函数和聚焦策略方面, 依赖于手动设计的规则和算法, 缺乏对复杂数据和场景的学习能力。

收稿日期: 2024-05-17

基金项目: 国家自然科学基金(52075152); 湖北省自然科学基金(2022CFB301); 湖北省重点研发计划(2022BBA0016)

本文提出一种基于 YOLOv5s 的视觉系统自动调焦方法,分两步实现焦平面位置的高精度快速检测。首先采集聚焦过程的图像样本组成数据库,利用 YOLOv5s 网络训练出深度学习模型。学习模型通过实时推理相机采集图像焦距信息,搜索到景深区间,并通过伺服电机驱动 Z 轴使相机到达焦平面附近,实现粗调焦。然后,利用清晰度评价函数和改进搜索算法找到焦平面位置实现精确调焦。

1 粗精结合的自动调焦方法

1.1 离焦成像原理分析

光学系统的成像受到失焦等因素的影响,使理想中的点像实际成像为弥散斑。这种失焦现象会导致图像质量下降,表现为清晰度降低、对比度减弱,以及目标细节和低频信息的损失^[8-9]。

如图 1 所示,保持像距 v 不变并改变物距 u ,当目标在一定范围内移动时都能获得清晰的图像,这个距离被称为景深 ΔL (ΔL_2 、 ΔL_1 分别为前后景深)。当保持物距 u 不变,将像面在一定范围内移动也能获得清晰的图像时,这个像面可移动的距离称为焦深 $\Delta\delta$ ($\Delta\delta_1$ 、 $\Delta\delta_2$ 分别为前后焦深)。因此,为了获得清晰图像,自动调焦系统通常采用两种方法:一种是通过改变物距 u 搜索景深区间 ΔL 的调焦方法,另一种是通过改变像距搜索焦深区间 $\Delta\delta$ 的变焦方法。

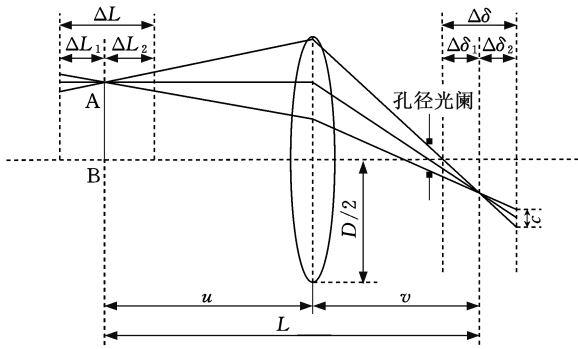


图 1 物方远心成像系统光路图

Fig.1 Optical path diagram of the object side telecentric imaging system

在上述模型中,景深近似计算式如下:

$$F = f/D \tag{1}$$

$$\Delta L_1 = \frac{FcL^2}{f^2 + FcL} \tag{2}$$

$$\Delta L_2 = \frac{FcL^2}{f^2 - FcL} \tag{3}$$

$$\Delta L = \Delta L_1 + \Delta L_2 = \frac{2f^2 FcL^2}{f^4 - F^2 c^2 L^2} \tag{4}$$

而焦深计算式如下:

$$\frac{\Delta\delta_1}{c} = \frac{\Delta\delta_2}{c} = \frac{v}{D} \tag{5}$$

$$\Delta\delta_1 = \Delta\delta_2 = \frac{vc}{D} \tag{6}$$

式中: F 为镜头的拍摄光圈值; f 为镜头焦距; D 为镜头通光直径; c 为容许弥散圆直径; L 为拍摄距离。

由式(6)可知焦深 $\Delta\delta$ 与焦距 f 相关,并且焦深很大程度上与相机的制造有关。由式(2)、式(3)可知景深 ΔL 区间可通过改变拍摄距离 L 来搜索得到。对于配备定焦镜头的工业相机,由于其焦距 f 为固定值,无法改变成像的视场范围,因此通常采用改变镜头与观测目标之间的相对距离的方法来搜索景深区间,从而获取清晰图像。

1.2 基于 YOLOv5s 的粗调焦方法

由于清晰的图像包含丰富的边缘和细节特征,而焦深范围内的模糊图像缺乏这些细节特征,因此可以将图像划分为景深范围内的清晰图像和离焦模糊图像。根据这些特征,可以判断当前图像是否属于景深内图像,从而实现景深区间的搜索。深度学习模型无需手动设计特征提取和处理步骤,能够从大量数据中学习复杂的非线性关系,实现端到端的学习。因此,本文引入深度学习来完成景深区间搜索任务,实现粗调焦过程。

1.2.1 YOLOv5s 网络结构分析

YOLOv5^[10-11] 作为一种兼顾速度和精度的单阶段目标检测算法,相比于两阶段算法如基于区域的卷积神经网络(R-CNN),舍弃了候选框提取阶段,直接对图像进行分类和回归,具有较快的检测速度。YOLOv5s 模型在保持较高检测精度的同时具有更快的目标检测速度^[12],更加符合自动调焦系统实时处理的需求,因此本文选用 YOLOv5s 模型。

如图 2 所示,YOLOv5s 网络结构主要由输入端、Backbone、Neck、Head 组成^[13]。相机实时采集的图像经过输入端预处理后被送入 Backbone 主干网络进行特征提取,然后通过 Neck 网络对提取的特征进行多尺度特征融合,接着在融合后的特征图上通过检测头部分 Head 进行目标检测和分类,最后通过后处理技术得到图像类别及其对应的置信度。

1.2.2 损失函数的选择

平板后盖中含有金属材料,当光线照射到平板后盖时,金属表面会产生反射和散射效应,使得光照在不同区域的分布出现不均匀。除此之外,在工业环境下图像还会受到复杂背景、温度波动等外界因素的影响,这些干扰不仅降低了图像的质量,还增加了目标检测任务中的噪声水平,从而影响模型的检测精度。

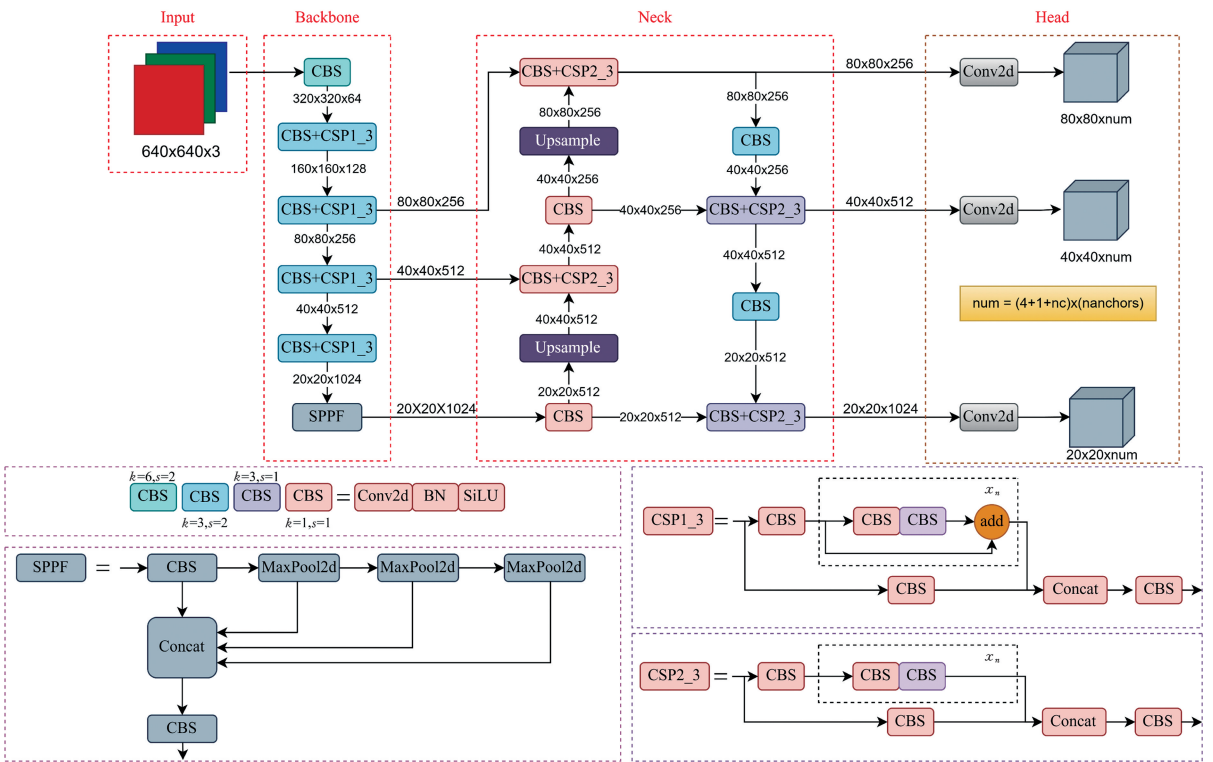


图 2 YOLOv5s 网络结构图

Fig.2 YOLOv5s network structure diagram

损失函数作为目标检测任务中的关键因素之一,直接影响模型的训练效果和检测性能。相比于模型默认的 CIoU_Loss 损失函数, Alpha-IoU 损失函数通过引入可调节参数 α ,能够在不同程度上调整对重叠区域和非重叠区域的关注度,通过调整 α 值,能够降低噪声对 IoU 的影响,提高模型对噪声的鲁棒性和检测的精度。经实验验证,当 $\alpha=3$ 时, Alpha-IoU 达到较好的检测效果,且不会额外增加训练推断时间,因此选用 α IoU_Loss 作为损失函数。 α IoU_Loss 的定义如下:

$$L_{\alpha\text{-CIoU}} = 1 - I^\alpha + \frac{\rho^{2\alpha}(b, b^{\text{gt}})}{C^{2\alpha}} + (\beta V)^\alpha \quad (7)$$

$$V = \frac{4}{\pi^2} (\arctan \frac{w}{h} - \arctan \frac{w^{\text{gt}}}{h^{\text{gt}}})^2 \quad (8)$$

$$\beta = \frac{V}{(1 - I) + V} \quad (9)$$

式中: I 为预测框和真实框的交并比; b, b^{gt} 分别为预测框与真实框中心点; ρ 为欧氏距离计算函数; C 为包含预测框和真实框的最小矩形的对角线距离; V 为预测框和真实框宽高比之间的差异; β 为权重参数; $(w, h), (w^{\text{gt}}, h^{\text{gt}})$ 分别为预测框和真实框的宽度和高度。

1.3 基于清晰度评价函数的精调焦方法

如果将正焦平面图像作为一个单独类别,利用深度学习完成精调焦过程,由于景深区间内图像的清晰度较高且差异较小,随着卷积层数的加深,景深区间图像特征被抽象到较高的层次,焦平面及其附近图像可能会被错误分类,所以提出基

于清晰度评价函数的精调焦方法。

由图 3 图像清晰度与成像位置的关系曲线可知,对于固定像距的成像系统,当目标物体位于景深 ΔL 范围内时,能够产生相对清晰的图像^[14],且图像清晰度会先增大后减小,曲线存在唯一峰值,该峰值对应的正焦位置即为精调焦需搜索位置。精调焦的控制原理为:伺服电机控制相机在景深区间内移动,获取各位置图像清晰度评价,并配合搜索算法移动物镜以搜索清晰度峰值位置,该处即为焦平面位置。

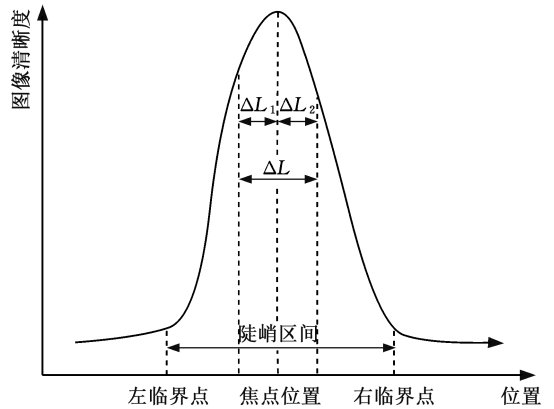


图 3 图像清晰度曲线

Fig.3 Image clarity curve

2 基于 YOLOv5s 的粗调焦算法实现

2.1 粗调焦网络模型训练

本文设计的自动调焦系统基于实际的平板后

盖精密尺寸视觉测量。图 4 为 lenovo、honor、realme 三种平板后盖在视觉测量中计算装配尺寸的目标区域细节图。后续特征提取及算法实现依赖于目标区域的清晰图像,因此选用图中三种平板的目标区域图像用于自动调焦成像,以完成搜索焦平面的前置工作。如图 5 所示, l 为到正焦平面的距离。图 5c 为三类平板的正焦平面图像,图 5a 和图 5e 分别为距离正焦面+3.0 mm 的正离焦图像和-3.0 mm 的负离焦图像,图 5b 为前景深区域与正离焦区域临界点位于+0.4 mm 附近的图像,图 5d 为后景深区域与负离焦区域临界点位于-0.4 mm 附近的图像。

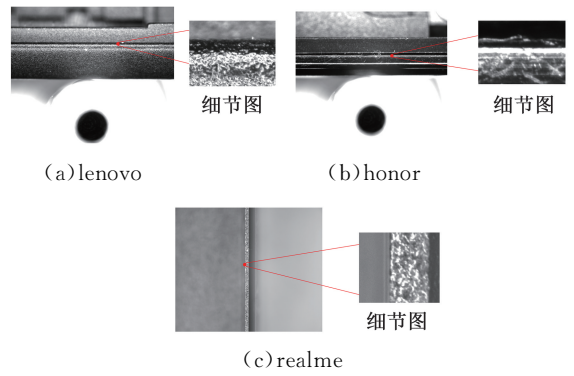


图 4 各类平板后盖待检测区域细节图

Fig.4 Details of the testing area for various types of tablet back covers

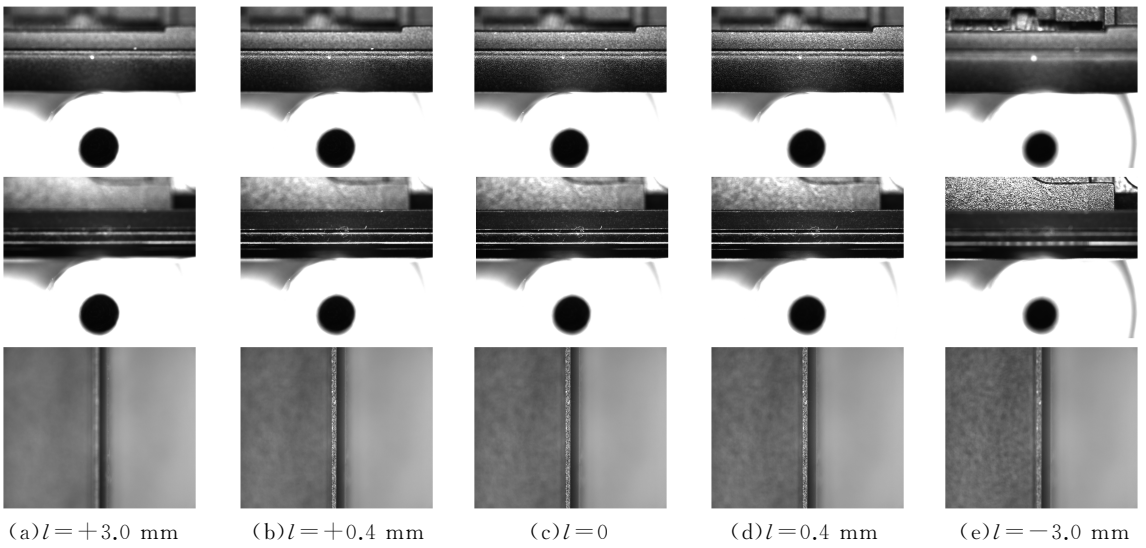


图 5 各离焦位置在相机上所成的图像

Fig.5 Images of defocused positions on the camera

在实际制作数据集过程中,首先将伺服电机的步距初始化为 0.1 mm,相机曝光时间设定为 30 ms,图像分辨率为 2448×2048 ,并依次从近到远采集 ± 4.0 mm 的图像。自动调焦系统会在新的线程中实时接收所采集的图像及其位置信息。对每一类平板重复上述操作,共采集 10 组数据,获取 2400 张图片。随后遍历所有图像的位置信息,将正离焦和负离焦区间的图像合并为离焦类,标签值为 0;将前景深、后景深及正焦面图像合并为清晰图像,标签值为 1。然后按 8 : 1 : 1 的比例随机选取训练数据、验证数据和测试数据。

模型训练的硬件环境包括 CPU Intel(R) Core(TM)i7-12700K 和 GPU RTX3070,软件环境为 Python3.10.9、CUDA12.1、Pytorch2.0.0、libtorch2.2.2。在训练阶段,通过微调超参数值(表 1)来提高模型的性能和训练速度。未提到的超参数与 YOLOv5 官方提供的默认参数相同。

2.2 评价指标

为了准确反映模型在每种平板后盖景深区间

表 1 实验参数

Tab.1 Experimental parameter

参数	值
Learning Rate	0.001
Optimizer Type	Adam
Momentum	0.937
IOU Threshold	0.6
Weight Decay	0.0005
Batch size	32

的搜索性能,采用不同的评价标准评价模型性能。

1) 分类任务评价指标^[15]:精度 P 、召回率 R 、 $F1$ 分数。

2) 目标检测任务评价指标:平均精度 u_{AP} 及均值平均精度 u_{mAP} ,计算公式如下:

$$u_{AP} = \int_0^1 P(R) dR \quad (10)$$

$$u_{mAP} = \frac{\sum_{i=1}^Q u_{AP}(i)}{Q} \quad (11)$$

其中, $P(R)$ 为 Precision-Recall 曲线函数; Q 为类

别总数; $u_{AP}(i)$ 为第 i 个类别的平均精度; u_{AP} 代表 Precision-Recall 曲线下的面积, 可以反映模型在单个类别上的检测性能; u_{mAP} 代表所有类别平均精度的平均值, 可以综合表达模型在不同类别上的检测准确度。

3) 速度和效率指标: 每秒处理的帧数 (FPS) 及推理时间, 用于评估实时系统的性能。

2.3 粗调焦性能对比实验

为了判断模型是否能完成粗调焦要求, 在自制的数据集上对 YOLOv5s- α IoU 和当前主流目标检测算法 SSD(单阶段多框检测器)^[16]、Fast R-CNN^[17]、YOLOv3-tiny^[18]、YOLOv4^[19]、YOLOv5s、YOLOv7-tiny^[20] 进行对比实验, 对比

不同算法的性能、精度、处理速度, 结果如表 2 和表 3 所示。

表 2 不同算法性能对比

Tab.2 Performance comparison of different algorithms

算法	honor			lenovo			realme		
	P	R	F1	P	R	F1	P	R	F1
SSD	83.4	81.2	82.3	82.3	85.2	83.7	80.3	79.8	80.0
Fast R-CNN	95.3	89.4	92.3	91.6	93.1	92.3	93.8	95.5	94.6
YOLOv3-tiny	88.6	89.9	89.2	86.2	88.7	87.4	91.2	91.5	91.3
YOLOv4	90.3	89.5	89.9	89.1	90.3	89.7	93.1	92.7	92.9
YOLOv7-tiny	94.1	93.2	93.6	93.8	94.0	93.9	95.7	94.6	95.1
YOLOv5s	96.5	95.8	96.1	96.6	97.8	97.2	98.5	97.3	97.9
YOLOv5s- α IoU	96.8	96.1	96.4	96.8	98.0	97.4	98.7	97.4	98.0

表 3 不同算法检测精度及效率对比

Tab.3 Comparison of detection accuracy and efficiency of different algorithms

算法	honor $u_{AP}/\%$	lenovo $u_{AP}/\%$	realme $u_{AP}/\%$	$u_{mAP}/\%$	帧率	推理时间/ms
SSD	82.5	83.4	80.3	82.1	67	14.93
Fast R-CNN	91.8	92.3	94.6	92.9	12	83.33
YOLOv3-tiny	88.1	86.3	89.4	87.9	216	4.63
YOLOv4	90.7	92.5	91.2	91.5	153	6.54
YOLOv7-tiny	93.6	94.1	95.2	94.3	182	5.49
YOLOv5s	97.0	97.3	98.6	97.6	203	4.93
YOLOv5s- α IoU	97.2	97.4	98.8	97.8	208	4.81

表 2 记录了不同算法在三种平板后盖上的 P 、 R 、 $F1$ 值, YOLOv5s- α IoU 算法针对这三种平板后盖的 $F1$ 分数分别达到了 96.4%、97.4%、98.0%, 优于其他检测算法。这表明, YOLOv5s- α IoU 相比于其他算法能更准确地区分处于离焦区间的模糊图像和处于景深区间的清晰图像。

表 3 中的数据表明, Fast R-CNN 算法虽然精度较高, 为 92.9%, 但它对单幅图像的推理时间长达 83.33 ms, 检测效率较低; SSD 算法推理时间较短, 检测效率有一定优势, 但是精度较低仅为 82.1%, 整体效果与 YOLO 系列算法差距较大。由于电机步距仅为 0.1 mm, 处于景深区间和离焦区间临界点的图像之间差异极小, 因此能够正确判断图像是否位于景深区间即精度是相对重要的评价指标。YOLOv3-tiny 算法由于参数量较少, 推理速度最快, 仅需 4.63 ms, 但在精度上相对较低; YOLOv5s 在检测精度上相比 YOLOv7-tiny 提高了 3.3%, 且在检测精度和速度上均优于 YOLOv4; YOLOv5s- α IoU 相比 YOLOv5s 在检测精度上提高了 0.2%, 帧数增加了 5 帧。综上所述, YOLOv5s- α IoU 在这些模型中有更好的表现, 能够加速收敛并提高检测精度。

3 基于清晰度评价函数的精调焦算法实现

3.1 清晰度评价函数性能分析

在精调焦过程中, 相机以步长 $\Delta S = 0.02$ mm 在景深区间 ± 0.4 mm 进行成像采集, 共采集 40

张图像。由于目标区域小于图像视场, 为了减小评价函数的计算量并提高算法的效率和响应速度, 在进行清晰度评价之前先提取 ROI 区域。选用目前三大类不同的调焦评价函数^[21] 即空域 Brenner、Laplace、Tenengrad 函数, 以及频域 DFT 函数和信息熵 Entropy 函数对图像进行清晰度评价, 得到图 6 所示的归一化评价曲线和表 4 所示各评价函数对单张图像评价所用的时间。

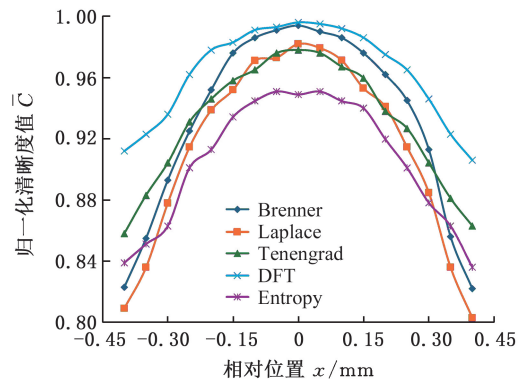


图 6 归一化后不同清晰度评价函数对景深区间图像评价曲线

Fig.6 Normalized evaluation curves of different clarity evaluation functions on depth interval images

综合图 6 和表 4 可以看出, 信息熵 Entropy 函数在精度和效率上都不如其他函数; Laplace 和 Tenengrad 函数在精度和效率相似, 但与 Brenner 函数还有一定差距; DFT 函数虽然精度最高, 但其评价时间均超过 1000 ms, 效率较低, 不满足系

表 4 不同清晰度评价函数评价单张图像所需时间

Tab.4 Time required for evaluating images with different clarity evaluation functions ms

函数	honor	lenovo	realme
Brenner	16	14	13
Laplace	125	104	81
Tenengrad	261	220	186
DFT	1283	1153	1023
Entropy	1134	1041	964

统实时性要求;Brenner 函数在精度上仅比 DFT 函数低 0.005,差距较小,同时处理图像所需的时间仅约 15 ms,远低于其他函数。因此选取 Brenner 函数作为精调焦过程中的图像清晰度评价函数。

3.2 精调焦算法流程设计

在选取合适的评价函数后,需要一定的调焦搜索算法来判断对焦状况,寻找焦平面位置。爬山搜索法^[22]原理简单,计算快捷,且根据上面的分析可知 Brenner 评价函数曲线接近单峰曲线,因此,可选取爬山法解决景深区间下的焦平面搜索问题。由于坐标轴运动产生的振动及噪声影响,本文对爬山算法进行了改进以提高搜索稳定性。改进后的搜索流程如图 7 所示。① 选取区间中点为起始点,计算当前位置图像的清晰度值 C_0 ;② Z 轴沿正方向移动 0.02 mm,待 Z 轴稳定 20 ms 后,计算该位置及下一位置图像的清晰度值

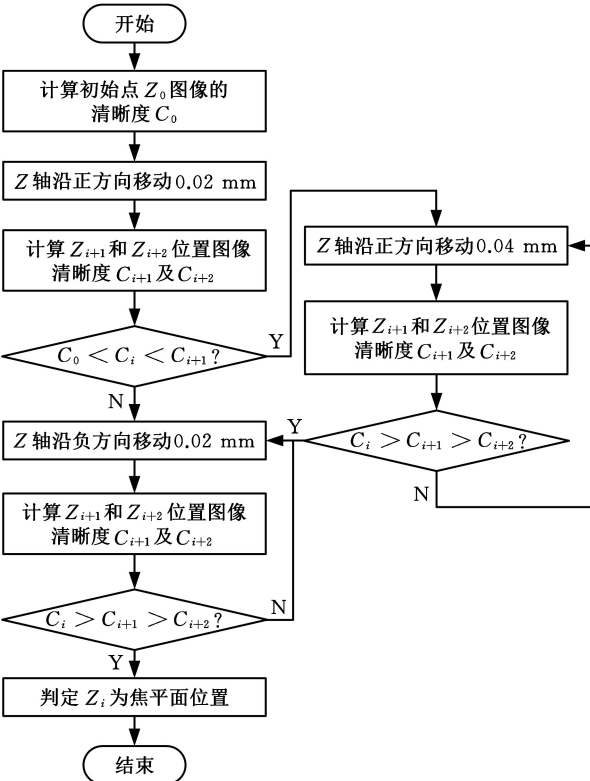


图 7 改进爬山搜索算法流程

Fig.7 Improvement of mountain climbing search algorithm process

C_i, C_{i+1} ;③ 若 $C_0 < C_i < C_{i+1}$,即三点的值连续上升,则说明搜索方向正确,增大步距为 0.04 mm 并继续沿正方向搜索,而不仅仅只使用两次结果来判定,这种改进能够有效提高搜索精度和稳定性;④ 若 $C_i > C_{i+1} > C_{i+2}$,说明已经越过焦平面,则改变搜索方向并减小步距为 0.02 mm,继续对比前后图像间的清晰度值;⑤ 若再次出现 $C_i > C_{i+1} > C_{i+2}$,则此时第 i 张图位置则为焦平面的位置。

4 系统设计及实验结果

4.1 系统设计

本文的自动调焦系统如图 8 所示,可以分为机械系统、控制系统和软件系统三大部分。机械系统包含电机、光栅、传感器等组件,确保了平板后盖在三维空间的精确定位;视觉系统作为机械系统核心部分,由 HIKROBOT 公司的 MV-CA050-12UM 相机、MVL-MY-08-65-MP 远心镜头(相机和镜头具体参数分别见表 5 和表 6)和 MV-LRSS-H-80-W 无影环形 LED 光源构成。

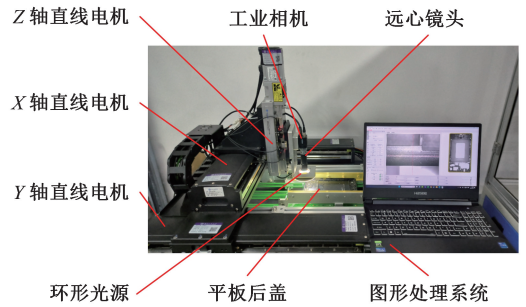


图 8 自动调焦系统

Fig.8 Automatic focusing system

表 5 工业相机参数

Tab.5 Industrial camera parameters

参数	MV-CA050-12UM
分辨率	2448×2048
传感器类型	CMOS
色彩类型	Black and white
像元尺寸	3.45 μm×3.45 μm
靶面尺寸	2/3"

表 6 远心镜头参数

Tab.6 Telecentric lens parameters

参数	MVL-MY-08-65-MP
放大倍率	0.8
工作距离	65 mm
景深	1.25 mm
物方分辨率	8.4 μm
物方远心度	<0.1%

控制系统以西门子 S7-1200PLC 为核心,外围包括启停电路、传感线路、驱动线路(X、Y 轴直线电机行程约 300 mm,重复定位精度不大于 2

μm ; Z 轴直线电机行程约 50 mm, 重复定位精度不大于 0.01 mm) 和控制线路, 内部则采用 LAD 与 STL 混合逻辑编程的可通信程序; 软件系统以 PC(GPU 为 RTX3070、CPU 为 i7-12700K, 使用 Visual Studio2017 搭载 QT5.12.12 及 libtorch 开发 UI 交互界面) 为载体, 完成图形处理工作。

视觉系统自动调焦流程如图 9 所示。其中:

$$\text{Sum}(Z_x) = \sum_{x=i}^{i+2} R(Z_x) \quad (12)$$

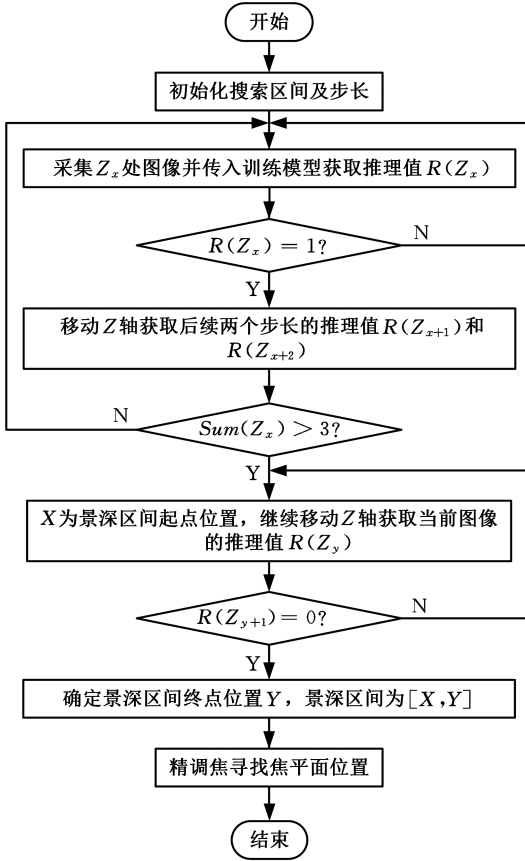


图 9 自动调焦流程图

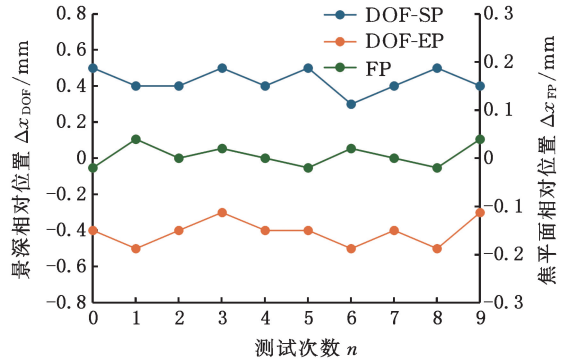
Fig.9 Automatic focusing flow chart

相机从初始位置由近及远采集图像, 并传入粗调焦网络模型获取清晰度推理值 $R(Z_x)$ 。为避免局部峰值影响, 当 $\text{Sum}(Z_x) > 3$ 时才确认进入景深区间, Z_x 为前景深起点, 从而有效排除局部极值和干扰带来的误判断, 准确可靠地搜索到景深区间起点。然后继续计算后续图像区间推理值, 直到出现 $R(Z_{y+1}) = 0$, 说明已经越过景深区间, 此时判定上一张图像的位置 Z_y 为后景深区间终点。景深区间为 $[X, Y]$, 完成粗调焦过程。随后进入图 7 的精调焦过程, 完成对焦平面的搜索。

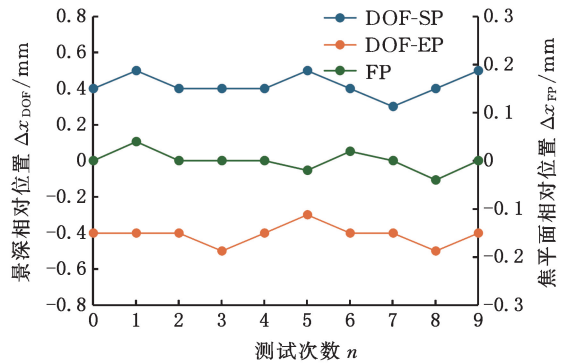
4.2 实验结果与分析

为了验证本文自动调焦系统的性能, 对本文自动调焦方法进行了可重复性和效率测试。

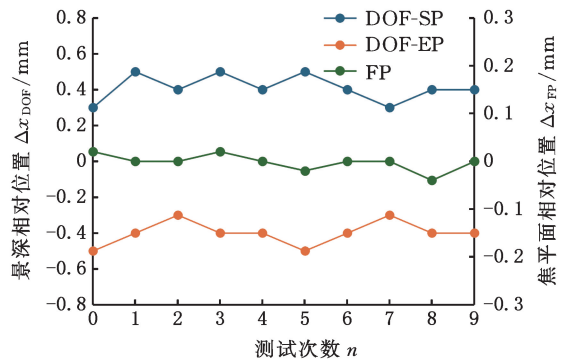
1) 可重复性实验。对本文方法进行 10 次可重复性测试, 搜索三种平板后盖的景深区间及焦平面位置, 绘制了图 10 所示的结果曲线。从图 10 中可以看出, 通过粗调焦得到三类平板后盖的 DOF-SP(景深区间起点)和 DOF-EP(景深区间终点)均有超过一半的数据准确定位到标准值 ± 0.4 mm 上, 其余数据的误差控制在 ± 0.1 mm 一个步长范围内, 误差范围较小。精调焦计算得到 FP(焦平面位置)与理论正焦平面的误差在 ± 0.04 mm 范围内, 由于 ± 0.04 mm 内的图像足够清晰, 满足后续处理要求, 因此, 使用本文方法能够将相机准确地移动到焦平面 ± 0.04 mm 的位置。自动调焦系统可重复性较好, 符合实验要求。



(a) honor 自动调焦统计结果图



(b) lenovo 自动调焦统计结果图



(c) realme 自动调焦统计结果图

图 10 自动调焦结果统计图

Fig.10 Statistical chart of automatic focusing result

2) 效率对比实验。表 7 所示为使用本文方法及基于 Brenner 清晰度评价函数和爬山搜索算法

的传统方法,对三种平板后盖进行 20 组调焦测试的结果,记录了不同平板后盖焦平面所需的平均时间及调焦失败次数。由表 7 可知,本文方法相比传统方法,调焦时间分别缩短了 46.52%、48.36%和 48.05%,在效率上占据优势。此外,由于传统方法步骤较多,容易受到系统噪声及振动等外部因素的影响,出现一次调焦失败情况,而本文方法在粗调焦和精调焦中均改进了判断条件,提高了系统稳定性,没有出现调焦失败。因此,本文提出的自动调焦系统相比传统方法有步骤少、效率高的优点,为基于图像的自动调焦技术提供一种新的思路。

表 7 自动调焦性能对比

Tab.7 Comparison of automatic focusing performance

方法	调焦时间/ms			平均 用时/ms	失败 次数
	honor	lenovo	realme		
本文方法	1583	1495	1436	1505	0
传统方法	2960	2895	2764	2873	1

5 结论

视觉检测系统的自动调焦是保证成像质量的关键技术之一,本文提出了一种基于 YOLOv5s 的快速调焦方法。首先,阐述了物方远心成像原理及粗精调焦方法,制作了位于离焦区间及景深区间图像数据集,利用 YOLOv5s 训练出的模型搜索景深区间,完成粗调焦。其次,对比不同清晰度评价函数的评价效果,选用了 Brenner 清晰度评价函数,并结合改进的爬山搜索策略完成精调焦,实现了焦平面的快速准确检测。使用精密视觉检测系统的测量对象(不同类型平板后盖)验证了本文方法的性能。实验结果表明:本文自动调焦方法能够在 1600 ms 内完成,调焦精度可达到 ±0.04 mm,能够满足视觉系统后续图像处理的速度和精度需求,在精密视觉检测领域具有实用意义和推广价值。

参考文献:

[1] 张怀利,李迎春,张廷华. 一种基于图像清晰度函数的调焦机构限区间变速控制方法[J]. 应用光学, 2021, 42(6):969-974.
ZHANG Huaili, LI Yingchun, ZHANG Tinghua, et al. Method of Variable Speed Control for Focusing Mechanism Limited Interval Based on Image Definition Function[J]. Journal of Applied Optics, 2021, 42(6):969-974.

[2] 姚曹,朱文越,徐文清,等. 结合跟星系统成像特征的自动调焦系统设计[J]. 光学精密工程, 2023, 31(14):2009-2018.

YAO Cao, ZHU Wenyue, XU Wenqing, et al. Design of Auto-focusing System Combining Imaging Characteristics of Follow Star System[J]. Optics and Precision Engineering, 2023, 31(14):2009-2018.

[3] 肖作江,朱海滨,徐志刚. 基于图像自准直自动调焦技术[J]. 光子学报, 2016, 45(10):47-51.
XIAO Zuojiang, ZHU Haibin, XU Zhigang. Based on Image Autocollimation Automatic Focus Technology[J]. Acta Photonica Sinica, 2016, 45(10):47-51.

[4] 李洋,王国名,王颖,等. 面向机器视觉测量的液体透镜调焦系统标定方法[J]. 红外与激光工程, 2022, 51(6):313-322.
LI Yang, WANG Guoming, WANG Ying, et al. Calibration Method of Liquid Lens Focusing System for Machine Vision Measurement[J]. Infrared and Laser Engineering, 2022, 51(6):313-322.

[5] 高宇轩,陈星明,代俊,等. 国内外调焦机构研究综述及发展趋势[J]. 红外, 2023, 44(4):20-32.
GAO Yuxuan, CHEN Xingming, DAI Jun, et al. Review and Development Trend of Domestic and Foreign Focusing Mechanism[J]. Infrared, 2023, 44(4):20-32.

[6] 潘宏亮,孙金霞,韩希珍. 图像清晰度评价与变步长融合调焦方法[J]. 红外与激光工程, 2023, 52(1):248-253.
PAN Hongliang, SUN Jinxia, HAN Xizhen, et al. Image Sharpness Evaluation and Variable-Step Fusion Focusing Method[J]. Infrared and Laser Engineering, 2023, 52(1):248-253.

[7] 罗其俊,葛宝臻. 基于圆边响应离焦估计的望远物镜自动调焦方法[J]. 中国光学, 2020, 13(4):760-769.
LUO Qijun, GE Baozhen. An Automatic Focusing Method of a Telescope Objective Lens Based on the Defocusing Estimation of a Circular Edge Response[J]. Chinese Optics, 2020, 13(4):760-769.

[8] 杨聚圃,杜佳林,李凡星,等. 基于深度学习的数字光刻自动检焦方法[J]. 光子学报, 2022, 51(6):261-273.
YANG Jupu, DU Jialin, LI Fanxing, et al. Deep Learning Based Method for Automatic Focus Detection in Digital Lithography[J]. Acta Photonica Sinica, 2022, 51(6):261-273.

[9] 谭伟,齐文雯,何红艳,等. 一种推扫型敏捷遥感卫星星上自动调焦技术[J]. 光学学报, 2020, 40(5):197-204.
TAN Wei, QI Wenwen, HE Hongyan, et al. An On-board Autofocusing Method for Scanning Agile Remote Sensing Satellite[J]. Acta Optica Sinica,

- 2020, 40(5):197-204.
- [10] 张学志,赵红东,刘伟娜,等. 基于改进 YOLOv5 的红外车辆检测方法[J]. 红外与激光工程, 2023, 52(8):243-252.
ZHANG Xuezhi, ZHAO Hongdong, LIU Weina, et al. An Infrared Vehicle Detection Method Based on Improved Yolov5[J]. Infrared and Laser Engineering, 2023, 52(8):243-252.
- [11] WANG Hai, CAI Yingfeng, CHEN Xiaobo, et al. Night-time Vehicle Sensing in Far Infrared Image with Deep Learning[J]. Journal of Sensors, 2015, 2016(Pt.1):3403451:1-8.
- [12] 刘彦磊,李孟喆,王宣宣. 轻量型 YOLOv5s 车载红外图像目标检测[J]. 中国光学(中英文), 2023, 16(5):1045-1055.
LIU Yanlei, LI Mengzhe, WANG Xuanxuan. Lightweight Yolov5s Vehicle Infrared Image Target Detection[J]. Chinese Optics, 2023, 16(5):1045-1055.
- [13] 薛康,刘琨,江俊峰,等. 基于 YOLOv5s 模型的光纤振动传感事件精准检测研究[J]. 光学学报, 2023, 43(2):262-272.
XUE Kang, LIU Kun, JIANG Junfeng, et al. Optical Fiber Vibration Sensing Detection with High Accuracy Based on YOLOv5s Model[J]. Acta Optica Sinica, 2023, 43(2):262-272.
- [14] 杨瑞宇,吕浩,龚晓霞,等. 红外成像系统中自动对焦功能优化方法研究[J]. 红外技术, 2020, 42(10):940-946.
YANG Ruiyu, LYU Hao, GONG Xiaoxia, et al. Optimizing Method of Autofocusing Function in Infrared Imaging System[J]. Infrared Technology, 2020, 42(10):940-946.
- [15] 李大湘,苏仲恒,刘颖. 基于改进 YOLOv4 的道路交通标志识别[J]. 光学精密工程, 2023, 31(9):1366-1378.
LI Daxiang, SU Zhongheng, LIU Ying. Road Traffic Sign Recognition Algorithm Based on Improved Yolov4[J]. Optics and Precision Engineering, 2023, 31(9):1366-1378.
- [16] 陈家栋,雷斌. 基于改进 SSD 轻量化的交通路口目标检测[J]. 传感器与微系统, 2022, 41(10):117-121.
CHEN Jiadong, LEI Bin. Traffic Intersection Target Detection Based on Improved SSD Lightweight [J]. Transducer and Microsystem Technologies, 2022, 41(10):117-121.
- [17] AHN H, LEE J. Classification of Vehicles Based on Faster R-CNN Suitable for Use in Actual Road Environments[J]. Korean Institute of Intelligent Systems, 2018, 28(3):210-218.
- [18] WANG Yanming, JIA Kebin, LIU Pengyu. Impolite Pedestrian Detection by Using Enhanced YOLOv3-tiny[J]. Computers, Materials and Continua(Tech. Science Press), 2020, 2(3):113-124.
- [19] 吴凤和,崔健新,张宁,等. 基于改进 YOLOv4 算法的轮毂表面缺陷检测[J]. 计量学报, 2022, 43(11):1404-1411.
WU Fenghe, CUI Jianxin, ZHANG Ning, et al. Surface Defect Detection of Wheel Hub Based on Improved YOLOv4 Algorithm[J]. Acta Metrologica Sinica, 2022, 43(11):1404-1411.
- [20] 薛珊,安宏宇,吕琼莹,等. 复杂背景下基于 YOLOv7-tiny 的图像目标检测算法[J]. 红外与激光工程, 2024, 53(1):269-280.
XUE Shan, AN Hongyu, LYU Qiongying, et al. Image Target Detection Algorithm Based on Yolov7-tiny in Complex Background[J]. Infrared and Laser Engineering, 2024, 53(1):269-280.
- [21] 翟永平,周东翔,刘云辉. 极小景深条件下显微镜大范围聚焦算法[J]. 光学学报, 2012, 32(4):189-198.
ZHAI Yongping, ZHOU Dongxiang, LIU Yunhui. Large Range Autofocusing Algorithm for Microscopy with Small Depth of Field[J]. Acta Optica Sinica, 2012, 32(4):189-198.
- [22] 秦富贞,陈玉杰,刘晓玲. 基于图像技术的光电测量设备智能调焦方法[J]. 激光杂志, 2020, 41(9):192-195.
QIN Fuzhen, CHEN Yujie, LIU Xiaoling, et al. Intelligent Focusing Method for Photoelectric Measuring Equipment Based on Image Technology [J]. Laser Journal, 2020, 41(9):192-195.

(编辑 袁兴玲)

作者简介:胡新宇,男,1975年生,教授、博士研究生导师。研究方向为机器视觉检测与智能控制。E-mail:19991012@mail.hbut.edu.cn。刘锡阳*(通信作者),男,2000年生,硕士研究生。研究方向为机器视觉、图像处理。E-mail:1047974853@qq.com。

本文引用格式:

胡新宇,刘锡阳,张骏巍,等.基于 YOLOv5s 的精密视觉检测系统快速调焦方法[J]. 中国机械工程,2025,36(4):864-872.

HU Xinyu, LIU Xiyang, ZHANG Junwei, et al. Fast Focusing Method for Precision Vision Detection System Based on YOLOv5s[J]. China Mechanical Engineering, 2025, 36(4):864-872.