

基于双目立体视觉的盲人避障技术研究^{*}

王新 谷亚东

(河南理工大学物理与电子信息学院 焦作 454000)

摘要:为保障盲人安全出行,针对盲道占用、破损及无盲道等问题,提出一种结合目标检测和双目立体视觉的盲人避障方法。首先,通过双目相机采集人行道信息,利用改进的YOLOv8s模型对人行道上的障碍物进行检测。然后,使用改进的立体匹配算法对障碍物进行匹配,该方法使用自适应阈值的FAST算法在尺度空间上寻找特征点,同时使用最小二乘法获取特征点的亚像素级坐标,并对特征描述符进行降维处理。最后,通过理想双目模型将二维像素坐标转化为三维空间坐标,获取障碍物的深度值。实验证明,在10 m的范围内,避障系统可以精确识别障碍物的种类,且FPS可以达到149.1;在测量障碍物深度时,其最大误差控制在5.6%以内,FPS可以达到3.8,满足盲人避障所需的实时性和距离准确度的要求。

关键词:目标检测;深度学习;立体匹配;双目测距;盲人避障

中图分类号:TP391;TN919.8 **文献标识码:**A **国家标准学科分类代码:**520.6040

Research on obstacle avoidance technology for blindbased on binocular stereo vision

Wang Xin Gu Yadong

(School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China)

Abstract: In order to ensure the safe travel of the blind, an obstacle avoidance method combining object detection and binocular stereo vision was proposed to solve the problems of blind path occupation, damage and absence. Firstly, the sidewalk information is collected by binocular camera, and the obstacles on the sidewalk are detected by the improved YOLOv8s model. Then, an improved stereo matching algorithm is used to match the obstacles, which uses FAST algorithm with adaptive threshold to find the feature points on the scale space, and uses least square method to obtain the sub-pixel coordinates of the feature points and reduce the dimension of the feature descriptors. Finally, the two-dimensional pixel coordinates are converted into three-dimensional spatial coordinates by the ideal binocular model, and the depth value of the obstacle is obtained. Experiments show that the obstacle avoidance system can accurately identify the types of obstacles in the range of 10 meters, and the FPS can reach 149.1; when measuring the depth of obstacles, the maximum error is controlled within 5.6%, and the FPS can reach 3.8, which meets the requirements of real-time and distance accuracy required by the blind to avoid obstacles.

Keywords: target detection;deep learning;stereo matching;binocular ranging;obstacle avoidance for the blind

0 引言

盲人群体是残疾人中较大的群体,据世界卫生组织(World Health Organazation, WHO)统计,我国是盲人数量最多的国家,盲人数量约有一千七百多万^[1],并且这一数字仍在逐年上升。盲人出行主要依赖导盲杖等辅助工具,但传统导盲工具作用有限,难以解决盲人出行问题^[2]。盲人出行时,避开前行道路上的障碍物是他们面临的最大挑

战,因此精准获取障碍物的距离和类别是实现盲人避障的首要条件。

目前,测距领域的研究主要依赖红外线传感器^[3]、超声波传感器^[4]、视觉传感器^[5]、激光雷达^[6]及毫米波雷达^[7]等传感器。相比激光雷达和毫米波雷达,视觉传感器不仅获取的环境信息更丰富,且成本更低,得到了广泛的应用。主流的视觉传感器分为单目相机^[8]和双目相机^[9],单目相机只有一个摄像头,无法直接测量深度,需要通过运动或其他

算法来估计深度。而双目相机有两个摄像头,利用同一物体在不同图像坐标系下的视差来计算深度,具有鲁棒性强、成本低的优点。随着人工智能技术的发展,在复杂多变的环境中,利用神经网络进行障碍物检测是一种高效且稳定的方法。由于避障设备计算能力有限,为了实时获取障碍物的距离和类别信息,需要对目标检测和立体匹配算法进行改进。

文献[10]提出一种改进的尺度不变特征变换(scale-invariant feature transform, SIFT)算法,使用 BRISK 算法生成二进制描述符,并使用汉明距离进行特征匹配,大大缩短了特征匹配时间,然而匹配成功的特征点数量过少,难以准确计算障碍物深度;文献[11]在 YOLOv8 中引入轻量级特征提取网络和特征融合网络,在不降低检测精度的同时将检测速度提高 37.8%;文献[12]将双目相机和 YOLOv5 模型相结合,通过 YOLOv5 模型检测二维目标,并根据目标中心位置计算深度,使移动机器人可以精准定位室内目标。然而,这些研究仅在室内进行,忽略了室外复杂环境对障碍物检测和测距的影响。

综上所述,为保障盲人安全出行,本文提出一种结合 YOLOv8s 和双目立体视觉的盲人避障方法。在本文提出的方法中,首先通过立体校正将双目相机校正为理想的双目结构;其次使用改进的 YOLOv8s 模型获得图像中障碍物的类别和位置信息;然后使用改进的立体匹配算法实现对障碍物的快速检测和匹配;最后,结合测距公式计算障碍物的深度和三维坐标信息。基于这些信息,系统能够在 0~10 m 的范围内规划出安全的行走路线,从而有效保障盲人安全出行。

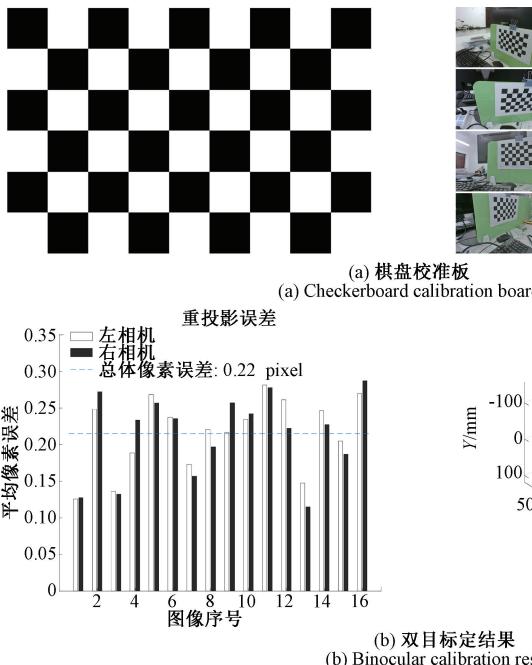


图 2 双目标定
Fig. 2 Binocular calibration

1 双目测距原理

在障碍物识别和测距过程中,传统的图像处理和匹配算法无法满足实时性和高精度的需求。因此,本文采用改进的 YOLOv8s 网络进行障碍物识别,并采用改进的立体匹配算法进行障碍物匹配,从而显著提升测距速度和精度。避障系统流程如图 1 所示。

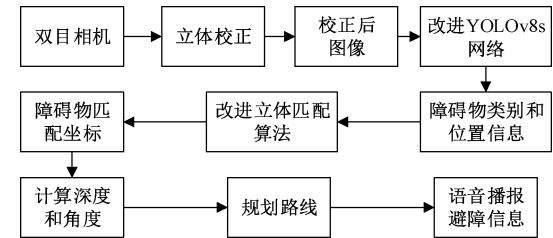


图 1 避障系统流程

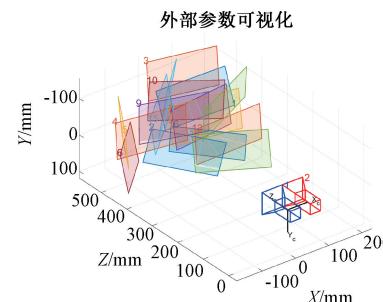
Fig. 1 Obstacle avoidance system flow

双目相机测距实质是二维平面到三维立体空间的映射,由于两个相机的三维空间坐标系不同,同一物体在两个相机像素平面上的成像位置也有所不同,利用视差公式,即可计算出目标物体与双目相机之间的距离^[13]。

由于相机的制造和安装误差,双目相机会产生横向和径向畸变,直接使用理想模型计算深度会产生较大误差,因此,使用前需要对双目相机进行校正,使用张正友标定法校正双目相机。使用双目相机从不同角度拍摄棋盘格图像,每个棋盘格的长度为 26 mm。然后,使用 MATLAB 中的相机标定工具对双目相机进行标定,标定结果如图 2 所示。



(a) 棋盘校准板
(a) Checkerboard calibration board



(b) 双目标定结果
(b) Binocular calibration results

图 2 双目标定

通过标定得到相机参数后,对相机进行校正,消除畸变,从而将实际双目系统模型转化为理想双目系统模型。理想双目系统模型如图 3 所示。

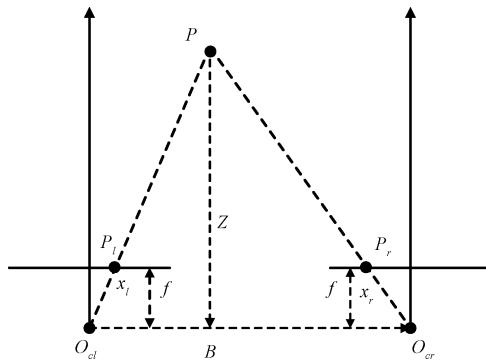


图 3 理想双目系统模型

Fig. 3 Ideal binocular system model

在理想双目系统模型中, P 是世界坐标中的一点, O_{cl} 和 O_{cr} 是左右相机的光心, x_l 和 x_r 分别是 P 在左右相机像素平面的横坐标, y 是纵坐标, B 为左右相机的基线, 根据三角形相似原理, 可以得到式(1)。

$$\frac{B - (x_l - x_r)}{B} = \frac{Z - f}{Z} \quad (1)$$

f 和 B 的值通过标定获取, $x_l - x_r$ 通过图像匹配得到, 通过式(2)即可计算点 P 的三维空间坐标。

$$\begin{cases} X = \frac{Bx_l}{x_l - x_r} \\ Y = \frac{By}{x_l - x_r} \\ Z = \frac{fB}{x_l - x_r} \end{cases} \quad (2)$$

2 改进多尺度立体匹配算法

SIFT 算法具有尺度不变性和旋转不变性, 在光照不均匀和噪声明显的情况下, 依然保持较高的鲁棒性^[14], 但运行速度较慢, 为了实现目标的快速匹配, 本文基于 SIFT 算法, 提出一种改进的多尺度立体匹配算法, 可以实现亚像素级匹配。

2.1 构建尺度空间

本文算法通过高斯金字塔构建尺度空间。通过对原始图像做不同尺度的高斯模糊和下采样得到高斯金字塔^[15], 金字塔共 O 组, 每组包含 S 层, O 根据式(3)计算得到, 高斯模糊的标准差 σ 通过式(4)计算, S 根据需要设定。

$$O = \log_2 \{\min(M, N)\} - 3 \quad (3)$$

$$\sigma(o, s) = \sigma_o \times 2^{\frac{s}{S}} \quad (4)$$

式中: M, N 是图像的宽和高, σ_o 是初始尺度(通常取 1.6), o 是组序号, s 是层序号, S 是每组的层数。

2.2 特征点检测

FAST 算法是一种高效的角点检测算法^[16], 它通过比

较以目标点为中心的邻域内像素的灰度值来判断该点是否为角点, 判断依据是一个固定的阈值。然而在复杂光照条件下, 固定的阈值难以保证提取特征点的数量和质量, 对后续测距有很大影响。为此, 本文使用一种基于灰度的自适应阈值^[17], 在不同的光照条件下都能保持较好的特征点检测性能。自适应阈值计算如式(5)所示。

$$T = \alpha \times \left(\sum_{i=1}^{16} I_i - I_{max} - I_{min} \right) / 14 \quad (5)$$

式中: α 为自适应参数, 通过实验寻求合适值, I_i 为 16 邻域第 i 个像素的灰度值, I_{max}, I_{min} 分别为 16 邻域灰度值的最大值与最小值。

使用改进的 FAST 算法在高斯金字塔的每层上提取角点, 为筛选出更加稳定的角点, 定义角点响应函数 V , 如式(6)所示。计算每个角点的响应值, 并利用非极大值抑制消除冗余角点, 仅保留局部区域响应值最大的角点作为特征点。 I_p 为角点灰度值。

$$V = \max \begin{cases} \sum I_i - I_p, I_i - I_p > t \\ \sum I_p - I_i, I_p - I_i < t \end{cases} \quad (6)$$

2.3 亚像素级坐标

在双目测距中, 距离与视差成反比, 像素级的误差对测距结果有很大影响, 为了获取更加精确的距离信息, 本文使用插值的方法获取特征点的亚像素级坐标。

构建一个二次多项式模型 R , 如式(7)所示。将特征点及其 8 邻域的灰度值作为输入, 利用最小二乘法得到完整表达式。

$$R = a_1 x^2 + a_2 y^2 + a_3 x + a_4 y + a_5 xy + a_6 \quad (7)$$

式中: x 和 y 是像素的横坐标和纵坐标, a_1 至 a_6 为二次多项式的系数。分别计算 x 和 y 的偏导, 并令它们的偏导数为 0, 就可以得到亚像素坐标相对于特征点的偏移量 x' , y' 。根据特征点的坐标和偏移量, 即可得到特征点的亚像素级坐标, 偏移量的计算如式(8)所示。

$$\begin{cases} x' = -\frac{2a_2 a_3 - a_4 a_5}{4a_1 a_2 - a_5^2} \\ y' = -\frac{2a_1 a_4 - a_3 a_5}{4a_1 a_2 - a_5^2} \end{cases} \quad (8)$$

2.4 特征点描述

SIFT 特征描述符由 128 个浮点数组成, 计算量庞大且耗时较长, 为了提高算法的实时性能, 通过对特征描述符降维来减少计算量, 提高匹配速度。

本文基于 GLOH 特征描述符^[18]对 SIFT 特征描述符进行改进, 提出一种基于距离分层的圆形描述符, 降低特征描述的计算量并减少匹配时间。具体步骤如下:

1) 定义一个以特征点为中心, 半径为 $9 \times scale$ 的圆作为特征描述符的计算范围, $scale$ 为关键点的尺度大小。

2) 根据圆内像素点和特征点的距离, 将圆划分内圈和外圈, 内圈和外圈的半径比为 1:3, 将外圈划分为 8 个等分

的子区域, 内圈单独 1 个子区域。整个特征描述区域被划分为 9 个子区域, 每个子区域内的像素数量相同。特征描述符如图 4 所示, R_1 、 R_2 是内圈和外圈的半径。

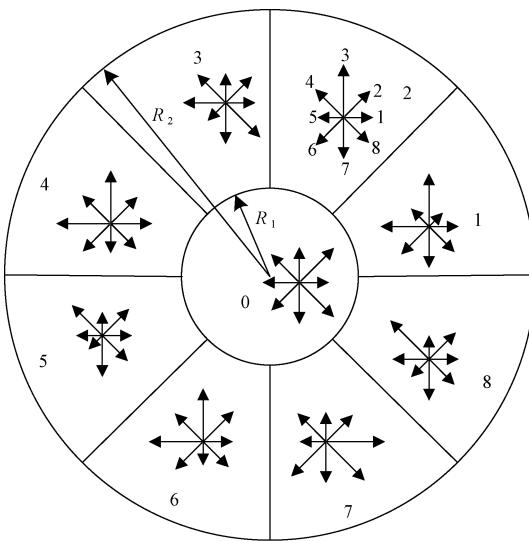


图 4 圆形特征描述符

Fig. 4 Circular feature descriptor

3) 根据像素点到特征点的距离对像素点的梯度进行高斯加权, 计算每个子区域 8 个方向上的梯度直方图, 并进行归一化。加权系数如式(9)所示。

$$W_{x,y} = \frac{l}{R_2} \times \frac{-((x - x_0)^2 + (y - y_0)^2)}{2 \times (1.5\sigma)^2} \quad (9)$$

式中: (x, y) 和 (x_0, y_0) 为所求像素点和特征点在描述区域的坐标位置, σ 为所在尺度, l 为所求像素点和特征点的距离, R_2 是特征描述区域的半径。

每个特征点描述符的维度为: 9 个分区 \times 8 个梯度方向 = 72 维。与 SIFT 算法相比, 每个特征点描述符减少了 43.7% 的数据量, 并保留了图像的平移、缩放、旋转等基本不变性特征。

2.5 特征点匹配

改进算法计算出的特征描述符是一组高维特征, 一般使用欧氏距离作为特征匹配的相似性度量指标。当匹配点对的最近距离值与次近距离值之比小于阈值时, 定义为粗匹配正确的点, 为得到更加精准的匹配结果, 在粗匹配的基础上使用交叉匹配提纯匹配点。

3 改进 YOLOv8s 模型

为了减少模型参数并提高目标检测准确率, 对 YOLOv8s 模型进行 3 点改进: 首先, 将 PConv^[19] 与 C2f 模块相结合构建 C2f-PConv 模块; 其次, 使用简化的加权双向特征金字塔网络 (bidirectional feature pyramid network, BiFPN)^[20] 作为特征融合网络; 最后, 将损失函数改为 WIoU^[21] 损失函数, 改进后的 YOLOv8s 模型结构如图 5 所示。

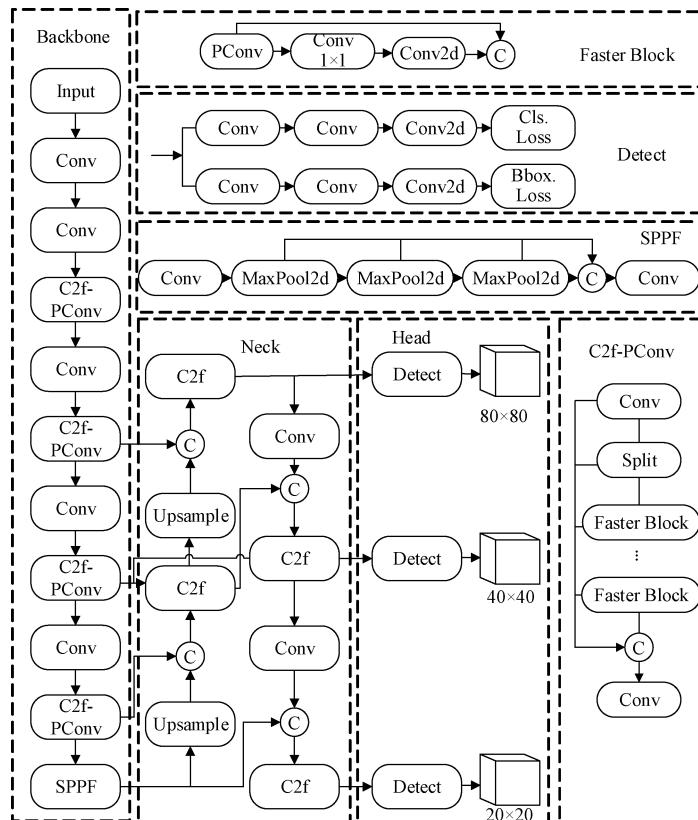


图 5 改进 YOLOv8s 结构图

Fig. 5 Improved YOLOv8s structure diagram

3.1 改进主干网络

YOLOv8s 主干网络中的 C2f 模块使用了大量的 Bottleneck 结构, 计算十分复杂, 将 PConv 引入 C2f 模块, 可以大大减少计算量。

PConv 只在部分通道执行卷积操作, 其余通道保持不变, 大大减少了计算量, PConv 模块结构如图 6 所示。 h 、 w 为输入特征图的宽和高, c 为总输入通道数和总输出通道数, c_p 为 PConv 参与卷积的通道数, k 为卷积核尺寸。如果 $c_p/c=1/3$, PConv 的计算量仅为普通卷积的 $1/9$ 。

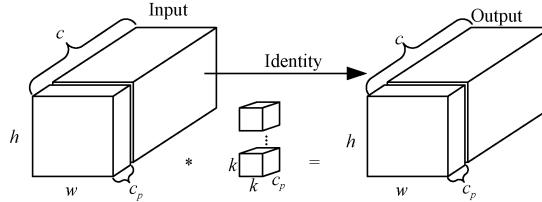


图 6 PConv 卷积

Fig. 6 PConv Convolution

3.2 改进特征融合网络

YOLOv8s 模型的特征融合网络在特征融合时存在贡献不均匀的问题, 不同尺度之间的特征不能得到充分利用。为解决这个问题, 本文在 YOLOv8s 模型中引入简化的 BiFPN 网络, 原版 BiFPN 网络结构如图 7(a)所示, 简化的 BiFPN 网络如图 7(b)所示。

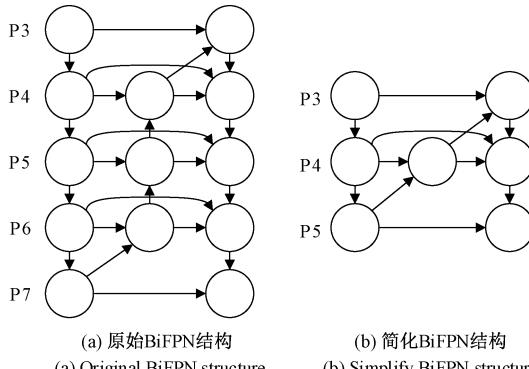


图 7 BiFPN 结构

Fig. 7 BiFPN structure

3.3 改进边界框损失函数

训练样本的标注框是识别障碍物的重要信息。由于障碍物重叠、遮掩等情况, 导致标注框与实际障碍物的位置出现错位, 这种标签被称为低质量标签。相比 CIoU 损失函数, WIoU 损失函数拥有动态非单调机制, 可以动态评估标签质量, 在训练时为低质量标签分配更小的梯度增益, 抑制模型对低质量标签的过度学习, 使模型聚焦在普通质量的标签上, 从而使模型有更好的泛化能力; 同时 WIoU 损失函数可以动态调整梯度增益, 使模型更加关注对目标定位贡献较大的边界框, 从而提高模型的定位精度。

4 避障实验结果与分析

4.1 立体匹配算法对比

为验证本文立体匹配算法的有效性, 使用 SIFT 算法、PCA-SIFT 算法和本文算法进行对比实验, 实验图像为在空旷区域和人行道区域拍摄的照片。双目相机型号为 HBV-1780-2, 图像大小为 720×2560 。

原始图像经过校正得到图 8 所示的实验图像, 实验图像首先通过 YOLOv8s 模型检测障碍物, 去除图像中无用的背景信息, 然后对障碍物进行特征匹配。

图 9 为 SIFT 算法匹配结果, 图 10 为 PCA-SIFT 算法匹配结果, 图 11 为改进算法匹配结果。图中灰色实线表示匹配正确, 黑色虚线表示匹配错误, 匹配成功率和算法时间如表 1 所示。

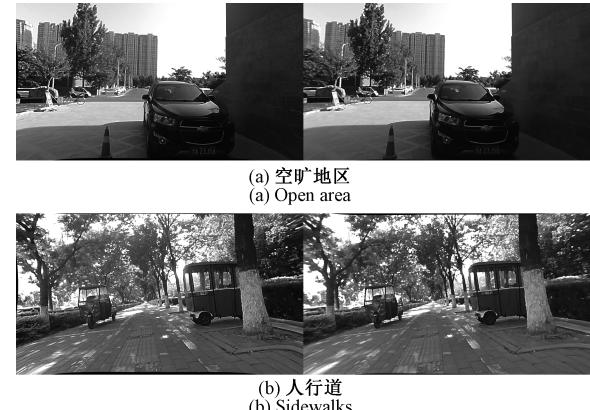
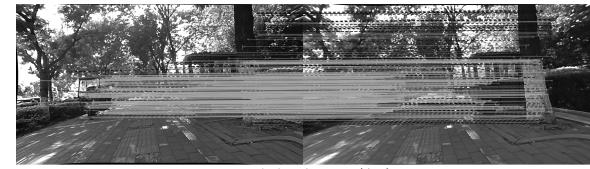


图 8 实验图像

Fig. 8 Experimental image



(a) 空旷地区-SIFT算法
(a) Empty area-SIFT algorithm



(b) 人行道-PCA-SIFT算法
(b) Sidewalk-PCA-SIFT algorithm

图 9 SIFT 特征匹配

Fig. 9 SIFT feature matching

由于 SIFT 算法和 PCA-SIFT 算法的特征提取部分是一样的, 仅在特征描述方面有所不同, 因此图 9 和图 10 中的匹配点位置高度相似, 但是 SIFT 算法的误匹配数量更多。而图 11 使用改进算法, 在得到更多匹配点的同时, 误匹配的数量也少于其他两种算法。

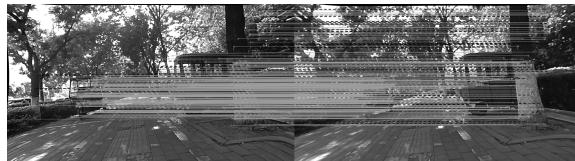
(a) 空旷地区-PCA-SIFT算法
(a) Open area-PCA-SIFT algorithm(b) 人行道-PCA-SIFT算法
(b) Walkway-PCA-SIFT algorithm

图 10 PCA-SIFT 匹配

Fig. 10 PCA-SIFT matches

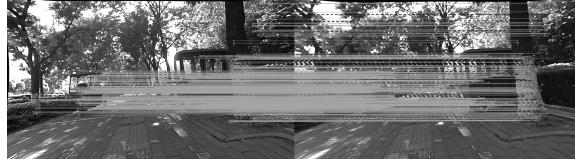
(a) 空旷地区-改进立体匹配算法
(a) Open area-improved stereo matching algorithm(b) 人行道-改进立体匹配算法
(b) Sidewalk-improved stereo matching algorithm

图 11 改进立体匹配算法匹配

Fig. 11 Improved stereo matching algorithm

表 1 立体匹配算法对比

Table 1 Comparison of stereo matching algorithms

算法	特征	匹配	匹配	算法
	匹配数	成功数	成功率/%	时间/s
SIFT(a)	327	233	71.3	0.296
PCA-SIFT(a)	292	227	77.7	0.275
改进算法(a)	473	374	79.1	0.154
SIFT(b)	530	388	73.2	0.371
PCA-SIFT(b)	468	360	76.9	0.348
改进算法(b)	604	486	80.4	0.265

由表 1 可知,改进算法在实验图像上的匹配点数量最多,速度最快。相比 SIFT 算法,改进算法的匹配成功率分别提高了 7.8% 和 7.2%,算法时间分别减少了 48.0% 和 28.6%;相比 PCA-SIFT 算法,匹配成功率提高了 1.4% 和 3.5%,算法时间分别减少了 44.0% 和 23.9%。PCA-SIFT 算法通过将特征向量投影到低维度的子空间中,有效降低了特征匹配所需的时间,然而,其特征检测时间并未得到

优化。相比之下,本文提出的算法巧妙结合了 FAST 算法,可以快速检测特征点,并使用 72 维特征描述符,保证特征匹配精准性的同时,显著提高了整体算法的速度。相较于 SIFT 和 PCA-SIFT 算法,本文算法在匹配数量上和时间上都展现出更优的性能。

4.2 障碍物检测结果

训练平台硬件配置为 i7-8700 CPU, NVIDIA RTX 4060, 8GB 显卡, 软件环境为 win10 系统, 64 位, python 3.11, pytorch 2.1.1 框架, Cuda 12.3 版本, 实验过程中训练批次大小设置为 16, 初始学习率为 0.001, 迭代次数为 150, 优化器采用 SGD。

本文实验使用自建数据集,包含各种尺寸的障碍物图片共 3 600 张, 经过缩放、裁剪、插值等操作将图像大小调整为 640×640。数据集包含 14 个种类, 分别是两轮车、汽车、猫、狗、垃圾桶、消防栓、红绿灯、路障、人、电线杆、挡车石、树、三轮车、斑马线。将数据集按照 6:2:2 的比例划分训练集、验证集和测试集, 并通过加噪、翻转、缩放等方法对训练集图片进行数据扩充, 得到 6 480 张图片。

本文采用平均精度(mAP)对模型的检测效果进行评估, 使用计算量(FLOPS)和参数量(Params)对模型的计算资源占用程度进行评估。mAP 的计算公式为:

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

$$AP = \int_0^1 P(R) dR \quad (12)$$

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \quad (13)$$

式中: TP、FP、FN 分别表示正确检测的数量、错误检测的数量、没有检测出的数量。

为讨论模型改进过程中每个模块对模型的影响, 在测试集中进行了消融实验, 结果如表 2 所示。

由表 2 可知, 实验 1 将 C2f 替换为 C2f-Faster 后, 计算量和参数量明显下降, 但平均精度也略微下降, 表明使用 C2f-Faster 模型可以实现网络的轻量化, 但也会影晌平均精度; 在实验 5 中, BiFPN 相比原特征融合网络, 增加了跨尺度连接, 使用不同的权重进行特征融合, 相比实验 1, 在略微增加计算量的同时, 平均精度增加了 2.2%; 实验 8 将损失函数改为 WIoU, 由于 WIoU 的动态梯度增益分配策略, 相比实验 5, 在不增加参数量和计算量的情况下, 平均精度提高了 0.9%。

在同样的配置下, 利用测试集对 YOLOv8s 和改进模型进行对比, box_loss 损失如图 12 所示, 结果如表 3 所示。

表 2 消融试验
Table 2 Ablation test

实验编号	C2f-Faster	BiFPN	WIoU	平均精度 mAP@0.5/%	参数量 Params/M	计算量 FLOPs/G	权重/MB	帧率/fps
1	×	×	×	79.0	11.13	28.5	22.5	134.1
2	√	×	×	77.4	9.68	24.2	19.1	153.7
3	×	√	×	80.7	11.20	31.1	24.3	122.6
4	×	×	√	79.3	11.13	28.5	22.5	135.9
5	√	√	×	79.6	9.75	24.4	19.3	146.6
6	×	√	√	81.0	11.20	31.1	24.3	123.5
7	√	×	√	78.3	9.68	24.2	19.1	155.4
8	√	√	√	80.5	9.75	24.4	19.3	149.1

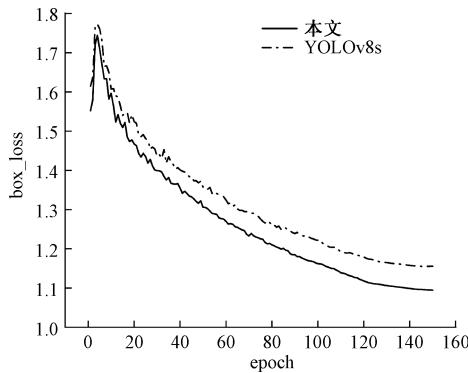


图 12 测试集损失

Fig. 12 Test set loss

表 3 不同模型对比
Table 3 Comparison of different models

模型	平均精度 mAP@0.5/%	参数量 Params/M	计算量 FLOPs/G	帧率/fps
YOLOv8s	79.0	11.13	28.5	134.1
本文	80.5	9.75	24.4	149.1

由表 3 可以看出,改进模型相比原始模型在参数量和计算量分别减少 12.4% 和 14.4% 的情况下,平均精度提高了 1.5%,并且 FPS 也提高了 11.2%。改进后的模型有更好的识别效果和检测速度,能够满足实时检测障碍物的需要。为更直观感受改进效果,使用双目相机采集图像对改进前后模型进行测试,结果如图 13 所示,左侧是 YOLOv8s 检测结果,右侧是本文模型检测结果。从图 13 可以看出,在不同光照条件下,改进模型都拥有更好的检测效果。

4.3 避障实验

为验证双目测距的准确性,针对图 8 中的几种障碍物从左向右进行编号:a1(路障)、a2(汽车)、b1(树)、b2(三轮车)、b3(树)、b4(三轮车)、b5(树),并使用软尺测量距离,实

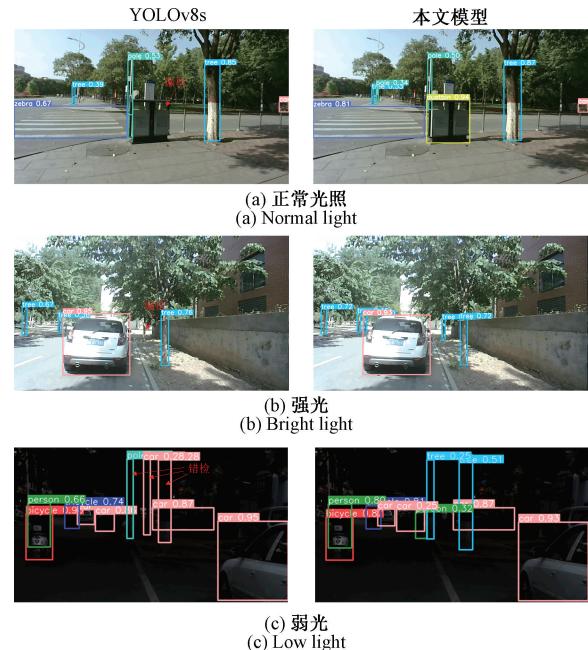


图 13 不同光照下的检测结果

Fig. 13 Test results under different lighting conditions

验估计距离和实际距离如表 4 所示。误差 e 使用式(14)计算, Z_c 是目标的估计距离, Z_m 是目标的实际距离, θ 为目标偏离角度,通过式(15)计算,正前方为 0°,范围是 -90°~90°。

$$e = \frac{\sqrt{(Z_c - Z_m)^2}}{Z_m} \times 100\% \quad (14)$$

$$\theta = \arctan \frac{X}{Z} = \arctan \frac{x_t}{f} \quad (15)$$

从表 4 可以看出,在 10 m 范围内,测距最大误差仅为 5.6%,满足盲人避障所需的要求。

在盲人行走过程中,本文所提方法可以精准识别出前进道路上的障碍物,并给出精确距离和角度,可以有效帮助盲人躲避障碍物,保障盲人出行安全。

表 4 测距结果				
编号	实际/m	估计/m	$\theta/(^\circ)$	$e/\%$
a1	3.51	3.62	-18.1	3.1
a2	3.03	3.17	-4.9	4.6
b1	7.99	7.62	-35.2	3.7
b2	6.52	6.84	-19.8	4.9
b3	9.41	9.94	6.1	5.6
b4	4.29	4.15	9.4	3.3
b5	3.09	3.17	27.0	2.5

5 结 论

为帮助盲人安全出行,本文提出一种基于 YOLOv8s 和双目立体视觉盲人避障方法。在障碍物检测方面,使用改进的 YOLOv8s 模型检测图像中的障碍物,在复杂环境下,仍然具有较高的检测速度和精度;在立体匹配方面,使用自适应阈值的 FAST 算法实现在尺度空间上快速提取特征点,并使用最小二乘法计算特征点的亚像素级坐标,提高测距的精度,使用圆形描述符对特征点进行描述,有效降低了计算量,提高算法速度。本文所提算法,在 10 m 范围内,可以精准且迅速的识别障碍物并进行测距,测距最大误差为 5.6%,满足盲人避障所需的实时性和距离准确度的要求。

参考文献

- [1] World Health Organization. World report on vision: CC BY-NC-SA 3.0 IGO[R]. Geneva: World Health Organization, 2020: 42.
- [2] KHAN A, KHUSRO S. An insight into smartphone-based assistive solutions for visually impaired and blind people: Issues, challenges and opportunities [J]. Universal Access in the Information Society, 2021, 20(2): 265-298.
- [3] CHANG W J, CHEN L B, CHEN M CH, et al. Design and implementation of an intelligent assistive system for visually impaired people for aerial obstacle avoidance and fall detection [J]. IEEE Sensors Journal, 2020, 20(17): 10199-10210.
- [4] QI L, LIU Y, GAO CH SH, et al. Hybrid pedestrian positioning system using wearable inertial sensors and ultrasonic ranging [J]. Defence Technology, 2024, 33: 327-338.
- [5] WEI B B, LIU J, LI AO, et al. Remote distance binocular vision ranging method based on improved YOLOv5[J]. IEEE Sensors Journal, 2024, 24(7): 11328-11341.
- [6] 李倩,陈付龙,郑亮,等. IMU 紧耦合的多激光雷达定位与建图方法[J]. 电子测量技术,2024,47(9):26-32.
- [7] LI Q, CHEN F L, ZHENG L, et al. IMU tightly coupled multi-lidar location and mapping method[J]. Electronic Measurement Technology, 2024, 47(9): 26-32.
- [8] 赵越坤,罗素云,魏丹,等. 基于毫米波雷达和视觉的目标检测方法[J]. 计算机技术与发展,2023,33(6): 35-40.
- [9] ZHAO Y K, LUO S Y, WEI D, et al. Target detection method based on millimeter wave radar and vision[J]. Computer Technology and Development, 2023, 33(6): 35-40.
- [10] 周东兴,韩东升. 基于单目视频的伪双目动态测距方法[J]. 电子测量技术,2021,44(3):38-44.
- [11] ZHOU D X, HAN D SH. Pseudo-binocular dynamic ranging method based on monocular video [J]. Electronic Measurement Technology, 2021, 44 (3): 38-44.
- [12] SUN X Y, JIANG Y ZH, JI Y F, et al. Distance measurement system based on binocular stereo vision [C]. IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2019, 252(5): 052051.
- [13] ZHONG B, LI Y B. Image feature point matching based on improved SIFT algorithm[C]. 2019 IEEE 4th International Conference on Image, Vision and Computing(ICIVC). IEEE, 2019: 489-493.
- [14] 袁宇乐,汤文兵. 基于改进 YOLOv8s 模型的电动车骑乘人员头盔佩戴检测[J]. 湖北民族大学学报(自然科学版),2024,42(3):355-360,367.
- [15] YUAN Y L, TANG W B. Helmet wearing detection of electric vehicle riders based on improved YOLOv8s model [J]. Journal of Hubei University for Nationalities(Natural Science Edition), 2024, 42(3): 355-360,367.
- [16] QIAN W J, HU CH H, WANG H ZH, et al. A novel target detection and localization method in indoor environment for mobile robot based on improved YOLOv5[J]. Multimedia Tools and Applications, 2023, 82(18): 28643-28668.
- [17] MU Q SH, WEI J, YUAN ZH G, et al. Research on target ranging method based on binocular stereo vision[C]. 2021 International Conference on Intelligent Computing, Automation and Applications (ICAA), 2021: 81-85.
- [18] HOSSEIN N Z, AGAHI H, MAHMOODZA D A. Image matching based on the adaptive redundant keypoint elimination method in the SIFT algorithm [J]. Pattern Analysis and Applications, 2021, 24(2):

669-683.

- [15] 朱宇韩,卜雄洙,裘恺丽,等.基于改进 AGAST 角点域特征的障碍检测算法[J].国外电子测量技术,2021,40(9):54-58.

ZHU Y H, BU X ZH, QIU K L, et al. Obstacle detection algorithm based on improved AGAST corner feature[J]. Foreign Electronic Measurement Technology, 2021, 40(9): 54-58.

- [16] 马哲伟,周福强,王少红.昏暗环境下自适应 ORB-SLAM2 算法研究[J].电子测量技术,2024, 47(6): 94-99.

MA ZH W, ZHOU F Q, WANG SH H. Research on adaptive ORB-SLAM2 algorithm in dim environment[J]. Electronic Measurement Technology, 2024, 47 (6): 94-99.

- [17] 行劳仪,徐成,高宏伟.高效高精度光照自适应的 ORB 特征匹配算法[J].电子测量与仪器学报,2023, 37(7):140-147.

XING F Y, XU CH, GAO H W. ORB feature matching algorithm for high efficiency and high precision illumination adaptive[J]. Chinese Journal of Electronic Measurement and Instrument, 2023, 37(7): 140-147.

- [18] ZHAO Z L, WANG F, YOU H J. Robust region feature extraction with salient MSER and segment distance-weighted GLOH for remote sensing image registration[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 2475-2488.
- [19] CHEN J R, KAO SH H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [20] YU CH SH, SHIN Y. SAR ship detection based on improved YOLOv5 and BiFPN [J]. ICT Express, 2024, 10(1): 28-33.
- [21] CHO Y J. Weighted intersection over union(wIoU): A new evaluation metric for image segmentation[J]. ArXiv preprint arXiv: 2107.09858, 2021.

作者简介

王新,教授,主要研究方向为信号处理、故障诊断、人工智能。

E-mail:wangxin@hpu.edu.cn

谷亚东(通信作者),硕士研究生,主要研究方向为图像处理、人工智能。

E-mail:208712436@qq.com