

# 基于双目视觉的自动跟踪系统设计

盛 晶, 王利恒\*

(武汉工程大学 电气信息学院, 武汉 430205)

**摘 要:** 设计了一种基于双目视觉的目标自动跟踪系统。该系统通过双目相机获取图像信息后,先利用Yolov5s实时检测算法快速识别目标,再利用核相关滤波法对目标进行跟踪,最后采用半全局立体匹配算法结合加权最小二乘法进行目标深度的测量,获取距离信息和位置。根据目标位置控制系统的速度和方向,使系统与目标保持相对方位和设定的距离,实现自动跟踪。从结果上看,采用了该系统的拍摄机器人可以执行危险任务,解决安全问题。

**关键词:** 双目视觉;实时检测;半全局立体匹配法;目标识别与跟踪

**中图分类号:** TN209;TP391.4 **文献标志码:** A **文章编号:** 1005-488X(2022)04-0292-06

## Design of Automatic Tracking System Based on Binocular Vision

SHENG Jing, WANG Liheng

(School of Electrical and Information Engineering, Wuhan Institute of Technology, Wuhan 430205, CHN)

**Abstract:** An automatic target tracking system based on binocular vision was designed. After the image information was acquired by the binocular camera of the system, the Yolov5s real-time detection algorithm was used to quickly identify the target, the kernel correlation filtering method was used for target tracking, and finally the semi-global stereo matching algorithm combined with the weighted least squares method was used to measure the target depth to obtain distance information and location. After the position of the target was obtained, the speed and direction of the system were controlled, so that the system could maintain the relative orientation and set distance from the target, and realize automatic tracking. From the results, robots could replace reporters to perform real-time tracking reports of dangerous tasks, solving the safety problems of reporters.

**Key words:** binocular vision; real-time detection; SGBM; target recognition and tracking

### 引 言

随着计算机技术的发展,图像处理能力大大增

强,数字技术在图形处理领域和光电领域的应用进一步扩展<sup>[1]</sup>。目标识别主要分为基于图像特征和基于深度学习两种途径。识别图像特征主要有颜色

收稿日期:2022-06-08

作者简介:盛 晶(1998—),女,硕士研究生,研究方向为人工智能与仪器仪表;(E-mail:1523460211@qq.com)

王利恒(1973—),男,博士,教授,主要研究方向为智能仪器仪表。(E-mail:958900865@qq.com)

\* 通讯作者

特征,形状特征,梯度特征和模式特征<sup>[2]</sup>,但无法从图片中提取更深入的高语义特征,若处于复杂的环境中识别准确率无法保证,所以提出了基于深度学习的识别算法,包括Fast R-CNN<sup>[3]</sup>,SSD<sup>[4]</sup>及YOLO系列<sup>[5]</sup>等。对于目标实时跟踪定位主要使用两种传感器,一种是距离传感器,如雷达<sup>[6]</sup>,激光<sup>[7]</sup>,红外<sup>[8]</sup>,这些方法只能获取固定目标与障碍物的距离信息<sup>[9]</sup>。一种是视觉传感器,可以做到无接触、直观快速、测量范围广以及对特定目标识别等<sup>[10]</sup>。张岚<sup>[11]</sup>设计了智能小车目标跟踪系统,利用摄像头识别目标,投影法对目标定位,但是容易受到相似目标和环境的影响。张乘龙<sup>[12]</sup>等人设计了一种基于KCF算法的跟踪系统,利用KCF对目标进行跟踪,若目标速度突变,不能快速识别,容易致使目标丢失。

针对以上问题,设计基于双目视觉的自动跟踪系统,利用同步同帧的双目相机进行图像采集和处理任务。在YOLO算法,核相关滤波法和半全局立体匹配法的基础上,以STM32控制器为处理核心,搭载树莓派,支持在嵌入式上使用python来编程。通过对目标识别检测模型以及双目测距模型的建立,实验仿真得出结果,从结果上看,该方法能满足对目标自动跟踪的要求。

## 1 系统总体设计

目标跟踪系统主要由目标识别和目标跟随两大功能模块组成。目标识别负责在繁杂的视场环境中识别分割出被跟踪的目标,并处理得到目标相对于跟踪器的方位、距离和相对速度等相关信息。目标跟随负责将目标识别处理好的信息作为输入变量,控制跟随器完成目标的实时跟踪。

### 1.1 系统的硬件组成

要实现目标的识别定位跟踪功能,系统硬件就需要支持图像采集和实时处理功能以及底盘驱动功能。图像采集和处理需要处理能力较强的硬件,这里选用树莓派作为图像采集和处理平台。而底盘的驱动与跟踪控制需要实时性好,硬件控制功能强的嵌入式处理器平台,这里选用STM32。两个平台通过通信完成信息交互。

视觉获取系统主要由同步同帧双目摄像头组成,实现对目标人物信息的获取。嵌入式计算系统主要由树莓派组成,实现对目标图像信息的处理。底层控制系统主要由STM32控制器,舵机,电机组成,实现底盘的运动对目标跟踪。硬件整体架构图如图1。通过双目相机进行标定图片和目标图片的

采集,利用树莓派进行图像处理,将结果通过串口传递给STM32控制模块,控制模块控制电机舵机对目标进行跟踪,同时将自身的状态信息传递回控制器,形成一个闭环控制,始终保持与目标在一定的距离范围内,从而实现对目标的自动跟踪。

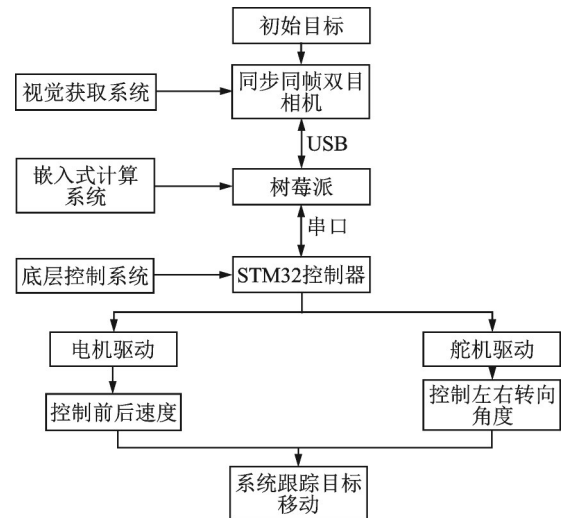


图1 硬件整体架构图

Fig.1 Overall hardware architecture

### 1.2 软件功能组成

要实现目标的识别定位跟踪功能,将其分为目标识别模块,目标方位检测模块以及双目测距模块。当摄像机拍摄目标图像,目标识别模块确定目标是否在图像中。目标检测模块对图像分割进行方位跟踪。双目测距模块测量与目标之间的距离。以上模块得到的处理信息送入控制系统实现自动跟踪功能。软件设计流程图如图2。

系统通过图像处理模块对图像预处理,将左相机拍摄的图像送入目标识别模块,使用YOLOv5s算法对目标进行识别检测。将识别到的第一帧目

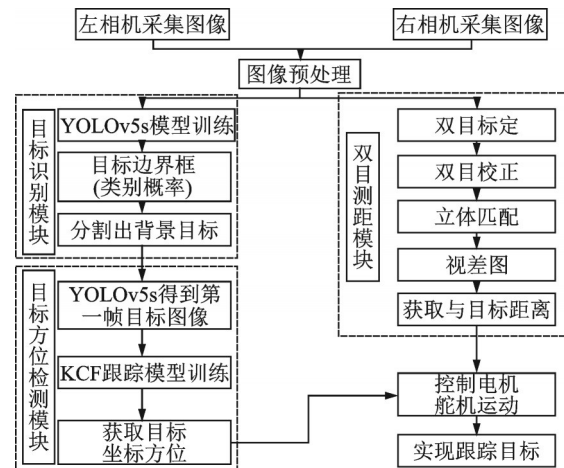


图2 软件设计流程图

Fig.2 Flow chart of software design

标图像送入目标方位检测模块,采用KCF核相关滤波算法,利用第一帧图像做正样本训练出测试样本集,再进行特征提取训练模型,对跟踪目标下一帧位置进行预测,根据预测坐标对目标进行跟踪;将双目相机拍摄的图像送入双目测距模块,根据标定结果对原始图像进行校正,然后对校正后的两张图像进行像素点匹配,获得深度图,对目标定位。以上实现对目标的自动跟踪。

## 2 目标方位检测

对目标实现方位检测,当摄像机拍摄到图像后,检测目标是否处于图像中,且使目标保持在图像中央,目标转动时,通过目标识别与方位检测模块处理得到的信息传入微控制器控制舵机转动,带动摄像机跟随目标转动,实现方位的检测功能。

### 2.1 目标识别算法设计

目标方位检测第一步就是先检测识别出目标。本系统选用一种速度快,实时性强的YOLOv5s识别算法。YOLOv5s内存大小为4.10 M,适合部署在嵌入式设备上。

#### 2.1.1 YOLOv5s 目标检测

YOLOv5s检测最大的优点是速度快。输入的图片使用马赛克方式进行数据增强和采取自适应瞄框计算,得到最佳瞄框值;主干网络里面包括Focus, CBH, CSP-1x, SPP四个模块,图片通过Focus模块进行自我复制然后进行切片操作,将 $4 \times 4 \times 3$ 变成 $2 \times 2 \times 12$ 的特征图,加快对候选区域的特征提取速度。切片图像通过CBH模块,经过32个卷积核的卷积操作,进行特征提取。在特征提取的过程中利用CSP1-x模块来优化网络结构中的梯度信息,使训练出的模型保持轻量化。接着SPP模块对不同尺寸的输入转变为相同大小的输出,解决输入图像尺寸不统一的问题;Neck网络采用FPN+PAN网络结构,通过上采样操作,将高层特征的信息和低层特征进行融合,计算出预测的特征图。通过下采样操作,将低层特征信息和高层特征信息融合,输出预测特征图;输出端采用CIoU作为回归损失函数<sup>[13-14]</sup>。YOLOv5目标检测流程如图3。

#### 2.1.2 目标识别

YOLOv5s将输入图像归一化然后划分成网格,目标的中心落入其中某一网格中,该网格预测和识别该目标。本系统使用PC进行试验,基于Windows10,使用Pytorch深度学习框架。得到结果图包括目标类型及置信度评分<sup>[15]</sup>,如图4。

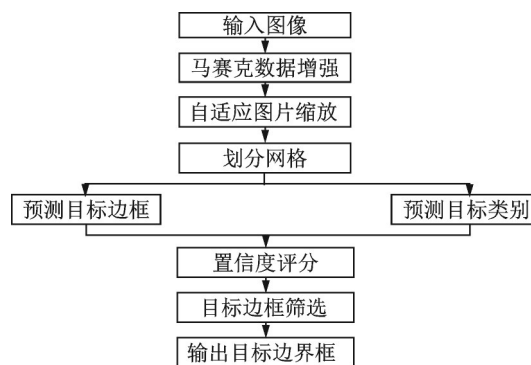


图3 YOLOv5目标检测流程图

Fig.3 Flow chart of YOLOv5 target detection

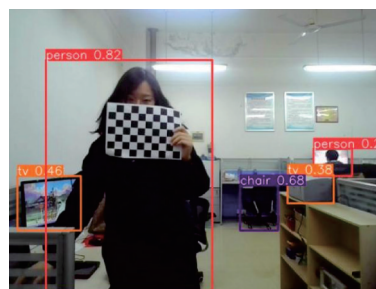


图4 目标识别图

Fig.4 Target recognition diagram

### 2.2 目标跟踪算法设计

本系统采用的是一种判别式目标跟踪算法,核相关滤波法,将目标跟踪问题转化为一个分类问题,即将前景目标和背景进行分类,找到物体当前帧的位置,预测物体在下一帧的位置。

#### 2.2.1 KCF算法原理

对目标实现方位跟踪,正样本通过使用循环矩阵获取大量的负样本,分别计算各样本的HOG特征,得到训练样本数据。接着对滤波器训练,通过岭回归,选择高斯核函数,计算训练样本权值 $\alpha$ ,用相关滤波器对样本进行训练。根据滤波器的响应值大小判别目标当前帧的位置,得到的最大响应位置即为目标最佳预测位置,就是目标所在方位<sup>[16]</sup>。

KCF算法滤波器的训练过程就是求解岭回归问题,通过岭回归的方式建立目标函数:

$$f(x_i) = w^T x_i \quad (1)$$

岭回归方程式由最小二乘以及正则项组成,如下:

$$\min_w \sum_i (f(x_i) - y_i)^2 + \lambda \|w\|^2 \quad (2)$$

$(f(x_i) - y_i)^2$ 为损失函数, $\lambda$ 为正则化参数, $\lambda \|w\|^2$ 为正则项,是为了排除因循环矩阵变换后而变形过度的虚拟样本。

对式(2)关于 $w$ 求导,可得闭式解:

$$w = (X^T X + \lambda I)^{-1} X^T y \quad (3)$$

利用循环矩阵对角化的性质得到  $w$  在傅里叶域表示:

$$w = F^{-1} \left( \frac{x\hat{y}}{\hat{x}\hat{x}^* + \lambda} \right) \quad (4)$$

引入高斯核函数上式求解得到:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda} \quad (5)$$

由此可用此滤波器来对输入图像进行相关性检测得到跟踪结果,样本响应为:

$$\hat{f}(z) = \hat{k}^{xz} \hat{\alpha} \quad (6)$$

样本矩阵为  $z$ , 响应值  $\hat{f}(z)$  最大位置为相关性最强的位置,也是预测的目标位置<sup>[17]</sup>。

### 2.2.2 目标跟踪

在整个跟踪测试中,选用室内场景,跟踪目标人物,目标运动时被跟踪模块检测到,然后就会对图像中的目标人物进行跟踪。效果图如图5。



图5 目标跟踪效果  
Fig.5 Target tracking effect

## 3 目标深度测量(双目视觉架构设计)

利用双目摄像机获取深度信息来调整姿态和位置。通过距离控制系统前进速度,将测得的深度送入控制器中,控制电机速度,使目标和系统保持在一定的距离。

双目立体视觉测距系统的设备包括黑白棋盘,双目摄像头,笔记本电脑。双目测距原理图如图6,根据三角测距原理,假设空间中存在一点  $P(X, Y, Z)$ ,

以左相机光心为中心作为世界坐标系的坐标原点,  $f$  为相机焦距,基线  $b$  为两相机光心之间的距离,推出深度  $Z$  和视差之间的关系为<sup>[18]</sup>:

$$Z = \frac{b \times f}{X_L - X_R} \quad (7)$$

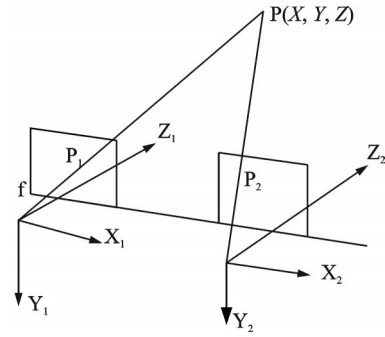


图6 双目视觉原理图  
Fig.6 Schematic of binocular vision

$X_L$  和  $X_R$  是两个成像点在左右两个像面上与左边缘的距离,  $(X_L - X_R)$  为双目视差,  $b$  和  $f$  确定,视差和深度成反比,求出视差就得到物体的深度  $Z$ 。

### 3.1 双目相机的标定

对相机的标定采用张正友标定法,两个摄像机在不同方位拍摄标定板,改变标定板在图像中的位置用来获得多组标定图。标定流程如图7。



图7 双目标定流程图  
Fig.7 Flow chart of double target determination

使用双目摄像头手动拍摄27组标定图,在MATLAB工具箱里面实现对摄像头的标定,求出相机的内外参数如表1,立体标定参数如表2。

### 3.2 立体校正

校正反映到图像上是把消除畸变后的两幅图像严格地行对应,使两幅图像的极线恰好在同一水

表1 双目相机内外参数

Tab.1 Internal and external parameters of binocular camera

标定参数	左相机	右相机
内参矩阵	$\begin{bmatrix} 481.917 & 177.78 & 0 & 320.333 & 407.5 \\ 0 & 481.820 & 47.423 & 203.930 & 706.12 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 487.824 & 089.98 & 0 & 303.363 & 434.05 \\ 0 & 487.463 & 760.05 & 226.329 & 101.41 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$
畸变参数	(0.088 545 18 -0.272 952 98 -0.007 838 82 0.010 007 36	(0.017 225 96 0.146 513 55 -0.003 273 82 0.001 365 99
$D_l, D_r$	0.408 977 63)	-0.349 431 61)

表 2 立体标定参数

Tab. 2 Stereo calibration parameters

旋转矩阵	$\begin{bmatrix} 0.994\ 364\ 77 & 0.004\ 801\ 66 & 0.105\ 903\ 99 \\ -0.002\ 858\ 58 & 0.999\ 824\ 93 & -0.018\ 491\ 75 \\ -0.105\ 974\ 24 & 0.018\ 084\ 81 & 0.994\ 204\ 41 \end{bmatrix}$
平移矩阵	(-4.256 104 3 - 0.049 748 71 0.261 941 99)

平线上,这样一幅图像上任意一点与其在另一幅图像上的对应点就必然具有相同的行号,只需在该行进行一维搜索即可匹配到对应点。在室内的环境下进行实验,校正后的结果如图 8。

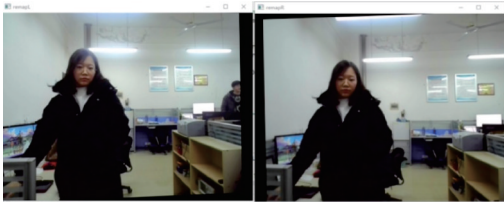


图 8 立体校正图

Fig. 8 Stereoscopic correction diagram

### 3.3 立体视觉匹配

SGBM (semi-global-block matching) 是一种用于计算双目视觉中视差的半全局匹配算法。通过双目相机获取目标图像,左相机获取的图像经过水平 sobel 算子处理之后用于计算 BT (Birchfield and Tomasi) 代价,右图直接用于计算 BT 代价,然后将代价组合,送入代价模块,进行代价聚合,通过视差计算得到视差图,利用 WLS (weighted least squares) 加权最小二乘法对图像进行平滑去噪,最后进行后处理得到最优视差图<sup>[7]</sup>。算法流程如图 9。

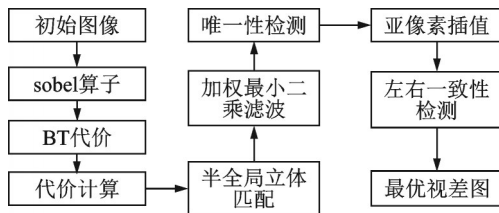


图 9 立体匹配算法流程图

Fig. 9 Flow chart of stereo matching algorithm

利用 rectify() 函数对图像进行校正,实现图像的共面,消除畸变,remap() 对图像进行裁剪,通过 SGBM 算法对矫正后的左右视图进行立体匹配,得到初始视差图,视差图结果如图 10。

使用 ReprojectImageTo3D() 函数将双目匹配获得的像素坐标转换为世界坐标系中的 3D 坐标,通过鼠标建立点击事件,测量结果如表 3,得到目标三维坐标以及目标与摄像头的测量距离。



图 10 双目视差图

Fig.10 Diagram of binocular parallax

表 3 目标测量结果

Tab.3 Target measurement results

点	像素坐标	世界坐标/mm	距离/m
1	(212, 230)	(-60.911, 8.600, 280.234)	0.287
2	(79, 406)	(-132.213, 105.882, 268.84)	0.318
3	(106, 412)	(-89.012, 82.919, 204.12)	0.238
4	(107, 436)	(-100.362, 105.388, 231.24)	0.273

## 4 目标跟踪控制

将上述软件处理模块得到的方位和距离信息送入微控制器中,利用微控制器实现目标跟踪控制。微控制器通过输入信息控制电机舵机对目标进行跟踪。当目标移动时,将上述模块得到的信息作为输入送入 STM32 控制器,控制器控制电机舵机控制模块和测速模块,然后产生控制信号 PWM 波控制电机前后舵机左右移动,减速电机控制速度,霍尔编码器测速,测量的速度通过串口回传到控制器,继续调节机器人速度,速度调节到一定值使目标与机器人保持一定的距离,与目标保持在同一距离同一速度上,实现对目标稳定的跟踪控制<sup>[20]</sup>。

目标移动时,利用舵机改变机器人的前进方向,控制器产生驱动信号,当产生转角差值时,驱动电路是电机转动,达到角度阈值,则电机停转,实现方位跟踪。方位跟踪控制流程如图 11。

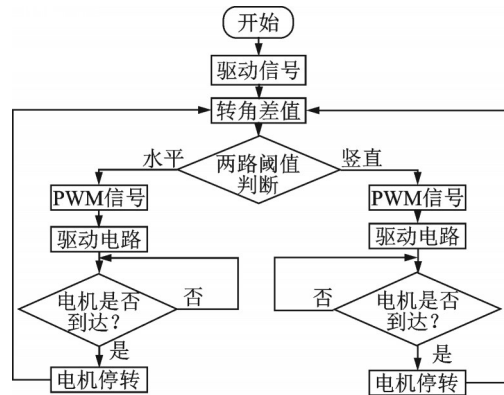


图 11 方位跟踪控制

Fig.11 Azimuth tracking control

目标移动时,利用电机驱动机器人前进,用PID控制算法驱动直流电机。通过两个闭环PID控制机器人移动距离与速度目标与机器人的距离发生变化,这时通过一个PID控制距离,将测得的

实际速度与目标速度比较,通过另一个PD算法调节速度,减速电机控制速度,将期望距离与实际距离做比较,形成一个闭环控制<sup>[8]</sup>。控制回路设计如图12。

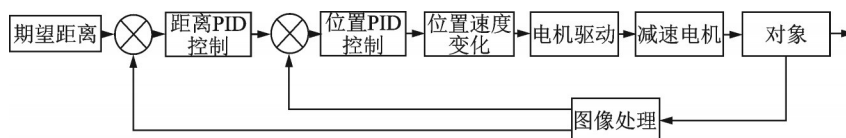


图12 距离跟踪控制

Fig.12 Schematic of distance tracking control

## 5 结束语

本系统利用YOLOv5s对目标进行检测,获取目标图像,使用KCF算法进行目标跟踪,控制摄像机左右移动使目标保持在图像的中央。利用双目视觉搭配SGBM算法,测出相机与目标的距离,控制相机移动速度,让相机与目标保持在一定的距离,保持目标锁定。整体结合起来,系统能够对目标人物保持快速、实时、稳定的跟踪,很好的代替记者进入危险场景实施对现场的实时报道,具有广泛的应用前景。

### 参 考 文 献

[1] 付 朋,陈新度,吴 磊.基于角点的双目视觉绝对定位研究[J].电子测量与仪器学报,2018,32(3):1-8.  
 [2] 尹宏鹏,陈 波,柴 毅,等.基于视觉的目标检测与跟踪综述[J].自动化学报,2016,42(10):1466-1489.  
 [3] Ren Shaoqing, He Kaiming, Girshick Ross, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J].IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39 (6): 1137-1149.  
 [4] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multi box detector [J].Lecture Notes in Computer Science, 2016: 21-37.  
 [5] 王婉婷,姜国龙,褚云飞,等.从RCNN到YOLO系列的物体检测系统综述[J].齐鲁工业大学学报,2021,35(5):9-16.  
 [6] Mac Huang J B, Yang X, et al. Hierarchical convolutional features for visual tracking[C].2015 IEEE International Conference on Computer Vision, Chile, 2015:2570-2574.

[7] Nam H, Han B. Learning multi- domain convolution-neural networks for visual tracking [C].2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA United States, 2016:4293-4302.  
 [8] Bertinetto L, Vlmadre J, Henrioues J F, et al. Fully-convolutional siamese networks for object tracking[C]. European Conference on Computer Vision, Amsterdam USA, 2016:850-865.  
 [9] 黄 刚.双目视觉系统实现羽毛球实时跟踪[J].电子测量与仪器学报,2021,35(6):117-123.  
 [10] 李 艳,唐达明,周莹亮,等.基于计算机视觉的轮椅跟随系统设计[J].计算机工程与应用,2021,57(10):163-172.  
 [11] 张 岚.智能小车目标识别跟踪系统的设计[J].中国新技术新产品,2015,22:20-22.  
 [12] 张乘龙,夏筱筠,柏 松,等.基于KCF跟踪算法的目标轨迹记录系统[J].计算机系统应用,2017,26(5):113-118.  
 [13] 马琳琳,马建新,韩佳芳,等.基于YOLOv5s目标检测算法的研究[J].电脑知识与技术,2021,17(23):100-103.  
 [14] Shang Yuanhang, Wu Gang. Vehicle identification algorithm based on UAV lightweight YOLOV5 snow background[J]. International Core Journal of Engineering, 2021, 7(12): 353-357.  
 [15] 李一鸣,王 潇.基于YOLOv5s模型的轧钢表面缺陷检测[J].制造业自动化,2021,43(11):117-119.  
 [16] 柳天宇,王克强.基于KCF算法的空中目标跟踪模拟[J].激光与红外,2021,51(10):1396-1400.  
 [17] 张 继.基于YOLOv4与KCF的无人机飞行障碍精准识别与实时跟踪研究[J].现代计算机,2021,15:35-41.  
 [18] 卢艳军,王诗宇,张太宁,等.针对目标跟踪过程的双目视觉测距方法研究[J].电光与控制,2020,27(1):60-63,101.  
 [19] 赵晨园,李文新,张庆熙.双目视觉的立体匹配算法研究进展[J].计算机科学与探索,2020,14(7):1104-1113.  
 [20] 竺春祥,鹿存跃.一种基于STM32的目标实时跟踪系统研究[J].半导体光电,2017,38(6):908-910.