

同行专家业内评价意见书编号: 20240854198

附件1

浙江工程师学院（浙江大学工程师学院） 同行专家业内评价意见书

姓名: _____ 杨玉林

学号: _____ 22160076

申报工程师职称专业类别（领域）: _____ 电子信息

浙江工程师学院（浙江大学工程师学院）制

2024年03月22日

一、个人申报

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

1. 对本专业基础理论知识和专业技术知识掌握情况

作为工程师学院机器人与智能制造项目毕业生，我对机器人的基础理论知识和专业技术知识有着扎实的掌握。

1) 基础理论知识

在机器人与智能制造领域的基础理论知识方面，我具备以下能力：

机器人学：熟悉机器人的运动学、动力学和控制理论，能够进行机器人系统的建模、仿真和控制。

人工智能：了解人工智能的基本原理，包括机器学习、深度学习和图像处理等领域，具备人工智能技术在机器人领域的应用能力。

自动化控制：具备自动化系统建模、控制算法设计和实时控制的理论基础，能够应用于智能制造系统中。

2) 专业技术知识

在机器人与智能制造领域的专业技术知识方面，我有着丰富的实践经验和深入的研究，具备以下能力：

机器人视觉：熟悉机器人视觉系统的原理和应用，能够进行视觉传感器的选择、标定和视觉导航算法的设计。

智能制造系统：了解智能制造系统的架构和关键技术，包括工业物联网、大数据分析和智能制造平台的建设与应用。

人机协作：具备人机协作机器人系统的设计与开发能力，包括安全性设计、任务规划和协作控制。

综上所述，我在机器人与智能制造领域的基础理论知识和专业技术知识方面有着全面的掌握，并且在实际工作中有着丰富的应用经验。我期待通过此次申请，为企业的智能制造和机器人技术发展贡献我的专业能力

2. 工程实践的经历

本人于2022年8月1日至2023年8月1日在杭州卫丰机器人公司进行专业实践，完成项目“面向机器人抓取的位姿识别与视觉伺服”，相关成果在第五届中国高校智能机器人创意大赛或二等奖，并发表EI会议论文一篇。

3. 在实际工作中综合运用所学知识解决复杂工程问题的案例

1) 案例简介

案例名称：面向机器人抓取的位姿识别与视觉伺服

主要研究目标：在装配场景中，机械臂的抓取目标常常是多种类、多实例且通常位于复杂纹理背景，这对视觉伺服的特征提取器提出了更高要求。利用目标检测神经网络实现抓取目标的识别、特征提取可以有效提升视觉伺服系统处理多种类、多实例目标的性能，提升应对复杂纹理背景的鲁棒性。本项目通过改进目标检测神经网络YOLOv3，将其应用于基于图像的视觉伺服中，完成多种类多实例目标物体的识别、定位、抓取、装配。

2) 技术难点：

模型训练。改进后的YOLOv3网络不仅要完成分类任务，还需要完成多种数据的回归拟合任务，训练难度增大。

数据获取。训练数据需要包含特征点位置、相机姿态角度等难以人工进行标注的数据。

3) 研究内容：

与传统方法相比，将深度学习引入视觉伺服方法能够解决目标识别以及特征提取鲁棒性差的问题。根据前文的研究现状，基于深度学习的视觉伺服方法主要有基于深度学习加后处理的

方式和端到端方式。前者的后处理步骤仍然依赖于传统机器视觉方式，仍有一定程度的鲁棒性问题，后者将特征提取与控制器完全黑箱化，不利于面对多目标的视觉伺服控制。为了解决这些缺陷，本文将深度学习用于视觉伺服的目标检测和特征提取部分，完成复杂背景环境下的多目标视觉伺服任务。

4) 技术路线:

在目标检测领域，YOLO系列算法以其快速准确的特点被广泛研究。近年来，相关研究将YOLO系列算法应用于视觉伺服中，一些研究仅仅将YOLO网络作为分离复杂背景的手段，分离出目标物体后仍需利用图像处理算法提取特征用于伺服。本项目在YOLOv3的基础上改进其输出层，增加特征点像素坐标、姿态旋转角度、特征点深度的回归预测，使得网络直接输出能用于视觉伺服的相关特征，而无需鲁棒性不高的后处理步骤。

5) 取得的成效:

这个项目通过将深度学习引入视觉伺服方法，解决了传统方法中出现的目标识别和特征提取鲁棒性差的问题。具体而言，本案例在YOLOv3的基础上改进了其输出层，增加了特征点像素坐标、姿态旋转角度、特征点深度的回归预测，使网络直接输出能用于视觉伺服的相关特征，消除了传统视觉处理中鲁棒性不高的后处理步骤，提高了系统的稳定性和准确性。

该项目取得了以下经济和社会效益:

i) 提升产业和企业竞争力

视觉伺服技术广泛应用于机器人、自动化生产线、医疗器械等领域，实现工业现场的智能化和自动化。通过将深度学习应用于视觉伺服技术中，可以提高生产线的自动化程度和效率，减少生产成本。对于企业而言，这意味着更快、更低成本的生产 and 更高的利润。

ii). 提高工作效率

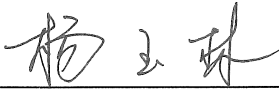
传统视觉伺服技术需要使用多种算法进行目标检测和特征提取，操作繁琐、效率低下。而该项目所提出的基于深度学习的视觉伺服方法能够直接输出能用于视觉伺服的相关特征，省去了复杂的处理步骤，大大提高了工作效率。

iii). 促进视觉伺服技术发展

传统视觉伺服技术的鲁棒性和准确性存在一定局限，无法处理多目标视觉伺服控制等复杂问题。而该项目基于深度学习的视觉伺服方法解决了这些问题，对于推动视觉伺服技术的发展起到了重要的作用。

(二) 取得的业绩（代表作）【限填3项，须提交证明原件（包括发表的论文、出版的著作、专利证书、获奖证书、科技项目立项文件或合同、企业证明等）供核实，并提供复印件一份】					
1. 公开成果代表作【论文发表、专利成果、软件著作权、标准规范与行业工法制定、著作编写、科技成果获奖、学位论文等】					
成果名称	成果类别 [含论文、授权专利（含发明专利申请）、软件著作权、标准、工法、著作、获奖、学位论文等]	发表时间/授权或申请时间等	刊物名称/专利授权或申请号等	本人排名/总人数	备注
Multi-Object Robot Visual Servo Based on YOLOv3	会议论文	2023年05月13日	Proceedings of 2023 IEEE 12th Data Driven Control and Learning Systems Conference	1/2	EI会议收录
一种基于改进秃鹰搜索算法的串联机械臂运动学标定方法	发明专利申请	2023年03月16日	申请号：202310255059.4	2/2	
中国高校智能机器人创意大赛二等奖	获奖	2022年08月28日	二等奖	1/4	

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

(三) 在校期间课程、专业实践训练及学位论文相关情况	
课程成绩情况	按课程学分核算的平均成绩： 87 分
专业实践训练时间及考核情况(具有三年及以上工作经历的不作要求)	累计时间： 1 年(要求1年及以上) 考核成绩： 86 分(要求80分及以上)
本人承诺	
<p>个人声明：本人上述所填资料均为真实有效，如有虚假，愿承担一切责任，特此声明！</p> <p style="text-align: right;">申报人签名： </p>	

浙江工业大学研究生学院

攻读硕士学位研究生成绩表

学号: 22160076	姓名: 杨玉林	性别: 男	学院: 工程师学院	专业: 电子信息	学制: 2.5年						
毕业时最低应获: 24.0学分	已获得: 24.0学分	入学年月: 2021-09			毕业年月: 2024-03						
学位证书号: 1033532024602156	毕业证书号: 103351202402600382		授予学位: 电子信息硕士								
学习时间	课程名称	备注	学分	成绩	课程性质	学习时间	课程名称	备注	学分	成绩	课程性质
2021-2022学年秋季学期	数值计算方法		2.0	94	专业选修课	2021-2022学年春季学期	自然辩证法概论		1.0	79	公共学位课
2021-2022学年秋季学期	中国特色社会主义理论与实践研究		2.0	87	公共学位课	2021-2022学年夏季学期	研究生英语基础技能		1.0	免修	公共学位课
2021-2022学年冬季学期	标准与知识产权		2.0	96	专业选修课	2021-2022学年夏季学期	研究生英语		2.0	免修	公共学位课
2021-2022学年冬季学期	工程伦理		2.0	90	公共学位课	2021-2022学年夏季学期	机器人智能控制		3.0	97	专业学位课
2021-2022学年秋季学期	研究生论文写作指导		1.0	93	专业学位课	2021-2022学年夏季学期	自动化前沿		2.0	85	跨专业课
2021-2022学年冬季学期	智能工业机器人		2.0	90	专业学位课	2021-2022学年春夏季学期	工程技术发展前沿		2.0	85	专业学位课
2021-2022学年春季学期	人工智能制造技术		2.0	88	专业学位课						

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

2. 备注中“*”表示重修课程。

学院成绩校核章: (60)

成绩校核人: 张梦依

打印日期: 2024-04-02

Multi-Object Robot Visual Servo Based on YOLOv3

Publisher: IEEE

[Cite This](#)[PDF](#)Yulin Yang; Shan Liu [All Authors](#)

82

Full

[Text Views](#)

Abstract

Document Sections

- 1 Introduction
- 2 Related Works
- 3 Method
- 4 Experiments and Results
- 5 Conclusions

Authors

Figures

References

Keywords

Metrics

Abstract:

Aiming at the low robustness of image feature extractor in Image-Based Visual Servo (IBVS), a robot visual servo method based on object detection neural network YOLOv3 is proposed. By improving the output layer of YOLOv3 and adding attitude angle of camera, the pixel coordinate and depth information of feature points, the robustness of the IBVS system based on point features is improved while it can cope with multi-type and multi-instance objects, and the problem of the image Jacoby matrix falling into singularity caused by excessive rotation angle error of the optical axis is avoided. The visual servo convergence is accelerated. Meanwhile, the network training data generation algorithm of the desired image is used to replace the traditional manual data annotation, which reduces the cost of data acquisition, and the data enhancement method ensures the generalization performance of the training model.

Published in: 2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS)

Date of Conference: 12-14 May 2023

DOI: 10.1109/DDCLS58216.2023.10166105

Date Added to IEEE Xplore: 07 July 2023

Publisher: IEEE

► ISBN Information:

Conference Location: Xiangtan, China

► ISSN Information:

► Funding Agency:

SECTION 1 Introduction

In the traditional industrial robot control, manual instruction is usually used to grasp a single object with known posture. Due to the lack of sensing ability to grasp the object, industrial robots are prone to interference from external uncertain factors in the process of teaching grasp, such as small changes in the pose of the object. In recent years, with the continuous development of computer vision algorithm and vision sensor hardware, it has become a feasible way to combine vision sensor with industrial robot, and realize multi-object grasping by combining machine vision technology and industrial robot control technology.

There are two ways for visual information to be introduced into the robot control system, "look-then-move" [1], [2] and "look-and-move" [2]. The former does not directly apply the visual information to the robot control closed-loop, but first uses the visual information to get the object position, and then controls the robot to move to the corresponding position. The latter directly applies the visual information to the closed-loop control of the robot, making the robot control more sensitive and rapid. In this way, the visual servo uses the visual information to control the end of the robot arm to reach the target pose. Compared with the "look-then-move" mode, the visual servo technology does not depend on the calibration accuracy of internal and external parameters of the camera, and has a certain robustness to the disturbance of the visual sensor.

Robot visual servo is the cross fusion of robot control and image processing. Hutchinson et al. [3]–[5] laid the theoretical foundation of visual servo, and relevant articles [6]–[9] also made a comprehensive



Multi-Object Robot Visual Servo Based on YOLOv3

Yulin Yang¹, Shan Liu²

1. Polytechnic Institute, Zhejiang University, Hangzhou 310015, China
E-mail: yulinyang@zju.edu.cn

2. State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering
Zhejiang University, Hangzhou 310027, China
E-mail: sliu@zju.edu.cn

Abstract: Aiming at the low robustness of image feature extractor in Image-Based Visual Servo (IBVS), a robot visual servo method based on object detection neural network YOLOv3 is proposed. By improving the output layer of YOLOv3 and adding attitude angle of camera, the pixel coordinate and depth information of feature points, the robustness of the IBVS system based on point features is improved while it can cope with multi-type and multi-instance objects, and the problem of the image Jacoby matrix falling into singularity caused by excessive rotation angle error of the optical axis is avoided. The visual servo convergence is accelerated. Meanwhile, the network training data generation algorithm of the desired image is used to replace the traditional manual data annotation, which reduces the cost of data acquisition, and the data enhancement method ensures the generalization performance of the training model.

Key Words: Visual Servo; Object Detection; Manipulators

1 Introduction

In the traditional industrial robot control, manual instruction is usually used to grasp a single object with known posture. Due to the lack of sensing ability to grasp the object, industrial robots are prone to interference from external uncertain factors in the process of teaching grasp, such as small changes in the pose of the object. In recent years, with the continuous development of computer vision algorithm and vision sensor hardware, it has become a feasible way to combine vision sensor with industrial robot, and realize multi-object grasping by combining machine vision technology and industrial robot control technology.

There are two ways for visual information to be introduced into the robot control system, "look-then-move" [1, 2] and "look-and-move" [2]. The former does not directly apply the visual information to the robot control closed-loop, but first uses the visual information to get the object position, and then controls the robot to move to the corresponding position. The latter directly applies the visual information to the closed-loop control of the robot, making the robot control more sensitive and rapid. In this way, the visual servo uses the visual information to control the end of the robot arm to reach the target pose. Compared with the "look-then-move" mode, the visual servo technology does not depend on the calibration accuracy of internal and external parameters of the camera, and has a certain robustness to the disturbance of the visual sensor.

Robot visual servo is the cross fusion of robot control and image processing. Hutchinson et al. [3-5] laid the theoretical foundation of visual servo, and relevant articles [6-9] also made a comprehensive classification and summary of the robot visual servo system. Depending on the errors of the visual features applied to the control loop, Visual Servo can be divided into Position-Based Visual Servo (PBVS) and

Image-Based Visual Servo (IBVS). The PBVS method obtains the Cartesian spatial pose of the object through the visual information and controls the robot to approach the object by the pose error. However, the IBVS method does not need to estimate the pose of the object from 2D image to 3D space, but directly controls the robot to approach the object with the feature error in the 2D image.

2 Related Works

2.1 Traditional Method of IBVS

Yang et al. [10] design a visual servo simulation platform by taking the object's four corners as the feature points whose initial position need be manually marked. Dong et al. [11] implemented a kinematic incremental visual servo system whose background of the scene is covered with white paper, and the feature points are marked on the object to reduce the impact of noise on feature extraction. Wang et al. [12] proposed a method to satisfied the visibility constraints caused by the camera's field of view (FOV) limitation. In the experiment, they also marked the object with dots of different colors to facilitate feature extraction. Zhu et al. [13] use ORB algorithm to accelerate feature extraction and matching, and used PROSAC algorithm to eliminate the wrong matching points, and realize visual servo for multi-instance of the same kind of object.

The advantage of these methods is to describe the whole image with a few features, which greatly simplifies the system model and controller design. However, the traditional machine vision method that feature extraction relies on is easy to be affected by the environment. It can only be realized in simulation or simple environment, and cannot be applied in complex industrial environment.

2.2 IBVS Based on Deep Learning and Post Processing

Based on mask R-CNN, Li et al. [14] firstly segment the object and the background, and then obtain the point features by using the traditional image processing algorithm. In order

*This work is supported by the Primary Research and Development Plan of Zhejiang Province, China under Grant 2020C01114.

经检索“Engineering Village”，下述论文被《Ei Compendex》收录。（检索时间：2023年10月19日）。

<RECORD 1>

Accession number:20233114457224

Title:Multi-Object Robot Visual Servo Based on YOLOv3

Authors:Yang, Yulin (1); Liu, Shan (2)

Author affiliation:(1) Polytechnic Institute, Zhejiang University, Hangzhou; 310015, China; (2) College of Control Science and Engineering, Zhejiang University, State Key Laboratory of Industrial Control Technology, Hangzhou; 310027, China

Source title:Proceedings of 2023 IEEE 12th Data Driven Control and Learning Systems Conference, DDCLS 2023

Abbreviated source title:Proc. IEEE Data Driven Control Learn. Syst. Conf., DDCLS

Part number:1 of 1

Issue title:Proceedings of 2023 IEEE 12th Data Driven Control and Learning Systems Conference, DDCLS 2023

Issue date:2023

Publication year:2023

Pages:893-898

Language:English

ISBN-13:9798350321050

Document type:Conference article (CA)

Conference name:12th IEEE Data Driven Control and Learning Systems Conference, DDCLS 2023

Conference date:May 12, 2023 - May 14, 2023

Conference location:Xiangtan, China

Conference code:190501

Sponsor:Beijing Information Science and Technology University; CAA; et al.; IEEE Beijing Section; Qingdao University; Technical Committee on Data Driven Control, Learning and Optimization (DDCLO)

Publisher:Institute of Electrical and Electronics Engineers Inc.

Number of references:23

Main heading:Manipulators

Controlled terms:Data acquisition - Image enhancement - Object detection - Object recognition - Visual servoing

Uncontrolled terms:Attitude angle - Depth information - Feature extractor - Image features - Image-based visual servo - Multiobject - Neural-networks - Objects detection - Output layer - Visual servo

Classification code:723.2 Data Processing and Image Processing

DOI:10.1109/DDCLS58216.2023.10166105

Funding details: Number: 2020C01114, Acronym: -, Sponsor: Primary Research and Development Plan of Zhejiang Province;

Funding text:*This work is supported by the Primary Research and Development Plan of Zhejiang Province, China under Grant 2020C01114.

Database:Compendex

Compilation and indexing terms, Copyright 2023 Elsevier Inc.

注:

1. 以上检索结果来自 CALIS 查收查引系统。
2. 以上检索结果均得到委托人及被检索作者的确认。





(12) 发明专利申请

(10) 申请公布号 CN 116277000 A

(43) 申请公布日 2023. 06. 23

(21) 申请号 202310255059.4

(22) 申请日 2023.03.16

(71) 申请人 浙江大学

地址 310058 浙江省杭州市西湖区余杭塘路866号

(72) 发明人 刘山 杨玉林

(74) 专利代理机构 杭州求是专利事务有限公司 33200

专利代理师 刘静

(51) Int. Cl.

B25J 9/16 (2006.01)

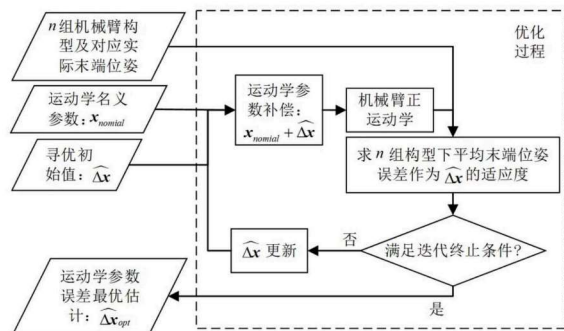
权利要求书2页 说明书9页 附图5页

(54) 发明名称

一种基于改进秃鹰搜索算法的串联机械臂运动学标定方法

(57) 摘要

本发明公开了一种基于改进秃鹰搜索算法的串联机械臂运动学标定方法。该方法由串联机械臂末端位置、姿态数据测量和优化求解运动学参数误差两部分组成,位姿测量数据输入优化算法,优化算法输出运动学参数误差估计值。该方法改进秃鹰搜索优化算法,针对机械臂运动学模型参数名义值与实际值存在误差而导致机械臂末端位姿存在误差的情况,进行了运动学标定,求解了运动学参数误差的最优估计并代入正向运动学补偿,提升串联机械臂末端绝对位姿精度。实例表明,在运动学标定问题中,改进的变参数秃鹰搜索算法相较于其他已经应用于该领域的搜索优化算法具有更强的全局寻优能力,更好的收敛性能,更高的时间效率。





获奖证书

HONOR CERTIFICATES



证书编号
ZNJQRZX2022010005

学校：浙江大学

学生：杨玉林 王凌枫 李明芮 麻航翔

指导老师：刘山

参赛作品：俄罗斯方块机器人

在2022年第五届中国高校智能机器人创意大赛俄罗斯方块专项赛决赛中荣获

二等奖

特发此证，以资鼓励！

中国高校智能机器人创意大赛组委会



2022年八月二十八日