

基于手势交互的机械臂轨迹示教系统

花傲阳

(上海理工大学机械工程学院, 上海 200082)

✉huaoyang@126.com



摘要: 为了解决传统机械臂示教方式上手难度高、效率低、人机交互系统不友好的问题, 结合机器视觉与神经网络, 提出一种基于手势交互的机械臂轨迹示教系统。该系统通过机器视觉定位指尖位置, 记录下指尖的运动轨迹并将其用于机械臂的轨迹示教; 通过卷积神经网络识别静态手势, 用于与机械臂进行如抓取、放置、开始示教等动作的人机交互。系统通过ROS(机器人操作系统)与机械臂通信, 驱动机械臂完成示教轨迹的复现。通过设计实验对系统的性能进行评估, 结果显示本示教系统具有编程难度低、易用性好、项目部署周期短的优点, 可用于复杂轨迹的示教工作, 极大地提升了机械臂示教工作的效率。

关键词: 机械臂; 轨迹示教; 机器视觉; 神经网络; 人机交互

中图分类号: TP391.4 **文献标识码:** A

Manipulator Trajectory Teaching System based on Gesture Interaction

HUA Aoyang

(College of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200082, China)

✉huaoyang@126.com

Abstract: Aiming at the problems of difficulty to learn, low efficiency and unfriendly human-computer interaction system in traditional manipulator teaching methods, this paper proposes to establish a gesture interaction-based manipulator trajectory teaching system combining machine vision and neural network. Fingertip position is located through machine vision, and the movement trajectory of the fingertip is recorded and used for teaching manipulator trajectory. Convolutional neural network is used to identify static gestures, which are used for Human-computer interaction with the manipulator's actions, such as grasping, placing, and teaching. The system communicates with the manipulator through ROS (Robot Operating System), and drives the manipulator to complete the reproduction of the teaching trajectory. The performance of the system is evaluated through experiment designing. Results show that this teaching system has the advantages of low programming difficulty, user-friendliness, and short project deployment cycle. It can be used for teaching complex trajectories, which greatly improves teaching efficiency of the manipulator.

Keywords: manipulator; trajectory teaching; machine vision; neural network; human-computer interaction

1 引言(Introduction)

近年来, 我国制造业发展经历了从制造大国向制造强国的转型升级。机械臂是现代工业生产管理过程中的重要组成部分, 由于机械臂具有工作精度高、重复能力强、能在恶劣的环境下进行操作等优点, 被广泛应用于零部件的焊接、喷涂、装配等工业任务中。目前, 机械臂主要包含在线示教及离线示教两种示教方式。在线示教主要通过示教盒进行逐点示教或直接拖动机械臂示教; 离线示教主要是用计算机辅

助设计软件制作三维模型, 并导入各机械臂厂家的仿真软件中进行轨迹示教^[1]。传统示教方法具有操作复杂、人机交互的友好性不高、不同品牌操作方式不同、示教精度依赖于工人熟练度等缺点, 因此提出一种简单高效的机械臂示教办法是当前机械臂研究的热点与难点。

随着机器学习和机器视觉技术的发展, 机械臂演绎编程为人机交互提供了新的解决办法, 是降低机械臂示教复杂性的重要途径^[1]。机械臂演绎编程是通过对人的示教动作进行观

测、学习,从而自动生成程序及运动轨迹的过程,具有难度低、人机交互便捷、操作灵活等优点。本文将手势交互与机械臂轨迹示教相结合,开发了一个利用手指进行轨迹示教,基于手势识别进行人机交互的机械臂示教系统。本系统通过相机获取图像,基于卷积神经网络的手势识别模块识别不同的输入手势以实现人机交互。基于机器视觉的指尖识别模块可以准确地识别指尖的位置,将其生成的轨迹用于机械臂的轨迹示教,此方法具有上手简单、人机交互友好、轨迹准确率高等优点,在示教复杂轨迹时能极大地提高工作效率。

2 示教系统的设计(Design of teaching system)

根据机械臂在日常生产中的应用场景,本示教系统由指尖识别、手势识别两个部分组成。首先通过相机获取图像数据,接着对图像进行滤波降噪处理。图像处理模块用于识别、跟踪指尖的轨迹,手势识别主要用于与机械臂进行人机交互。最后根据指尖的路径、示教过程中的手势以及机械臂手眼标定后得到的各项参数生成机械臂的运动轨迹。系统的流程图如图1所示。

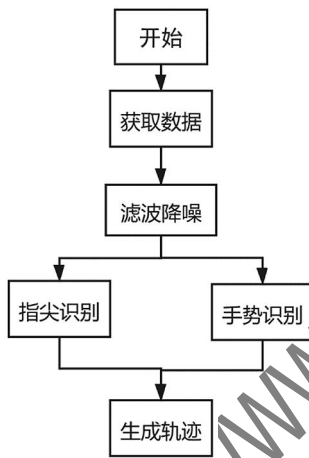


图1 系统流程图

Fig.1 System flow chart

3 运动指尖识别(Movement fingertip recognition)

本系统是从视频中获取动态的指尖点轨迹,因此需要将移动的手指作为前景从静态的背景中分割出来,常见的运动物体的分割方法有帧间差分法以及背景消除法。由于帧间差分法不能提取运动物体的完整区域,存在较大的空洞,移动速度的快慢也会影响运动物体的提取质量,而背景消除法对运动速度不敏感且稳定性高,因此更加适合运动物体的准确提取。

3.1 背景消除

背景消除法是把当前图像与背景图像作差分,通过设置差分后图像的阈值,将视频中运动的物体作为前景提取出来^[2]。背景消除的流程如图2所示。



图2 背景分割的流程

Fig.2 Background segmentation process

若 $f_t(x,y)$ 为 t 时刻的一帧图像, $f(x,y)$ 为背景模型, $D_X(x,y)$ 为背景消除后的图像, T 为固定阈值,那么在 t 时刻背景分割的公式为:

$$D_X(x,y) = \begin{cases} 1 & |f_t(x,y) - f(x,y)| > T \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

运动物体的提取效果十分依赖背景模型的建立,如何获取一个高质量的背景模型,是背景消除法的关键和难点。常见的背景消除法有基于混合高斯模型背景的减除法(GMM)、基于邻近算法(KNN)的背景分割算法、平均背景法等。由于平均背景法具有工作原理简单、可靠性高等优点,故本系统采用平均背景法进行背景消除。平均背景法是通过累加足够帧数的图片,计算每个像素的平均值及其平均差得到背景的统计模型,通过设置合理的高低阈值,来分割图像的前景与背景。实验中原图像与分割后得到的前景图像如图3所示。



(a)背景分割前的手指

(b)背景分割后的手指

图3 背景分割前后的图像

Fig.3 Image before and after background segmentation

3.2 指尖检测

从视频中获取前景之后,由于前景图片存在一定的噪声,因此首先需要对其进行开运算、闭运算、滤波等操作以减少噪声,随后对指尖部位进行定位。目前常用的手部定位提取方法有基于肤色分割的手部提取、基于深度相机的深度阈值提取。本系统利用背景差分所得的前景图像,基于轮廓查找及凸包算法对指尖定位。其主要算法流程如图4所示。

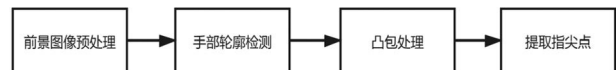


图4 指尖检测流程

Fig.4 Fingertip inspection process

3.2.1 图像的预处理

为了抑制图像中的噪声,本系统采用中值滤波、形态学算法来处理图像。中值滤波是将像素替换为围绕这个像素的矩形框内的中值像素,常用于去除图像中的椒盐噪声。图像形态学算法主要使用开运算、闭运算,开运算是先对图像进行

行腐蚀然后膨胀,主要用于消除图像中小的噪声;闭运算是将特征区域先进行膨胀然后腐蚀处理,可以填充特征区域内部的细小空洞。

3.2.2 基于手部轮廓检测与凸包算法的指尖提取

在基于中值滤波及形态学处理之后,下一步就是提取手部轮廓,进而利用凸包算法找到指尖点。本系统利用OpenCV(计算机视觉公开库)中的findContours函数进行轮廓查找,通过提取最长的轮廓线来确定手掌区域;接着用凸包算法将找到的轮廓用点连接起来构成一个凸多边形;最后通过寻找凸多边形各个点中纵向像素值的最小值来确定指尖点的位置。手的轮廓图与手部的凸多边形如图5所示。

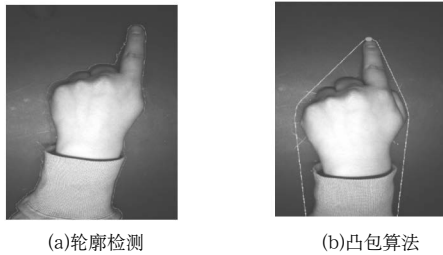


图5 轮廓检测与凸包算法

Fig.5 Contour detection and convex hull algorithm

4 基于神经网络的手势识别(Gesture recognition based on neural network)

人机交互功能是决定一个系统“友善”程度的重要指标,常见的人机交互方式包括鼠标键盘输入、触屏点击、语音识别等。考虑到在机械臂示教过程中使用鼠标键盘等设备进行人机交互所带来的不便,本系统开发了一套基于机器视觉和卷积神经网络的手势识别模块作为人机交互的接口。通过手势可以控制机械臂示教的开始与结束,以及物体的抓取与放置功能,提升了本系统的人机交互能力与操作的便携性。

4.1 手势识别方案设计

本系统的手势识别模块共分为四个步骤:(1)通过相机采集图像,并将手势识别区域中的手势分割出来。(2)通过图像滤波处理去除噪点,向下采样制作训练用的数据集。(3)训练神经网络。(4)采集图像信息。(5)输出手势结果。根据系统在实际工作中的需要,本系统共设计了四个不同的手势分别对应停止示教、开始示教、抓取与放置四个功能。四个手势如图6所示。每个手势经过分割、滤波、下采样的操作以后,得到 32×32 像素的灰度图片。



图6 手势集的设计

Fig.6 Gesture set design

4.2 卷积神经网络的设计与训练

本系统采用LeNet-5网络结构进行手势识别。LeNet-5是一种经典的七层卷积网络结构,用于手写数字识别且取得了令人瞩目的成绩,是最早用于图像识别的卷积神经网络^[3]。该网络由两个卷积层、两个池化层,以及三个全连接层组成。整个网络的结构图如图7所示。

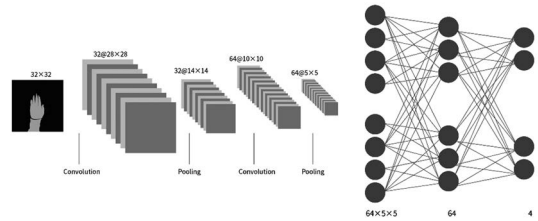


图7 基于LeNet-5的网络结构

Fig.7 Network structure based on LeNet-5

卷积神经网络相对于传统图像处理的手势识别算法具有提取图像特征效率高、分类准确率高、分类速度快等优点,一般包含输入层、卷积层、池化层以及全连接层四个部分。卷积层和池化层是用于提取图像特征的核心模块^[4]。卷积层通过卷积核与所对应的图像区域进行卷积提取图片的特征。本系统的神经网络包含两个卷积层,卷积核的数量分别为32个、64个。由于卷积核的大小对分类的结果具有较大的影响,采用较大的卷积核会导致手部特征丧失,较小的卷积核又会降低网络训练、识别的效率^[4],因此为了保留尽可能多的图形特征同时兼顾网络的运算速度,本系统对两个卷积层分别使用 5×5 、 3×3 大小的卷积核,提高了网络的性能^[2]。

卷积运算提取到的图像特征是线性的,但真正的样本往往是非线性的,为此引入非线性激活函数来解决,使得每个像素点可以用0—1的任何数值来代表,模拟更为细微的变化。激活函数一般具有非线性、连续可微、单调性等特性。本系统卷积层采用ReLU激活函数。ReLU激活函数的表达式为:

$$f(x) = \max(0, x) \quad (2)$$

池化层主要用于压缩图像的大小,提取图像的主要特征,简化网络的计算。池化对输出的上层进行下采样的过程,会减少参数矩阵的尺寸并保留有用的信息,从而减少最后全连层中的参数数量。池化常用的方式一般有两种,分别为最大池化与均值池化,本文采用卷积核大小为 2×2 的最大池化采样,抑制了网络的过拟合,提高了模型的非线性映射能力^[5]。

全连接层居于整个网络的尾部,是卷积神经网络的“分类器”,在卷积层与第一层全连接时,卷积、池化所产生的特征图会被展开为一维的特征向量,后面的全连接层的每个

神经元都与所有上层神经元相连^[6]，将前面层级所提取到的特征进行非线性组合得到输出。全连接层的表达式为：

$$f(x) = w \times x + b \quad (3)$$

式中， x 为输入的值， w 为每个输入所对应的权重系数， b 为偏置项，是一个常数。全连接层连接所有特征输出至输出层，输出层使用归一化指数函数输出分类的结果标签^[2]。

5 实验验证(Experimental verification)

为了对本示教系统性能进行评估，本节设计了传统示教方法与使用本系统示教方法的对比实验。通过记录机械臂的轨迹、示教所需的时间以及使用者在示教过程中的满意程度等，从效率、精度、易用性等三个维度与传统的用示教盒示教作对比，以验证本系统的优越性，同时发现它的不足之处。

5.1 示教系统平台搭建

实验的硬件组成包含一部相机、一台机械臂、一台预装Ubuntu 16.04操作系统以及ROS(机器人操作系统)的笔记本电脑。在使用系统进行示教前需要先完成相机标定、机械臂与相机间的手眼标定以及示教系统与机械臂的通信^[7]。手势示教的流程是首先加载卷积神经网络训练好的手势模型，接着由工业相机实时采集示教的画面，操作者用右手在示教区域进行示教，系统会实时捕获手指指尖的位置。左手在手势识别区域进行人机交互，通过手势来控制机器人完成开始示教、停止示教、物体的抓取与放置等操作。搭建实验平台的设计图如图8所示。

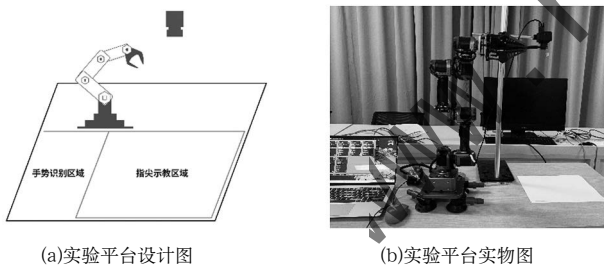


图8 实验平台的设计图与实物图

Fig.8 Design drawing and physical drawing of the experimental platform

5.2 实验设计

本实验的方案为：在实验台上平铺一张A4大小的白纸，在纸上画一个矩形，分别使用传统的示教盒以及本示教系统对同一轨迹进行轨迹示教，两种示教方法如图9所示。示教结束后，机械臂会复现示教的轨迹，与此同时，机械臂末端的笔会在白纸上记录下机械臂的运行轨迹。最后，分别对这两种示教方法进行评估，通过测量在特定点位的位置偏移量来评估示教的精度，通过测量示教所需的时长来评估示教的效率，通过不同操作者的调查问卷对示教系统的易用性进行评价^[8]。

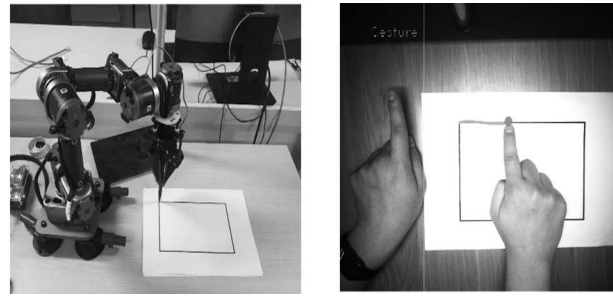


图9 使用不同的方法进行示教

Fig.9 Teaching with different methods

5.3 实验结果分析

本次实验共安排了四名同学分别用两种示教方式进行轨迹示教，在轨迹上取若干个采样点，记录在每个采样点实际轨迹与理想轨迹的平均偏移量、最大偏移量以及每次示教所需的总时长。使用示教盒示教产生的相关数据详情如表1所示。

表1 使用示教盒示教产生的数据

Tab.1 Data generated by teaching with the teaching box

内容	学生1	学生2	学生3	学生4	平均值
总时长/s	23.5	38.2	27.3	43.9	33.2
平均误差/mm	1.7	0.9	0.5	1.1	1.05
最大误差/mm	3.8	2.5	1.7	2.3	2.58
易用性/min	7	3	6	2	4.5

使用本系统示教产生的相关数据详情如表2所示。

表2 使用本系统示教产生的数据

Tab.2 Data generated by using this system to teach

内容	学生1	学生2	学生3	学生4	平均值
总时长/s	10.3	7.9	13.5	20.9	13.5
平均误差/mm	4.2	4.5	2.7	2.2	3.4
最大误差/mm	7.9	6.5	4.4	3.9	5.67
易用性/min	7	7	9	8	7.75

通过对比两种编程方式所用的总时长以及最终实际轨迹与理想轨迹之间的偏差，可以得出以下结论：传统的方法需要在逐点示教之后才能生成一段完整的轨迹，对新手操作者来说易用性不高且操作烦琐。本系通过手势即可与机器人交互，借助滑动手指即可完成示教，具有操作简单、效率高等优点。另一方面，本系统无论在平均误差还是最大误差方面，对比传统方法都是不尽人意的^[9]，除了操作者由于操作产生的随机误差以外，相机标定时所带来的相机标定误差、用手眼标定来确定机械臂与相机位置关系所产生的位置误差、机械臂生产制造以及日常磨损产生的误差等都使得规划的轨迹误差偏大^[4]。因此，后期研究可以采用精度更高的相机标定算法、手眼标定算法及机械臂误差补偿算法来减少系统误差。

6 结论(Conclusion)

使用传统的示教方法进行示教存在费时费力、专业基础要求高、人机交互性能差等问题,这些缺点阻碍了机械臂进一步的发展。为了解决上述问题,本文基于机器视觉及机器学习提出了一种简单高效的机械臂示教系统。该系统利用背景消除法、轮廓检测法以及凸包算法识别并追踪手指的位置,将手指的位置点转换为机械臂的运动轨迹。为了便于人机交互,本系统搭建了基于LeNet-5的视觉手势交互模块,设计了四种交互手势用于与机械臂进行人机交互。该网络具有95.8%的识别准确率,友好的人机交互接口降低了机器人示教的上手难度,提高了示教的效率。最后通过与传统的示教盒示教作比较,验证了本系统在保证轨迹精度的同时具有更高的工作效率以及更简洁的操作逻辑。

参考文献(References)

- [1] 李致豪.基于视觉手势交互的协作机器人示教系统研究[D].长春:长春工业大学,2020.
- [2] 李炳臻,刘克,顾佼佼,等.卷积神经网络研究综述[J].计算机时

(上接第62页)

5.2 社会可行性分析

自2007年以来,我国行业主管部门出台了一系列有关北斗产业的政策文件,如《关于深圳北斗卫星导航系统应用产业化实施方案》《关于在行业推广应用北斗卫星导航系统的指导意见》等,都表明我国正在大力推动北斗产业发展。另一方面,在网络时代下,企业每天接受的物流信息也是联系的、动态的、多样的、复杂的,物流信息系统作为企业信息系统的一部分不可或缺。

5.3 经济可行性分析

Oracle商用费用比较高,有按CPU数和按用户数(Named User Plus)两种购买方式,而商用车载北斗每年也有一定的服务费用。这对于一般的中小型企业是一种前期投入,但是对具有一定规模的企业,适当投入成本更有利于组织协调管理。因此,需要企业根据自身定位,权衡考虑^[9]。

6 结论(Conclusion)

我国BDS性能已达到世界先进水平,在诸多领域广泛应用,服务用户规模达到亿级。当前物流车辆导航系统以GPS/BD双模定位为主,BDS正在逐步取代GPS在我国运输行业中的地位。基于BDS的物流车辆管理信息系统的使用,实现了物流企业的智能化、信息化,实时掌握人、车、货动态,进行数据分析,优化人车匹配,合理调配车辆,进一步降低企业成本。随着5G技术的应用,实现低延迟、低功耗、万物互联,车辆运输各个环节将更加协调高效,企业的经济效益将进一步提升。

参考文献(References)

- [1] WANG C, LIAO T Y, YANG C C. Logistics quality

代,2021(04):8-12.

- [3] 周忠祥,胡晋,王越,等.机器人演示编程技术的研究进展[J].自动化博览,2020,37(06):48-57.
- [4] 张宏达.基于球杆示教笔的机器人轨迹演示编程系统研究[D].哈尔滨:哈尔滨工业大学,2020.
- [5] 段中兴,白杨.结合深度学习的机器人示教系统设计[J].计算机测量与控制,2020,28(11):164-169.
- [6] 誉洪生.面向工业装配演示编程的动作识别[D].杭州:浙江大学,2017.
- [7] 叶必鹏.基于视觉的多机器人室内协同SLAM算法的研究与实现[D].哈尔滨:哈尔滨工业大学,2018.
- [8] HU H S, ABDUI S. Static hand gesture recognition using neural networks[J]. Artificial Intelligence Review, 2014, 41(2):147-181.
- [9] SUN P, FREUND R M. Computation of minimum-volume covering ellipsoids[J]. Operations Research, 2014, 52(5):690-706.

作者简介:

花傲阳(1995-),男,硕士生.研究领域:图像处理.

management practices and performance of international distribution centre operators[J]. International Journal of Shipping and Transport Logistics, 2021, 13(3):300-326.

- [2] CHEN J, ZHAO X, LIU C, et al. Evaluating the latest performance of precise point positioning in Multi-GNSS/RNSS: GPS, GLONASS, BDS, Galileo and QZSS[J]. The Journal of Navigation, 2021, 74(1):247-267.
- [3] 沈玮烨,闫星臣.冷链物流仓储管理信息系统设计[J].软件工程,2020,23(09):43-45.
- [4] 祝伟.基于BD/GPS/GIS的物流车辆管理系统的设计与实现[D].长春:吉林大学,2016.
- [5] 汪吟凡.基于Oracle数据库的ITRF成果数据管理应用研究[D].湘潭:湘潭大学,2020.
- [6] 王婷婷.基于北斗的车载物流定位系统研究[D].西安:西安电子科技大学,2019.
- [7] 陆承佳.基于Python的网络爬虫在物流信息追踪中的应用[J].信息与电脑(理论版),2020,32(12):133-136.
- [8] 魏婷.保险业呼叫中心运营平台的开发与实现[D].北京:北京化工大学,2019.
- [9] 陈越千,郑琰.药品仓库辅助管理信息系统设计[J].软件工程,2020,23(10):49-51.

作者简介:

朱先涛(1998-),男,本科生.研究领域:物流信息系统设计.薛亮(1978-),女,博士,讲师.研究领域:物流系统规划与管理,物流信息系统设计.本文通讯作者.