

基于 Leap Motion 的虚拟课堂手势交互方法^①



胡发丽, 高全力, 王西汉, 李庆敏

(西安工程大学 计算机科学学院, 西安 710048)

通信作者: 胡发丽, E-mail: 17865318790@163.com

摘要: 随着虚拟现实技术的飞速发展, Leap Motion 等体感传感器出现并被广泛地应用在人机交互中. 针对 Leap Motion 体感控制器在识别范围边缘识别率低且识别速度慢的问题提出了一种基于深度神经网络的 Leap Motion 手势交互方法. 该方法在定义的交互手势基础上, 设计了三维交互系统并应用到虚拟场景中. 系统首先通过 Leap Motion 进行数据捕捉, 对获取到的红外图像采用深度神经网络进行特征提取并实现对手势的分类识别, 然后结合 Leap Motion 获取的手部坐标前后帧的变化来判断动态手势, 最终结合动态手势完成虚拟场景中的交互功能. 经过实验验证, 本文手势识别方法无论是在识别速度还是识别精度上都优于 Leap Motion 自带的手势识别方法, 同时在 Leap Motion 识别范围边界处仍能保持较高的识别率.

关键词: 虚拟课堂; 虚拟现实; 手势交互; Leap Motion; 深度神经网络

引用格式: 胡发丽, 高全力, 王西汉, 李庆敏. 基于 Leap Motion 的虚拟课堂手势交互方法. 计算机系统应用, 2022, 31(8): 160-168. <http://www.c-s-a.org.cn/1003-3254/8626.html>

Gesture Interaction Method in Virtual Classroom Based on Leap Motion

HU Fa-Li, GAO Quan-Li, WANG Xi-Han, LI Qing-Min

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract: With the rapid development of virtual reality technology, somatosensory sensors such as Leap Motion appear and are widely used in human-computer interaction. This study proposes a Leap Motion gesture interaction method based on a deep neural network to resolve the problem that the Leap Motion somatosensory controller has a low recognition rate and a slow recognition speed at the edge of its recognition range. In addition to the defined interactive gestures, a three-dimensional interactive system is designed and applied to a virtual scene. Specifically, the system captures data with Leap Motion, uses the deep neural network to extract features from the acquired infrared images, and implements gesture classification and recognition. Then, the changes in the hand coordinates between two adjacent frames acquired by Leap Motion are utilized to determine dynamic gestures. Finally, the interaction function in the virtual scene is fulfilled by investigating the dynamic gestures. Experimental verification shows that the proposed gesture recognition method is superior to the built-in gesture recognition method of Leap Motion in both recognition speed and recognition accuracy. Moreover, it still maintains a high recognition rate at the edge of Leap Motion's recognition range.

Key words: virtual classroom; virtual reality; gesture interaction; Leap Motion; deep neural network

虚拟现实技术自问世以来就受到了人们的普遍关注, 在教育、军事、医疗等方面得到广泛的应用. 疫情以来, 线上教学成为人们关注的主要教学方式, 但线上

教学多以视频语音为主, 缺乏课堂沉浸感, 虚拟现实技术的出现, 为远程沉浸式教学提供了强有力的技术支撑. 但目前的虚拟课堂的交互方式主要有键盘、鼠标

^① 基金项目: 国家科技支撑计划子课题 (2018YFB1004501); 陕西省重点产业创新链项目 (2020ZDLGY07-05); 国家自然科学基金青年项目 (61902300)

收稿时间: 2021-10-30; 修改时间: 2021-12-02; 采用时间: 2021-12-21; csa 在线出版时间: 2022-06-01

以及手柄等,这些交互方式形式单一,难以完全满足人们沉浸式体验的需求^[1].为增强虚拟课堂的人机交互体验,人们不断探寻新的交互方式如手势、语音、人脑控制等.其中,手势在传达信息方面承担着重要的角色,是人体与外界交流的主要方式之一,不同的手势代表着不同的信息.因此,可以利用手势交互达到在虚拟地理课堂中自然交互的目的,体验全新的课堂模式.

1 相关工作

20世纪80年代开始,人们已经开始通过数据手套上的传感器来获取人的手势信息.这种方式具有识别数据精准、速度快的优点,但数据手套价格十分昂贵,并且穿脱不方便,影响使用者的操作体验^[2].随着虚拟现实技术的飞速发展,Leap Motion和Kinect等体感传感器出现并被应用在手势识别中,这类设备使用计算机图像处理技术,对传感器获取到的图像或视频进行特征提取和手势识别,为手势交互提供了新的解决方案^[3,4].

林书坦等人^[5]基于Leap Motion构建了一套手写数字系统,通过建立手部模型来判定用户所指示的数字,该方法具有较高的识别率.2019年,金童等人^[6]基于Leap Motion捕捉手势的轨迹,构建了一套可以用于空中书写和正确笔画顺序演示功能的汉字书写和识别系统,该系统解决了中小学生和外国友人汉字学习时的枯燥感,极大提高了交互体验,但手势识别率有待提高.同年,李菲菲^[7]基于Leap Motion构建虚拟健身房交互框架并将其应用于Unity 3D构建的健身房虚拟场景中,使得虚拟健身房的交互体验更加逼真有趣.然而,该方法在手势识别速度上并没有达到理想的效果.2020年,林莹莹等人^[8]针对陶瓷制作条件严苛的局限性问题,采用Leap Motion开发了能够使用裸手交互的、具有良好沉浸感的虚拟现实陶艺体验系统,但该方法对识别范围要求较高,边缘处操作很不流畅.虽然Leap Motion对手部能够进行精准高效的追踪,但是由于Leap Motion的识别范围有限,仅在传感器上方25~600 mm的区域内有效.因此当手部超出范围时,这类方法会出现识别精度下降的问题.

近年来,深度学习在图像分类、目标检测等计算机视觉领域得到了飞速的发展^[9-15],因此许多研究人员将深度学习应用在手势识别领域.卢迪等人^[16]提出一种基于深度学习的手势识别算法,该算法采用改进的目标检测框架YOLOv4-tiny,能够对视频图像中的手

势进行准确的识别.王粉花等人^[17]在YOLOv3-tiny的基础上进行算法改进,较大地提高了手势识别速度.常建红^[18]提出一种改进的Faster R-CNN的手势识别算法,该算法通过两阶段目标检测方式提高了常规手势的识别准确率,然而在检测速率上可能慢于一阶段目标检测方法.胡鹏程^[19]采用LSTM来构建循环神经网络用于识别时序数据,实现了动态的手势识别,但LSTM的网络较深计算复杂非常耗时.吴晓凤等人^[20]通过修改Faster R-CNN框架的关键参数并提出扰动交叠率算法,进一步提高了识别准确率.然而,这类基于深度学习的算法通常采用RGB图像,对复杂背景的识别鲁棒性相对较弱.同时多数算法仅仅识别手势所包含的语义信息,无法获取当前手的空间位置.

本文针对Leap Motion在识别范围边缘识别率低以及识别速度慢的问题,将Leap Motion精确手部空间位置追踪与基于神经网络的手势识别相结合,提出了一种基于神经网络的手势交互方法.该方法设计和定义了一套交互手势,通过基于神经网络的目标检测框架识别Leap Motion采集到的红外图像,然后在手势识别的基础上设计了三维交互系统并将其应用到虚拟场景中.经实验证明,该方法能够有效提高手势识别率和识别速度,并且在Leap Motion识别范围边缘仍具有较高的识别率,实现在虚拟课堂中自然地进行人机交互.

2 Leap Motion的手势识别原理

Leap Motion是一款用于手部姿态捕捉的体感控制器,该控制器具有体积小、售价适中、手部姿态识别精度高的优点^[21].设备内部有两个用于识别并接收红外光反射回馈的球面摄像头,另有4个负责提供光源的红外LED,其内部结构如图1所示.

光源透过顶层的滤光片后会被障碍物(例如人手)反射回摄像头,再经由内部集成的算法计算出主要部位(如关节)的相关坐标数据.这些数据可以经由Leap Service转发给用户的应用,做进一步处理.Leap Motion最顶层的滤光片起到了很重要的作用,它可以过滤除波长在940 nm以外的其他杂波,只容许自带红外LED产生的红外光波进出,所以Leap Motion控制器并不需要其他的滤波器,就能得到一个去除掉复杂背景的红外成像,使得获取到的图像背景简单利于识别^[22].



图1 Leap Motion 内部结构

Leap Motion 可以实时检测跟踪传感器上方 25–600 mm 的扇形区域内手部的空间位置和方向, 并建立手的立体模型^[23], Leap Motion 识别范围如图 2 所示。

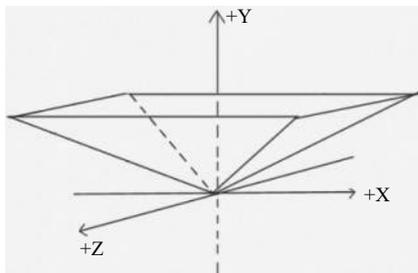


图2 Leap Motion 识别范围

Leap Motion 是基于双目视觉的手势识别设备, 通过两个摄像机来提取包括三维位置在内的信息进行手势的综合分析判断, 建立手部的立体模型。由于需要进行立体匹配, 复杂的立体模型需要大量的数据, 计算相对来说比较复杂^[24,25]。

3 手势交互方法设计

3.1 手势设计与定义

Leap Motion 设备仅支持预定义手势, 如: 挥手、点击、画圈等, 这些预定义手势并不能满足本文交互场景的实际需求, 因此本文自定义手势进行交互。手势设计需要从用户体验、交互过程以及交互系统等方面综合考量。本文对 30 名测试者进行手势设计调研, 通过大量的用户手势实验反馈, 最后选定 8 个手势进行交互。本文手势交互过程分为两步: 第 1 步为单一手势识别; 第 2 步为手势运动判断。手势交互开始时将进行开始确认, 若要切换手势则需进行切换操作。交互过程中的单一手势分别为左手握拳、右手握拳、左手五指伸开、右手五指伸开、左手食指伸出其他手指握拳、右手食指伸出其他手指握拳, 系统识别出手势后将进

入运动判断阶段, 左手握拳右滑、右手握拳左滑、左手五指伸开向上抬起、右手五指伸开向下放下、左手食指伸出其他手指握拳点击、右手食指伸出其他手指握拳点击。手势设计对照如表 1 所示。

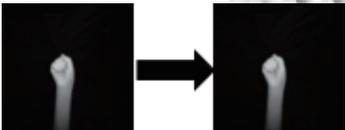
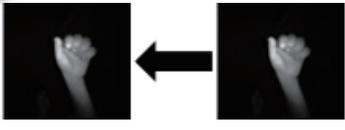
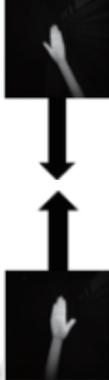
3.2 手势识别方法

本文手势识别方法采用 YOLOv5 目标检测框架并在其基础上针对手势识别进行一定改进。手势识别方法结构主要由输入端、backbone、neck 和 prediction 四个部分组成, 其结构图如图 3、图 4 所示。

其中输入端输入的图像为 Leap Motion 采集的红外图像, 利用 Mosaic 数据增强技术对图片进行拼接。Mosaic 数据增强方法基本原理是读取原始数据集中 4 张图片, 对 4 张图片进行翻转、色域变换以及缩放等操作, 然后按照 4 个方向的位置拼接在一起。通过 Mosaic 数据增强, 能够丰富物体的检测背景, 提高网络的鲁棒性。由于手势操作需在 Leap Motion 有效范围内进行, 操作时手势放置位置距离设备较近, 因此大部分手势数据集的手势宽高比在 1:1.4 左右, 小目标检测相对较少。为了避免使用通用锚框参数在手势数据集上造成漏检误检的情况, 达到最优检测效果的同时减少先验的数量, 提高检测速度。本文使用 K-means++ 算法聚类手势数据集的 anchor boxes, 在 19×19, 38×38 两个尺度上各聚类生成 3 个锚框, 共 6 个锚框。相比之前的 9 个锚框能够有效减少锚框数量, 提升检测速度。由于不同图片具有不同的长度和宽度。因此需将原始图像统一缩放到标准大小, 并将其输入到网络中, 即自适应图像缩放法。本文采用缩减黑边的方式, 对原始图像自适应的添加最少的黑边, 提高目标检测和推理的速度。

手势识别的 backbone 部分由 Focus 和 CSP 组成。其中 Focus 主要实现对图片进行切片操作, 切片操作是指将数据切分为 4 份, 每份数据相当于 2 倍下采样得到, 然后在 channel 维度进行拼接, 最后进行卷积操作。这样做可以最大程度地减少信息损失有利于进行下采样操作。CSP 模块中共有两种 CSP 结构 (见图 4), CSP1_X 结构带有残差结构应用于 backbone, 能够有效防止网络加深时所引起的梯度消失, 从而使得到的特征粒度更细。另一种 CSP2_X 结构将主干网络的输出分成了两个分支, 通过 concat 操作使网络对特征的融合能力得到加强, 保留了更丰富的特征信息, 该结构应用于 neck 部分。

表1 手势设计对照表

序号	手势	手势示例	运动情况	操作示例	手势意义
1	右手“OK”手势		—	—	开始
2	右手食指和中指伸出		—	—	切换手势
3	右手握拳		左滑		物体左转
4	左手握拳		右滑		物体右转
5	左手五指伸开		向上抬起		物体放大
6	右手五指伸开		向下放下		物体缩小
7	左手食指伸出其他手指握拳		点击		选中
8	右手食指伸出其他手指握拳		点击		更改样式

手势识别方法的 neck 部分采用 FPN+PAN 的结构,用于生成特征金字塔.特征金字塔会增强模型对于不同缩放尺度对象的检测,丰富了特征图的表达能力,加强了网络特征融合的能力.根据观察,通过 Leap Motion 采集到的红外图像中,手的大小处在一定范围之间,因此不需要针对较小目标进行检测.在本文中,为了提高检测速度, prediction 部分只针对 19×19, 38×38 两个尺度进行.

预测部分的损失函数由 classification loss (分类损失函数) 和 bounding box regression loss (回归损失函数) 两部分构成.采用 Focal Loss 评价目标框和预测框的类损失和置信度损失,采用 GIoU_Loss 评价 bounding box 的位置损失. Focal Loss 损失考虑的是目标检测中正负样本严重不均衡的问题.公式如式 (1) 所示:

$$FL(pt) = -(1 - pt)^{\gamma} \log(pt) \tag{1}$$

函数 $GIoU$ 是源自 IoU 的一种边框预测的损失计算方法, 定义如式 (2):

$$GIoU = IoU - \frac{C - (A \cup B)}{C} \quad (2)$$

其中, IoU 可以由式 (3) 计算得到:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

B 是预测框和真实框, $GIoU Loss = 1 - GIoU$.

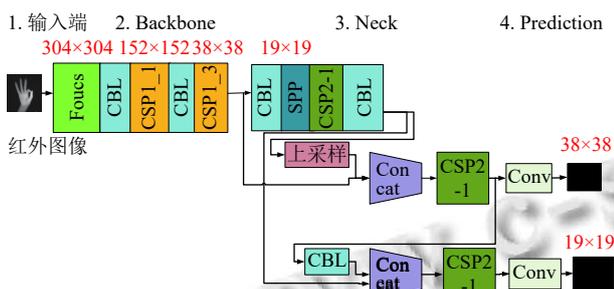


图3 手势识别方法结构图

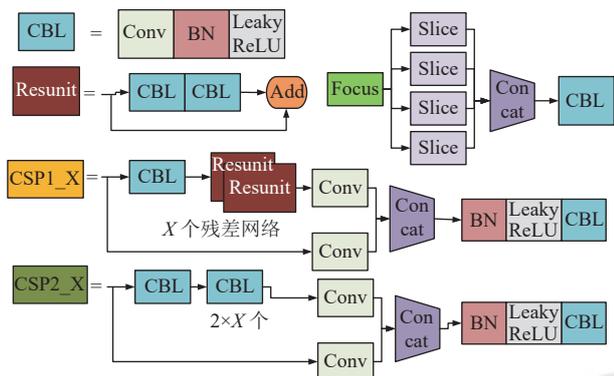


图4 部分组件详细结构图

3.3 动态手势判断方法

考虑到基于深度卷积神经网络的手势识别方法只能识别图像信息, 并不能体现手部的空间位置信息, 而 Leap Motion 能够获得手部结构的关节点信息, 故本文将结合 Leap Motion 获取到的手部关节点信息来进行动态手势的判断. 手的运动是通过对比当前帧与之前特定帧得到的, 通过对比前后帧的数据来判定手的运动变化, 本文主要通过判断手部关节点的位移来判断手势的运动情况.

Leap Motion API 提供手部模型的关节点坐标, 通过计算两个时刻关节点之间的欧氏距离, 计算公式如式 (4), 可以描述手势运动的空间移动信息.

$$P = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2 + (z_i - z_c)^2} \quad (4)$$

其中, P 为关节点位移, (x_i, y_i, z_i) 为关节点在 i 时刻的空间坐标, (x_c, y_c, z_c) 为关节点在 c 时刻的空间坐标.

3.4 手势交互流程

本文的手势交互过程分为单一手势识别、动态手势运动判断两个部分. 首先通过 Leap Motion 进行数据采集, 并将获取到的红外图像输入到手势识别方法中进行特征提取进而实现对手势的识别与分类, 单一手势的识别结果将作为判断动态手势的重要条件. 手势的运动情况通过对比前后帧位置变化进行判定, 图 5 为手势交互流程.

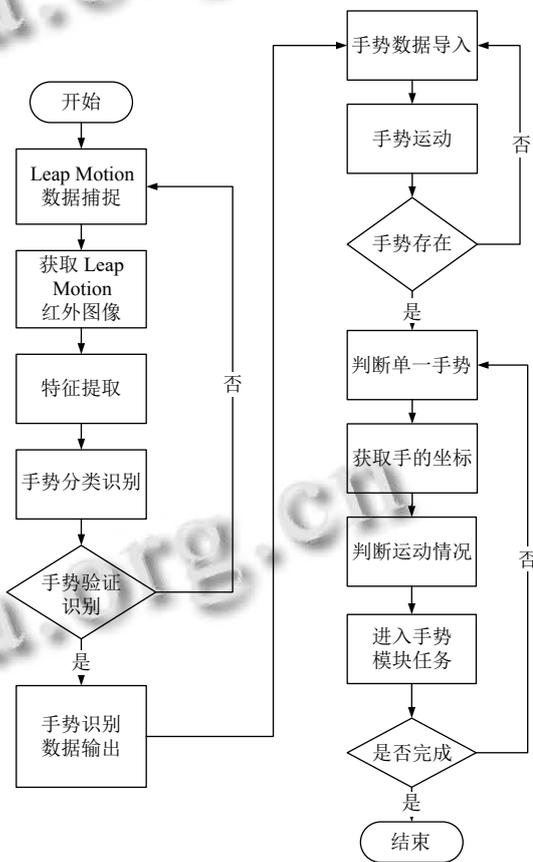


图5 手势交互流程

其中手势模块任务分为 6 个模块, 分别为判断是否左手握拳右滑、判断是否右手握拳左滑、判断是否左手五指伸开向上抬起、判断是否右手五指伸开向下放下、判断是否左手食指伸出 (其他手指握拳) 点击、判断是否右手食指伸出 (其他手指握拳) 点击, 若判断结果为真, 则做出对应的物体左转、物体右转、物体放大、物体缩小、选中物体、更换物体风格的相应操作.

4 实验设计

4.1 实验平台

本文实验在 PC 机上完成, PC 配置为 Windows 10 操作系统, 处理器为 Intel(R)Core(TM)i7-9700k, 内存 32 GB, GPU 为 NVIDIA GeForce GTX 1080, GPU 内存为 24 GB, Unity 版本为 2018.4.16, 以及一台 Leap Motion.

4.2 实验数据

手势识别的实验数据为 Leap Motion 采集到的红外图像. 数据集共采集 26 000 张, 包括 8 种手势, 手势数据采集过程如图 6 所示, 手势示例如图 7 所示. 经过

多次实验测试对比, 决定将其中 80% 的图片用来训练, 20% 的图片用来测试. 同时利用 Leap Motion 官方网站提供的开发文件来获取手部坐标数据.



图6 手势数据集采集过程



图7 手势示例

4.3 数据集标签的标定

通过使用 LabelImg 工具对红外图像中的手势进行逐一标注, 并生成相对应的.txt 文件, 文件内包含手势类别、归一化的标注框的中心坐标 (x, y) , 归一化的标注框的宽和高 (w, h) . 归一化公式如式(5)~式(8)所示:

$$x = \frac{x_{\max} + x_{\min}}{2width} \quad (5)$$

$$y = \frac{y_{\max} + y_{\min}}{2height} \quad (6)$$

$$w = \frac{x_{\max} - x_{\min}}{width} \quad (7)$$

$$h = \frac{y_{\max} - y_{\min}}{height} \quad (8)$$

4.4 数据增强与数据预处理

本文在采取一定数量的数据之后对数据集进行数据扩增. 利用数据增广技术如平移、缩放、对比度变换, 添加噪声等, 将数据集扩充到 26 000 张, 同时对数据进行预处理工作: 利用 Mosaic 数据增强技术对图片进行拼接; 使用 K-means++ 算法进行聚类生成锚框, 以及采用缩减黑边的方式, 对原始图像自适应的添加最少的黑边等.

4.5 实验过程

本文首先对 YOLOv5 手势识别算法和本文改进过

的手势识别方法进行实验对比, 然后在虚拟场景下进行手势交互实验. 本文的交互场景为虚拟地理课堂三维系统, 系统使用 unity 游戏开发引擎, 能够使学生沉浸式体验高仿真虚拟课堂, 本文选取《认识地球》一课, 通过完成相对应的手势操作实现地球仪的放大、缩小、旋转、点击地球仪选中地图板块、点击地球仪更换地球仪风格的操作, 如图 8 所示.

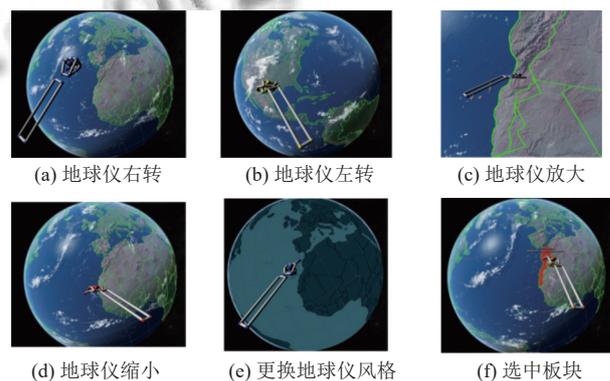


图8 虚拟场景交互

5 实验结果及分析

5.1 手势识别效果分析

实验使用平均精度均值 (mean average precision, mAP) 作为评价标准, 它综合考虑了查准率和查全率.

平均精度均值首先需要考虑计算每一个物体类别的平均精度 (AP), AP 是计算某一类 P-R 曲线下的面积, mAP 则是计算所有类别 P-R 曲线下面积的平均值. 本文对比了 YOLOv5 手势识别算法与本文手势识别方法的 mAP 值、处理时间以及 P-R 曲线如表 2、表 3、图 9、图 10 所示.

表 2 不同手势识别方法平均 mAP 对比

手势识别方法	mAP
YOLOv5 手势识别算法	0.989
本文改进的手势识别方法	0.993

表 3 不同手势识别方法处理时间对比 (ms)

手势识别方法	预处理时间	推理时间
YOLOv5 手势识别算法	0.2	5.8
本文改进的手势识别方法	0.1	5.6

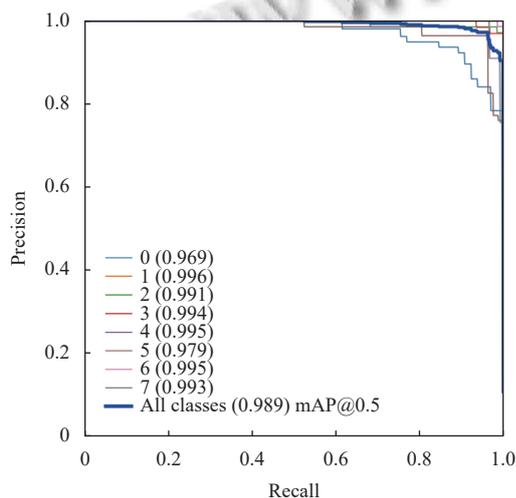


图 9 YOLOv5 手势识别算法的 P-R 曲线

通过分析表 2、图 9、图 10 可以发现, 本文手势识别方法的平均精度值比 YOLOv5 算法的平均精度值高, 单个手势的识别精度 (即 P-R 曲线下的面积) 也比 YOLOv5 算法的识别精度高. 说明本文使用 K-means++ 聚类算法重新生成的锚框大小比通用的锚框大小更适合本文的手势数据集, 从而使网络对目标的检测变得更加准确, 减少了漏检和误检的次数.

从表 3 可以发现, 本文改进过的手势识别算法的数据的预处理时间以及推理时间更短. 这是由于本文仅在 19×19 和 38×38 两个尺度上进行预测, 锚框数量由原来的 9 个变为 6 个, 锚框数量的减少大大提高了识别速度, 从而更好实现实时识别的效果.

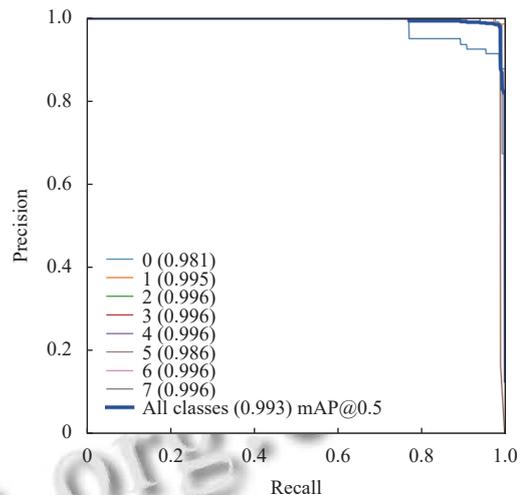


图 10 本文手势识别方法的 P-R 曲线

评估分类器性能好坏的方法是观察 P-R 曲线, 如果 Recall 值在增长的同时, Precision 的值能够保持在一个很高的水平, 说明分类器的性能较好, 相反性能比较差的分类器可能会损失很多 Precision 值才能换来 Recall 值的提高. 从图 9、图 10 可以看出, 本文手势识别算法的 P-R 曲线比 YOLOv5 手势识别算法的 P-R 曲线更稳定, 即在 Recall 值不断增加的情况下, Precision 依然能保持较高的值. 而 YOLOv5 手势识别算法在 Recall 值不断增加的过程中损耗了更多的 Precision 的值. 因此本文改进的手势识别方法分类效果更好, 有明显的改善效果.

5.2 手势交互效果分析

为了验证本文手势交互方法的有效性, 本文设计了两个实验方案: 方案一为本文的手势交互方案; 方案二为使用 Leap Motion 自带的手势交互方案. 随机选取 10 名被试者进行实验, 实验者在未被告知系统差别的情况下, 单个手势分别通过两个方案进行实验, 记录每种手势交互 20 次所需时间以及每种手势的识别成功率.

通过观察图 11、图 12 不同手势交互方法完成交互所需时间以及每种手势的识别成功率可以发现, 使用 Leap Motion 自带的手势交互方法所需要的时间更长、每种手势的识别成功率更低.

Leap Motion 是基于双目摄像头进行深度成像, 根据双目信息推算出深度, 然后从深度图得到手部模型, 最后利用关键点信息识别手势. 由于 Leap Motion 是对手部关键点进行检测, 所以对于手指重叠的复杂情况

不能准确捕获,建立的模型与实际情况存在较大偏差,出现两只手指识别成一只手指的情况,尤其在范围边界,手势识别率下降,出现剧烈抖动,影响操作体验。

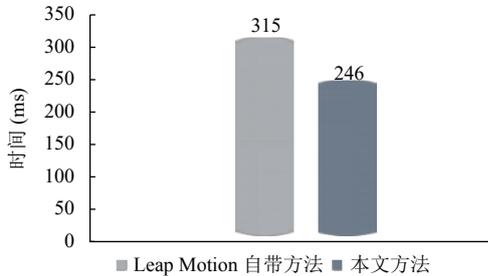


图 11 不同手势交互方法完成手势所需时间

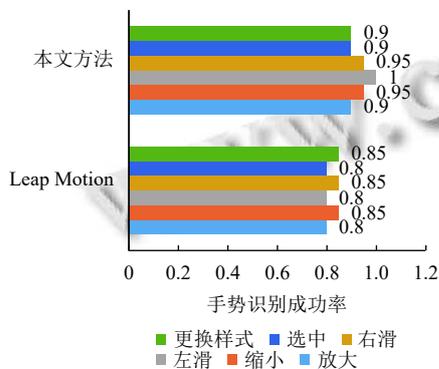


图 12 不同手势交互方法手势识别成功率

而本文基于神经网络对手势进行识别,是对手部的红外图像直接进行识别和分类,不会出现手部模型匹配不准而出现手部抖动的现象,识别过程更流畅。

经过多次实验发现,采用 Leap Motion 自带手势识别系统进行交互的最佳交互区域在 Leap Motion 上方 90–430 mm,超过最佳交互区域的范围会出现识别不准确的现象,而采用本文的手势识别交互方法在范围边界仍可以进行交互。为更好地对比不同的手势识别交互方法在不同识别范围内手势识别的准确率,本文记录了两种交互方法在不同识别区域范围内完成 8 个手势,每个手势完成 5 次的手势识别情况,对其结果进行对比,如表 4 所示。

表 4 不同交互区域手势识别情况

手势识别方法	最佳交互区域范围内		最佳交互区域范围外	
	正确个数	错误个数	正确个数	错误个数
Leap Motion 自带方法	35	5	23	17
本文方法	38	2	32	8

从表 4 中可以看出不论是在最佳识别区域范围内还是最佳识别区域范围外,本文的手势识别交互方法都

优于 Leap Motion 自带的手势识别方法,具有较好的识别效果,进一步验证了本文手势识别交互方案的优越性。

6 结束语

本文提出了一种基于深度神经网络的 Leap Motion 手势交互方法,设计出一套交互手势,将神经网络运用到手势识别当中,并在此基础上设计了三维交互系统将其应用到虚拟场景中。通过实验对比发现,本文的手势识别交互方法无论在识别精度还是识别速度上都优于 Leap Motion 自带识别系统,并且在 Leap Motion 识别范围边界处仍能保持较高的识别率。未来,会进一步增加复杂动态手势的研究,为用户提供更加丰富的交互体验。

参考文献

- 黄俊,景红. 基于 Leap Motion 的手势识别在虚拟交互中的研究. 计算机应用研究, 2017, 34(4): 1231–1234. [doi: 10.3969/j.issn.1001-3695.2017.04.062]
- 刘瑜兴,王淑侠,徐光耀,等. 基于 Leap Motion 的三维手势交互系统研究. 图学学报, 2019, 40(3): 556–564.
- Avola D, Bernardi M, Cinque L, et al. Exploiting recurrent neural networks and leap motion controller for the recognition of sign language and semaphoric hand gestures. IEEE Transactions on Multimedia, 2018, 21(1): 234–245.
- Liu FL, Zeng W, Yuan CZ, et al. Kinect-based hand gesture recognition using trajectory information, hand motion dynamics and neural networks. Artificial Intelligence Review, 2019, 52(1): 563–583.
- 林书坦,尹长青. 基于 LeapMotion 的数字手势识别. 电脑知识与技术, 2015, 11(35): 108–109.
- 金童,蒋玉茹,邱伟. 基于 Leap Motion 的汉字空中书写及识别系统. 电子技术与软件工程, 2019, (23): 47–49.
- 李菲菲. 基于 Leap Motion 的手势交互及在虚拟场景中的应用研究 [硕士学位论文]. 济南: 山东大学, 2019.
- 林莹莹,蔡睿凡,朱雨真,等. 基于 Leap Motion 的虚拟现实陶艺体验系统. 图学学报, 2020, 41(1): 57–65.
- 彭玉青,赵晓松,陶慧芳,等. 复杂背景下基于深度学习的手势识别. 机器人, 2019, 41(4): 534–542.
- 谷学静,周自朋,郭宇承,等. 基于 CNN-LSTM 混合模型的动态手势识别方法. 计算机应用与软件, 2021, 38(11): 205–209. [doi: 10.3969/j.issn.1000-386x.2021.11.032]
- 张维,林泽一,程坚,等. 动态手势理解与交互综述. 软件学报, 2021, 32(10): 3051–3067. [doi: 10.13328/j.cnki.jos.006217]

- 12 解迎刚, 王全. 基于视觉的动态手势识别研究综述. 计算机工程与应用, 2021, 57(22): 68–77. [doi: [10.3778/j.issn.1002-8331.2105-0314](https://doi.org/10.3778/j.issn.1002-8331.2105-0314)]
- 13 Adithya V, Rajesh R. A deep convolutional neural network approach for static hand gesture recognition. Procedia Computer Science, 2020, 171: 2353–2361.
- 14 Wu BX, Yang CG, Zhong JP. Research on transfer learning of vision-based gesture recognition. International Journal of Automation and Computing, 2021, 18(3): 422–431.
- 15 王粉花, 张强, 黄超, 等. 融合双流三维卷积和注意力机制的动态手势识别. 电子与信息学报, 2021, 43(5): 1389–1396. [doi: [10.11999/JEIT200065](https://doi.org/10.11999/JEIT200065)]
- 16 卢迪, 马文强. 基于改进 YOLOv4-tiny 算法的手势识别. 电子与信息学报, 2021, 43(11): 3257–3265. [doi: [10.11999/JEIT201047](https://doi.org/10.11999/JEIT201047)]
- 17 王粉花, 黄超, 赵波, 等. 基于 YOLO 算法的手势识别. 北京理工大学学报, 2020, 40(8): 873–879.
- 18 常建红. 基于改进 Faster RCNN 算法的手势识别研究 [硕士学位论文]. 保定: 河北大学, 2020.
- 19 胡鹏程. 基于深度学习的手势识别系统研究与实现 [硕士学位论文]. 北京: 北京邮电大学, 2018.
- 20 吴晓凤, 张江鑫, 徐欣晨. 基于 Faster R-CNN 的手势识别算法. 计算机辅助设计与图形学学报, 2018, 30(3): 468–476.
- 21 高宇, 何小海, 吴晓红, 等. 基于 Leap Motion 的动态手势识别方法. 计算机系统应用, 2019, 28(11): 208–212. [doi: [10.15888/j.cnki.csa.007172](https://doi.org/10.15888/j.cnki.csa.007172)]
- 22 Zhang QX, Deng F. Dynamic gesture recognition based on leapmotion and HMM-CART model. Journal of Physics: Conference Series, 2017, 910(1): 012037.
- 23 Lu W, Tong Z, Chu JH. Dynamic hand gesture recognition with leap motion controller. IEEE Signal Processing Letters, 2016, 23(9): 1188–1192.
- 24 Potter L E, Araullo J, Carter L. The leap motion controller: A view on sign language. Proceedings of the 25th Australian Computer-human Interaction Conference: Augmentation, Application, Innovation, Collaboration. Adelaide: ACM, 2013. 175–178.
- 25 Zaiqi IA, Pentiu ŞG, Vatavu RD. On free-hand TV control: Experimental results on user-elicited gestures with Leap Motion. Personal and Ubiquitous Computing, 2015, 19(5–6): 821–838.

(校对责编: 孙君艳)