

基于智能穿戴设备的人体姿态识别方法和系统

邢云涛¹ 陈吉² 凌云志²

(1. 杭州唯灵医疗科技有限公司 浙江杭州 310000; 2. 维灵(杭州)信息技术有限公司 浙江杭州 310000)

摘要: 人体姿态识别是计算机视觉和人机交互领域的一个重要研究方向,它可以应用于许多场景,如动作捕捉、虚拟现实、智能家居等。随着硬件技术的发展,基于智能穿戴设备的姿态识别系统近年来受到广泛关注。与传统的基于摄像头的系统相比,穿戴式设备具有更好的可穿戴性和随时随地的监测能力。近年来,研究者们提出了多种基于传感器融合的人体姿态估计方法,利用加速度计、陀螺仪、磁力计等传感器,结合机器学习算法实现了更加精准的姿态识别。同时,相关系统也在可穿戴设备、虚拟现实等应用场景中得到了广泛应用和验证。
关键词: 智能; 识别; 穿戴设备; 人体姿势

前言

人体姿态识别技术是计算机视觉和人机交互领域的一个核心研究方向,它可以广泛应用于动作捕捉、虚拟现实、智能家居等诸多场景。随着硬件技术的发展,基于智能穿戴设备的姿态识别系统近年来受到了广泛关注。与传统的基于摄像头的系统相比,穿戴式设备具有更好的可穿戴性和随时随地的监测能力,能够更好地融入人类日常生活。研究者们提出了多种基于传感器融合的人体姿态估计方法,利用加速度计、陀螺仪、磁力计等传感器,结合机器学习算法实现了更加精准的姿态识别。这些系统不仅在可穿戴设备领域广泛应用,在虚拟现实等新兴技术中也得到了验证和推广。

1. 背景技术和发展现状

人体姿态识别是计算机视觉和人机交互领域的一项核心技术,其主要目标是准确获取和分析人类的动作和行为信息。这项技术在虚拟现实、智能家居、医疗康复等领域都有广泛应用前景。随着硬件设备的快速发展,基于可穿戴设备的姿态识别系统近年来引起了广泛关注。相比传统的基于摄像头的系统,这类基于传感器的方法具有更好的可穿戴性和随时随地的监测能力,能够更好地融入人类的日常生活。

在技术层面上,人体姿态识别系统通常由硬件设备和软件算法两部分组成。硬件设备主要包括各类传感器,如加速度计、陀螺仪、磁力计等,用于采集人体运动信息。软件算法则负责对这些传感器数据进行分析 and 处理,实现对人体姿态的精准识别。研究人员提出了多种基于传感器融合的姿态估计方法,利用机器学习等先进算法,在识别准确性、鲁棒性等方面实现了大幅提升。这些基于可穿戴设备的姿态识别技术已经在诸多应用场景中得到了验证和推广。在虚拟现实领域,它可以实现更加沉浸式的交互体验;在医疗康复中,它可用于患者的运动状态监测和辅助训练;在运动训练领域,它则可帮助教练精准评估运动员的动作质量。此外,在智能家居、工业生产等场景下,这项技术也展现出广阔的应用前景。

2. 嵌入式可穿戴设备上的优化

2.1 系统功耗优化

可穿戴设备往往受制于有限的电池容量,因此降低系统功耗是优化嵌入式实现的首要任务。针对这一问题,研究人员提出了多种技术方案。一是针对传感器数据采集进行优化,如利用压缩感知等技术,减少不必要的数据采集,从而降低功耗;二是在算法层面进行优化,如采用更加高效的机器学习模型,或使用动态调整模型复杂度的技术,在保证识别精度的前提下尽量减少计算资源的消耗;三是在系统架构层面进行优化,如利用异构计算资源如 GPU/FPGA 等实现功能分工,最大程度降低整体功耗。综合运用这些技术手段,可以有效提升可穿戴设备的续航能力。

2.2 实时性优化

除了功耗,实时性也是嵌入式姿态识别系统需要重点解决的问题。

在许多应用场景下,如虚拟现实、运动监测等,对系统响应速度有很高的要求。为此,优化算法计算复杂度、采用并行计算架构、进行软硬件协同设计等都是可行的优化策略。此外,还可以利用边缘计算技术,将部分计算任务下放到末端设备,减轻中央处理器的负担,从而大幅提升系统的实时性能。

2.3 存储优化

针对嵌入式设备有限的存储资源,如何高效利用存储空间也是一个需要关注的问题。一方面,可以采用模型压缩、量化等技术,减小机器学习模型的存储占用;另一方面,动态加载所需模型或数据,按需使用存储资源也是一种行之有效的方法;此外,利用外部存储设备如 SD 卡等,将部分数据和模型存储于此,也能有效扩充系统的存储容量。通过上述优化手段,可以充分利用有限的存储资源,满足嵌入式系统的实际需求。

3. 多模态信息融合应用于人体姿态识别

3.1 多模态信息的互补性

单一传感器数据往往存在缺失、噪声等问题,限制了识别系统的鲁棒性。而将视觉、音频等多种信息源进行融合,可以充分发挥各自的优势,互补弥补彼此的不足。例如,结合 RGB 图像中的外观信息和深度图像中的 3D 结构信息,能更准确地捕捉人体关键点的位置;加入语音信号中蕴含的情感状态特征,则可以增强对复杂动作意图的理解。综合利用这些信息源,可以大幅提高识别的精度和可靠性。

3.2 多模态融合的建模方法

如何有效地融合多种异构数据源,是多模态融合技术的核心问题。现有的方法主要包括:特征级融合、决策级融合以及端到端的深度融合。特征级融合是将不同模态的特征进行拼接或级联,再输入到分类器中;决策级融合则是分别训练单一模态的分类器,再采用投票、加权等方式综合决策。而端到端的深度融合方法可以直接学习模态间的交互关系,无需进行手工特征提取,能更好地挖掘多模态信息的潜在联系。研究人员正在不断探索更加高效的多模态融合架构,以进一步提升识别性能。

4. 基于智能穿戴设备的人体姿态识别模型的训练方法

如图 1 所示,本方法将会在多个连续的预设动作姿态下,获取通过智能穿戴设备采集的用户的生理数据。生理数据包括 ECG 信号和 ACC 信号。预设动作姿态是预先设置在连续时间段内的人体动作状态,其可以是静止状态,也可以是多个连续的动作状态,包括慢跑、下楼梯、上楼梯、步行、坐姿、站立和平躺等。生理数据是在用户按照预设动作姿态分别执行对应的动作的同时获取,每个生理数据和每个预设动作姿态分别对应一个时间戳。另外,为支持模型的有效训练,该生理数据并不限于是一组或几组数据,而是又海量数据组成的生理数据集;为了提升模型的泛化能力,采集该生理数据的人群应尽可能的丰富,可以但不限于包括健康人群、亚健康

人群和有心脏疾病、呼吸疾病的病患人群。ECG 信号通过智能穿戴设备的电极采集,其包括用户心脏电活动的电压与时间的关系图;这些电极检测在每个心动周期(心跳)期间由心肌去极化随后复极化引起的微小电变化。正常心电图模式的改变发生在许多心脏异常中,包括心律失常(如心房颤动和室性心动过速)、冠状动脉血流不足(如心肌缺血和心肌梗死)和电解质紊乱(如低钾血症和高钾血症);ACC

通过智能穿戴设备的加速度计采集,包括用户在预设时间段内,在静止坐标系中的速度变化率;坐标加速度不同,坐标加速度是在固定坐标系中的加速度,例如,静止在地球表面的加速度计将测量由地球重力引起的加速度,直线向上(根据定义) $g \approx 9.81\text{m/s}^2$ 。相比之下,自由落体(以大约 9.81m/s^2 的速度向地球中心下降)的加速度计将测量为零。

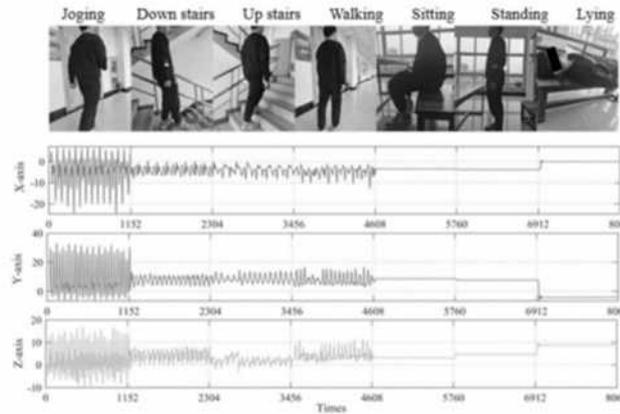


图 1

基于预设动作姿态、生理数据及其匹配关系构建训练数据集,并按照预设规则对状态信息进行预处理,得到优化后的训练数据集。优化后数据集包括多个携带状态标签的生理数据;由于是在集信号的同时标记对应人体姿态与运动,因此,可以按照预设动作姿态和生理数据的时间戳,通过预设动作姿态分别对各个生理数据添加标签;根据各个添加标签之后的生理数据,构建训练数据集,该训练数据集由多个携带以姿态结果标签的生理数据组成。为了便于输入模型进行训练,以及提高模型识别准确率。首先需要分别获取 ACC 信号在 X、Y 和 Z 轴的分量信号,并通过低通滤波器分别对分量信号进行滤波得到滤波后信号,并对滤波后信号进行滑窗处理,得到优化之后的 ACC 信号;其次,通过小波变换对 ECG 信号进行去基线漂移降噪得到降噪后信号,并对降噪后信号进行滑窗处理,得到优化之后的 ECG 信号;将优化之后的 ACC 信号和优化之后的及其对应的姿态标记信号,保存为优化后的训练数据集。另外,小波变换是一种多尺度分析的工具,可以用于对信号进行去噪和特征提取。在实施例中,处理 ECG 信号时,小波变换常被用于去除信号中的基线漂移和噪声。

优化后的训练数据集,对混合深度学习网络进行训练,得到用于人体姿态识别模型。将该训练数据集中各个元素转换为 $(n, 256, 4)$ 的矩阵格式之后输入该混合深度学习网络,其中, n 为样本数量,256 为 2s 的数据长度,4 是通道数量(ACC 的 3 个通道+ECG 的 1 个通道)。该混合深度学习网络以训练数据集为输入,通过多个卷积层、池化层以及双层全连接层的处理,以每个生理数据的标签为监督信息进行有监督训练,模型在学习过程中能够通过输入样本和其对应的标签进行关联,从而学习到输入(生理数据)和输出(生理数据对应的人体姿态)之间的映射关系。以 epochs 数量 80 为训练结束条件,训练过程中姿态标记共区分 7 类,作为模型的 labeling。在接收到优化之后的训练数据及之后,CNN 结构通过第一卷积层提取训练数据集中各个生理数据的特征,得到多个特征图,并通过第一池化层对特征图进行下采样,得到第一优化特征以及通过第二卷积层提取优化特征图的特征,得到多个第二特征图,并通过第二池化层对第二特征图进行下采样,得到第二优化特征;其中,在深度学习网络中,通常会交替使用卷积层和池化层,形成深层次的特征提取和抽象过程。这有助于网络学习输入数据的层次结构和抽象特征。CNN 的主

要特点是通过卷积层和池化层来有效提取输入数据的局部特征,并且具有共享权重和局部连接的结构

CNN 结构包括双层结构,其第一层结构包括第一卷积层和第一池化层,第二层结构包括第二卷积层和第二池化层,第一层结构和第二层结构分别具备 64kilters 和 128kilters 个卷积核;另外,在卷积层和池化层之间还通过激活函数连接,valid 表示不使用填充,Same 表示使用填充,该类技术特征均未本领域的常规手段,因此,在本实施例中不再具体展开说明。可以理解本方案中采用的双层结构可以具备更好的特征学习和特征表示能力,进而提升模型的整体训练效果。通过 LSTM 结构,接收多个存在时序关联的第二优化特征,并通过长短时记忆层和门控机制,对多组第二优化特征进行联合提取,得到包含用户多个生理数据之间的关联隐藏信息的第三优化特征;LSTM(Long Short-Term Memory)是一种循环神经网络(RNN)的变体,传统的 RNN 在处理长序列时,容易出现梯度消失或梯度爆炸的问题,导致难以捕捉和学习长期依赖关系。

结语

智能穿戴设备的人体姿态识别技术正在快速发展,为各领域提供了新的解决方案。未来,随着传感器技术的不断进步,以及深度学习等算法的不断创新,这一领域必将迎来更多突破性的发展。我们可以期待这项技术在未来为人类生活带来更多便利和 benefit。同时,在隐私保护、能耗优化等方面也需要持续的研究与创新。总之,这个领域充满前景,值得我们持续关注与探索。

参考文献:

- [1]数字化体能训练研究——以 Garmin 智能穿戴为例[A].谭茜尹;李明.第十三届全国体育科学大会, 2023.
- [2]“识变与求变”:新发展阶段乒乓球体能训练的新问题与新思路[A].景佳欣;王森;乔亚楠.第二届陕西省体育科学大会, 2024.
- [3]新规则下跆拳道体能训练方案探究[A].郭金涛;卢秀栋;刘广松.第二届陕西省体育科学大会, 2024.
- [4]山西省 U16 男子空手道专项体能训练实践策略[A].王一斐;杨舒晴;李宁.第二届陕西省体育科学大会, 2024.
- [5]基于可穿戴式惯性传感器的人体运动跟踪方法综述[J].张肇豪;何百岳;杨旭升;张文安.自动化学报, 2019(08).