**人体动作识别项目说明文档**

|  |  |
| --- | --- |
| 学院： | 信息工程学院 |
| 专业： | 电子信息工程 |
| 姓名： | 刘媛媛 |
| 学号： | 1211002016 |

# 研究现状

## 1.1研究背景

人体动作识别项目的研究背景根植于对人类行为理解和技术应用的持续追求。人体动作识别技术在智能监控、视频检索、人机交互、运动分析等诸多领域发挥了至关重要的作用。在安防监控领域，人体动作识别技术有助于预防违法犯罪等异常事件的发生，能够准确识别出正在进行的危险行为，从而及时发出预警，有效保障公共场所的安全；在视频检索领域，人体动作识别技术能够协助相关人员高效完成视频检索任务，从大量视频数据中快速定位到特定动作，检索到所需目标；在人机交互领域，人体的动作则是人机交互的桥梁，通过对人体动作的准确识别，能够让机器更加正确地理解人们要表达的意图，从而更精准地完成相应的工作。

随着信息技术和人工智能的迅猛发展，人们对于如何通过技术手段精确捕捉、分析和理解人体动作的需求日益增长。这一研究领域融合了多个学科的知识和技术，包括计算机视觉、机器学习、传感器技术以及人体运动学等。在传感器技术方面，随着传感器的小型化、多样化和成本的降低，惯性传感器、摄像头等能够高精度地捕捉动作数据的设备得到了广泛应用，为动作识别的实时性和准确性提供了坚实的技术基础。随着计算机技术与网络技术的快速发展，多媒体信息呈爆炸式增长，其中视频作为重要的信息载体，越来越受人们青睐更多信息通过视频进行传播。

算法设计是人体动作识别研究的核心，机器学习和深度学习方法在此扮演着重要角色。通过这些方法，研究人员能够处理从传感器获取的大量数据，并从中提取出有意义的动作特征，实现对复杂动作的自动识别和分类。此外，随着深度学习技术的不断进步，特别是卷积神经网络（CNN）和循环神经网络（RNN）的应用，使得对动作模式的识别更加精确和高效。

## 1.2国内外研究现状

## 1.2.1动作特征提取

能否提取出适当特征来充分表达人体行为信息，对后续的识别准确性至关重要。动作特征提取的主要任务是根据人体动作的特征对不同动作进行识别和分类。人体动作识别的研究现状主要集中在动特征提取和分类器设计两个方面。

在动特征提取方面，国内外研究者提出了多种不同的特征提取方法，包括基于颜色、形状、运动和深度等方面的特征。其中，基于运动的特征提取方法被广泛应用于人体动作识别中，其主要包括光流、运动轨迹、骨骼关节点等特征。国内外的研究者通过在实验数据集上的大量实验验证，证明了这些特征在人体动作识别中的有效性。

在分类器设计方面，国内外的研究者主要采用支持向量机、神经网络、决策树等机器学习方法来进行人体动作的分类。近年来，深度学习方法在人体动作识别领域的应用越来越普遍，其中卷积神经网络是最常用的深度学习方法之一。通过深度学习方法，研究者可以自动学习并提取人体动作的高层次特征，从而获得更好的分类性能。

Yamato等人通过选取图像轮廓信息的方法表示人体行为特征，首先对RGB图像二值化处理，再利用边缘检测的算法提取出目标轮廓像素，最后将提取的轮廓信息同标准模板库中的动作模板进行匹配给出最终结果；Carlsson等人则是对边缘检测算法进行了优化，提高了边缘信息提取的准确度；minchisescu设计了对图像信息分块采样的方案，在时间维度上采样图像信息也取得了不错的结果。总体来说，基于RGB图像提取动作特征的方法有一定的优势，但同时也有一个很大的缺陷，就是极其容易受光线、遮挡的影响，使得识别精度降低严重。

## 1.2.2动作分类方法

获取动作特征值信息之后，就需要选择或者设计分类算法完成动作分类的任务。人体动作识别中的动作分类算法是研究重点之一，其主要任务是将提取到的人体动作特征进行有效分类，以识别出不同的动作。国内外的研究者们在动作分类算法方面做出了一系列的研究和探索。

Bobick等人提出了基于运动能量图的方法，其思路是依据视图的变化来实现动作识别；Karpathy等人对帧数据的局部时空信息进行学习，提出了通过在时域上对CNN网络连通性进行扩展方法识别率有提升显著；Srivastava首先将输入数据转化为固定的长度，然后通过堆叠多个LSTM单元完成分类任务，其实质是通过使用多层的LSTM神经网络来实现分类的任务，这使得动作识别效果得到大幅度提升，算力降低。

存在问题：上述研究人员的方法均是倾向于对网络模型进行改进，又或者是帧数据像素级别的处理来提高识别精度。而即使是利用了骨架关节点的信息，也没有考虑关节之间的夹角信息及关节之间的距离的长短对最终动作识别精度的影响。

# 问题定义

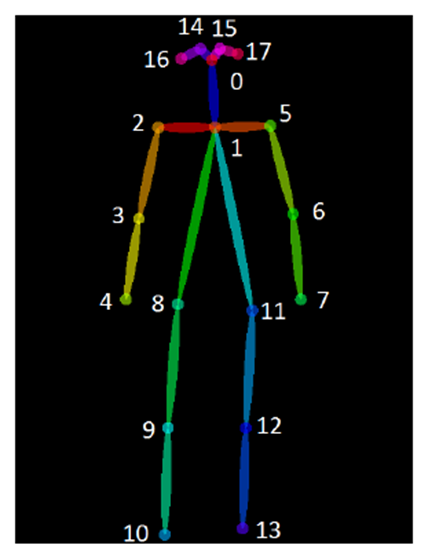
人体动作识别项目是通过计算机视觉技术识别和分析人体在运动中的各种动作，包括但不限于行走、跑步、跳跃、踢腿等。该项目旨在利用人体关键点检测和动作识别算法，实时地监测和识别人类的动作，为健康管理、体育训练、游戏互动等领域提供数据支持和智能应用。在这个项目中，主要涉及到的问题包括数据采集和标注、动作特征提取和建模、模型训练和优化以及实时动作识别和应用等方面。

首先，数据采集和标注是人体动作识别项目中的第一步，需要通过摄像头或传感器等设备获取人体运动的视频或数据，并对其进行标注，标注方式多种多样，如关键点标注、骨架标注等。其次，动作特征提取和建模是关键环节，通过提取人体关键点或骨架的空间位置、运动轨迹、角度变化等特征，建立数学模型描述各种动作。然后，模型训练和优化是项目中的核心任务，通过深度学习、机器学习等方法对大量的标注数据进行训练，优化动作识别模型的准确率和鲁棒性。最后，实时动作识别和应用是项目的最终目标，通过将训练好的模型应用到实际场景中，实现对人体动作的实时监测、分析和识别，为健康管理、运动训练、虚拟现实等应用提供支持。

通过解决以上问题，将能够实现对人体动作的全面识别和分析，为人们的健康管理和运动训练提供更加智能和便捷的支持。

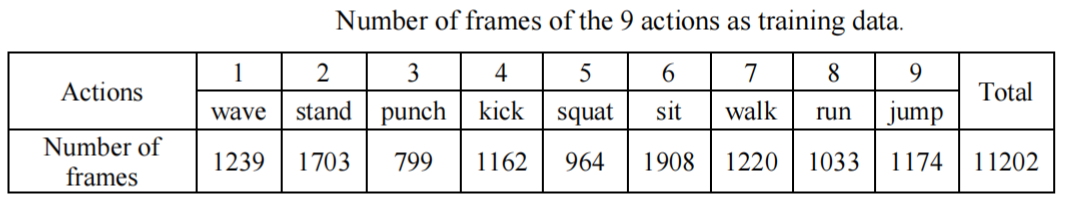
# 数据集介绍

## 3.1数据采集

本文的人体关节点获取部分，直接使用的OpenPose训练好的模型，因此本系统运行时使用笔记电脑本内置摄像头即可。关节点数据模型采用COCO数据集格式，其数据集格式说明如下：

0号位置是鼻子,1号位置是颈部，2号位置是左肩膀，5号位置是右肩膀，3号位置是左胳膊肘，6号位置是右肘，4号位置是左手腕，7号位置是右手手腕，8号位置是左髋关节，11号位置是右髋关节，9号位置是左腿膝盖，12号位置是右腿膝盖，10号位置是左脚踝，13号位置是右脚踝，14号位置是左眼，15号位置是右眼，16号位置是左耳，17号位置是右耳和18号背景信息点。通过OpenPose预先训练好的模型，获取这19个骨骼关节点的x轴信息，y轴信息，还有每一个点的置信度c，代表该关节点的识别准确程度。本系统在特征提取方面只需要x轴信息，y轴信息，没有加入置信度。

本项目收集了9种数据格式的视频数据分别是['stand','walk','run','jump','sit','squat','kick','punch','wave']，每个视频的长度从0.8秒到2分钟不等，并且每个视频仅限包含一种类型的操作。例如，在一个视频中，我踢了0.8秒；在另一个视频中，不停地挥舞着手臂长达2分钟。这些视频是以640x480的大小和10帧/秒的帧速率记录的，数据集分布如下表所示：



## 3.2数据集相关获取方式

KTH数据集：2004年发布，包含6类人体行为：行走、慢跑、奔跑、拳击、挥手和鼓掌，每类行为由25个人在四种不同的场景（室外、伴有尺度变化的室外、伴有衣着变化的室外、室内）执行多次，相机固定。该数据库总共有2391个视频样本。视频帧率为25fps，分辨率为160×120，平均长度为4秒。（官网：<https://www.nada.kth.se/cvap/actions/）>

Weizmann动作检测数据集：数据同样是固定镜头下的10个典型动作的视频，同时数据集提供了一些带有其他物体的动作作为干扰，可以测试模型的鲁棒性。官方同时提供了去除背景的程序，但是数据集的数据量比较少的90组常规数据和21组鲁棒测试数据，对于目前的模型训练来说显得有些不足，不过对于本来就需要用小数据的模型比如迁移学习或者One-shortLearning来说或许是适合的数据集。（官网：<http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html）>

HMDB51动作检测数据集：51个类别可以被分为5个大类：常见的面部动作(smile，laugh，chew，talk)、复杂的面部动作(smoke，eat，drink)、常见的肢体动作(climb，dive，jump)、复杂的肢体动作(brushhair，catch，drawsword)、多人交互肢体动作(hug，kiss，shake,hands)。

# 项目总结

在人体动作识别的小组项目中，每个成员各司其职，有的致力于数据集的整理，有的专注于模型的选择，还有的负责算法分支。通过各自的努力，我们克服了很多困难，如：如数据集的选择、动作类型的定义以及模型算法选择等问题，逐步完善和优化我们的人体动作识别系统。在与老师的一次次汇报交流中，我们逐渐明晰了我们的系统构建，将所收集到的理论知识与相关算法知识结合起来，运用到我们的系统中，从而解决大部分问题，在这个过程中，我学会了倾听和尊重他人观点的重要性，有效地表达自己的想法，以及如何协调协作。这个项目的研究使我学到了丰富的知识并成功应用，也培养了我独立思考和团队合作的能力，在一次次交流合作中成长。