第二章 人体行为识别和3D姿态估计研究概述

2.1 人体行为识别相关研究

2.1.1 传感器数据采集

2.1.2 特征提取与特征表示

2.1.3 人体行为识别方法的研究与比较

　 2.1.4 人体行为识别面临的挑战

2.2 3D人体姿态估计相关研究

2.2.1 视觉数据采集

2.2.2 特征提取与特征表示

2.2.3 3D姿态估计方法的研究与比较

2.2.4 3D姿态估计面临的挑战

2.3 本章小结

第三章 基于长短期记忆网络和卷积神经网络的人体行为识别

3.1 时间递归神经网络

3.2 长短期记忆网络

3.3 卷积神经网络

3.4 基于长短期记忆网络的人体行为识别模型

　 　3.4.1 模型框架

　　 3.4.2 网络结构设计

3.5 基于卷积神经网络的人体行为识别模型

　 　3.5.1 模型框架

　　 3.5.2 网络结构设计

3.6 实验过程与分析

　　 3.6.1 实验数据集

　　 3.6.2 实验环境与设置

　　 3.6.3 试验结果与分析

3.7 本章小结

第四章 基于深度卷积和长短期记忆神经网络的行为识别

4.1 基于深度卷积和长短期记忆神经网络的人体行为识别模型

　 　4.1.1 模型框架

　 　4.1.2 网络结构设计

4.2 实验结果与分析

　 　4.2.1 实验数据集

　　 4.2.2 实验环境与设置

　　 4.2.3 试验结果与分析

4.3 本章小结

第五章 基于前馈神经网络的3D人体姿态估计

5.1 前馈神经网络

5.2 基于前馈神经网络的3D姿态估计模型

　　 5.2.1 模型框架

　　5.2.2 网络结构设计

5.3 实验过程与分析

　　 5.3.1 实验数据集

　 　5.3.2 实验环境与设置

　　5.3.3 试验结果与分析

5.4 本章小结

第六章 总结与展望

6.1 总结

6.2 展望

参考文献

附录

致谢

**2.1 人体行为识别相关研究**

人类行为识别旨在，给出一组关于他或她自己和周围环境的前提下，识别一个人的行为。通过利用从各种资源检索而来的信息，例如环境[90]或穿戴的传感器[91,92]，来完成识别的任务。有些方法已经采用，在不同的身体部位佩戴专用的运动传感器，如腰部、手腕、胸部和大腿，实现良好的分类效果[93]。这些传感器通常令普通用户感到不舒适，所以，不能为行为监测提供长期的解决方案。

智能手机正在为以人为中心的应用带来新的研究机会，用户是场景信息丰富的数据来源，而手机是第一手的感应工具。最新的设备都带有嵌入式内置传感器，如麦克风、双摄像头、加速计、陀螺仪等。使用带有惯性传感器的智能手机，是人体行为识别的替代解决方案。这些大众化的设备为自动且全面地监控日常生活的行为（ADL）提供了一个灵活可负担得起同时提供电话服务的一体化解决方案。因此，在过去的几年中，一些研究工作提出，使用智能手机来实现理解人体行为。例如，采用Android智能手机上的嵌入式三轴加速度计来进行人体行为识别的始作俑者之一[94]；以及其他相关的报道[95,96]。

本节人体行为识别相关研究的内容包括：数据采集、特征提取、特征表示、识别方案与方法、以及当前人体行为识别面临的一些挑战和难题等，以下从各个方面详细描述。

**2.1.1 传感器数据采集**

人体行为识别实验中用到的数据集主要是利用智能手机上相关的加速度计和陀螺仪传感器收集而来。在数据采集的时候，一般手机会绑在收集人员或者实验志愿者的腰部或腿部。加速度计和陀螺仪信号已经被很多的研究和应用报道，对人体识别相关任务非常高效。因为，加速度计信号是一个三轴加速度计测定手机相对地球表面运动方向的改变值，由重力和身体运动部分组成；而陀螺仪信号是手机围绕某个轴向的旋转角速率值。

**2.1.2 特征提取与特征表示**

从传感器中收集而来的信号数据常常伴有噪音，所以一般要进行数据预处理，清除和过滤冗余和无用的特征或数据。通过应用噪声过滤器消除不需要的特征之后，这些数据被划分成2.5秒时间左右大小的滑动窗口，且滑动窗口之间的数据实现50%重叠(128读数/每窗口)。特征提取是在时域和频域上进行的。数据集中的每一行，即每个滑动窗口，共有561个具有时域和频域的特征向量、用户的行为标签和主题等。附录中的表1列出应用在时域和频域信号特征的所有度量。为了简化性能评估，数据集也被随机划分为两组，其中70％的数据被选择用于训练，其余30％用于测试。

**2.1.3 人体行为识别方法的研究与比较**

**2.1.4 人体行为识别面临的挑战**

**3.6 实验过程与分析**

**3.6.1 实验数据集**

本实验考虑选择两种公开数据集来训练和测试实验中的人体行为识别模型，最终来评估模型的表现效果。这两种数据集分别是Opportunity 行为识别数据集[84]和UCI人体行为识别数据集[85]，它们是不同场景下与人体行为相关的信号，在数据收集过程中都用到了三轴加速度计。这些信号数据通过一系列64个样本大小的滑动窗口分割而成，每个连续分割的窗口间的数据都有50%的数据分别与前后窗口重叠。之后，对这些数据实现0均值和标准方差归一化，归一化后的数据才能作为CNN等深度学习模型的输入。

1. Opportunity行为识别数据集

Opportunity行为识别数据集是关于家庭环境或厨房中早餐场景下的人体活动。该数据集使用了多个穿戴式传感器来收集64Hz频率下每天不同主题活动行为的信号，这些主题活动行为包括：开关冰箱、开关洗碗机、站着饮水、清洁、准备喝咖啡等，一共17种不同的行为，总共6个小时的记录。目前，该数据集可以在UCI机器学习库中下载和使用，而且已经有诸多第三方出版物[86-87]在使用。更重要的是，它还被做为大型公开行为识别竞赛的数据集，有大量的参与者使用不同的算法模型(如附录中表2所示)在运动模型识别中取得了不错的成绩。当然，现在可以从网站[88]下载这个公开的数据集。

1. UCI人体行为识别数据集

本实验一部分数据集来自UCI存储库的公开智能手机数据集。该数据集包含加速度计和陀螺仪的数据，由30名受试者进行6种不同的活动。这些活动包括：走路、上楼、走下楼、坐、站立和躺。21名随机受试者的数据为训练集，剩余9个数据用于测试。原始加速计和陀螺仪XYZ信号要进行均值为零的标准化(即减去均值后除以标准偏差)，对每个行为案例产生一组128个归一化值组成的向量。人体行为识别数据集已经被作为公共使用，并且呈现形式为原始的惯性传感器信号。

**3.6.2 实验环境与设置**

基于CNN架构的算法模型都是在基于Linux系统的服务器上运行，GPU配置为Nvidia Tesla M40，内存大小为64GB。

**参考文献**

[84] Chavarriaga R, Sagha H, Calatroni A, et al. The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition, Pattern Recognition Letters. 34(15):2033–2042, 2013.

[85] Anguita, Davide, et al. A Public Domain Dataset for Human Activity Recognition using Smartphones, in: ESANN, 2013.

[86] Zeng, M.; Nguyen, L.T.; Yu, B.; Mengshoel, O.J.; Zhu, J.; Wu, P.; Zhang, J. Convolutional Neural Networks for human activity recognition using mobile sensors. In Proceedings of the 6th IEEE International Conference on Mobile Computing, Applications and Services (MobiCASE), Austin, TX, USA, pp. 197–205, 6-7 November 2014.

[87] Gordon, D.; Czerny, J.; Beigl, M. Activity recognition for creatures of habit. Pers. Ubiquitous Comput. 2014, 18, 205-221.

[88] Opportunity Dataset. 2012. Available online: <https://archive.ics.uci.edu/ml/datasets/OPPORTUNITY+Activity+Recognition>, 19 November 2015.

[89] Yang, J.B.; Nguyen, M.N.; San, P.P.; Li, X.L.; Krishnaswamy, S. Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition. In Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI), Buenos Aires, Argentina, pp. 3995–4001, 25-31 July 2015.

[90] R. Poppe. Vision-based human motion analysis: An overview. Computer Vision and Image Understanding, 108(1-2):4-18, 2007.

[91] P. Lukowicz, J.A. Ward, H. Junker, M. Stäger, G. Tröster, A. Atrash, and T. Starner. Recognizing workshop activity using body worn microphones and accelerometers. Proceedings of the 2nd Int Conference Pervasive Computing, pages 18-22, 2004.

[92] D.M. Karantonis, M.R. Narayanan, M. Mathie, N.H. Lovell, and B.G. Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Transactions on Information Technology in Biomedicine, 10(1):156-167, 2006.

[93] R. Nishkam, D. Nikhil, M. Preetham, and M.L. Littman. Activity recognition from accelerometer data. In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence, pages 1541-1546, 2005.

[94] J.R. Kwapisz, G.M. Weiss, and S.A. Moore. Activity recognition using cell phone accelerometers. SIGKDD Explorations Newsletter, 12(2):74-82, 2011.

[95] T. Brezmes, J.L. Gorricho, and J. Cotrina. Activity recognition from accelerometer data on a mobile phone. Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living, pages 796-799, 2009.

[96] W. Wu, S. Dasgupta, E.E. Ramirez, C. Peterson, and G.J. Norman. Classification accuracies of physical activities using smartphone motion sensors. Journal of Medical Internet Research, 14(5), 2012.

**附录**

表 1. 特征向量计算的度量

|  |  |
| --- | --- |
| 函数 | 描述 |
| mean | 平均值 |
| std | 标准差 |
| mad | 绝对值中位数 |
| max | 数组最大值 |
| min | 数组最小值 |
| sma | 信号大小范围 |
| energy | 均方和 |
| iqr | 四分差 |
| entropy | 信号熵值 |
| arCoeff | 自回归系数 |
| correlation | 相关系数 |
| maxFreqInd | 最大频率分量 |
| meanFreq | 频率信号加权平均 |
| skewness | 偏频信号 |
| kurtosis | 频率信号峭度 |
| energyBand | 频率间隔的能量 |
| angle | 向量夹角 |

表 2 . 基准分类方法在Opportunity数据集上的运用

|  |  |
| --- | --- |
| Method | Description |
| LDA | 线性判别分析。基于正态分布特征的高斯分类器，且所有类具有相同的协方差矩阵。 |
| QDA | 二次判别分析。与LDA类似，这种技术也基于正太分布的特征，但类协方差可能不同。 |
| NCC | 最近邻分类器。测试样本与质心的欧氏距离，每一类样本都用于分类。 |
| 1NN | K近邻算法。计算测试样本和训练样本之间欧氏距离的懒惰算法，最频繁发生的k-最近样本是输出。 |
| 3NN | 与1NN类似，使用3个近邻。 |
| UP | Parma大学提交至OPPORTUNITY竞赛。使用均值、方差、最大值和最小值的模式比较。 |
| NStar | 新加坡大学提交至OPPORTUNITY竞赛。使用单一近邻和归一化数据的kNN算法。 |
| SStar | 新加坡大学提交至OPPORTUNITY竞赛。使用标量数据的SVM算法。 |
| CStar | 新加坡大学提交至OPPORTUNITY竞赛。使用最近邻和支持向量机的融合kNN算法。 |
| NU | Nagoya大学提交至OPPORTUNITY竞赛。使用均值、方差和熵值的C4.5决策树算法。 |
| MU | Monash大学提交至OPPORTUNITY竞赛。决策树嫁接算法。 |
| CNN | Yang[89]等提供的结果。对主题1、2、3使用平均性能计算的值。 |