

基于神经网络的行人步态识别系统研究

苏航 汤武惊

(中山大学深圳研究院, 广东 深圳 518057)

摘要: 步态识别作为一种核心的识别方法, 其识别精度和成本直接影响该技术在相关领域的监测效率和结果, 现有的关于步态分析的研究仍存在识别成本高、识别精度低等缺点, 该文在已有研究的基础上提出了基于步态的行为识别方法, 并基于该方法开发出一套简洁、灵活的图像数据采集系统, 利用傅里叶变换、四分位差等理论降低步态特征中的维度并简化行为参数的提取步骤, 并通过应用装置及模式识别试验进行验证。试验结果验证了该方法的有效性, 该方法识别精度提高至98.7%, 机器学习性能度量 Equal Error Rate (EER) 低至1.96%, 在提高识别精度的同时有效提高了计算效率。

关键词: 物联网; 步态识别; 特征融合; 终端设备

中图分类号: TP 391

文献标志码: A

DOI:10.13612/j.cnki.cntp.2022.21.028

0 引言

随着互联网技术的不断发展, 物联网技术的应用及发展成为当前不可阻挡的潮流, 行为识别作为底层部分, 在网域步态分析流程中起着关键作用。在行为识别过程中, 要求计算机能够协助用户执行多种类型的识别操作, 以提高用户的处理效率, 即在进行特征数据采集的同时可以对目标对象进行行为识别。当前, 人工智能领域的行为识别技术主要使用光流信息来确定目标对象在图像帧中的时间信息以及空间信息, 从而解析目标对象的行为状态, 因此计算量较大且过程较复杂, 该处理流程大大提高了运算设备的计算成本且还

会降低识别效率。基于以上弊端, 该文提出了一种基于步态的行为识别方法, 并在此基础上提供了相关的终端设备及储存介质, 以解决现有行为识别研究中存在的成本高且效率低的问题, 该方法将目标帧特征数据导入池化融合网络中, 并通过神经网络等算法输出目标图像数据对应的特征数据, 再确定所述目标视频数据中目标对象与环境对象之间的相互位置关系, 进而确定特征的步态行为数据^[1]。

1 研究方法

该文的步态识别数据分析流程如图1所示, 由于采用以上方法需要确定步态数据的时间信息和空间信息, 因此需

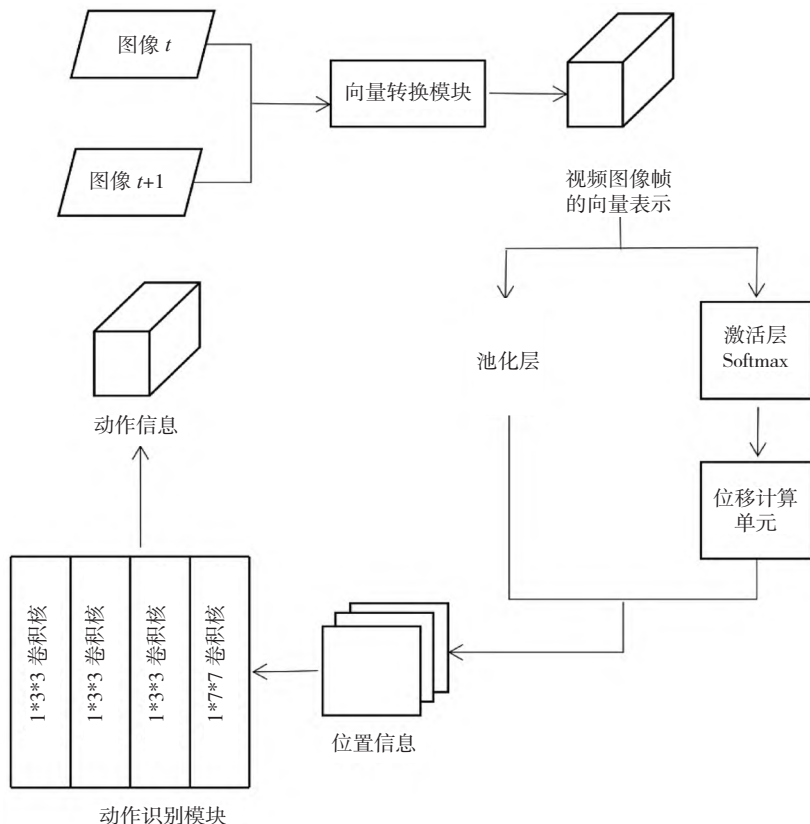


图1 动作特征提取流程图

基金项目：深圳市中央引导地方科技展专项资金资助项目“复杂背景下步态识别算法研究”（项目编号：2021Szzup173）；深圳市科技计划资助“重2022021 基于边缘计算与人工智能的智慧城市多维感知预警系统研发”（项目编号：JSGG20210802153003011）。

要使用预设的帧间动作提取网络,以确定相邻图像间的动作特征信息;需要使用池化融合网络,以确定目标特征对应的融合特征数据,这就使得到的数据为混有其他无关信息的初始数据,要精确识别目标对象的步态特征信息,就需要从中找出较稳定且能够表征步态的特征信息,在提取特征前,需要对初始数据进行筛选,先从中去除可能造成影响的干扰信号,例如噪声、杂波以及光晕等影响识别精度的无关信息,在进行识别前,对其进行预处理,再开始提取目标特征。因此需要通过特征组合,并使用神经网络算法,最后通过模式分类识别特征。

1.1 特征融合

目前,绝大多数基于步态的行为识别方法均需要进行特征分类,该文的帧间动作特征数据提取单元可将所述目标视频数据导入预设的帧间动作提取网络,从而得到帧间动作特征数据。所述帧间动作特征数据用于确定所述目标图像数据中相邻的目标对象图像帧之间的动作特征信息。在确定目标对象的动作特征后,需要确定特征点坐标对应的空间信息,即位移数据,该文使用的行为识别方法如下:根据所述位移相关矩阵确定各个所述特征点坐标在所述2个连续的图像帧间的最大位移距离,并根据所得到的最大位移距离确定目标对象的位移矩阵,该方法的原理是利用所描述特征点的最大位移量建立目标物体在2D平面上的位移场,并通过激活 Softmax 对所述位移场进行池化降维,得到一维置信度张量,最后对得到的一维置信度张量进行融合,构建用于表达三维空间的位移矩阵,而融合特征数据单元将所述帧间动作特征数据导入池化融合网络,输出所述目标视频数据对应的融合特征数据。该文使用的特征组合流程如图1所示。

该动作特征提取流程主要分为4个步骤:1)设定原始特征空间。在初始状态下,对特征原始数据设定对应空间($X=\{x|t, t+1\}$),该空间内包括动作提取网络的输入数据为2个单独的视频图像帧,即图像 t 以及图像 $t+1$ 。2)特征向量转换。 X 空间中包括的2个视频图像帧是帧序号相邻的2个视频图像帧,电子设备可以通过向量转换模块对上述2个视频图像帧进行向量转换,将其转换为用图像帧向量模式表示的信息。3)池化降维处理。通过池化层对转换后的向量模式信息进行降维处理,并采用激活层和位移计算模块确定2个视频图像帧对应的向量标识之间的位移信息,从而通过动作识别单元确定2个视频图像帧间的动作信息。4)步态特征识别。提取降维后得到位移和动作信息,即2个二维坐标点,采用4层卷积核过滤杂波等干扰信息,从而进行二维步态识别。其中,动作识别单元具体可以由多个卷积层构成(图2),可以包括基于 $1*7*7$ 的卷积核构成的第一卷积层、基于 $1*3*3$ 的卷积核构成的第二卷积层、基于 $1*3*3$ 的卷积核构成的第三卷积层以及基于 $1*3*3$ 的卷积核构成的第四卷积层。

由于上述帧间动作识别流程由各提取模块间各迭代求解输出,因此获得的各个动作特征信息是离散的,为了在后面流程中便于动作特征识别,需要在此基础上对特征数据进行特征融合,该文提供的终端设备可以将帧间动作特征数据导入上述池化融合网络中,以进行池化降维处理和特征融合操作,从而输出对应的融合特征数据。其中,特征融合的数据

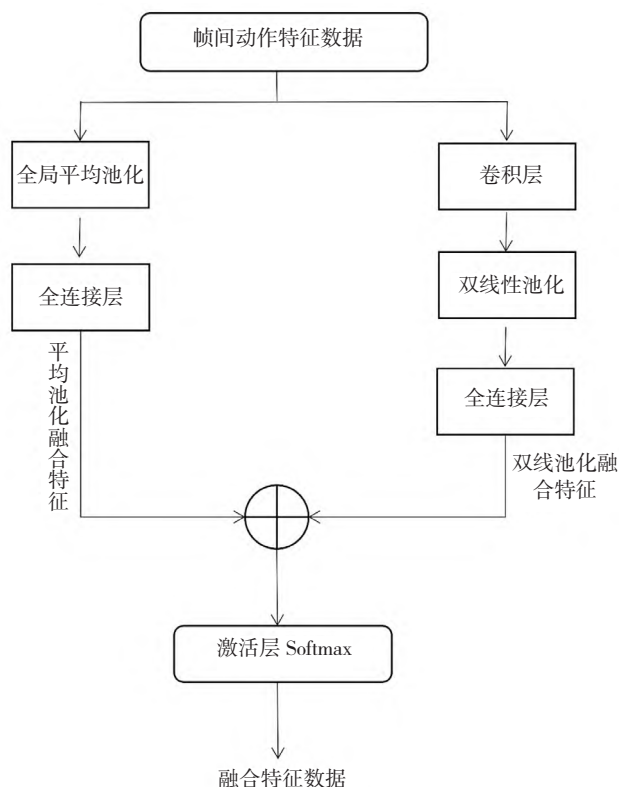


图2 模式分类流程图

据计算方式如公式(1)所示^[2]。

$$Maxpool = \sum_{i=1}^N Action_i \cdot Action_{i+1}^T \quad (1)$$

式中: $Maxpool$ 为融合特征数据; $Action_i$ 为第 i 个图像帧对应的动作数据信息; N 为所述目标视频数据中的总帧数。

1.2 模式分类

现有研究特征识别流程中的模式分类大多数使用神经网络算法和贝叶斯网络^[3],该文在此基础上,采用同源双线性池化网络,其具体计算流程如下:通过计算不同空间位置特征的外积,从而生成1个对称矩阵,然后再对该矩阵做平均池化,以获得双线性特征,它可以提供比线性模型更强的特征表示,并且可以以端到端的方式进行优化。传统的全局平均池化(GAP)只捕获一阶统计信息,而忽略了对行为识别有用的更精细的细节特征,针对这个问题,拟借鉴细粒度分类中所使用的双线性池化方法与GAP方法融合^[4],使对相似度较高的行为可以提取更精细的特征,从而得到更好的识别结果,进而得到所述目标对象的行为类别。模式分类流程如图2所示。

该文采用双向端与端间的训练过程并结合神经网络算法,可以在一定程度上降低模式分类过程中的训练难度,其最终的计算结果可以通过激活函数(Activation Function)输出。

2 数据采集

该文根据所提方法并结合当前时常需求提供了一种终端设备,其包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,所述处理器执行所述

计算机程序时实现数据采集、特征融合以及数据库训练的功能(即最终的步态识别系列过程)。

2.1 硬件系统

该文基于识别方法提供的程序系统可以应用于智能手机、服务器、平板电脑、笔记本电脑、超级移动个人计算机(Ultra-Mobile Personal Computer, UMPC)以及上网本等能够对视频数据进行行为识别的终端设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,其特点是当处理器运行该电脑程序时,能够满足各种权利需求,其核心组件如图3所示。

通过该硬件系统采集的数据会经过其核心的算法进行特征、模式分类计算并拟合,终端设备在得到拟合后的特征数据后,会将其导入链接层,确定拟合数据与行为类别训练数据库间的置信度,最后选取置信度最高的特征数据作为该目标对象的行为识别结果,由于该过程存在的行为识别数据长度足够长,因此该识别数据中包括多个类型的动作特征,会导致计算时效增长,从而导致识别结果严重滞后,而该文提供的终端设备可以根据各个行为特征出现的先后次序,按时间信息输出1个行为序列,该序列包括多个元素,而其中的每一个元素都会对应一个行为类别^[5],以此类推,可以高效识别目标对象的步态行为信息,其系统结构如图4所示。

现有研究中存在的3D卷积神经网络的末尾使用的全局

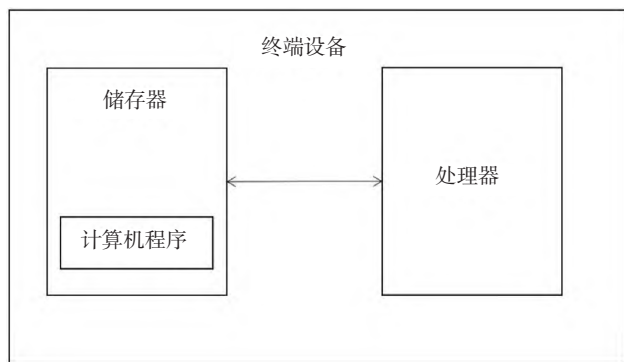


图3 硬件设备结构图

平均池化层在一定程度上影响时间信息的丰富性。针对该问题,该文选择深度双向转换器(Transformer)来替代全局平均池化。从输入视频中采样的 K 帧通过3D卷积编码器编码,得到的特征图(Feature Map)在网络的最后不使用全局平均池化,而是将特征向量分割成固定长度的tokens序列^[6],再将学习的位置编码添加到提取的特征中,以保证特征位置信息得以保存。融合位置编码后利用深度双向转换器中的Transformer块对时间信息进行建模,通过深度双向转换器的多头注意力机制得出的特征向量融合了时间信息,然后将这些向量进行连接,通过多层感知机进行特征维度变换,再通过计算对比损失完成端到端的训练^[7]。最终得到一个泛化性能良好的预训练模型。

2.2 数据采集试验

该文的数据采集试验选取人的3种行为步态信息(漫步、原地踏步以及跑步)进行识别,并征集60名志愿者参与试验,分别为适龄儿童10名、中青年10名以及老年人10名(男女各30人),保障人数均衡且覆盖面足够广泛,为了保证识别结果的准确性,分别选择试验场地的白天(光线足够)、晚上(光线暗淡)2个时间点进行数据采集,如图5所示。

图5(a)为步态特征映像图,利用机器识别采用卷积神经网络、sobel算子进行边缘检测识别计算,并针对平面卷积进行差分近似计算;图5(b)展示了整个计算过程的耗时,即经过池化层降维处理后的特征识别计算效率提高。试验设计了3种步态模式,共设置2组,每组各30人,每种模式重复采集10次,形成1800组数据样本,且每次步态模式转换时需要志愿者停顿2s~3s,以便设备切换并提取不同步态的模式信号,将所有的数据整理好,建立一个目标对象的步态数据库,数据采集如图6所示。

3 结果与分析

通过以上试验对形成的数据样本进行检索识别,从识别结果上可以看出,该识别装置对3种步态识别精度为(适龄儿童女/男)97%/98.1%、(中青年女/男)98%/98.6%以

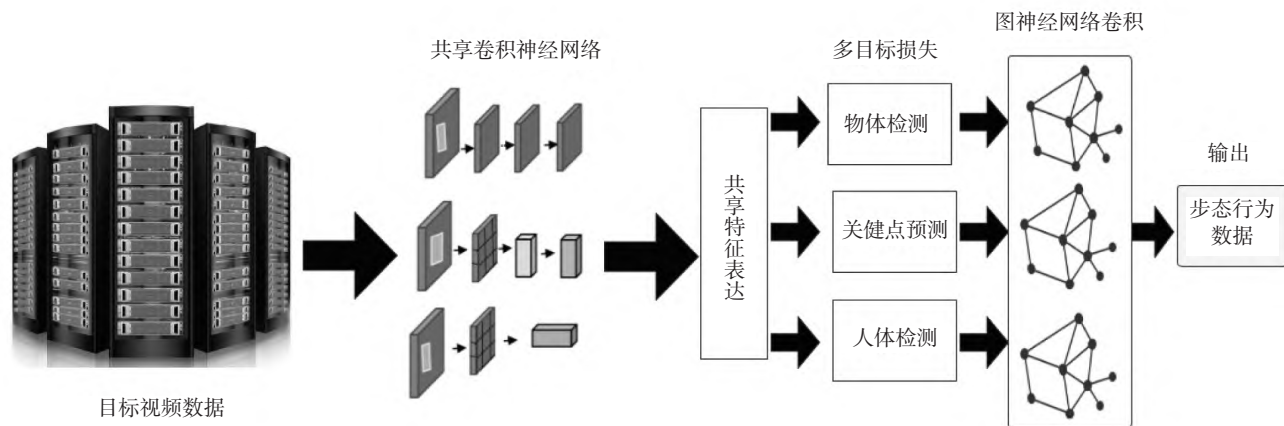


图4 系统结构示意图



(a) 步态特征映像图

```
1 测试完成,耗时: 0:04:20.835781
2 ===Rank-1 (Exclude identical-view cases) ===
3 NM:82.883,BG:71.161, CL:51.159
4 ===Rank-1 of each angle (Exclude identical-view cases) ===
5 NM:[76.74 85.82 90.92 87.22 81.63 74.80 78.98 88.27 92.55 84.29 70.51]
6 BG:[67.96 79.90 81.14 73.44 68.98 59.59 64.18 76.02 80.31 72.89 58.37]
7 CL:[46.63 59.59 60.20 58.77 52.14 44.29 49.80 54.80 51.63 48.88 36.02]
```

(b) 特征识别耗时图

图5 步态数据采集处理



图6 步态采集数据解析示意图

及(老年人女/男)98.8%/98.5%,比现有的采用贝叶斯算法BayesNet 识别方式高,详细分析结果见表1。

表1 数据采集结果分析表

组别	漫步	原地踏步	跑步	时间/s	容错率/%
1	10	6	5	10.0/7.0/7.0	2.00
2	12	8	4	12.0/8.5/6.0	2.90
3	15	9	7	15.0/9.0/7.0	1.00
参照				14.6/9.9/7.6	2.46

注:参照为参照现有贝叶斯法识别步态检测时间。

在表1中,组别1为适龄儿童,组别2为中青年,组别3为老年人群体。由于试验分为3种年纪的6组试验,因此为了让结果更具通用性,使用同一组男女识别数据的平均值,且取其中10组的值进行计算。

4 结语

根据试验结果可知,该方法具有可行性,且使用该方法集成的数据采集系统简洁、灵活,识别精度提高至98.7%,机器学习性能度量低至1.96%,在提高识别精度的同时有效提高了计算效率。

该文基于现有研究存在的问题提供了一套检测识别方法,并基于该方法提供了1套终端设备,通过试验证明了该方法及试验设备的可行性和灵活性,在一定程度上为该领域的后续研究提供借鉴,但是该文局限于试验数据样本较少,存在训练结果与识别结果有较小偏差的情况,因此需要在后

续研究中基于多种场景、各类群体以及多种气象条件进行数据采集及样本训练,以提高设备识别精度,扩大方法的实际应用前景。

参考文献

- [1] 王孝元.一种基于深度神经网络与机器视觉的步态识别方法及系统:CN110909612A[P].2020-03-24.
- [2] 王凯,孙咏梅,张泓,等.体域网中基于特征组合的步态行为识别[J].中国科学:信息科学,2013,43(10):1353-1364.
- [3] 李超.基于卷积神经网络的人体行为分析与步态识别研究[D].杭州:浙江大学,2019.
- [4] 马昕,田皓宇,李贻斌.一种基于广义卷积神经网络的骨架异常步态识别方法及系统:CN114582023A[P].2022-06-03.
- [5] 靳丹,刘晓光,刘英挺,等.基于特征组合的步态行为识别方法[P].中国专利:CN103886341A[P].2014-06-25.
- [6] Wang Xing. Application and Technology Analysis of Internet of Things in Logistics Warehouse Management System[J]. China Computer & Communication, 2018, 40(2): 4..
- [7] 钱兴,张晓明.基于轻量卷积神经网络的步态识别系统[J].中北大学学报(自然科学版),2022,43(5):474-480.
- [8] 连旭.基于步态特征的车辆前方行人识别系统研究[J].时代汽车,2017(16):89-91.