

DOI:10.19651/j.cnki.emt.2211121

基于堆叠惯性信号的跳台滑雪动作识别^{*}鲍文霞¹ 董震¹ 王年¹ 杨先军²

(1. 安徽大学电子信息工程学院 合肥 230039; 2. 中国科学院合肥物质科学研究院 合肥 230031)

摘要: 动作识别是滑雪运动状态智能监测的关键环节之一。本文以跳台滑雪运动为研究对象,通过堆叠的方式将不同惯性传感器、不同关节的数据进行融合生成结构化的数据,从而利用深度卷积神经网络实现跳台滑雪动作的识别。首先对采集到的跳台滑雪运动过程中的不同传感器、人体不同关节的惯性传感数据进行归一化处理映射至 $[0,1]$ 之间,然后通过颜色映射将各类数据堆叠生成图像,接着利用 Resnet 等二维卷积神经网络对跳台滑雪中的动身至助滑、直线助滑、曲线助滑、起跳及早期飞行、稳定飞行及落地共 5 类动作的堆叠惯性信号图像进行识别。实验结果表明,对 9 次跳台滑雪数据融合后生成的 2 250 幅堆叠惯性信号图像进行识别,召回率和准确率达到 93.8% 和 91.7%;同时分析了单个类别惯性传感器对各关节数据融合后的识别结果的影响。本文提出的不同传感器、不同关节堆叠惯性信号融合和动作识别方法能够为跳台滑雪运动的智能化分析提供支撑。

关键词: 跳台滑雪;动作识别;卷积神经网络;惯性数据

中图分类号: TP212.6 **文献标识码:** A **国家标准学科分类代码:** 510.1099

Action recognition of ski jumping based on stacked inertial signals

Bao Wenxia¹ Dong Zhen¹ Wang Nian¹ Yang Xianjun²

(1. School of Electronic Information Engineering, Anhui University, Hefei 230039, China;

2. Hefei Institute of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

Abstract: Motion recognition is one of the key links in the intelligent monitoring of ski jumping. This paper takes ski jumping as the research object, and fuses the data of different inertial sensors and different joint nodes to generate structured data by stacking, to realize the recognition of ski jumping movements by using deep convolutional neural network. Firstly, the collected inertial sensing data of different sensors and different human body points in the process of ski jumping are normalized and mapped to between $[0,1]$. And then using color mapping to stack all kinds of data to create an image. Then use two-dimensional convolutional neural networks such as Resnet, to identify 5 types of movements in ski jumping: start-to-slip, straight-line assist, curve-assist, take-off and early flight, stable flight and landing. The experimental results show that the 2 250 stacked inertia signal images generated by 9 times of ski jumping data fusion are recognized, and the recall rate and accuracy are 93.8% and 91.7%, respectively. At the same time, the influence of a single class inertial sensor on the recognition result of the fusion data of each joint node is analyzed. The proposed method of stacking inertial signal fusion and action recognition of different sensors and different joints can provide support for intelligent analysis of ski jumping.

Keywords: ski jumping; action recognition; convolutional neural network; inertial data

0 引言

让科技服务于人类一直是人工智能发展的重要指导思想。人体动作识别作为人工智能领域的热门研究方向之一,对于评估人体状态,解析人体运动有着重要意义。按照传感器的不同,人体动作识别一般可以细分为基于 RGB

视频的动作识别^[1-3]、基于深度图序列的动作识别^[4-6]以及基于可穿戴惯性传感器的动作识别^[7-8]等。每一种传感方式都有其独特的优点和应用条件。摄像机为特定视场提供丰富的视觉动作信息,但是它们的图像数据因为场景角度及照明等具有局限性。深度相机利用红外光为特定视野提供 3D 深度动作信息,但是深度相机只能部署在室内。可

收稿日期:2022-08-19

^{*} 基金项目:国家重点研发计划课题项目(2020YFF0303800)资助

穿戴惯性传感器需要佩戴在身体上,它们以所佩戴关节部位的人体质心、加速度等信号的形式提供运动姿态信息,而没有视角及场景的限制,可靠稳定。可穿戴惯性传感器具多种类型,用途广泛,目前已被广泛应用于人体动作识别^[9-10]。基于可穿戴惯性传感器的独特优点,我们利用惯性传感器来对跳台滑雪动作识别方法进行研究。

利用惯性传感数据进行动作识别,首先一般需要对所采集到的原始数据进行滤波、对齐等预处理操作,然后人工设计时频域特征选择,再利用传统的机器学习的方法进行动作识别^[11]。相比于传统的机器学习算法,深度学习方法在进行识别任务时可以自动的从数据中学习更高级的特征,减少了为每个问题开发新的特征抽取的任务,并且能够端到端的完成识别任务。但现有的深度学习模型大都只能接收结构化数据,而非结构化的几何形态学数据则很难直接输入神经网络。基于此,本文提出一种基于堆叠惯性信号的跳台滑雪动作识别方法,对跳台滑雪运动过程中的惯性传感数据进行映射和堆叠生成结构化的二维图像,然后再利用深度卷积神经网络进行动作识别。该方法在实现不同传感器和关节数据融合的同时也获得了适合深度卷积神经网络的输入数据,为滑雪运动智能分析提供了一种新的途径。

1 数据获取

跳台滑雪是以滑雪板为工具,在专设的跳台上通过助滑坡获得速度,比较空中飞行距离和动作姿势的一种雪上竞技项目^[12]。本文所用惯性数据均由惯性测量单元所采集,在每个惯性测量单元中均包含 3D 加速度计、3D 陀螺仪和 3D 磁力计,每个组件构成了一个微机电系统(MEMS)。其中,加速度计用来测量加速度、或速度(速率)相对于时间的变化;陀螺仪对围绕一个平面的姿态角、以及角速度(旋转速率)进行测量;磁力计测量给定位置的磁场方向和强度,与罗盘指针的工作原理相仿,可指示沿地平面的航向角或方向^[13-14]。将惯性测量单元佩戴在运动员跳台滑雪过程中的重要关节处进行数据采集,本文主要对微机电系统进行复杂的非线性运算所得到的数据类型中的 3 类惯性数据进行实验,分别是:运动员的质心数据、各关节加速度数据以及各关节角速度数据。

数据采集地点为吉林北大壶滑雪场,采集时间为 2021 年 2 月 26 日,采集场地等级为 K50,雪道结构为 K90,采集现场如图 1 所示。所采集数据为 4 名专业跳台滑雪运动员进行 9 次跳台滑雪的数据,每人的起跳次数分别为 2 次、3 次、3 次和 1 次。

所采集的 3 类惯性数据的采样频率为 240 Hz,数据以 Excel 文件形式保存。在 3 类惯性数据中,质心类型数据含有 X、Y、Z 三维数据;加速度类型数据含有:T8、T12、L3、L5、LLL(left lower leg, LLL)、LF(left foot, LF)、RLL(right lower leg, RLL)、RF(right foot, RF)共 8 个关节,



图 1 采集现场图像

角速度类型数据含有:T8、T12、L3、L5、LK(left knee, LK)、LA(left ankle, LA)、RK(right knee, RK)、RA(right ankle, RA)共 8 个关节,每个关节都含有三维数据,各关节具体位置如图 2 所示。在三维坐标下,人体正前方为 X 轴正方向;Y 轴根据右手定则,指向侧面,与 X、Z 轴正交;Z 轴为重力竖直方向,指向上方时为正。

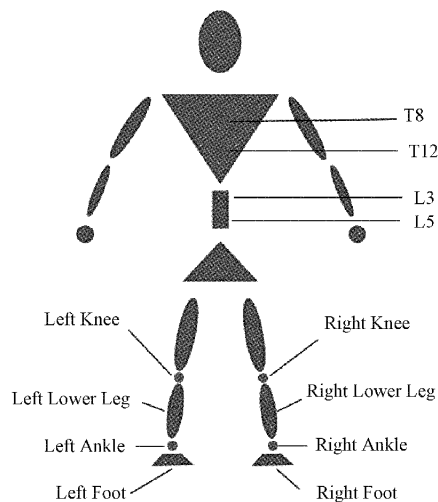


图 2 佩戴惯性测量单元的关节具体位置

将所采集原始数据中不属于跳台滑雪运动阶段的多余数据去除,对运动阶段数据进行 5 个动作阶段的划分,并进行归一化处理。本文中跳台滑雪的 5 个动作阶段分别为:动身至助滑阶段、直线助滑阶段、曲线助滑阶段、起跳及早期飞行阶段、稳定飞行及落地阶段。

2 基于堆叠惯性信号图像的动作识别方法

2.1 堆叠惯性信号图像的生成

对利用运动员 9 次跳台滑雪所采集到的惯性类型数据进行 5 个动作阶段划分,对于每次跳台滑雪,各动作阶段含有的数据帧数范围都在 350~600 帧之间。为了保留原始数据的运动规律并且保证送入卷积神经网络中的图像数量,对每个动作阶段按规律随机抽取 100 帧数据,例如每两帧去除一帧或每 5 帧去除两帧等,每跳数据的每个阶段重复操作 50 次,对 9 次跳台滑雪数据文件按同样的操作进行处理,共得到 $50 \times 5 \times 9 = 2\,250$ 份全类型数据文件,后续将文件按照 3 种数据类型进行拆分,共得到 $2\,250 \times 3 = 6\,750$

份数据文件。

对得到的数据文件进行归一化处理。所采集原始数据的取值在整个运动阶段内连续变化,变化区间并不集中,为了保留原始数据间的关系和规律并能够进行后续处理,我们对采集到的9次跳台滑雪数据进行归一化处理。

下面公式中, X^* 为对原始数据进行归一化处理后的数据, X 为原始数据, X_{max} 和 X_{min} 分别为原始数据集中每种数据每个维度数值中的最大值和最小值。对原始数据进行归一化处理,原始数据的数据值的范围映射到了 $[0,1]$ 之间,在保留了原来数据中存在的大小关系的同时,消除了量纲和数据取值范围的影响。

$$X^* = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

在对数据进行归一化处理将数据映射到 $[0,1]$ 之间以后,利用不同颜色的渐变表示不同数值的大小,其中最小值用蓝色表示,最大值用红色表示,中间刻度即0.5数值用黄色表示,利用颜色的渐变表示数值的变化。

图3所示为:将质心、加速度、角速度3类惯性数据进行数据级融合的动身至助滑阶段的堆叠惯性信号图像;图4所示为:速度类型数据全关节动身至助滑阶段的堆叠惯性信号图像。

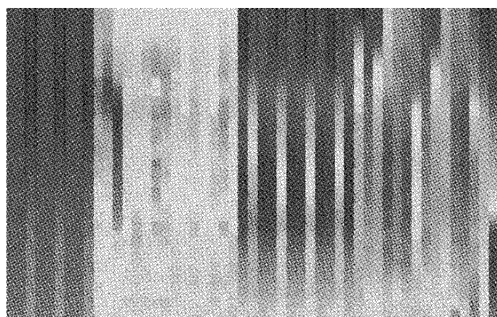


图3 3类惯性数据动身至助滑阶段堆叠惯性信号图像

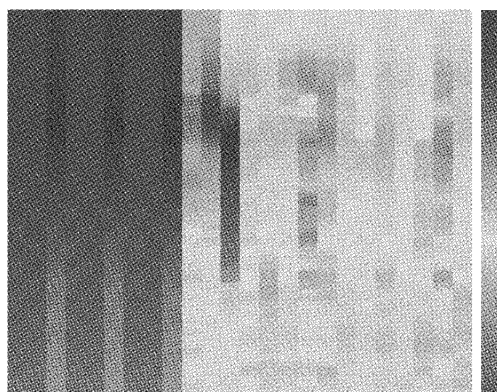


图4 速度类型数据动身至助滑阶段堆叠惯性信号图像

对9次跳台滑雪数据都进行变换处理。9次跳台滑雪数据共得到2250份含有3类惯性数据的文件,转换成2250幅3类惯性数据进行数据级融合的堆叠惯性信号图

像;将文件按照3种数据类型进行拆分后,各类型惯性数据分别得到2250份数据文件,转换共生成 $2250 \times 3 = 6750$ 幅堆叠惯性信号图像。对9次跳台滑雪数据进行处理生成的图像分类保存,共得到 $2250 + 6750 = 9000$ 幅数据图像,其中3类惯性数据进行数据级融合的堆叠惯性信号图像像素大小为 6275×1900 ,单类惯性数据生成的堆叠惯性信号图像像素大小为 2950×1900 。

2.2 基于卷积神经网络的动作识别

将惯性传感数据转换成堆叠惯性信号图像后,就可以利用基于图像分类的深度卷积神经网络(CNN)对跳台滑雪动作进行识别,例如深度残差网络ResNet34。ResNet34的特点是能够通过增加相当的深度来提高准确率,其网络大部分是由图5所示的残差单元组成,其内部使用了跳跃连接(shortcut),即图5中标注“X identity”的曲线,缓解了深度神经网络中增加深度带来的性能退化问题,使得深层网络的性能优于浅层网络。许多参数相同的残差单元连接起来组成残差块,而许多残差块组合起来再加上预处理层和最后的全连接分类层就组成了ResNet网络。

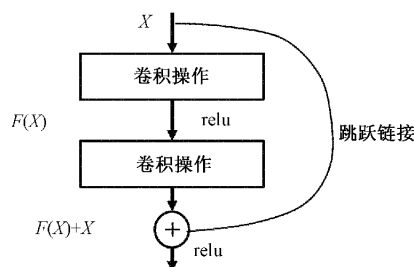


图5 Resnet网络中的残差单元

在传统CNN中的一个网络单元中,网络单元的输入为 X ,输出为 $H(X)$,即这个网络单元完成了非线性映射 H 。而ResNet残差单元的作用不是将输入特征 X 映射为 $H(X)$,而是将其映射为 $H(X) - X$,将 $H(X) - X$ 记作 $F(X)$,即该残差单元计算原映射与输入特征的差值 $F(X)$ 。图5中用2个卷积操作(weight layer)叠加来完成从输入特征 X 到残差 $F(X)$ 的映射,将 $F(X)$ 与上图中右侧的跳跃连接进行元素级别的加法操作,最终该网络块的输出为 $F(X) + X$,即上图中整个模块的输出仍为 $H(X)$ 。从上述操作可看出,跳跃连接的引入只在原来的基础上增加了元素加法操作,并没有引入大量的额外计算和可学习参数,同时保证了网络识别的精度。

实验所采用卷积神经网络模型内部所含总参数量(Params)以及浮点运算次数(FLOPs)大小,代表网络模型的计算空间复杂度及计算时间复杂度,是反映网络模型运算效率的关键指标。由表1可以看出Resnet网络具有较低的计算时间复杂度,而其内部特有的残差单元保证了其具有较高的识别精度。

最后将堆叠惯性信号图像以不同数据类型、不同运动阶段划分好,用第一名运动员S1的两跳数据、运动员S2的

表 1 各网络模型参数量比较

所用网络	总参数量/M	浮点运算次数/G
AlexNet	61.10	0.7
VggNet16	138.36	15.5
GoogleNet	6.8	1.6
MobileNetV2	3.4	0.6
ResNet	25.56	3.9

三跳数据、运动员 S3 的前两跳数据共 7 跳数据生成的 7 000 幅图像作为训练集;用运动员 S3 的第 3 跳数据、S4 的一跳数据共 2 跳数据生成的 2 000 幅图像作为预测集,送入卷积神经网络,得到动作阶段识别结果。

3 实验结果及分析

实验使用 python 作为编程语言,使用精确率(Precision)、召回率(Recall)、F1 得分(F1 score)以及准确率(Accuracy)作为评价指标来综合评价深度学习算法的性能:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1\ score = \frac{2TP}{2TP + FP + FN} \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

上述公式中,精确率代表正确预测的正样本数与所有预测为正样本的数量的比值,召回率代表正确预测的正样本数与真正样本总数的比值,精度和召回率是呈负相关的。F1 得分即是综合考虑精度和召回率的一个指标,是基于精度与召回率的调和平均定义的。准确率指的是正确预测的样本数与总测试样本数的比值,反映的是模型算法整体的性能,TP 为真阳性样本数;TN 为真阴性样本数;FP 为假阳性样本数;FN 为假阴性样本数。

将利用跳台滑雪数据所得到的 3 类惯性数据进行数据级融合及只含有单类惯性数据生成的 5 个动作阶段的堆叠惯性信号图像送入 ResNet34 等卷积神经网络中,根据每个动作阶段图像特征规律的不同,利用卷积神经网络提取图像特征进行动作识别,记录下跳台滑雪 3 类惯性数据进行数据级融合及单类惯性数据的动作识别精度。除用 ResNet34 卷积神经网络进行动作识别外,用 4 种经典卷积神经网络 AlexNet、VggNet16、GoogleNet、MobileNetV2 进行识别精度的比较。

对得到的分类结果进行分析,评估每种类型数据在每个运动阶段的特点及差异性,并对所得到的实验结果进行分析比较,分析各类型数据对整体识别精度的影响及原因。

3 类惯性数据进行数据级融合动作识别实验结果如

表 2 所示。由表 2 可以看出,对跳台滑雪 3 类惯性数据进行数据级融合生成堆叠惯性信号图像进行动作识别时,整体的识别精度偏低,主要原因是由于在跳台滑雪整个运动过程中不同类型数据的变化规律不同,各类型数据在每个运动阶段具有不同特征,某些数据类型各运动阶段的特征差异性大,而有些数据类型各运动阶段的变化规律区分度低,降低了整体的识别精度。

表 2 数据级融合数据实验结果

所用网络	精确率	召回率	F1 得分	准确率
AlexNet	84.1	86.5	85.3	89.1
VggNet	84.5	88.6	86.5	87.9
GoogleNet	83.9	75.3	79.3	90.0
MobileNet	86.4	84.0	85.2	92.2
ResNet	85.9	93.8	90.5	91.7

接下来对各类型惯性数据分别进行动作识别并进行分析。

质心数据动作识别实验结果如表 3 所示。同表 3 可以看出,5 种卷积神经网络对质心类数据不同动作的识别准确率范围为 78.5%~86.5%,主要原因是由于在跳台滑雪整个运动阶段中,相较于动身至助滑、起跳及早期飞行、稳定飞行及落地这 3 个动作阶段而言,直线助滑阶段至曲线助滑阶段的过程中运动员的动作变化幅度小,质心变化较小,因此会造成质心类型数据总体分类准确率的降低^[15-16]。

表 3 质心类型数据实验结果

所用网络	精确率	召回率	F1 得分	准确率
AlexNet	80.2	76.8	78.5	90.0
VggNet	78.5	76.8	77.7	90.0
GoogleNet	80.0	75.0	77.4	88.8
MobileNet	86.5	90.0	88.2	91.4
ResNet	80.7	86.6	83.5	90.3

速度数据动作识别准确度如表 4 所示。由表 4 可以看出,加速度类型数据利用 5 种卷积神经网络对各动作阶段进行识别时,整体的识别精度较高,分类识别准确率基本稳定在 88%以上。

表 4 加速度类型数据实验结果

所用网络	精确率	召回率	F1 得分	准确率
AlexNet	89.8	80.6	84.9	93.3
VggNet	99.1	97.7	98.4	95.0
GoogleNet	94.3	87.4	90.7	96.0
MobileNet	88.4	91.7	90.0	91.9
ResNet	92.3	86.6	89.4	95.2

主要原因是在跳台滑雪的整个运动过程中,动身至助

滑、直线助滑、曲线助滑、起跳及早期飞行以及稳定飞行及落地这5个动作阶段的关节加速度变化规律相差较大。例如在起跳及早期飞行阶段,运动员发力起跳时部分关节加速度数值大,而在飞行时动作幅度小加速度数值小,数值变化规律为先大后小;在直线助滑阶段,运动员主要通过助滑坡获取初速度,在坡道内动作幅度相对较小,加速度数值差值较小;而在稳定飞行及落地阶段,前半过程的加速度数值小,后半过程需要调整姿势进行落地,加速度数值大,运动变化规律和起跳及早期飞行阶段相反,差异性大^[17]。由于不同运动阶段的加速度数据差异性大,所形成的堆叠惯性信号图像相似程度低,使得阶段识别精度高。

由此可以看出,利用跳台滑雪运动员加速度类型数据对于5个分运动阶段的识别具有更高的精度。

角度数据动作识别准确度如表5所示。从表5可以看出,对角速度类型数据生成的堆叠惯性信号图像利用5种经典卷积神经网络模型进行阶段识别时,动作识别精度基本保持在50%~65%之间,整体精度不高,分类精度相较于速度类型数据偏低。

表5 角速度类型数据实验结果

所用网络	精确率	召回率	F1得分	准确率
AlexNet	54.1	59.6	56.7	70.5
VggNet	52.9	51.4	52.1	53.4
GoogleNet	50.7	67.6	58.0	64.8
MobileNet	60.0	57.7	58.8	57.5
ResNet	65.5	55.3	60.0	57.7

主要原因是在跳台滑雪的整个运动过程中,动身至助滑、起跳及早期飞行这2个阶段的角速度在段内前半部分变化较大,变化规律较相似,类间区分度低;而稳定飞行及落地阶段的角速度在段内后半部分变化较大,总体上导致了较低的分类准确率。因此,利用跳台滑雪运动员角速度类型数据进行5个动作阶段的识别的效率相对较低,识别精度低于速度类型数据。

对比各类型惯性数据利用卷积神经网络进行动作识别实验所用时长。质心类型数据因为只含有X、Y、Z三维数据,所含数据量少,实验用时最短;加速度及角速度类型数据分别含有8个关节点,每个关节点都含有三维数据,所含数据量相近,实验用时基本相同。在同一类型惯性数据中对比5种卷积神经网络实验所用时长,VggNet网络因为网络模型最复杂,所用时长最长,其它4种网络所用时长相近。

4 结 论

对采集到的跳台滑雪惯性数据中的3类惯性数据进行归一化及数据分割等处理,将一维惯性数据通过映射和堆叠从而转换成二维的惯性信号图像,实现了不同传感器、不同关节点惯性数据的融合,并利用Resnet等5种卷积神

经网络分别进行动作识别对比试验。实验表明基于堆叠惯性信号的跳台滑雪动作识别方法的精确率、召回率、F1得分以及准确率均取得了较好的结果。同时本文分析了单个类别惯性传感器数据识别结果,其中加速度类型数据及质心类型数据的识别精度最高,有利于提升融合后数据的识别精度;角速度类型数据的识别精度最低,对融合后识别精度的影响最大。根据本文的研究结论,在后续的研究工作中将进一步探讨不同惯性传感器加权融合的方式,进一步提高跳台滑雪动作识别的精度。

参考文献

- [1] 洪耀球. 基于深度信念网络的视觉人体动作识别[J]. 计算机科学, 2021, 48(S2): 400-403.
- [2] 龚捷, 罗聪, 罗琴. 基于注意力机制和残差网络的动作识别模型[J]. 电子测量技术, 2021, 44(14): 111-116, DOI: 10.19651/j.cnki.emt.2106848.
- [3] 赵挺, 曹江涛, 姬晓飞. CNN A BLSTM network 的双人交互行为识别[J]. 电子测量与仪器学报, 2021, 35(11): 100-107.
- [4] QIN X, CAI R, YU J, et al. An efficient self-attention network for skeleton-based action recognition[J]. Scientific Reports, 2022, 12(1): 4111-4111.
- [5] 宋轶航, 胡静, 徐超, 等. 基于深度信息的特征学习与动作识别方法[J]. 计算机应用研究, 2021, 38(11): 3446-3450, DOI: 10.19734/j.issn.1001-3695.2021.01.0067.
- [6] 唐心宇, 宋爱国. 人体姿态估计及在康复训练情景交互中的应用[J]. 仪器仪表学报, 2018, 39(11): 195-203, DOI: 10.19650/j.cnki.cjsi.J1803879.
- [7] MESHKAT, FARZANEH, ABDALI M, et al. A new in-air handwritten persian characters recognition method based on inertial sensor position estimation and convolutional neural network[J]. Journal of Ambient Intelligence and Humanized Computing, 2022: 1-16, DOI: 10.1007/s12652-022-03770-8.
- [8] HSIEH C Y, HUANG H Y, LIU K C, et al. Multiphase identification algorithm for fall recording systems using a single wearable inertial sensor[J]. Sensors, 2021, 21(9): 3302-3302.
- [9] 耿宏杨, 郇战, 梁久祯, 等. 面向复杂动作的运动状态转移识别模型[J]. 小型微型计算机系统, 2021, 42(11): 2323-2330.
- [10] 段晨, 廖福元. 基于惯性传感器和深度学习的上肢动作识别研究[J]. 国外电子测量技术, 2019, 38(11): 57-61, DOI: 10.19652/j.cnki.femt.1901608.
- [11] 吴建超, 刘昱. 基于脚部惯性传感数据的人员运动速度识别方法[J]. 传感技术学报, 2018, 31(12): 1828-1833.
- [12] ELFMARK O. Kinematic determination of the aerial phase in ski jumping[J]. Sensors, 2022, 22(2):

- 540-540.
- [13] 冉孟元,刘礼,李艳德,等.基于惯性传感器融合控制算法的聋哑手语识别[J].计算机科学,2021,48(2):231-237.
- [14] ZHU C, LUO L, MAI J, et al. Recognizing continuous multiple degrees of freedom foot movements with inertial sensors[J]. IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society, 2022; 30: 431-440.
- [15] 曹峰锐,伍颢,唐伟棣,等.跳台滑雪起跳和早期飞行阶段生物力学研究进展[J].北京体育大学学报,2022,45(1): 35-44, DOI: 10. 19582/j. cnki. 11-3785/g8. 2022. 01. 004.
- [16] ETTEMA G. Imitation jumps in ski jumping: Technical execution and relationship to performance level[J]. Journal of Sports Sciences, 2020, 38 (18): 1-6.
- [17] ALEXANDER J, WOLFRAM M, MANFRED S. Optimization of the flight technique in ski jumping: The influence of wind[J]. Journal of Biomechanics, 2019,88: 190-193.

作者简介

鲍文霞,博士,教授,主要研究方向为仪机器学习与图像视觉等。

E-mail:bwxia@ahu.edu.cn

董震,硕士研究生,主要研究方向为深度学习与人体姿态估计。

E-mail:13685693574@163.com