

一种基于双流融合3D卷积神经网络的动态头势识别方法

谢佳龙,张波涛,吕 强

(杭州电子科技大学自动化学院,浙江杭州 310018)

摘 要: 目前基于视觉的动态头势识别算法泛化能力弱、识别率低,头戴式传感器的方法经济性、便携性差.针对以上问题,提出了一种无需头戴设备的动态头势识别算法.这种基于双流融合3D卷积神经网络的方法用头部动作生成稠密光流,并将原始数据和光流数据并行输入构建的动作特征提取器,最后进行特征融合.结果表明所提算法比人工特征提取方法和C3D模型有更高的准确率、更好的泛化能力,在无需头戴传感器的情况下有近似头戴式传感器的识别率.

关键词: 深度学习; 机器视觉; 人机交互; 动作识别; 动态头势; 双流网络

中图分类号: TP183 **文献标识码:** A **文章编号:** 0372-2112 (2021) 07-1363-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20201183

A Dynamic Head Gesture Recognition Method Based on 3D Convolutional Two-Stream Network Fusion

XIE Jia-long, ZHANG Bo-tao, LÜ Qiang

(School of Automation, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China)

Abstract: Present vision based on dynamic head gesture recognition algorithms usually have disadvantages in generalization and recognition rate, and head-mounted sensors are expensive and inconvenient. In view of the above problems, a dynamic head gesture recognition algorithm without head-mounted sensors is proposed. Using this method based on two-stream 3DCNN(3D Convolutional Neural Network), the dense optical flow is generated by head movements, the original data and dense optical flow are put into the motion feature extractor in parallel, and finally, features are fused. Experimental results show that the proposed algorithm has higher recognition accuracy and better generalization than the artificial feature extraction and C3D(Convolutional 3D) methods, and its recognition rate is as good as those head mounted sensors.

Key words: deep learning; computer vision; human-computer interaction; action recognition; dynamic head gesture; two-stream network

1 引言

人机交互是实现用户和机器之间信息交换的渠道,目前手柄、触摸板、键盘等仍是常用人机交互方式^[1]. 肢体语言是通过人体部位的协调运动进行交互的一种沟通方式,其中头部动作在面对面交流中扮演着重要的角色,能携带特定的语义,传递副语言线索,如态度、情绪等^[2]. 在日常交流中,人们通常以“点头”表示赞同,“摇头”表示否定^[3,4]. 为使机器人更流畅地与人交流,必须从当前交互方式过渡到一种更加本能

和自然的交互方式^[5,6].

头势交互往往有着巨大的应用价值,可帮助残障人士控制轮椅、机器人等设备^[7],增加他们日常生活的自主权^[8]. 在虚拟现实,可提升人机交互的准确性与智能性^[4]. 此外,在测谎、头颈损伤康复、智能控制等领域也有着广泛的应用^[9].

截至目前,头部动作的捕捉主要依靠头戴式设备或头贴式传感器,通过传感器获取头部动作电信号实现头势理解,例如:Jackowski^[10]使用FSM-9传感器采集头部动作信号,并通过头部俯仰角和偏航角信息实

现了机械臂控制; Rudigkeit^[11]采用九轴 MEMS 运动传感器采集头部动作信号并建立头势识别模型, 获得了约 93.56% 的识别率. 以上基于传统手工计算特征的方法常受到头戴式传感器精度、特征模型、使用人员的影响. 在机器学习领域, 文献[1, 9]分别通过隐马尔可夫模型和随机森林对头势电信号进行分类, 使识别率提高到约 99.00%. 然而, 以上方法需一直头戴传感设备, 且头戴式设备价格高昂, 便携性与交互体验较差^[4].

针对以上方法的缺陷, 一些研究人员尝试用单纯的机器视觉来获取头部的动作信息, 例如: Bankar^[12]采用 Camshift 算法结合无迹卡尔曼进行头势识别, 但识别率仅 63.33%; 其后, 一些研究者尝试结合学习算法提高识别率, 例如: 文献[6]通过肤色检测与 HMM 算法进行头势识别, 达到了 87% 的识别率; 文献[13]则采用多视图响应模型结合 HMM 算法进行头势识别, 达到了 88.1% 的识别率; 文献[14]利用 HOF 结合极限学习机进行分类, 获得了 95.63% 的识别率. 然而, 基于以上方法识别率较低, 难以达到交互的目的.

综上所述, 目前基于头戴式传感器的方法经济性、便携性、舒适性较差, 而机器视觉的方法识别率相对较低. 同时, 使用深度学习还需克服以下问题: 网络训练需大量带标签的头势数据. 移动设备受计算能力所限, 面临实时性与识别率的冲突.

针对以上问题, 本文提出了一种无需头戴设备的动态头势识别算法. 该方法利用光流获取动作信息, 去除部分背景噪音, 构建了双流融合 3DCNN (Two-Stream 3DCNN) 网络从 RGB 视频、光流视频中提取时域动作特征, 并进行融合与分类. 所提算法可有效提取、融合时域动作特征, 比人工特征提取方法、C3D 模型有更高的识别率、更好的泛化能力, 无需头戴设备有近似头戴式设备的识别率.

2 头势的动作特征提取与融合策略

2.1 基于光流的动作信息提取与增强

光流 (Optical Flow, OF) 法是指空间运动物体在观测成像面上像素运动的瞬时速度, 通过光流计算可更好地获取到运动信息^[15], 避免背景和肤色等因素的影响. Farneback^[16]将前后两帧中所有像素点的移动估算加入光流, 提出了稠密光流法. 本文将利用稠密光流算法改善现有方法在头势识别中存在的动作信息不连续、干扰较多的缺陷.

2.2 特征提取器的构建

为更好地提取 RGB 和光流视频的动作特征, 本文采用深度学习来弥补手动特征提取存在的特征提取困难、分类效果差的缺点. 在图像处理领域, 与传统的人

工特征选取、分类器相比, 卷积神经网络 (Convolutional Neural Networks, CNN) 针对姿势、光照、背景与杂波等噪音有较好的鲁棒性^[17]. 普通 CNN 只含 2D 卷积, 适于 2D 图像, 但对含时间关系的图像序列, 不能提取其时间维度上的动作特征. 针对该问题, 文献[18]改良 CNN 网络实现了初步的动作特征提取; 文献[19]提出了 3DCNN 结构, 把 2D 卷积核在时间维度上进行扩展, 从而得到 3D 卷积核, 可同时提取时域和空域中的运行特征; Tran^[20]等对 3DCNN 中 3D 卷积核大小与识别结果的关联性做了进一步深入分析, 并提出适用于识别含有时间关系的 C3D 网络, 证实了 3D 卷积可有效提取动作特征, 但目前尚未发现其被用于头势识别.

三维卷积将一个三维核与多个通道叠加形成的立方体进行卷积, 卷积层中的特征映射被连接到前一层中的相邻帧, 从而捕获运动信息^[19]. 3D 卷积过程如式(1).

$$v_{ij}^{xyz} = \text{ReLU}(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}) \quad (1)$$

其中, R_i 为 3D 卷积核在时间维度上的大小, P_i 、 Q_i 分别是 3D 卷积核在空间维度上 x 、 y 方向的大小, p 、 q 、 r 分别是当前时间维度、 x 、 y 方向上的位置, w_{ijm}^{pqr} 是位于第 m 个特征图连接的卷积核在 (p, q, r) 点上的值, v_{ij}^{xyz} 为第 i 层网络的第 j 个特征映射在位置 (x, y, z) 的卷积值, b_{ij} 是当前网络偏置, $\text{ReLU}()$ 是激活函数.

2.3 双流通道特征融合

双流通道融合方式和位置会影响分类效果^[21]. 文献[21]提出了“交叉图融合法”与“3D 卷积融合法”, 通过串联两种方法可获得融合特征, 如式(2).

$$\begin{cases} Z_{t,i,j,2d} = x_{t,i,j,d} \\ Z_{t,i,j,2d-1} = x'_{t,i,j,d} \end{cases} \quad (2)$$

其中, Z 是融合特征图, $Z \in R^{T \times H \times W \times 2D}$, T 为特征图时域大小、 H 和 W 为空域大小、 D 为特征图数, $x_{t,i,j,d}$ 、 $x'_{t,i,j,d}$ 分别是第 1、2 通道的特征图, t 、 i 、 j 分别是时间轴、图像 x 、 y 轴位置, d 是特征图通道位置.

对交叉融合得到的新特征图 Z 进行 3D 卷积融合, 如式(3).

$$Y = Z * F + b \quad (3)$$

其中, Y 是融合特征图, F 为 $3 \times 3 \times 3$ 的卷积核, b 为偏置常数.

将卷积结果 Y 输入过滤器大小为 $3 \times 3 \times 3$ 的 3D 最大池化, 形成 Fusion 模块, 如图 1.

在网络越深的位置融合, 识别率越高^[21]. 因此为更好地提取头势特征, 本文在最后一个卷积层加入 Fusion 模块用于 RGB 通道特征和光流通道特征的融合.

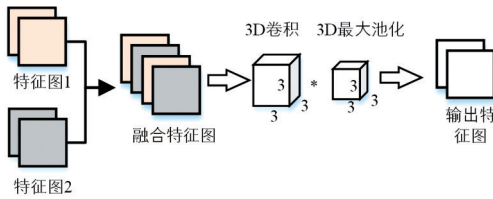


图1 Fusion模块原理图

3 基于 3D 卷积神经网络的动态头势识别

3.1 基于双流 3DCNN 的头势识别策略

为提高头势识别的准确率和实时性,设置双流 3DCNN 通道分别对 RGB、光流视频做特征提取. 本文借鉴 C3D 结构设计了特征提取 3DCNN 通道提取动作特征,再以 Fusion 模块获取高维时空融合特征,整体流程如图 2. 该算法不仅提取了 RGB 视频在时域和空域上的动作特征,还提取了光流视频中更高层次的动作特征,同时保证了空间特征和时域特征的有效利用.

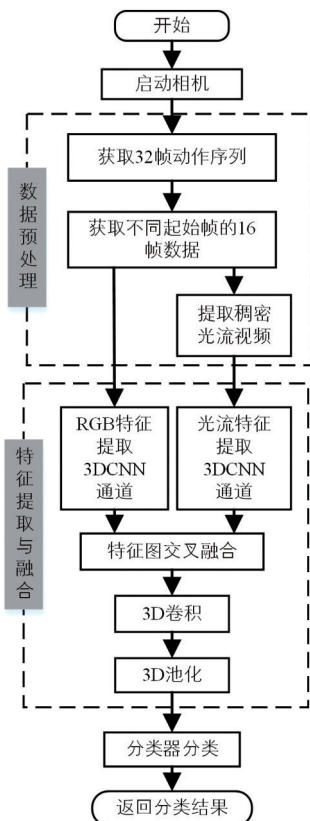


图2 基于双流融合 3DCNN 的动态头势识别策略

3.2 面向动态头势识别任务的网络结构构建

本文所提双流融合 3DCNN 借鉴文献[19~21]中网络结构进行设计,网络结构如图 3. 该结构用 2 个 3×3×3 的卷积核取代 1 个 5×5×5 的卷积核获取更大的感受野,增加网络深度,减少网络参数量^[18]. 故本文模型中的卷积核均为 3×3×3,步长为 1. 为提高网络计算速度,把原

始 C3D 网络的特征图数量减半,但网络深度不变,组成单个特征提取 3DCNN 通道. 最终 RGB 通道和光流通道网络的卷积核个数均设置为 32、64、128、128、256、256、256 和 256, Fusion 输出的特征图数量为 256.

为最大化提取时间维度上的特征,保持时间维度尺度不变,故设置 RGB 通道和光流通道的第一个 3D 池化操作,即池化 1 和池化 6 大小为 1×2×2,其余池化滤波器均为 2×2×2,所有池化滤波器步长为 1,两个全连接层有 2048 个神经元连接.

3.3 动态头势识别网络的优化函数选择

网络的激活函数采用 ReLU (Rectified Linear Unit) 函数^[22],其优势在于能有效解决在反向传播时梯度爆炸的问题. 此外,在图 3 的 fc1 层和 fc2 层后引入 Dropout^[23]减弱神经元节点间的联合适应性,防止过拟合,增强网络泛化能力,设置如式(4).

$$\text{Dropout} = \begin{cases} 0.5, & \text{training} \\ 1, & \text{testing} \end{cases} \quad (4)$$

权重优化采用自适应优化 Adam 算法^[24],该算法的特点是参数少、所需内存少和计算高效等,能更快获得最优权重. 并在网络最后采用 Softmax 分类器进行分类,如式(5).

$$S_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (5)$$

其中, k 是类别数, z_i 是输入, S_i 是类别 i 的概率.

最后采用交叉熵损失函数计算损失,如式(6).

$$L(x) = - \sum_{i=1}^n y_i \log(\hat{y}_i) \quad (6)$$

其中, n 是类别数, y_i 为输入 x 下的真实值, \hat{y}_i 是对 x 的预测值, $L(x)$ 是输入数据 x 下的损失.

4 实验结果与分析

以下将所提算法与其他算法进行对比测试:(1)与 C3D 对比测试;(2)与现有头势识别方法对比测试.

4.1 数据集构建

本文采集了多人在不同光源、角度、背景等情况下点头、摇头、左偏和右偏 4 类头势数据. 相机参数:分辨率为 640px×360px,帧速率为 15frame/s 和 30frame/s. 用户与相机距离为 50cm~100cm,期间相机位姿不变.

为解决数据不足造成的过拟合问题,本文通过随机轻微旋转、裁剪、亮度等手段进行数据扩充,并采集非训练集和验证集中用户的数据为测试集 1,戴口罩的数据为测试集 2,以此验证网络泛化能力,其中每个视频数据共 32 帧,数据如表 1. 在视频输入网络前自动按比例缩放视频并居中裁剪为 112px×112px,并随机从 32 帧的前 16 帧中选取起始帧,并取起始帧往后 16 帧作网

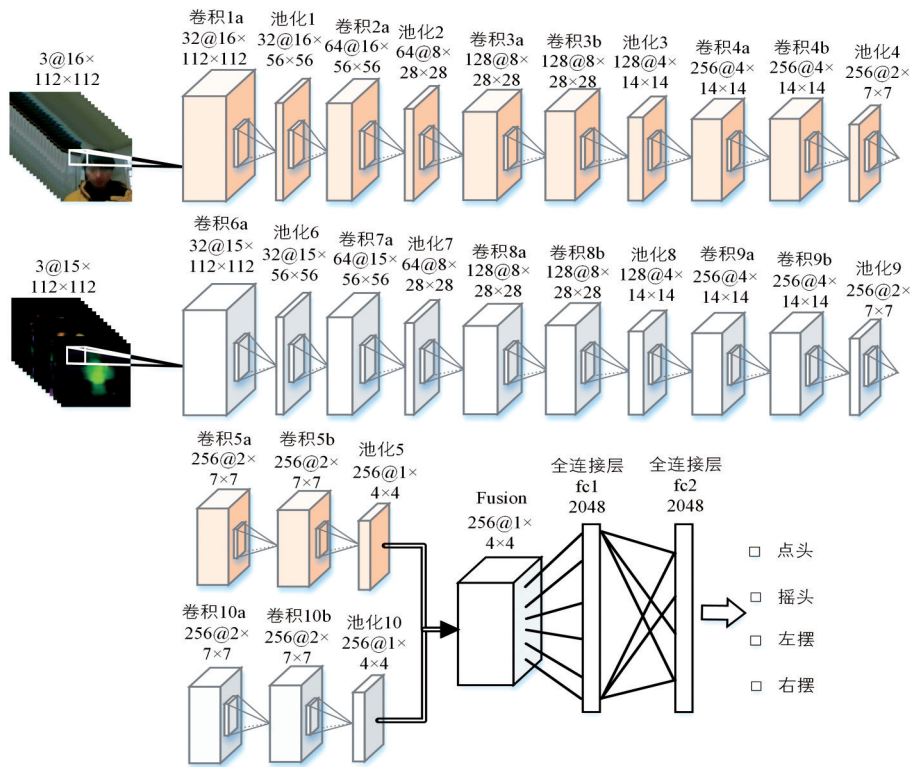


图3 双流融合3DCNN网络结构

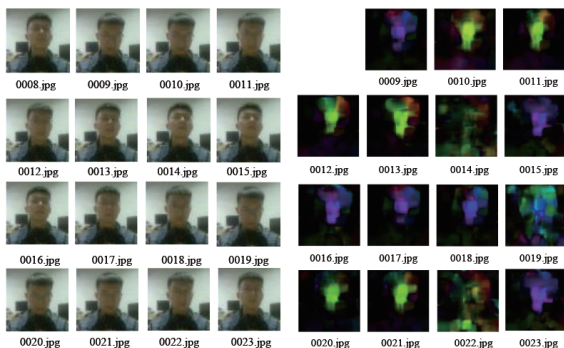
表1 动态头势数据集

类别	训练集	验证集	测试集1	测试集2
点头	724	208	96	80
摇头	732	204	108	80
左偏	656	204	112	80
右偏	796	260	103	80

网络输入,以保证动作起点的随机性,其中可包含1或2个头部运动周期。

4.2 提取稠密光流

稠密光流可提取视频中丰富的动作信息,表征人体运动的部位及运动方向。图4是“点头”的RGB和稠密光流样本序列。



(a) “点头” RGB 样本序列 (b) “点头” 光流样本序列

图4 “点头”的RGB序列与光流序列

4.3 动态头势分类网络模型的训练与测试

实验配置:操作系统为Ubuntu18.04 64位,配置Python3.6、Tensorflow1.12搭建网络;GPU为单个NVIDIA GeForce GTX2080Ti,内存16GB,CPU为单个6核AMD Ryzen 5 2600X.采用mini-batch方法,batch size设为30,初始学习率为0.0001,迭代次数为1000次,每批次输入数据在数据集随机选取.为保障对比测试的一致性,C3D模型输入RGB视频,并设置与本文一样的超参数。

本文方法与C3D模型在验证集上的准确率变化如图5所示.由图可见,训练时会出现不同幅度的振荡,但会随着训练次数的增加而不断减小.此外,本文方法收敛更快、准确率更高,波动更小.由此可见,双流融合3DCNN网络更适于提取头部动作特征,且能节约更多训练时间与计算资源。

为测试所提算法性能,可视化“点头”的特征图.提取图3中卷积2a、3b、5b、7a、8b和10b在池化后的特征数据.此外,在空域上分别可视化卷积3b、8b池化后的RGB特征、光流特征、Fusion融合特征及fc2的输出特征,如图6。

由图6可知,在时域和空间域上,RGB通道可提取轮廓、亮度等特征,光流通道提取了头部运动特征.随层数增加,网络提取出高层次特征,通过特征融合,特征图数量增多,特征提取更充分.特征图表明:双流融合3DCNN

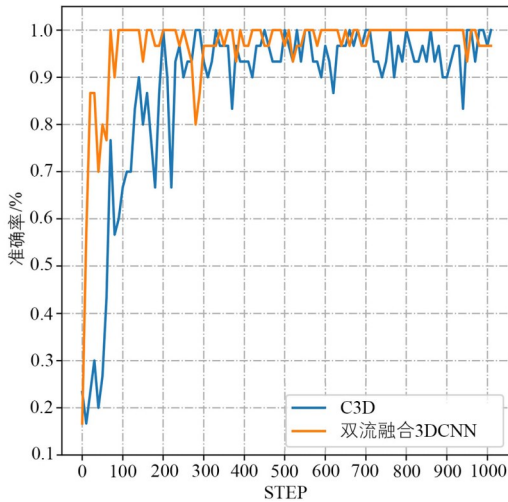
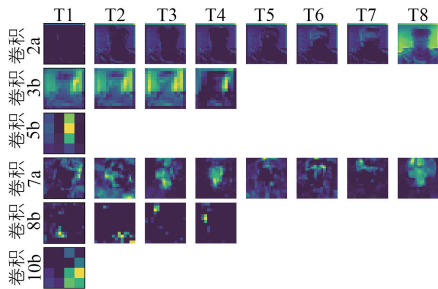
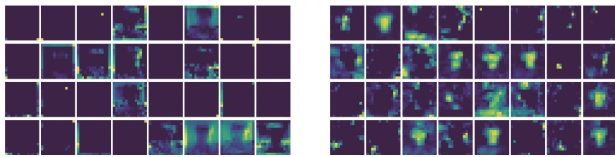


图5 本文方法与C3D模型在验证集上的准确率

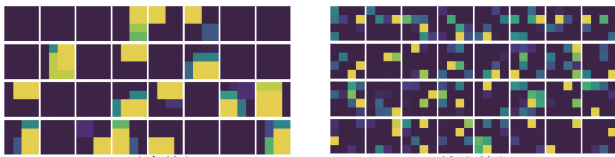


(a) 时域特征



(b) RGB特征

(c) 光流特征



(d) 融合特征

(e) 输出特征

图6 特征图可视化

网络可有效地提取头势的时域和空域动作特征。

4.4 实时性与平均识别率对比测试

以下实验旨在验证本文方法对识别率、泛化能力及实时性的提升。与C3D对比测试,统计其参数量、识别率与实时性。

由表2可见,本文方法在验证集的识别率比C3D高1.83%,在测试集1上高2.43%,而测试集2上高5%,且参数量比C3D模型减少50%,计算速度快9ms。以上表明双流融合3DCNN有更好的泛化能力,更小的参数体积以及更好的实时性。

本文方法与C3D在单类别下的识别率如图7。由图可见,所提算法与C3D模型对摇头、左偏与右偏之间存

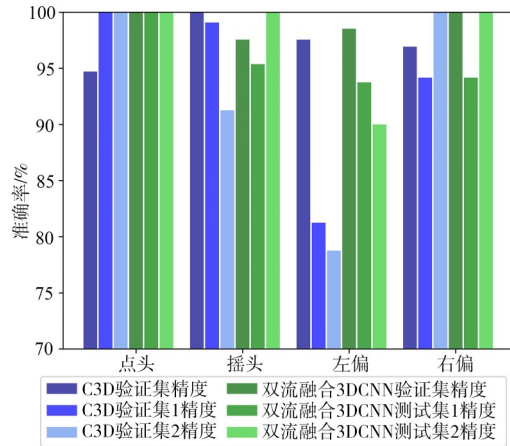


图7 单类别下在本文方法与C3D模型下的识别率

表2 本文方法与C3D准确率对比

模型	参数量	验证集	测试集1识别率	测试集2识别率	识别耗时
C3D ^[20]	78M	97.26%	93.32%	92.50%	94ms
本文方法	30M	99.09%	95.75%	97.50%	85ms

在误判。从动作特点上可知,点头与其余3个动作有较强区分性,而其余3个动作相似度较高,导致其识别率下降。实验结果表明:所提算法具有比C3D模型更高识别率,动作特征提取能力更强。

4.5 双流融合3DCNN与现有方法对比测试

针对动态头势识别问题,以下将本文方法与现有方法进行对比测试,包括:基于头戴传感器设备的方法与基于机器视觉的方法。因基于头戴式设备的方法所使用的数据为电信号,基于视觉的方法使用的是图像信号,且公开文献显示尚无公共头势图像数据集可用。故本文沿袭现有基于视觉类型方法的对比方式对测试结果进行展示,如表3所示。

由表3可见,头戴式设备可使头部动作识别获得较

表3 本文所提算法与现有方法识别率对比

采取设备	算法	类别数	识别率
头戴式设备	FSM-9 ^[10]	4类	81.80%
	GY-87+HGR-DT ^[9]	6类	94.63%
	GY-87+HGR-RF ^[9]	6类	99.17%
相机	Camshift with UKF ^[12]	/	63.33%
	Eye Detection+SVM ^[25]	3类	91.10%
	MS-ConvLSTM ^[3]	4类	73.25%
	肤色检测+HMM ^[2]	3类	87.00%
	MVM+HMM ^[13]	2类	88.10%
	HOF+ELM ^[14]	7类	95.63%
	C3D ^[20]	4类	92.50%
	双流融合3DCNN	4类	97.50%

高识别率,而基于纯视觉的识别率通常较差.本文所提算法在基于视觉下比多数头戴式设备的方法具有更高的识别率,仅比GY-87+HGR-RF略低.

5 总结

本文提出了一种基于双流融合3DCNN的动态头势识别算法,该算法比人工特征提取的方法及C3D模型具有更高的准确率、更好的泛化能力.在无需头戴设备的情况下有近似头戴式设备的识别率,达到了近似头戴式设备的识别率,具有优良的便携性、识别率和泛化能力.

参考文献

- [1] ZHAO J B, ALLISON R S. Real-time head gesture recognition on head-mounted displays using cascaded hidden Markov models[A]. Proceedings of the 2017 IEEE International Conference on Systems, Man and Cybernetics [C]. Banff, AB, Canada: IEEE, 2017. 2361 – 2366.
- [2] VADIRAJ S K, RAO A, GHOSH P K. Automatic identification of speakers from head gestures in a narration[A]. Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing [C]. Barcelona, Spain: IEEE, 2020. 6314 – 6318.
- [3] SHARMA M, AHMETOVIC D, JENI L A, et al. Recognizing visual signatures of spontaneous head gestures[A]. Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision [C]. Nevada, USA: IEEE, 2018. 400 – 408.
- [4] ZHAO J B, ALLISON R S. Comparing head gesture hand gesture and gamepad interfaces for answering Yes/No questions in virtual environments [J]. Virtual Reality, 2020, 24: 515 – 524.
- [5] FUJIE S, EJIRI Y, NAKAJIMA K, et al. A conversation robot using head gesture recognition as para-linguistic information[A]. Proceedings of the RO-MAN 2004 13th IEEE International Workshop on Robot and Human Interactive Communication [C]. Kurashiki, Japan: IEEE, 2004. 159 – 164.
- [6] NG P C, SILVA L C D. Head gestures recognition[A]. Proceedings of the 2001 International Conference on Image Processing [C]. Thessaloniki, Greece: IEEE, 2001. 266 – 269.
- [7] SOLEA R, MARGARIT A, CERNEGA D, et al. Head movement control of powered wheelchair [A]. Proceedings of the 23rd International Conference on System Theory, Control and Computing [C]. Sinaia, Romania: IEEE, 2019. 632 – 637.
- [8] JACKOWSKI A, GEBHARD M, THIETJE R. Head motion and head gesture-based robot control: A usability study[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2018, 26(1): 161 – 170.
- [9] WU C W, YANG H Z, CHEN Y A, et al. Applying machine learning to head gesture recognition using wearables [A]. Proceedings of the 2017 IEEE 8th International Conference on Awareness Science and Technology [C]. Taichung, China: IEEE, 2017. 436 – 440.
- [10] JACKOWSKI A, GEBHARD M, GRÄSER A. A novel head gesture based interface for hands-free control of a robot[A]. Proceedings of the 2016 IEEE International Symposium on Medical Measurements and Applications [C]. Benevento, Italy: IEEE, 2016. 1 – 6.
- [11] RUDIGKEIT N, GEBHARD M, GRÄSER A. An analytical approach for head gesture recognition with motion sensors [A]. Proceedings of the 2015 9th International Conference on Sensing Technology [C]. Auckland, New Zealand: IEEE, 2015. 1 – 6.
- [12] BANKAR R, SALANKAR S. Improvement of head gesture recognition using Camshift based face tracking with UKF[A]. Proceedings of the 2019 9th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing [C]. Nagpur, India: IEEE, 2019. 1 – 5.
- [13] LU P, ZHANG M, ZHU X, et al. Head nod and shake recognition based on multi-view model and hidden Markov model[A]. Proceedings of the International Conference on Computer Graphics, Imaging and Visualization [C]. Beijing, China: IEEE, 2005. 61 – 64.
- [14] HONG T, LI Y W, WANG Z Y. Real-time head action recognition based on HOF and ELM[J]. IEICE Transactions on Information and Systems, 2019, 102(1): 206 – 209.
- [15] 罗会兰, 童康, 孔繁胜. 基于深度学习的视频中人体动作识别进展综述[J]. 电子学报, 2019, 47(5): 1162 – 1173.
- [16] LUO Hui-lan, TONG Kang, KONG Fan-sheng. The progress of human action recognition in videos based on deep learning: A review [J]. Acta Electronica Sinica, 2019, 47(5): 1162 – 1173.(in Chinese)
- [17] FARNEBÄCK G. Two-frame motion estimation based on polynomial expansion [A]. Proceedings of the 13th Scandinavian Conference on Image Analysis [C]. Halmstad, Sweden: Springer, 2003. 363 – 370.
- [17] LECUN Y, HUANG F J, BOTTOU L. Learning meth-

- ods for generic object recognition with invariance to pose and lighting[A]. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C]. Washington, USA: IEEE, 2004. 97 – 104.
- [18] SIMONYAN K, ZISSERMAN A. Two-stream convolutional networks for action recognition in videos[J]. Advances in Neural Information Processing Systems, 2014, 27: 568 – 576.
- [19] JI S, XU W, YANG M, et al. 3D convolutional neural networks for human action recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 35(1): 221 – 231.
- [20] TRAN D, BOUTDEV L, FERGUS R, et al. Learning spatiotemporal features with 3D convolutional networks [A]. Proceedings of the 2015 IEEE International Conference on Computer Vision [C]. Santiago, Chile: IEEE, 2015. 4489 – 4497.
- [21] FEICHTENHOFER C, PINZ A, ZISSERMAN A. Convolutional two-stream network fusion for video action recognition [A]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition [C]. Las Vegas, USA: IEEE, 2016. 1933 – 1941.
- [22] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines [A]. Proceedings of the 27th International Conference on Machine Learning [C]. Haifa, Israel: ACM, 2010. 807 – 814.
- [23] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929 – 1958.
- [24] KINGMA D P, Ba J. Adam: A method for stochastic op-

timization [A]. Proceedings of the 3rd International Conference on Learning Representations [C]. San Diego, USA: Conference Track Proceedings, 2015. 1884 – 2021.

- [25] SUNI S S, GOPAKUMAR K. A real time decision support system using head nod and shake [A]. Proceedings of the 2016 International Conference on Circuit, Power and Computing Technologies [C]. Nagercoil, India: IEEE, 2016. 1 – 5.

作者简介



谢佳龙 男, 1997年6月出生于四川省南充市. 现为杭州电子科技大学自动化学院研究生. 研究方向为机器人控制与人机交互.
E-mail: jolen_xie@hotmail.com



张波涛(通信作者) 男, 1982年9月出生于山东省潍坊市. 现为杭州电子科技大学副教授. 主要研究方向为机器视觉、机器人运动规划与控制等.
E-mail: billow@hdu.edu.cn



吕强 男, 1977年7月出生于辽宁省抚顺市. 现为杭州电子科技大学教授, 主要研究方向为多机器人合作控制、群体智能等.
E-mail: lvqiang@hdu.edu.cn