

基于手部关键点和肤色的手势识别^①



李 光¹, 王长元², 薛鹏翔³

¹(西安工业大学 兵器科学与技术学院, 西安 710021)

²(西安工业大学 计算机科学与工程学院, 西安 710021)

³(西安工业大学 理学院, 西安 720021)

通讯作者: 王长元, E-mail: cyw901@163.com

摘 要: 基于传统的手部轮廓特征提取不能应对飞行模拟环境下的脸部肤色、遮挡、光照影响, 以及传统的傅里叶描述子特征容易受到背景、手的姿态变化, 且对手势描述能力有限等问题, 对传统的手部分割和特征提取方法改进. 本文首先对采集的数据集进行肤色处理, 然后结合调用的手部关键点模型检测出手部 22 个特征点, 采用八向种子填充算法进行图像分割. 接着对手部轮廓和关键点连接骨架进行傅里叶描述子算法特征提取, 最后通过支持向量机算法对提取的手势特征数据集进行训练、识别. 实验结果表明, 本文方法具有较好的手部分割, 特征提取不易受到背景、手的姿态变化的影响, 能够很好地应对在飞行模拟环境下的复杂背景下的干扰, 识别准确率能够达到 98%. 本文研究在传统的手势识别算法中有一定的提高作用, 在手部交互技术领域有很重要的应用价值.

关键词: 手部关键点; 肤色分割; 特征提取; 傅里叶描述子; 支持向量机

引用格式: 李光, 王长元, 薛鹏翔. 基于手部关键点和肤色的手势识别. 计算机系统应用, 2021, 30(1):180-185. <http://www.c-s-a.org.cn/1003-3254/7787.html>

Gesture Recognition Based on Key Points of Hand and Skin Color

LI Guang¹, WANG Chang-Yuan², XUE Peng-Xiang³

¹(Northwest Institutes of Advanced Technology, Xi'an Technological University, Xi'an 710021, China)

²(School of Computer Science and Engineering, Xi'an Technological University, Xi'an 710021, China)

³(College of Science, Xi'an Technological University, Xi'an 710021, China)

Abstract: The traditional hand contour feature extraction can not deal with the effect of facial skin color, occlusion, lighting in the flight simulation environment. The traditional Fourier descriptor features are susceptible to background and hand posture changes, and have limited ability to describe gestures, etc. Hence, this study proposes a method to improve the traditional hand segmentation and feature extraction methods. Firstly, skin color processing is performed on the collected data set, and then the 22 key points of the hand are detected in combination with the called hand key point model, and an eight-way seed filling algorithm is used for image segmentation. Then, the contours of the hand and key points are connected to the skeleton to extract the features of the Fourier descriptor algorithm. Finally, the Support Vector Machine (SVM) algorithm is used to train and recognize the extracted gesture feature data set. The experimental results show that the method in this study has good hand segmentation, feature extraction is not easily affected by changes in the background and hand posture. Hence, it can well cope with interference in a complex background in a flight simulation environment, with the recognition accuracy reaching 98%. The research presented in this paper has a certain role in improving the traditional gesture recognition algorithm and has very important application value in the field of hand interaction technology.

Key words: hand key points; skin color segmentation; feature extraction; Fourier descriptor; SVM

① 基金项目: 国家自然科学基金 (52072293)

Foundation item: National Natural Science Foundation of China (52072293)

收稿时间: 2020-06-03; 修改时间: 2020-06-30, 2020-07-27; 采用时间: 2020-07-31; csa 在线出版时间: 2020-12-31

在计算机领域不断发展的今天,飞行模拟器^[1,2]以其安全性、经济性、可靠性、高逼真度等特点,特别是混合现实技术的发展给操作员提供了很大的沉浸感,已经成为飞行员训练和飞机研发过程中不可或缺的试验仿真设备,飞行模拟作为虚拟交互的典型应用之一,对手部自然交互^[3,4]要求比较高,需求也极为迫切。由于接触式测量方式价格成本高昂,人的动作容易受到硬件设备的约束,阻碍人手完成自然的肢体动作,在自然性和精度方面受到限制,所以目前基于计算机视觉手部交互方式在各行各业都有重大意义。

飞行模拟环境下的手势需要操纵油门、操作杆等,其传统的肤色分割容易受到面部肤色的影响,其手部轮廓特征太单一,而且采用传统的傅里叶描述子又容易受到背景、手的姿态变化等影响,在其飞行模拟操作下都会受到遮挡影响,从而影响其识别精度。

基于计算机视觉的手势识别中,手势分割中文献^[5]用肤色分割法和 SVM 对手势进行识别;基于运动的分割方法,例如帧间差分法^[6];运用 Kinect 基于深度的图像^[7]等。特征提取中用手势图的灰度、颜色等信息^[8],这些特征容易获得,但是容易受到环境光照影响;统计特征中包括梯度方向直方图(HOG)^[9]等;基于形状不同的表示方式,如傅里叶描述子^[10]、基于边界特征点^[11]、不变矩^[12]等。本文就是基于肤色和手势轮廓进行特征提取,结合特征点连接骨架进行傅里叶描述子特征提取。

手势识别算法有很多种,例如支持向量机、基于模板^[13]、基于卷积神经网络^[14]、隐马尔科夫模型^[15]、动态时间规整^[16]等。针对本文的小样本和快速计算选用支持向量机进行分类。

1 主要创新点

本文方法的创新点是首先根据不同光照环境通过最大类间方差法^[17]来求取二值化图像中的最佳阈值从而适应光照变化,其次在传统的肤色分割上结合手部特征点进行区域分割,解决飞行模拟环境下的脸部肤色、遮挡、光照影响;然后进行手部轮廓和手部特征点连接骨架的傅里叶描述子特征提取,解决傅里叶描述子特征易受到背景、手的姿态变化影响的问题,并且为了实现缩放不变性,对变换系数做了归一化处理。最后,用支持向量机对数据进行训练、识别。本文工作流程图如图 1 所示。

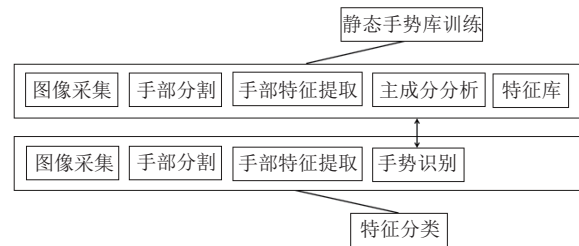


图 1 工作流程图

手部特征点主要是基于美国 CMU Perceptual Computing Lab 开源的手部关键点检测模型^[18,19]。手部关键点检测模型共输出 22 个关键点,其中包括手部的 21 个点,第 22 个点表示背景。通过 22 个关键点能够快速准确地定位出手部的位置,并且能够很好地分割出手部区域。如图 2,为手部标注的 21 个特征点。手部关键点结合肤色轮廓特征对手势遮挡具有抵抗作用,并且结合手势轮廓能够很好地表示手势。本文通过此方法提取手势特征,再通过 SVM 分类识别方法进行飞行模拟环境下静态手势模型训练和手势识别。



图 2 手部标注的 21 个特征点

2 手部特征提取

2.1 改进的手部区域分割算法

首先分割出手势的区域,后面才能进行特征提取。2015 年,文献^[9]使用 Kinect 研究了基于深度图像的 3D 手势识别,识别率达 76.6%。但其识别率并不高,并且还有 Kinect 摄像头的限制。肤色是人体手部最显著的特征。本文采用常规的肤色检测,对手部进行肤色信息提取,但是常常会有一部分的面部肤色等干扰,结合特征点对手部分割,提高其准确性;而且传统肤色处理容易受到不同光照的影响,本文根据不

同光照环境通过最大类间方差法来求取二值化图像中的最佳阈值从而适应光照变化,以提高其鲁棒性。

在诸多彩色模型中,肤色在 YCbCr 颜色空间中有很好的聚类特性,并且 YCbCr 颜色空间对光照变化不

敏感,所以可以利用这种特性来区分手势区域,其中 Y 表示亮度,分量 Cb 表示蓝色分量和亮度值之差,分量 Cr 反映了 RGB 输入信号红色部分与亮度值之间的差异.如图 3 所示.

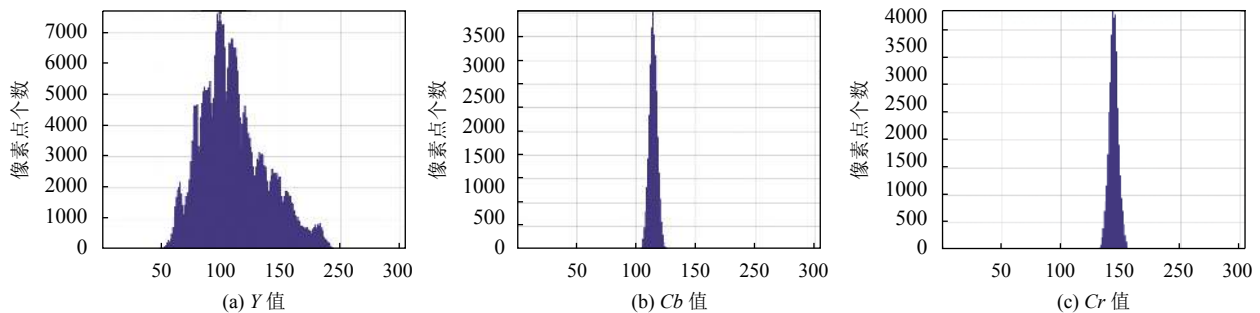


图3 YCbCr分量直方图

在 YCbCr 空间,肤色像素点会聚集到一个椭圆区域.先定义一个椭圆模型,然后将每个 RGB 像素点转换到 YCbCr 空间对比是否在椭圆区域,是的话判断为皮肤.本文将从视频中获取的帧图像从 RGB 空间转换到 YCbCr 空间以实现肤色的检测. RGB 空间与 YCbCr 空间之间的转换公式为:

$$\begin{cases} Y = 0.99R + 0.5870G + 0.1140B \\ Cb = -0.1687R - 0.3313G + 0.5000B + 128 \\ Cr = 0.5000R - 0.4187G - 0.0813B + 128 \end{cases} \quad (1)$$

本文采用基于 YCbCr 空间的 OTSU 自适应阈值算法作为图像分割算法.

OTSU 算法^[13](大津法或最大类间方差法)使用的是聚类的思想,把图像的灰度数按灰度级分成 2 个部分,使得两个部分之间的灰度值差异最大,每个部分之间的灰度差异最小,通过方差的计算来寻找一个合适的灰度级别来划分.所以可以在二值化的时候采用 OTSU 算法来自动选取阈值进行二值化.

设 t 为设定的阈值, w_0 为分开后前景像素点数占图像的比例, u_0 为分开后前景像素点的平均灰度, w_1 为分开后背景像素点数占图像的比例, u_1 为分开后背景像素点的平均灰度.

图像总平均灰度为:

$$u = w_0 \times u_0 + w_1 \times u_1 \quad (2)$$

从 L 个灰度级遍历 t ,使得 t 为某个值的时候,前景和背景的方差最大,则这个 t 值便是我们要求得的阈值.其中,方差的计算公式如下:

$$g = w_0 \times (u_0 - u) \times (u_0 - u) + w_1 \times (u_1 - u) \times (u_1 - u) \quad (3)$$

此公式计算量较大,可以采用:

$$g = w_0 \times w_1 \times (u_0 - u_1) \times (u_0 - u_1) \quad (4)$$

OTSU 算法被认为是图像分割中阈值选取的最佳算法,计算简单,不受图像亮度和对比度的影响.因此,使类间方差最大的分割意味着错分概率最小.

本文选用的基于 YCbCr 肤色模型和手势特征关键点的分割方法能够很好地解决面部肤色影响.图像分割效果如图 4 所示.

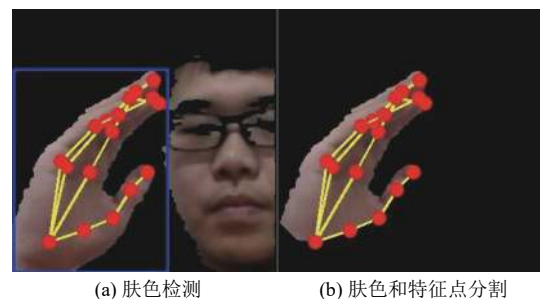


图4 手势部分图像分割

2.2 傅里叶描述子特征提取

物体的形状识别是模式识别的重要研究方向,广泛用于图像分析、机器视觉和目标识别的应用中.在计算机中物体的形状表示方式有多种:边界曲线上点的坐标序列,边界上的点到物体中心的距离以及边界上的点随弧长变化的切线方向等.傅里叶描述子的基本思想是:假定物体的形状是一条封闭的曲线如图 5 所示,沿边界曲线上的一个动点 $P(l)$ 的坐标变化 $x(l) + iy(l)$ 是一个以形状边界周长为周期的函数,这个周期函数可以用傅里叶级数展开表示,傅里叶级数中的一系列系数 $z(k)$ 是直接和边界曲线的形状有关的,称为傅里叶描述子.

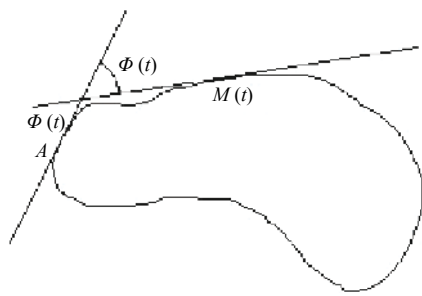


图5 假定的封闭曲线

此函数是连续的周期的, 所以可以通过傅里叶级数来表示:

$$\phi(t) = \sum_{K=0}^{\infty} a_k \exp(-jkt) \quad (5)$$

式中, a_k 就是傅里叶描绘子.

傅立叶描述子算法优点可以很好地描述轮廓特征, 并且只需少量的描述子即可大致代表整个轮廓. 傅里叶描述子与轮廓形状的尺度、方向和起始点有关, 对傅立叶描述子进行简单的归一化操作后, 即可使描述子具有平移、旋转、尺度不变性, 即不受轮廓在图像中的位置、角度及轮廓的缩放等影响, 是一个鲁棒性较好的图像特征.

对手部的轮廓特征提取如图6所示.

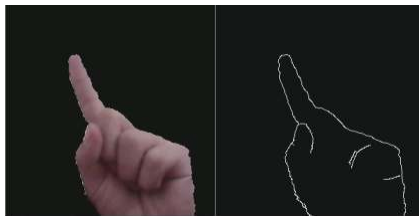


图6 手部边缘轮廓

傅里叶描述子提取物理特征虽说计算快速简单, 易受到背景、手的姿态变化, 且对手势描述能力有限等问题. 本文在手势轮廓的基础上增加了手势关键点的轮廓特征如图7所示, 这样提取的手势特征能够更好的表示手势, 并且具有一定的鲁棒性.

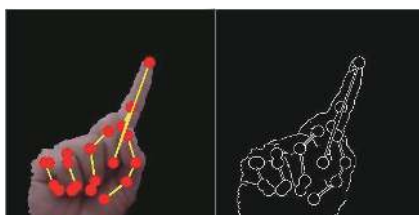


图7 结合特征点连接骨架的手势轮廓

3 支持向量机算法

本文通过在飞行模拟环境下的计算需求和样本条件下, 选择支持向量机作为分类算法. 支持向量机 (Support Vector Machine, SVM) 是一种非常重要和广泛的机器学习算法, SVM 的算法出发点是尽可能找到最优的决策边界, 这个边界距离两类数据点最远. 更进一步的, 是距离两类数据点的边界最远, 所以定义最接近边界的数据点定义为支持向量. 最后, 我们的目标变为找到这样一个直线 (多维叫超平面), 它与支持向量有最大的间隔. 使得模型的泛化能力尽可能的好, 因此 SVM 对未来数据的预测也是更加准确的如图8所示, 寻找最大化 margin.

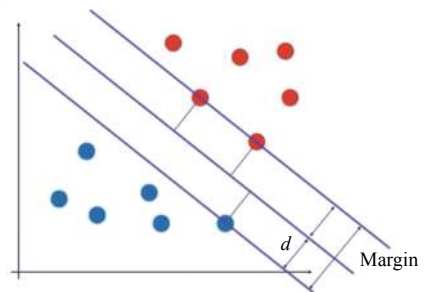


图8 分类模型图

设这个平面用 $g(x)=0$ 来表示, 其法向量用 w 表示, 点与平面实际距离为 r 点, 与平面的距离可以用 $g(x)$ 的绝对值来度量 (称为函数间隔).

$$\begin{aligned} \min_{w,b} & \frac{1}{2} W^T W \\ y_i(W^T X_i + b) & \geq 1, \\ i & = 1, \dots, l \end{aligned} \quad (6)$$

惩罚函数亦称处罚函数, 是一类制约函数. 对于约束非线性规划它的制约函数称为惩罚函数, 系数称为惩罚因子. 引入惩罚因子 C 的 SVM (软间隔支持向量机) 的目标函数为:

$$\min_{w,b,\xi} \frac{1}{2} W^T W + C \sum_{i=1}^l \xi \quad (7)$$

SVM 是一种有坚实理论基础的新颖的小样本学习方法. 它基本上不涉及概率测度及大数定律等, 因此不同于现有的统计方法. 从本质上看, 它避开了从归纳到演绎的传统过程, 实现了高效的从训练样本到预报样本的“转导推理”, 大大简化了通常的分类和回归问题. 模型图如图9.

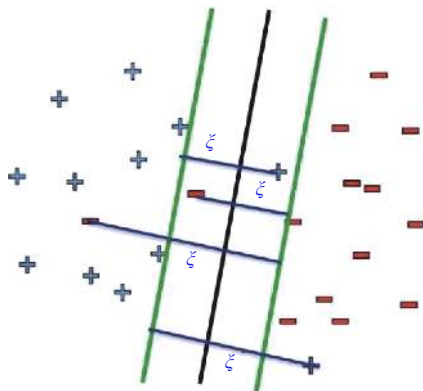


图9 引入 C, ξ 的分类模型图

4 实验分析

本文使用作者实验室的项目中飞行模拟环境中的采集的手势数据集, 该数据集中所有的数据在室内的自然光下收集, 没有任何特殊光照, 收集的设备是普通摄像机, 一共8种手势(推动油门、手持操作杆、自然状态、基本的指令手势等)。数据集共3200张, 每个手势400张, 共10个人参与, 参与者均是模拟飞行环境中的随机采取手势图片, 每人采集数目不等。同时利用旋转平移等操作对得到的手势库进行扩充。然后对整个手势库中的每张照片中的手势轮廓线计算傅里叶描述子并保存, 保存格式为“x_i”, 表示手势_x的第i张图片。训练集与测试集按3:1的比率分配。测试集800张, 训练集2400张。图10经过肤色处理的手势数据集片段; 图11是经过关键特征点的手势数据集片段。

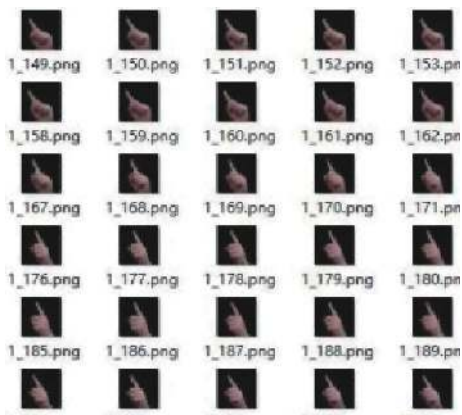


图10 经过肤色处理的手势数据集

本文提出的算法在 Python 中实现, 手势识别准确率在 98%。实验结果见表 1。

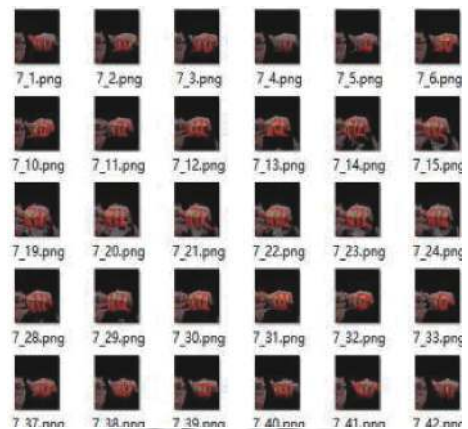


图11 经过关键特征点的手势数据集

表1 手势识别实验结果

测试集(张)	正确识别数	手势(种)	识别率(%)
800	784	8	98

4.1 实验结果对比

表2为使用本文方法与其他文献作对比, 文献[20]提出用 YCbCr 色彩空间检测肤色对输入的图像分割出感兴趣的手势区域, 然后再通过深度学习的方法训练出手势识别的模型。文献[9]从图像的3D数据出发, 以 Kinect 作为输入设备, 提出了一个基于深度图像的3D手势识别系统框架。文献[21]采用表面肌电(sEMG)信号进行手势识别, sEMG采集的趋势之一是带有等距电极的多通道臂带。即通过逐步进行包装器选择特征和减少维数提高sEMG的分类性能。考虑到40个维度, 对于带有高斯核的支持向量机和LMNN技术的结合, 平均精度为94%。而本文采用肤色结合特征点对手部分割、特征提取, 并用SVM向量机针对飞行模拟环境下的手势进行识别, 准确率达到98%。

表2 本文与其他文献对比

方法	识别率(%)
文献[9]	76.6
文献[20]	94.6
文献[21]	94
本文	98

5 结论与展望

本文结合现有的美国 CMU Perceptual Computing Lab 开源的手部关键点检测模型, 使用肤色进行二值化处理, 结合手部关键点准确分割手势, 接着对图像进行腐蚀膨胀处理, 在对手势轮廓特征提取的基础上加上

手部关键点特征经过 SVM 训练和测试实验后可知,此算法可以在飞行模拟环境中,面对背景复杂,遮挡影响,光照影响,能够有效地进行特征提取,极大地提高了手势识别效率.对于静态手势识别平均识别率能够达到98%.论文不足点:当面部和手部有重合时,特征提取时有一定的影响.因此本文在手势识别的完善上还有一定的提升空间,接下来将进一步研究.

参考文献

- 1 黄晶,宋宏川,钟明华,等.飞行模拟器最新发展研究.情报交流,2019,(1):69-73.
- 2 罗斌,姚鹏,翁冬冬,等.基于混合现实的新型轻量级飞行模拟器系统.系统仿真学报,2009,21(17):5406-5410.
- 3 周来.面向虚拟现实飞行模拟训练的视觉手交互技术研究[博士学位论文].南京:南京航空航天大学,2012.6.
- 4 周来,郑丹力,顾宏斌,等.虚拟现实飞行模拟训练中的视觉交互技术研究.航空学报,2013,34(10):2391-2401.
- 5 Zhou K, Wan Y, Huang HL. Research on recognition and application of hand gesture based on skin color and SVM. Journal of Computational Methods in Sciences and Engineering, 2020, 20(1): 269-278. [doi: 10.3233/JCM-193731]
- 6 陈宁,范英豪,白冰.融合帧间差分法和ViBe的运动目标检测算法.浙江科技学院学报,2020,32(1):32-37.
- 7 常晓锋,冯晓毅.基于背景减法和时空熵的运动目标检测新方法.计算机仿真,2008,25(3):235-238. [doi: 10.3969/j.issn.1006-9348.2008.03.063]
- 8 张圆圆.基于视觉的手势识别技术及其应用研究.计算技术与自动化,2015,34(1):131-135. [doi: 10.3969/j.issn.1003-6199.2015.01.028]
- 9 Le V, 朱煜, Nguyen A. 深度图像手势分割及HOG-SVM手势识别方法研究.计算机应用与软件,2016,33(12):122-126. [doi: 10.3969/j.issn.1000-386x.2016.12.030]
- 10 彭正初.基于傅里叶描述子的物体形状识别的研究[硕士学位论文].哈尔滨:哈尔滨工业大学,2016.
- 11 Wang Q, Gao X, Wang F, et al. Feature point matching method based on consistent edge structures for infrared and visible images. Applied Sciences, 2020, 10(7): 2302. [doi: 10.3390/app10072302]
- 12 Tang ZJ, Zhang HY, Pun CM, et al. Robust image hashing with visual attention model and invariant moments. IET Image Processing, 2020, 14(5): 901-908. [doi: 10.1049/iet-ipr.2019.1157]
- 13 张晓燕.基于肤色和模板匹配的静态手势识别.电子世界,2019,(15):202.
- 14 郝禹哲,张玉金,田海越,等.基于卷积神经网络的手势识别方法.传感器与微系统,2020,39(5):48-50.
- 15 丁泽宇,弓伟.基于隐马尔可夫模型的动态手势识别研究.电脑与电信,2019,(5):33-36.
- 16 田英冬,张元,韩燮.基于动态时间规整算法的虚拟沙画手势识别方法.中国科技论文,2017,12(20):2361-2367. [doi: 10.3969/j.issn.2095-2783.2017.20.015]
- 17 梁金明,魏正曦.OSTU算法的改进研究.四川理工学院学报(自然科学版),2010,23(5):543-545.
- 18 AIHGF. OpenPose 基于 OpenCV DNN 的手部关键点检测. <https://www.aiuai.cn/aifarm1111.html>.
- 19 Simon T, Joo H, Matthews I, et al. Hand keypoint detection in single images using multiview bootstrapping. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 4645-4653.
- 20 杨洋,郑紫微,孙兹昂,等.基于肤色分割与深度学习的手势识别.数据通信,2019,(5):30-33. [doi: 10.3969/j.issn.1002-5057.2019.05.007]
- 21 Junior JJAM, Freitas MLB, Siqueira HV, et al. Feature selection and dimensionality reduction: An extensive comparison in hand gesture classification by sEMG in eight channels armband approach. Biomedical Signal Processing and Control, 2020, 59: 101920. [doi: 10.1016/j.bspc.2020.101920]