

文章编号: 2095-2163(2021)07-0212-05

中图分类号: TP391.4

文献标志码: A

基于特征融合的手势识别

陈忠辉¹, 王等准¹, 万广¹, 方洪波¹, 黄以卫², 谢本亮¹

(1 贵州大学 大数据与信息工程学院, 贵阳 550025; 2 贵州财经大学 西部现代化研究中心, 贵阳 550025)

摘要: 在手势识别的过程中, 手势的多样性和复杂性会对识别的可靠性和准确性带来较大影响。基于视觉的手势识别通常采取单一特征用于手势分类, 但是单一特征无法较好地描述整个图像。因此本文提出多种特征融合的方法, 分别提取改进后的梯度方向直方图(HOG)特征和MB-LBP特征, 并进行特征融合, 结合支持向量机(SVM)分类器完成手势图像的识别。实验结果表明, 提取的融合特征包含手势图像的局部区域梯度信息和图像的纹理信息, 可以更加全面地描述图像的手势特征。相较于单一特征识别方法而言, 基于特征融合的方法有着更高的识别率。

关键词: 手势识别; MB-LBP; 改进 HOG 特征

Gesture recognition based on feature fusion

CHEN Zhonghui¹, WANG Dengzhun¹, WAN Guang¹, FANG Hongbo¹, HUANG Yiwei², XIE Benliang¹

(1 College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China;

2 West China Modernization Research Center, Guizhou University of Finance and Economics, Guiyang 550025, China)

[Abstract] In the process of gesture recognition, the diversity and complexity of gestures will have a greater impact on the reliability and accuracy of recognition. Vision-based gesture recognition usually adopts a single feature for gesture classification, but a single feature cannot describe the entire image well. Therefore, this paper proposes a variety of feature fusion methods, respectively extracting the improved histogram of gradient (HOG) features and MB-LBP features, and performing feature fusion, combined with SVM classifier to complete the recognition of gesture images. Experimental results show that the extracted fusion features include the local area gradient information of the gesture image and the texture information of the image, which can describe the gesture features of the image more comprehensively. Compared with the single feature recognition method, the method based on feature fusion has a higher recognition rate.

[Key words] gesture recognition; MB-LBP; improved HOG features

0 引言

手势是一种十分重要的沟通交流方式, 若能对手势进行成功识别就必将为人类与机器、设备和计算机的交互开创出崭新局面。随着计算机技术、尤其是增强现实^[1], 辅助驾驶^[2]等技术的飞速发展, 通过手势识别进行人机交互^[3]的需求越来越广泛。手势识别已然成为当前的热点研究方向之一, 因此手势识别技术的研究具有十分重要的现实意义。

研究可知, 基于视觉的手势识别包括着3个基本步骤: 手势分割^[4]、手势特征提取^[4]、手势识别^[5]。其中, 特征提取对手势识别的准确率有着至关重要的作用。当前, 常见的手势特征通常包括了形状特征和纹理特征。具体来说, 形状特征主要有: 梯度直方图^[6](HOG)和Haar-like^[7]特征; 纹理特征

有局部二值图^[8](LBP)和Gabor^[9]特征。这些特征均已被应用于基于视觉的手势识别中。

翁汉良等人^[10]将指尖个数和方向作为手势特征, 结合轮廓长度和面积等几何特征完成手势识别。杨学文等人^[11]根据手势主方向建立二维手势直角坐标系特征, 利用空间手势坐标点分布特征方法对手势进行初步识别, 再利用类-Hausdorff距离模板^[12]匹配的思想识别最终的手势。肖宇^[13]采用方向梯度直方图(HOG)来获得手势特征, 然后使用随机森林作为分类器实现对手势的识别。卢梦圆等人^[14]分别提取手势图像的梯度方向直方图(HOG)和局部二值模式(LBP)两种特征, 并进行特征融合, 结合支持向量机算法完成手势识别。赵倩楠等人^[15]采用串行融合方法将手势指尖特征和手部HOG特征相结合, 并采用支持向量机(SVM)^[16]作

基金项目: 国家自然科学基金(61562009); 贵州大学人才引进项目(2015-29); 半导体功率器件教育部工程研究中心开放基金项目(ERCMEKFJJ2019-(06)); 贵州财经大学人才引进项目(校聘发[2020]4号)。

作者简介: 陈忠辉(1995-), 男, 硕士研究生, 主要研究方向: 图像处理、目标检测; 谢本亮(1978-), 男, 博士, 副教授, 主要研究方向: 图像处理、计算机视觉、目标检测。

收稿日期: 2021-04-06

为分类器, 获得了较高的识别准确率。

由于单一特征描述手势图像信息存在局限性, 本文提出一种基于特征融合的手势识别方法。对预处理后的手势图像分别提取改进后的梯度方向直方图(HOG)特征和 MB-LBP^[17]特征并进行特征融合, 基于 SVM 分类器完成手势图像的识别。实验结果表明, 相较于单一特征识别, 本文算法对手势图像有着更高的识别准确率。

1 特征融合的手势识别

1.1 算法流程

基于特征融合的手势识别步骤可表述为: 对采集的手势图像进行预处理操作, 归一化尺寸并去除背景区域对分类结果的影响; 分别提取手势图像的改进后的 HOG 特征和 MB-LBP 特征并进行特征融合; 结合支持向量机(SVM)分类器进行分类和识别。算法流程如图 1 所示。

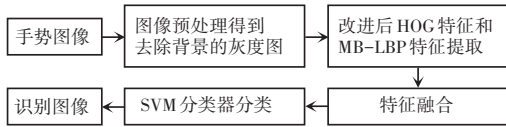


图 1 算法流程

Fig. 1 Algorithm flow chart

1.2 图像预处理

图像预处理是图像识别中的重要过程。经过预处理后的图像质量会得到很大提高, 从而得到更好的分类效果。本文对手势图像预处理过程如图 2 所示。

步骤 1 为提高图像特征提取、训练和预测速度, 将图像尺寸设置为 128 像素×128 像素, 并将图像由彩色 RGB 格式转换为 HSV 格式图像, 部分手势的 RGB 图像如图 2(a) 所示。

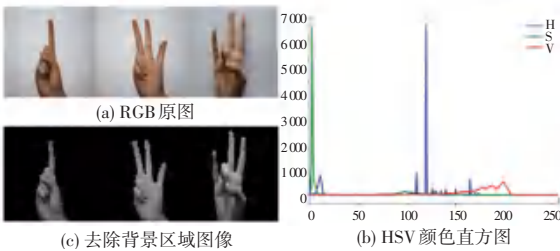


图 2 图像预处理

Fig. 2 Image preprocessing

步骤 2 为去除背景区域对手势图像识别的影响, 利用 HSV 颜色直方图(见图 2(b))将手势区域与背景区域进行分割, 分割阈值为:

$$\begin{cases} 0 \leq H \leq 180 \\ 0 \leq S \leq 43 \\ 46 \leq V \leq 255 \end{cases} \quad (1)$$

若手势图像某点像素的像素值处于阈值之内, 则将其像素值设为 0, 否则将其设为 255。由此可以获取背景区域已去除的二进制图像。将二值图像用作掩模, 并对灰度图像进行形态“与”运算, 得到去除背景区域的手势图像如图 2(c) 所示。经过图像预处理后, 手势图像去除了背景区域, 同时保留了手势的外观和形状等信息, 避免识别结果受手势变形的影响。

1.3 改进梯度特征提取

HOG 是一种描述图像本身梯度方向的特征, 其主要思想是通过计算图像局部区域的梯度直方图来表示图像局部的纹理信息和形状大小。当图像是用频域表示时, 低频区域表示图像的能量, 中频区域表示图像的纹理细节, 高频区域表示图像的边缘和噪声。传统 HOG 特征提取算法使用 Sobel 算子提取图像水平和垂直方向梯度信息。Sobel 算子使用整数阶微分可以完成图像高频边缘区域检测, 但同时也会造成中频区域信息模糊化, 增加图像噪声强度。相较于传统整数阶微分的边缘检测算子, 分数阶微分^[18]的边缘检测算子检测图像高频边缘区域同时可保留图像中频信号, 达到增强图像纹理信息的目的, 使提取的特征信息更加丰富。

由传统整数阶微分运算法则, 推导出一元函数分数阶微分差分表达式为:

$$\frac{\partial^a f(t)}{\partial t^a} = f(t) + (-a)f(t-1) + \frac{(-a)(-a+1)}{2} f(t-2) + \dots + \frac{\Gamma(-a+1)}{m! \Gamma(-a+m+1)} f(t-m) \quad (2)$$

其中, $a \in (0, 1)$ 表示微分步长; t 为分数阶微分上限; $\Gamma(n)$ 为 Gamma 函数。

由于图像是二维空间, 所以一元函数分数阶微分差分表达式推广到二维空间, 得到二维分数阶微分在水平和垂直方向上差分表达式分别为:

$$\frac{\partial^a f(x, y)}{\partial x^a} = f(x, y) + (-a)f(x-1, y) + \frac{(-a)(-a+1)}{2} f(x-2, y) \quad (3)$$

$$\frac{\partial^a f(x, y)}{\partial y^a} = f(x, y) + (-a)f(x, y-1) + \frac{(-a)(-a+1)}{2} f(x, y-2) \quad (4)$$

根据式(2)、式(3),可以在 x, y 方向上构造的分数阶微分算子分别为: $[\frac{a^2 - a}{2}, -a, 1]$, $[\frac{a^2 - a}{2}, -a, 1]^T$ 。

其中, $a \in (0, 1)$ 表示微分变量, $a < 0$ 时表示积分, $a > 0$ 时表示微分, $a = 0$ 时既不是积分, 也不是微分。

改进的 HOG 特征值的提取方法主要是把分数阶微分应用到边缘检测算子中, 通过梯度优化算子获取更多特征信息。改进后的 HOG 特征提取包括以下步骤:

(1) 采用分数阶微分算子提取图像在水平方向和垂直方向上的梯度。

(2) 将梯度方向等间隔量化到 k 个区间, 并将图像均匀划分为 $m \times m$ 个元胞。统计每个元胞内 k 个方向的投影, 生成梯度方向直方图。本文设置 $k = 9$, 像素的梯度方向为 $0^\circ, 20^\circ$, 特征向量为 9 维。

(3) 把相邻 2×2 元胞单元组合成一个区域块 (block), 区域块内归一化梯度直方图。

(4) 串联区域块内 (block) 获取的梯度直方图特征, 得到改进后的 HOG 特征向量。

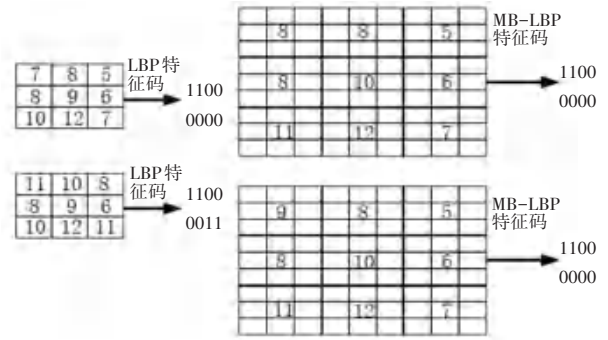
1.4 MB-LBP 特征提取

LBP 特征是一种用来描述图像局部纹理特征的算子, 传统 LBP 算子采取的是将中心点像素与领域采样的像素值进行比较, 这是将单个像素与单个像素之间做出比较, 因此 LBP 对噪声较为敏感。且当拍摄时不够明亮, 亮度不够均匀, 图像传感器长期工作, 温度过高等因素都会导致图像富含噪声, 此时 LBP 特征存在明显不足。为提高特征描述子的抗噪能力, 采用对噪声不敏感的 MB-LBP 提取手势图像的纹理特征。

传统 LBP 和 MB-LBP 的特征提取原理如图 3 所示。图 3(a) 为传统 LBP 特征提取编码图, 图 3(b) 为 MB-LBP 特征提取编码图。MB-LBP 特征提取尺寸为 3×3 个块 (block), 每个块中含有 3×3 个像素, 每块取对应 3×3 区域块内像素平均值。假设 MB-LBP 特征提取的第一个区域块正好对应 LBP 特征提取的 3×3 区域块。由图 3 可以看出, 当 LBP 特征提取的区域块内某些像素点受噪声影响像素值发生改变时, LBP 特征提取编码发生改变, 但 MB-LBP 特征提取编码没有受到噪声影响。且当 MB-LBP 的区域块取值为 1×1 时, MB-LBP 相当于传统非统一模式的 LBP。因此 MB-LBP 不仅关注单个像素点, 而且关注块与块之间的像素。

将预处理后的手势图像进行 MB-LBP 特征提取, 将获得的 MB-LBP 特征图分为互不重叠的 $L \times L$ 个区域块 (block), 计算每个区域块 (block) 的直方图并串联成一个直方图特征。该直方图特征向

量即为手势图像的纹理特征向量。



(a) 传统 LBP 特征提取 (b) MB-LBP 特征提取编码图

图 3 LBP 和 MB-LBP 原理图

Fig. 3 Schematic diagram of LBP and MB-LBP

将预处理后的手势图像进行基于分数阶微分的 HOG 特征和 MB-LBP 纹理特征提取, 采用简单的串联连接对 2 种特征进行融合, 将获得的 2 种特征联合作为手势图像融合特征。

1.5 分类器建立

支持向量机 (SVM) 是一种基于结构风险最小化原则的分类方法, 在解决小样本、非线性的数据及高维数据方面有着显著优势。因此, 研究中采用支持向量机算法 (SVM) 作为本文分类算法。选取稳定性和准确性均较好的 Liner 核函数作为 SVM 核函数。采用网格搜索法对模型参数进行寻优, 确定 SVM 的惩罚因子 $C = 1$ 。

2 实验结果与分析

2.1 实验环境

实验计算机配置信息是 Intel core CPU i5-7500 处理器, 8 GB 内存, NVIDIA GTX1060 显卡, 操作系统为 Win7。为了减少实验误差, 本文对比的算法识别率为 10 次实验下的平均识别率 avr , 具体为:

$$avr = \frac{\sum_{i=1}^{10} P_i}{S} \times 100\% \quad (5)$$

其中, P_i 代表第 i 次实验正确预测样本数, S 为预测样本总数。

采用百度 AI Studio 手势数据集进行训练与测试, 该数据集中手势分为 10 类, 共计 2 602 张手势图像, 对应数字 0~9, 部分手势图像如图 4 所示。采用分层采样方法, 并按 7:3 比例划分训练集和测试集。

2.2 实验参数

2.2.1 微分变量大小选择

微分变量的大小选择对改进后的 HOG 特征提

取和手势图像识别准确率有着较大影响。不同微分变量下识别率如图 5 所示。图 5 中, 识别率均为不同块尺寸和分块方式下最优识别率。从图 5 可以看出, 当微分大小为 0.5 时可以取得较好的实验结果。



图 4 部分手势图像

Fig. 4 Part of the gesture images

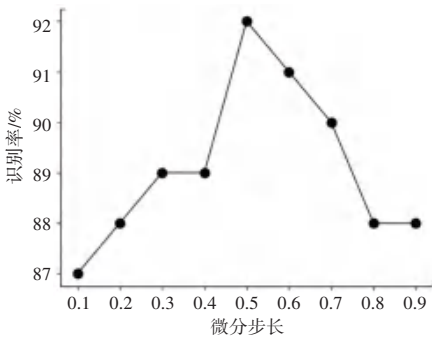


图 5 不同微分变量下识别率

Fig. 5 Recognition rate under different differential variables

2.2.2 块的尺寸选择和分块方式

MB-LBP 块的尺寸选择和 MB-LBP 特征图分块方式对手势图像的纹理特征有较大影响。MB-LBP 块的尺寸和分块方式与识别率关系的仿真结果如图 6 所示。由图 6 可以看出当 MB-LBP 特征图分块数目确定时, 不同尺寸块的 MB-LBP 下的识别率不同。根据图 6 可得, 本文 MB-LBP 块的尺寸设置为 2×2 , 分块数目设置为 4×4 。

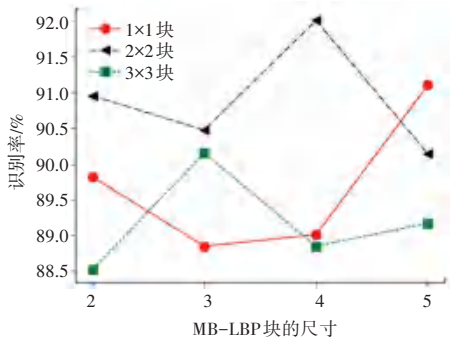


图 6 MB-LBP 块的尺寸和分块方式与识别率的变化关系

Fig. 6 The relationship between MB-LBP block size and block method and recognition rate

2.3 识别结果与分析

为验证本文算法对手势图像的识别性能, 将本文算法与单一特征的 HOG 特征匹配算法、LBP 特征

匹配算法的分类准确率进行比较, 不同算法的识别结果如图 7 所示。10 种手势平均识别准确率见表 1。

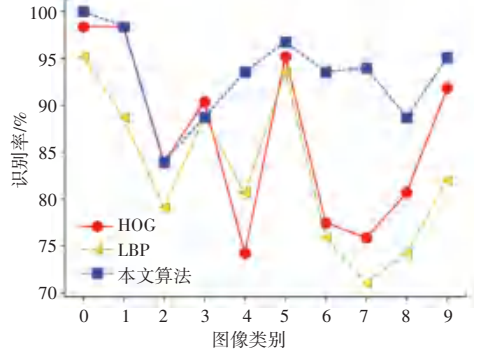


图 7 算法比较

Fig. 7 Algorithm comparison

表 1 算法平均识别率

Tab. 1 The average recognition rate of the algorithm

方法	测试/张	识别/张	平均识别率/%
HOG	619	536	86.59
LBP	619	513	82.88
本文算法	619	571	92.25

由表 1 分析可知, 使用单一特征的 HOG 特征匹配算法和 LBP 特征匹配算法的识别准确率分别为 86.59% 和 82.88% 相对较低, 且 HOG 特征匹配算法识别率高于 LBP 特征匹配算法识别率。相较于使用单一特征的 HOG 特征匹配和 LBP 特征匹配算法, 本文算法在识别准确率上分别提高 5.66% 和 9.37%, 达到 92.25%, 满足工业应用对手势图像识别率的要求。

3 结束语

针对单一特征描述手势图像存在局限性问题, 提出一种基于特征融合的手势识别方法。先后提取手势图像改进的分数阶微分的 HOG 特征和 MB-LBP 纹理特征, 将 2 种特征融合后结合 SVM 分类器进行分类识别。实验结果表明本文算法可以对手势图像实现精确识别, 识别准确率达到 92.25%, 优于其他特征识别方法。但是由于手势的多样性, 在未来的工作中将会对遮挡引起的不可识别的手势做进一步研究, 增强手势识别的多样性。

参考文献

[1] 袁庆曙, 王若楠, 潘志庚, 等. 空间增强现实中的人机交互技术综述[J]. 计算机辅助设计与图形学学报, 2021, 33(3): 321-332.

[2] 周天彤, 徐飞林, 张旖帆, 等. 基于 unity 和 kinect 的交警手势识别仿真系统的设计和实现[J]. 计算机测量与控制, 2016, 24(5): 156-159.