

基于肤色模型的动态手势分割与识别方法研究^{*}曹国华^{1,2} 刘福迪¹ 马国庆^{1,2} 刘 丽^{1,2}

(1. 长春理工大学机电工程学院 长春 130022; 2. 长春理工大学重庆研究院 重庆 401135)

摘 要: 为解决现有动态手势识别方法在复杂环境下因相近肤色、光照变化等因素造成势识别率不高、鲁棒性欠佳等问题,提出一种基于椭圆肤色模型的动态手势分割与识别方法。首先利用 YCrCb 颜色空间中的 Cr 分量结合 OTSU 阈值分割算法分割出手部区域。其次,针对手势复杂程度以及手指粗细不同的情况下直接运用形态学处理可能导致细节丢失,影响识别的准确性问题,对传统 Canny 算法进行改进,并结合形态学处理对手部边缘填充;然后通过结合 Kalman 与改进的 CamShift 算法来对手势进行跟踪,完成动态手势分割;最后通过 BP 神经网络对分割后的动态手势进行识别,通过优化算法在 GPU 上的实现,利用 GPU 的并行处理能力加速图像处理、特征提取及神经网络前向传播等计算密集型任务。这部分优化措施显著提升了动态手势识别方法的实时性能,使其能够更好地适应于各类对实时性要求高的应用场景。实验结果表明:此方法在应对复杂背景及光照环境变化时具有较强的鲁棒性及抗干扰能力,平均识别率可达 94.67%。

关键词: 机器视觉;动态手势分割;椭圆肤色模型;Otsu 阈值分割算法;彩色-深度图像

中图分类号: TP391;TN98 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Study on dynamic gesture segmentation and recognition
using skin color modelsCao Guohua^{1,2} Liu Fudi¹ Ma Guoqing^{1,2} Liu Li^{1,2}

(1. School of Mechanical and Electrical Engineering, Changchun University of Science and Technology, Changchun 130022, China;

2. Chongqing Research Institute, Changchun University of Science and Technology, Chongqing 401135, China)

Abstract: In order to solve the problems of low potential recognition rate and poor robustness caused by similar skin tones and lighting changes in complex environments, a dynamic gesture segmentation and recognition method based on elliptical skin color model was proposed. Firstly, the Cr component in the YCrCb color space combined with the OTSU threshold segmentation algorithm was used to segment the hand region. Secondly, in view of the problem that the direct application of morphological processing may lead to the loss of details and affect the accuracy of recognition in the case of different gesture complexity and finger thickness, the traditional Canny algorithm was improved, and the morphological processing was combined with the filling of hand edges. Then, by combining Kalman and the improved CamShift algorithm to track the gestures, the dynamic gesture segmentation is completed. Finally, the segmented dynamic gestures are recognized by the BP neural network, and the implementation of the optimization algorithm on the GPU is used to accelerate the computation-intensive tasks such as image processing, feature extraction and forward propagation of the neural network by using the parallel processing capability of the GPU. This optimization measure significantly improves the real-time performance of the dynamic gesture recognition method, making it better suitable for various application scenarios with high real-time requirements. Experimental results show that the proposed method has strong robustness and anti-interference ability in response to complex background and lighting environment changes, and the average recognition rate can reach 94.67%.

Keywords: machine vision; dynamic gesture segmentation; oval skin tone model; OTSU threshold segmentation algorithm; color-depth image

0 引 言

手势识别技术是人机交互领域的关键研究方向之一,

利用计算机分析和识别手势信息,使人机交互近似于人与人之间的沟通交流,让使用者有更加自然舒适的实际体验^[1]。相比于表面肌电信号与惯性信号的手势识别效果,

收稿日期:2024-08-21

* 基金项目:吉林省自然科学基金(YDZJ202301ZYTS423)、重庆市自然科学基金(CSTB2022NSCQ-MSX1120)、吉林省自然科学基金(YDZJ202301ZYTS263)、重庆市自然科学基金(cstc2021jcjy-msxmX0989)项目资助

计算机视觉动态手势识别的鲁棒性更高,更能适用于工业装配、生产等领域^[2]。从国内外对手势识别技术的研究历程来看,研究者在手势识别方法上的探索一直处于稳步提升的状态,主要研究方法可以分为基于接触式外设穿戴设备的手势识别和基于计算机视觉的手势识别^[3]。

近几年,深度学习技术的快速发展为手势识别领域提供了新的研究思路,并取得了许多显著的研究成果,基于深度学习的手势识别方法受到越来越多研究者的关注^[4]。牛雅睿等^[5]利用 Ghost 模块设计轻量级主干特征提取网络并引入加权双向特征金字塔网络改进特征融合网络,减少网络的参数量和计算量并且提高了网络精度,但是对复杂场景或极端光照条件下的手势识别能力较弱。黄杰^[6]将手势识别技术应用在手语识别领域,采用基于注意力机制的三维卷积神经网络来提取视频中手势的时空特征,有效提高了大规模手语孤立词的识别精度。晁人傑^[7]将三维信息压缩到二维中并且,并且使用小型 2D 卷积网络进行手势识别,采用注意力机制与 FPN(feature pyramid network)进行特征融合,大幅度减少网络参数量并提高网络对手势关键信息的获取能力。Sharma 等^[8]为解决复杂背景对于手势识别精度以及准确率的影响,在采用 YCbCr 颜色空间过滤手部肤色基础上使用 MediaPipe 来区分特定的手势区域,并通过动态模式分解(DMD)来完成时空特征提取,实现复杂背景下的手势识别,在 Ego hand 数据集的识别率 97.85%。Wu^[9]提出了一种双通道卷积神经网络模型用于手势识别,该算法是对预处理后的手势图像采用 Canny 边缘检测算法提取手势边缘特征,并将手势图像和边缘图像分别作为神经网络的两个输入通道进行学习和训练,提高了模型的泛化能力,但是仅能处理具有简单背景的手势图像。Noble 等^[10]将电容传感器阵列组成手势识别系统,通过系统采集手势数据对 MLP 分类器进行训练,平均准确率为 96.87%。邢益良等^[11]提出一种基于傅里叶算子的手势识别方法,通过候选窗口最大轮廓傅里叶描述子匹配度和置信度分割出手掌区域,根据手掌轮廓计算其傅里叶描述子得到手势轮廓特征值,最后通过 BP(backpropagation neural network)人工神经网络进行手势识别,可满足简单背景下手势识别的性能要求,但光照条件的变化会直接影响图像的质量和手势的可见性,过亮或过暗的光照都可能导致手势轮廓的模糊或丢失,使傅里叶描述子的提取变得困难,从而影响手势识别效果。唐文权等^[12]提出了一种基于 YCbCr 颜色空间的改进三帧差分法的动态手势识别方法。赵鸿图等^[13]提出一种复杂背景下多特征结合的深度学习手势识别,通过在 YCrCb 颜色空间采用椭圆肤色模型和 Otsu(nobuyuki otsu method)阈值肤色识别算法获取手势肤色特征,经算法判断后,对手势肤色图像采用改进 Canny 边缘检测算法获得手势边缘特征,并提出一种边缘填充方法对手势边缘图像处理,得到手势轮廓完整的手势边缘图像,但在某些情况下,边缘填充可能无法完美地模拟

原始图像中边缘的自然过渡。这可能导致填充区域与原始图像区域之间存在明显的界限或色差,影响图像的整体视觉效果和后续处理的效果。

上述手势识别方法取得了不错的识别效果,但各自算法仍存有不足之处。本文的研究目的使解决复杂环境下因相近肤色、光照变化等因素造成识别率不高、鲁棒性欠佳等题。研究方法为基于肤色特征与手势边缘特征相结合来对手势进行分割,然后通过结合 Kalman 与改进 CamShift (continuously adaptive mean shift)算法来对手势进行跟踪,完成动态手势分割,最后通过 BP 神经网络对分割后的动态手势进行识别,使其对于复杂背景下的手势识别场景具有一定社会使用价值。算法的主要创新包括如下 3 个方面:1)通过分析现有肤色识别算法在获取肤色特征时容易受到环境中光照变化以及肤色背景等因素影响,因此提出一种基于 YCrCb 颜色空间的肤色融合算法,对 YCrCb 颜色空间中的 Cr 分量单独进行 OTSU 处理,因为 OTSU 算法是对图像的灰度级进行聚类,所以有效避免光照影响,可获得效果更佳的手势肤色特征;2)结合形态学处理以及改进后的边缘检测算法对手部分割结果进行优化,确保手势细节信息的准确提取;3)运用基于 Kalman 滤波的改进 Camshift 算法进行动态手部跟踪,可增强对手势目标的识别能力,进而提高神经网络模型的性能以及效率。最终试验结果表明,本文算法对复杂背景情况下的动态手势有较好的识别效果,对复杂背景情况下的 10 类手势的识别效果较为理想,平均识别率均在 95.2%以上,平均识别时间均在 4.32 ms 以内。

1 手势分割

1.1 基于椭圆肤色模型的手部区域粗分割

在 YCbCr 色彩空间中,Y、Cb、Cr 分别代表亮度、蓝色色度、红色色度分量。YCbCr 颜色空间对于亮度变化的敏感性较低,因此为减少光照变化对手势识别或图像处理任务的影响,将 RGB 图像转换为 YCbCr 空间使亮度与色度信息分离^[14]。并且手部区域肤色在 YCbCr 颜色空间中聚类效果较好。YCbCr 肤色模型分割流程如图 1 所示。RGB 颜色空间到 YCbCr 颜色空间的转换公式如式(1)所示。

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \frac{1}{256} \begin{pmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{pmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

在 YCrCb 空间中 Cr 分量常用作于肤色检测方面,相对于背景区域来讲,肤色区域的 Cr 分量值通常相对较高,可以根据这一特征来提取肤色区域。OTSU 可自动确定最佳的分割阈值,不需要再次手动调节,并且 OTSU 算法是对图像的灰度级进行聚类,又有效避免了光照影响^[15]。对图 2 中所示类肤色背景下各种手势进行肤色分割,提取

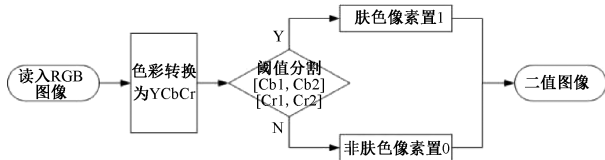


图 1 YCbCr 肤色分割流程

Fig. 1 YCbCr skin tone segmentation process

图像中的手部区域。查看是否能抵抗类肤色干扰问题,并观察在分割不同手势时是否会根据手势复杂程度的不同造成手部细节丢失等问题。提取结果如图 3 所示。



图 2 不同背景下复杂手势

Fig. 2 Complex gestures under different backgrounds



图 3 手部分割结果

Fig. 3 Hand segmentation results

通过运用 YCrCb 空间中 Cr 分量结合 OTSU 法对手部进行分割结果来看,在手势形状复杂或者手指较细时,直接使用形态学处理来平滑分割后的手部区域,可能会导致手势细节的丢失,影响后续识别的准确性。为提高手势分割效果,图像还需进一步处理。

1.2 改进的 Canny 边缘检测分割算法

Canny 算子是找到最优边缘的一种边缘检测算法,具有最优检测、最优定位和最优响应的特点,是实际应用中常见的边缘检测算法。手势相对复杂时,直接使用 Canny 算子会导致手势边缘细节信息丢失。Canny 算子的缺点是高斯滤波对图像的平滑处理会影响手势边缘的提取。此外,人为设定的双阈值也会导致边缘信息的丢失^[16]。针对 Canny 算子存在的不足之处,对其改进以获得效果更佳的手势边缘图像。具体改进如下:

1) 采用自适应加权中值滤波替换高斯滤波,可以避免自适应中值滤波的局限性,同时保留更多的边缘信息。

2) 为了提高检测的准确性,本文使用 3×3 的 Sobel 算子,并在 Sobel 算子水平和垂直梯度模板的基础上,采用 4 个方向的梯度模板对图像边缘进行计算,增加了 45° 和 135° 方向的检测,改进的梯度模板公式为:

$$S_0 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (2)$$

$$S_{45} = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} \quad (3)$$

$$S_{90} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (4)$$

$$S_{135} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix} \quad (5)$$

改进后的梯度计算方法综合考虑了当前像素 4 个方向周围 8 邻域的梯度幅值,对边缘的定位更加准确,也减少了边缘丢失现象和伪边缘的出现^[17]。增幅计算公式如下:

$$G(x,y) = \sqrt{G_0^2(x,y) + G_{45}^2(x,y) + G_{90}^2(x,y) + G_{135}^2(x,y)} \quad (6)$$

式中: (x,y) 为手部图像中的像素点; $G_0(x,y)$ 、 $G_{45}(x,y)$ 、 $G_{90}(x,y)$ 和 $G_{135}(x,y)$ 分别为在像素点 (x,y) 处 X、Y、 45° 和 135° 方向上的梯度分量^[18]。

3) 传统 Canny 算法需要人为设置高低阈值,自适应性差且检测边缘不够准确,因此本文采用将采用局部自适应阈值的方法代替传统 Canny 算子中非极大值抑制技术的双阈值的全局固定方法来实现对手势边缘的检测,避免因手势复杂程度不同而造成的细节丢失。利用 3×3 矩阵在手势图像内平滑移动,将 3×3 平滑框内像素点的灰度平均值作为自适应阈值 T_d , 其计算公式如下:

$$T_d = \frac{1}{9} (\sum M(x,y)) \quad x = 1,2,3 \quad y = 1,2,3 \quad (7)$$

式中: $M(x,y)$ 为 3×3 平滑框内 (x,y) 处的灰度值。

对于每个矩阵的中心像素,计算其梯度幅值并与自适应阈值进行比较:如果梯度幅值大于阈值,则为边缘点,标记为 1。如果梯度幅值很小(小于阈值的一半),则不是边缘点,标记为 0。如果梯度幅值在阈值的一半和阈值之间,则参考前一个中心像素的标记;如果前一个点是边缘点,则当前点也视为边缘点;否则不是。通过这种方式,遍历整个图像,根据每个像素的梯度幅值和自适应阈值来判断其是否为边缘点。

4) 融合形态学处理技术提取完整的手势轮廓边缘。在对手部图像进行分割提取后,准确获取轮廓边缘细节信息,为了解决背景噪声影响导致获取到的轮廓信息不完全以及有空洞等问题。本文将图像形态学处理技术与改进的 Canny 算子相融合。首先,对局部自适应阈值处理后的手势部分进行灰度腐蚀操作,以消除部分噪声。其次,再通过孔洞填充操作和腐蚀操作进一步消除外部因素对轮廓边缘的影响。最后再次提取图像的边缘轮廓,直至得到完整、连续且光滑的手部边缘图像。实验结果如图 4 所示。



图 4 手部轮廓分割结果

Fig. 4 Hand contour segmentation result

2 动态手势跟踪

针对 Camshift 算法对光照和噪声比较敏感等问题,提出将 Camshift 算法进行改进。通过将改进后的 Camshift 算法与 Kalman 滤波相结合,来完成手势跟踪^[19]。其中,为了确保 Camshift 跟踪算法的可靠性,采取了二次搜索策略来调整搜索窗口的位置和大小。在优化 Camshift 算法的基础上,引入卡尔曼滤波技术可以提高改进后 Camshift 算法的跟踪效率和鲁棒性。

2.1 CamShift 算法的改进

为了增强 Camshift 跟踪算法的性能,对其进行了两方面的改进。首先,将原本基于 RGB 颜色模型的跟踪转换为使用 HSV 颜色模型。因为 HSV 色彩模型中的色调(Hue)、饱和度(Saturation)和亮度(Value)3个独立分量,使得算法能够更准确地锁定并追踪手势的色彩特征,即便在光线突变或背景颜色相近的复杂场景中,也能保持高度的稳定性与精确度。其对于光照和背景变化的抗干扰能力更强,能够更好地捕捉目标手势的特征。其次,优化了颜色概率分布的计算方式。相比于传统的 Camshift 算法会在每一帧图像的全局范围内计算所有像素点的颜色概率分布,改进后的算法只计算比当前跟踪窗口稍大(例如 1.1 倍)的局部区域内的颜色概率分布。这样一来,不仅减少了计算量,提高了算法的实时性,而且通过排除背景噪声的干扰,进一步提升了跟踪的准确性,在目标发生快速移动或姿态变化时,迅速调整搜索区域,确保跟踪的连续性和实时性。

综上所述,改进后的 Camshift 跟踪算法通过使用 HSV 颜色模型和局部搜索策略,有效提高了手势跟踪的鲁棒性和实时性。具体的改进型 Camshift 跟踪算法框图如图 5 所示。

2.2 结合 Camshift 和 Kalman 滤波

Kalman 与 CamShift 算法相结合对手势进行预测跟踪,从原理上讲可以认为是两种算法的线性拟合。如式(8)所示。

$$X_{k+1} = \alpha X_{k+1,k} + (1 - \alpha)Y \quad (8)$$

式中: X_{k+1} 、 $X_{k+1,k}$ 、 Y 分别为在 $k+1$ 时刻跟踪目标、Kalman 滤波器、Camshift 算子的预测位置。 α 为比例因子。

比例因子 α 会根据当时目标受干扰程度的不同而发生

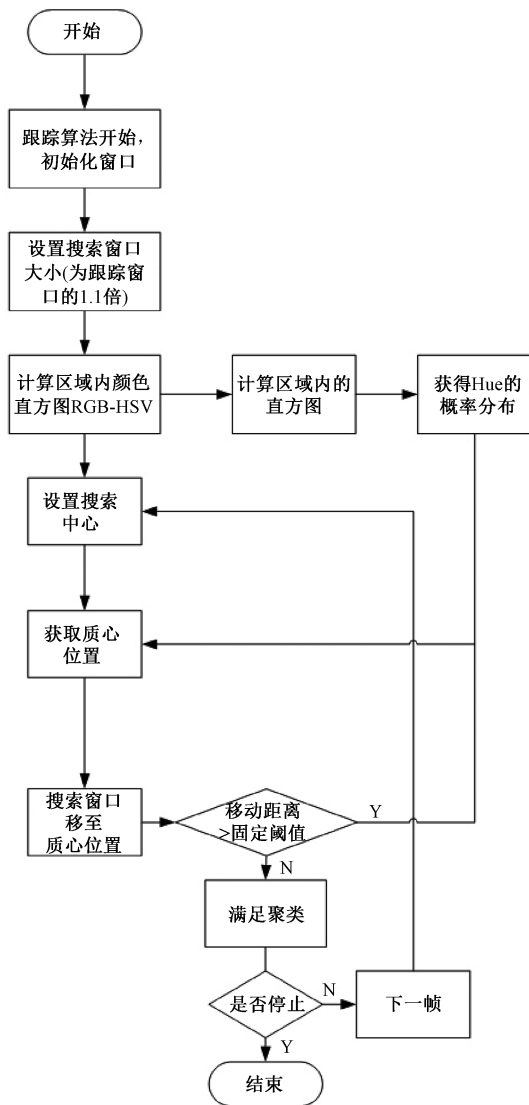


图 5 改进 Camshift 算法流程图

Fig. 5 Improved Camshift algorithm flowchart

改变。通过动态调整比例因子 α ,能够在不同干扰条件下保持较高的跟踪性能。整个组合算法的流程如图 6 所示,通过交替使用 Kalman 预测和 Camshift 搜索,以及对预测位置的加权融合,实现了对运动目标的稳定跟踪。

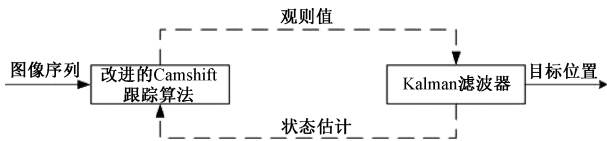


图 6 Camshift 与 Kalman 结合算法

Fig. 6 Camshift combines algorithms with Kalman

3 实验结果与分析

利用 YCrCb 颜色空间中的 Cr 分量结合 OTSU 法阈值分割算法分割出手部区域,并使用改进后的 Canny 算法融合形态学处理技术提取完整的手势轮廓边缘,去除干扰

影响^[20]。接下来结合 Kalman 与 CamShift 算法对动态手势进行跟踪,提取出具有运动和肤色双重特征的动态手势,最后通过 BP 神经网络对分割后的动态手势进行识别。

3.1 手势分割实验

利用 YCrCb 颜色空间中的 Cr 分量结合 OTSU 算法分割出手部区域,并将分割结果运用形态学处理,最终手势分割结果(二值化后)如图 7 所示。结果表明:面对不同光照及类肤色背景影响下采用 YCrCb 颜色空间中的 Cr 分量结合 OTSU 法阈值分割算法都能很好的对手势进行分割^[20]。结合图 8 所示不同方法手势分割时间对比得出,采用 YCrCb 颜色空间中的 Cr 分量结合 OTSU 法阈值分割算法在分割时间处于不同方法平均值的基础上,提升了手势分割准确率,进而提高后续动态手势识别的效率。

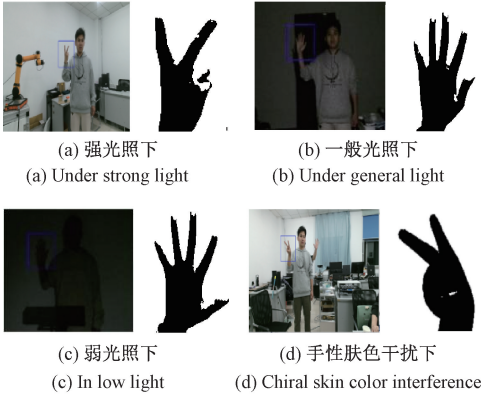


图 7 不同光照下手势分割效果

Fig. 7 Gesture segmentation under varied lighting

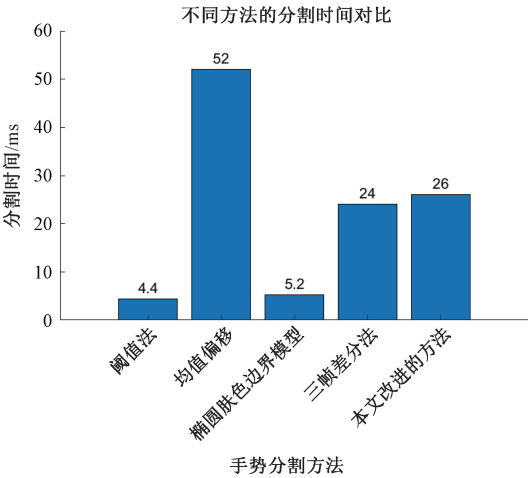


图 8 手势分割时间对比

Fig. 8 Gesture segmentation time comparison

3.2 动态手势跟踪实验

根据改进的跟踪算法,针对不同光照环境及类肤色背景进行了实验。图 9 所示为算法在不同背景环境影响下对第 1、8、15、23 帧的跟踪结果。可以看出在不同背景环境影响下,此算法对运动手势均可进行准确有效的跟踪;通过多

次实验发现利用 GPU 的并行处理加速改进的 CamShift 与 Kalman 算法进行跟踪效果可达 16.3 ms/帧左右。

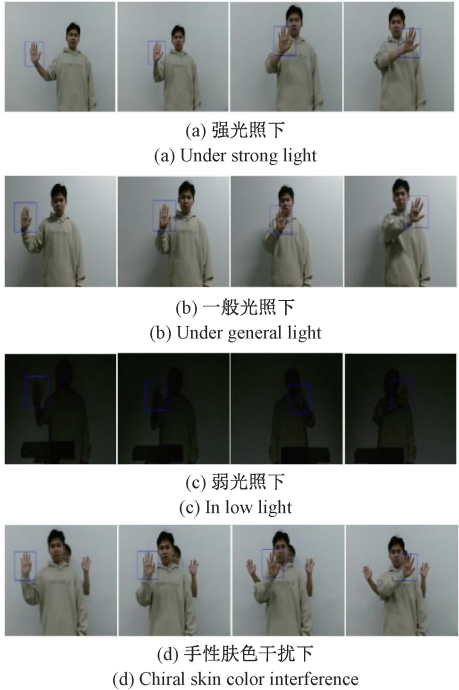


图 9 不同干扰下跟踪效果

Fig. 9 Tracking under different interferences

3.3 手势识别实验

综上所述,本实验成功构建了一个高效的动态手势识别系统,如图 10 所示。该系统基于一台配备 64 位 Windows 11 操作系统的 PC 机,核心硬件包括 AMD Ryzen 7 5800U 处理器(集成 Radeon Graphics,主频 1.90 GHz,八核架构)与 NVIDIA GeForce RTX 3050 Laptop GPU,辅以 Kinect v2 相机以捕捉精准的手势数据。开发环境则选用了 Python 3.7 编程语言,搭配 PyCharm 作为代码编辑器,并依托 TensorFlow-GPU 2.7.0 与 Keras 2.7.0 这两个强大的深度学习框架,同时利用 CUDA 11.3.1 与 cuDNN 11.3 优化 GPU 性能,实现高效计算。

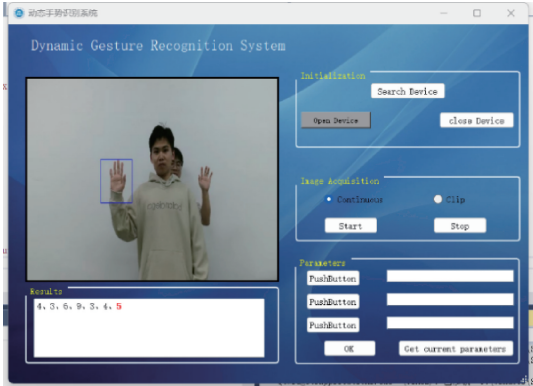


图 10 动态手势识别系统

Fig. 10 Dynamic gesture recognition system

为评估所设计系统在动态手势识别任务中的性能,调用动态手势数据库 CGD(chaLearn gesture data) 作为测试基准数据集来进行评估,该数据及涵盖了丰富的动态手势样本。为了全面评估算法效果,从 CGD 数据库中精心挑选了包含 3 000 个样本的测试集,覆盖了从 1~9 的多种动态手势,确保了测试的全面性和代表性。

为追求结果的可靠性与稳定性,实施了 11 次独立的实验,每次实验均随机抽取测试集中的部分样本进行识别,有效减少了偶然误差。应用所设计的动态手势识别系统与三帧差分法结合 BP 神经网络进行对比,分别进行动态手势识别。通过对比分析两种方法在识别率上的表现(如图 11 所示),发现优化后的算法在识别精度上展现出显著优势。具体实验结果汇总于表 1。

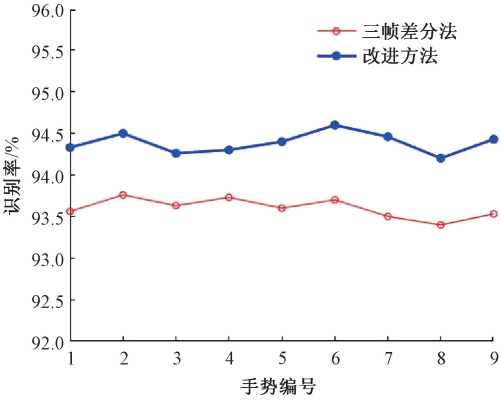


图 11 手势识别率对比

Fig. 11 Comparison of gesture recognition rates

表 1 不同分割算法最终识别情况对比

Table 1 Comparison of the final recognition of different segmentation algorithms

手势	图片数量	三帧差分法正确次数	跟踪帧时间/ms	改进方法正确次数	跟踪帧时间/ms
1	3 000	2 805	13.12	2 830	13.06
2	3 000	2 814	13.10	2 835	13.01
3	3 000	2 803	12.96	2 828	12.78
4	3 000	2 808	12.99	2 829	12.82
5	3 000	2 814	12.72	2 832	12.67
6	3 000	2 805	13.12	2 838	12.90
7	3 000	2 802	13.11	2 834	13.03
8	3 000	2 811	13.33	2 826	13.20
9	3 000	2 808	12.10	2 833	11.75

由表 1 及图 9 可以看出,改进的算法整体识别效果较好,平均识别率达到 94.41%,而传统的三帧差分法平均识别率只有 93.60%,改进的方法较传统检测法识别率提升 1%左右,平均每帧消耗时间也较三帧差分法减少了 0.15 ms 左右。

为检测改进方法在复杂环境中的识别效果,由 8 个人在不同光照及背景下,首先每人针对每个手势录入 5 次,用于训练样本,然后每人针对每个手势录入 10 次作为测试样本,测试样本总数为 720 个。对不同光照及背景情况进行了实验,实验结果如表 2 所示。

表 2 背景影响下识别情况对比

Table 2 Comparison of recognition situations under the influence of background

背景	光照强		光照一般		光照暗	
	准确率/%	识别时间/ms	准确率/%	识别时间/ms	准确率/%	识别时间/ms
类肤色背景	95.2	4.28	96.9	4.35	97.2	4.38
单一背景	95.9	4.20	97.9	4.25	96.3	4.30

由表 2 中识别率可知改进的方法针对不同环境背景下的平均识别率均在 95.2% 以上,平均识别时间均在 4.29 ms 以内。由此可见改进的方法具有较强的鲁棒性,在复杂背景及光照影响条件下的识别效果都较好,可以应用于复杂背景下的人机交互中。

4 结 论

针对人机交互过程中动态手势识别易受无关肤色、光

照变化等因素影响,手势识别率不高、鲁棒性欠佳等问题,提出了一种基于椭圆肤色模型的动态手势分割、识别方法。首先利用椭圆肤色模型去除光照影响,接下来运用 Canny 算子结合形态学处理对分割后的手部区域进行优化,去除类肤色背景影响。通过结合 Kalman 与改进的 CamShift 算法来对手势进行跟踪,完成动态手势分割。最后利用 BP 神经网络完成手势识别。经过实验及结果分析,该方法能够满足人机交互中动态手势识别需求,平均

识别率达到 94.41%，平均每帧消耗的时间在 12.80 ms 左右。另外，所改进的方法还具有鲁棒性强、准确性高等优点，能够很好地应对外接光照变化及复杂背景情况，同时克服手势发生短暂停留的问题。后续将对算法进行进一步优化，提升方法准确性，提高运行效率，推进人机交互发展进程。

参考文献

- [1] 刘翔,刘新妹,李传坤,等.一种基于特征位移的手势识别方法[J].电子测量技术,2024,47(8):141-147.
LIU X, LIU X M, LI CH K, et al. A gesture recognition method based on feature displacement[J]. Electric Measurement Technology, 2024, 47(8): 141-147.
- [2] 赵世昊,周建华,伏云发.注意力机制 CNN 结合肌电特征矩阵的手势识别研究[J].电子测量与仪器学报,2023,37(6):59-67.
ZHAO SH H, ZHOU J H, FU Y F. Gesture recognition of attention mechanism CNN combined with electromyography feature matrix[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(6):59-67.
- [3] ZHOU K, WAN Y, HUANG H. Research on recognition and application of hand gesture based on skin color and SVM[J]. Journal of Computational Methods in Sciences and Engineering, 2020, 20(1): 269-278.
- [4] 钟惠英.基于深度学习的动态手势识别方法[J].电子测量技术,2020,43(2):128-132.
ZHONG H Y. Dynamic gesture recognition method based on deep learning[J]. Electric Measurement Technology, 2020,43(2):128-132.
- [5] 牛雅睿,武一,孙昆,等.基于轻量级卷积神经网络的手势识别检测[J].电子测量技术,2022,45(4):91-98.
NIU Y R, WU Y, SUN K, et al. Gesture recognition detection based on lightweight convolutional neural networks[J]. Electric Measurement Technology, 2022,45(4):91-98.
- [6] 黄杰.基于深度学习的手语识别技术研究[D].合肥:中国科学技术大学,2018.
HUANG J. Research on sign language recognition technology based on deep learning[D]. Hefei: University of Science and Technology of China, 2018.
- [7] 晁人傑.基于 2D 卷积网络的动态手势分割与识别方法研究[D].西安:西安电子科技大学,2023.
CHAO R J. Research on dynamic gesture segmentation and recognition method based on 2D convolutional network[D]. Xi'an: Xidian University,2023.
- [8] SHARMA B, PANDA J. Spatiotemporal features representation with dynamic mode decomposition for hand gesture recognition using deep neural networks[J]. Signal, Image and Video Processing, 2024, 18(4): 3745-3759.
- [9] WU X Y. A hand gesture recognition algorithm based on DC-CNN[J]. Multimedia Tools and Applications, 2020, 79(13): 9193-9205.
- [10] NOBLE F, XU M, ALAM F. Static hand gesture recognition using capacitive sensing and machine learning[J]. Sensors, 2023, 23(7): 3419.
- [11] 邢益良,雷华军.基于傅里叶描述子的手势识别方法[J].工业控制计算机,2024,37(5):77-79,82.
XING Y L, LEI H J. A gesture recognition method based on Fourier descriptors[J]. Industrial control computers, 2024,37(5):77-79,82.
- [12] 唐文权,徐武,文聪,等.复杂背景下基于肤色检测的动态手势分割与识别[J].科学技术与工程,2019,19(33):330-335.
TANG W Q, XU W, WEN C, et al. Dynamic gesture segmentation and recognition based on skin tone detection in complex backgrounds[J]. Science, Technology and Engineering,2019,19(33):330-335.
- [13] 赵鸿图,李豪,梁梦华.复杂背景下多特征结合的深度学习手势识别[J].电子测量技术,2023,46(23): 77-84.
ZHAO H T, LI H, LIANG M H. Deep learning gesture recognition based on the combination of multiple features in complex backgrounds[J]. Electronic Measurement Technology, 2023, 46(23): 77-84.
- [14] 王新,王赛.基于改进 YOLOv5 算法的交警手势识别[J].电子测量技术,2022,45(2):129-134.
WANG X, WANG S. Traffic police gesture recognition based on improved YOLOv5 algorithm[J]. Electronic Measurement Technology, 2022,45(2):129-134.
- [15] OMAR E Z, AL-TAHHAN F E. A novel hybrid model based on integrating RGB and YCrCb color spaces for demodulating the phase map of fibres using a color phase-shifting profilometry technique[J]. Optik, 2024, 306: 171792.
- [16] 李涛.基于 FPGA 改进的 Canny 算法边缘检测系统[J].仪表技术与传感器,2024(8):65-70,92.
LI T. Improved edge detection system of Canny algorithm based on FPGA[J]. Instrumentation Technology and Sensor, 2024(8):65-70,92.
- [17] 王嘉俊.水面目标图像的边缘检测方法研究[D].镇江:江苏科技大学,2020.
WANG J J. Research on edge detection method of

- surface target image [D]. Zhenjiang: Jiangsu University of Science and Technology, 2020.
- [18] 崔莹,赵磊,李恒,等. 基于小波去噪与改进 Canny 算法的带钢表面缺陷检测[J]. 现代电子技术, 2024, 47(4):148-152.
- CUI Y,ZHAO L,LI H, et al. Strip surface defect detection based on wavelet denoising and improved Canny algorithm [J]. Modern Electronics, 2024, 47(4):148-152.
- [19] 陈瑞东,秦会斌. 多特征融合与 Kalman 滤波的 CAMShift 跟踪算法[J]. 计算机仿真, 2024, 41(3): 200-205,236.
- CHEN R D,QIN H B. CAMShift tracking algorithm based on multi-feature fusion and Kalman filtering[J]. Computer Simulation,2024,41(3):200-205,236.
- [20] 徐玥,周辉. 简单背景下基于 OpenCV 的静态手势识别[J]. 计算机科学,2022,49(S2):393-398.
- XU Y,ZHOU H. Static gesture recognition based on OpenCV in simple background[J]. Computer Science, 2022,49(S2):393-398.

作者简介

曹国华(通信作者),教授,博士生导师,主要研究方向为智能检测技术与装备。

E-mail:caogh@cust.edu.cn

刘福迪,硕士,主要研究方向为智能检测技术与装备。

E-mail:2366636354@qq.com

刘丽,硕士,实验师,主要研究方向为智能检测技术与装备。

马国庆,副教授,博士生导师,主要研究方向为智能检测技术与装备。