

面向高性能嵌入式人工智能计算平台的人机交互手势控制识别

Dongwei Fu*

重庆财经学院 重庆

【摘要】 本文主要介绍一种基于嵌入式 AI 计算平台的手势控制识别方法，利用红外摄像头和 AI 处理单元对手势进行识别和跟踪。使用红外摄像头作为检测工具可以实现低成本、易用等特点。该算法首先采用基于颜色模型的手势识别方法对输入的手势图像进行颜色转换并提取相应的肤色特征；然后通过对肤色特征进行加权，获得更丰富的颜色信息；最后通过计算颜色特征与肤色特征的相似度实现手势的识别。本文设计了一种基于红外摄像头和 AI 处理单元的手势控制识别系统。本文采用基于深度学习的目标检测算法，利用卷积神经网络（CNN）对目标进行检测，在不增加硬件成本的情况下实现手势识别与跟踪。在训练过程中，利用 CNN 提取手势图像的颜色特征和颜色信息进行训练；在测试过程中，使用相同数量和尺寸的红外摄像头对训练好的模型进行测试。此外，为了实现低成本、易用的特点，本文采用了基于深度学习的目标检测算法来验证手势识别的结果。

【关键词】 高性能嵌入式 AI 计算平台；人机交互手势；控制识别

【收稿日期】 2025 年 6 月 1 日

【出刊日期】 2025 年 7 月 8 日

【DOI】 10.12208/j.hcivr.20250003

Human-computer interaction gesture control recognition for high-performance embedded ai computing platform

Dongwei Fu*

Chongqing Finance and Economic College, Chongqing

【Abstract】 This article mainly introduces a gesture control recognition method based on embedded AI computing platform, which uses infrared cameras and AI processing units to recognize and track gestures. Using infrared cameras as detection tools can achieve low-cost, easy-to-use and other characteristics. In this algorithm, firstly, the Gesture recognition method based on color model is used to convert the color of the input gesture image and extract the corresponding skin color features; Then, by weighting the skin color features, richer color information is obtained; Finally, the recognition of gestures was achieved by calculating the similarity between color features and skin tone features. This article designs a gesture control recognition system based on infrared cameras and AI processing units. This article adopts a deep learning based object detection algorithm. The algorithm uses Convolutional Neural Network (CNN) to detect the target, and realizes Gesture recognition and tracking without increasing the hardware cost. In the training process, CNN is used to extract the color features and color information of gesture images for training, while in the testing process, infrared cameras of the same number and size are used to test the trained model. In addition, in order to achieve the characteristics of low cost and ease of use, this paper uses the target detection algorithm based on deep learning to verify the results of Gesture recognition.

【Keywords】 High performance embedded AI computing platform; Human machine interaction gestures; Control identification

1 简介

在科技飞速发展的今天，人们对人机交互的要求

*通讯作者：Dongwei Fu

注：本文于 2023 年发表在 Advances in Computer and Communications 期刊 4 卷 3 期，为其授权翻译版本。

也日益提高。传统的人机交互方式已无法满足用户的需求,人们需要一种更加自然、便捷、精准的人机交互方式。手势控制作为一种自然友好的人机交互方式,不仅降低了使用门槛,更提升了用户体验。在日常生活中,手势控制已经成为一种常见的人机交互方式。

手势控制是指利用手动手段进行控制。手势控制包含两个方面:一是控制对象,即物理对象;二是控制手段,是指人与机器之间的沟通。它又包括两类:一是基于视觉的手势识别技术;二是基于听觉的手势识别技术^[1]。

在手势识别过程中,需要对图像进行处理,并根据处理结果提取相应的特征进行识别。在此基础上,通过分析数据之间的相关性,实现手势的分类与跟踪。目前常见的手势识别平台包括基于视觉和基于听觉的两种。

1.1 手势控制

近几十年来,随着计算机软硬件技术的发展,越来越多的计算机平台采用了手势控制技术。这主要是因为手势控制技术简单、易操作、安全等优点,越来越受到人们的重视。基于视觉的手势控制平台基于计算机视觉,主要通过分析手势图像来实现手势的识别和跟踪。这种方法在手势识别方面的效率远高于基于听觉的方法。在视觉平台中,常用的算法包括灰度共生矩阵(GLCM)、边缘信息和颜色特征。其中,灰度共生矩阵是一种非线性变换,经过图像处理后可以获得大量的特征信息,从而达到良好的分类效果。在深度信息方面,虽然也有一些算法可以实现手势识别,但效率较低。相比之下,颜色特征更适合彩色图像。因此,本文采用将颜色空间转换为灰度空间的方法来实现手势识别^[2]。

手势控制平台目前主要在 Android 平台上进行设计和开发。目前手势控制平台主要存在两个问题:一是由于 Android 平台自身的限制,导致很多算法无法实现;二是 Android 平台缺乏相关的硬件资源来支撑手势控制平台的运行。

1.2 手势识别

深度学习作为新兴的人工智能技术,其发展速度远超人们的预期。基于深度学习的手势识别平台能够有效克服传统算法的不足,在准确率、实时性和鲁棒性方面均有更佳表现。然而,深度学习在处理大量数据时,计算复杂度高、训练时间长等问题难以满足实际需求。同时,随着手势识别算法的不断发展,其在复杂环境下的性能也得到了大幅提升,例如利用多层感知

器和循环神经网络提高分类准确率,利用卷积神经网络提高卷积计算效率等。

针对现有手势识别平台的不足,本文提出了一种基于深度学习技术与嵌入式 AI 计算平台相结合的手势识别方案。该方案利用深度学习算法对图像进行处理,将图像划分为若干子区域,提取关键信息,并将其作为训练数据集,实现手势分类与跟踪^[3]。

1.3 本文的工作

本文对人机交互手势识别进行了深入研究,将深度学习技术应用于手势识别,利用卷积神经网络进行手势识别,并采用改进的动态时间规整算法优化识别结果。本文的主要工作如下:(1)搭建了高性能嵌入式 AI 计算平台,在该平台上对视频图像进行预处理,并通过深度学习算法实现手势识别。(2)构建了用于手势识别的卷积神经网络模型,并优化了网络参数。(3)设计并实现了一种改进的动态时间规整算法,优化了手势识别的结果。(4)结合图像预处理、特征提取和分类三个阶段,设计了一种适用于手势识别的动态时间规整算法,并将其应用于手势识别平台。(5)利用 Python 语言对卷积神经网络模型进行优化,提高了模型在嵌入式平台上的运行效率。(6)使用卷积神经网络对手势识别平台进行测试,并对测试结果进行分析。测试结果表明,本文设计的手势控制识别平台在嵌入式平台上运行效率高,能够实现手势的快速识别,且具有较高的鲁棒性。

2 手势识别算法

手势识别算法主要包括检测、分割、识别和跟踪四个部分。其中,检测是指在手势图像中找到感兴趣的目标;分割是指将感兴趣的目标区域从背景中分离出来;识别是指将分割出的目标区域与数据库中对应的模板进行比较,判断目标是否属于待识别的手势;跟踪是指在空间中跟踪手势,目前有很多跟踪算法,例如 Kinect 和卡尔曼滤波器。Kinect 作为红外摄像头,分辨率高、功耗低,无需外界光源,可以避免光源对手势识别的干扰。它可以提取手势在空间中的运动轨迹,实现手势的检测和跟踪。卡尔曼滤波器在运动目标跟踪方面取得了良好的效果,但其无法自动获取运动目标的位置信息。针对这一问题,本文采用滑动窗口估计位置信息来解决该问题。滑动窗口大小是卡尔曼滤波器设计中最重要参数之一,它直接影响整个算法的性能^[4]。

2.1 手势图像采集

本文利用 Kinect 实现手势检测与跟踪。Kinect 使用三个红外摄像头采集手势图像，分别为 YCbCr 颜色空间、RGB 颜色空间和 HSV 颜色空间。YCbCr 颜色空间是通过 YCbCr 色度键从 RGB 颜色空间映射而来的。RGB 是最常见的颜色空间类型。HSV 颜色空间是通过扩展 RGB 颜色空间而得到的一种颜色分类方法，直观易懂，广泛应用于工业检测、医疗诊断等方面。HSV 颜色空间将图像分为 HSV 颜色空间中的三个分量，分别为 L、H、V，通过比较这三个分量的差异来对图像进行分类。在 HSV 颜色模型中，L 和 H 分别代表红色和蓝色，V 代表白色。在 RGB 颜色空间中，H 和 V 分别代表红色和蓝色，它们在 RGB 颜色空间中都是中性色，在图像处理中没有区别。HSV 颜色模型对彩色图像具有良好的适应性，同时也具有一定的灵活性。

2.2 手势分割

手势分割主要包括肤色检测和手势定位两部分。在肤色检测中，可以采用 RGB 颜色空间作为颜色空间，该颜色空间包含了人眼所能感知的所有内容，能够较好地表达人眼对颜色的感知。由于受光照条件等因素的影响，肤色通常会存在一定的偏差，因此本文采用 YCbCr 颜色空间进行肤色检测。在手势定位部分，本文采用自适应阈值对手势图像进行二值化，然后基于二值化结果进行形态学运算得到手势区域。手势定位是整个手势识别过程中最重要的步骤之一，它需要将手势图像与背景图像进行比较，提取出感兴趣的区域。

3 手势追踪与手势控制平台架构

在这个平台中，首先需要通过红外摄像头获取手势图像，然后利用红外摄像头将手势图像传输到云端，通过云端的边缘计算单元对这些图像进行处理，最终得到手势的位置、方向、速度等信息。同时，平台还需要对这些手势信息进行跟踪，并根据跟踪结果实现对手势的识别和控制。

本平台中，摄像头采用 1 个像素分辨率为 640×480 的红外摄像头和 1 个像素分辨率为 640×480 的彩色摄像头。利用红外摄像头捕捉手部图像。为了实现手势跟踪和控制功能，平台首先需要捕捉手势图像，并通过边缘计算单元将手势图像传输到云端。然后，在云端通过边缘计算单元对手势图像进行处理。如果手势图像中存在手势遮挡或其他干扰因素，则需要通过边缘计算单元去除这些干扰因素。在云端处理后，平台可以使用预先定义的算法对手势进行识别和控制。

要实现手部运动轨迹的跟踪，首先需要利用 OpenCV 库中的 Kinect (Kinectv2) 模块实现手部运动轨迹的跟踪。Kinect 可以从环境中检测并获取手势和运动信息，并利用这些信息进行手部识别和跟踪。

由于该平台仅需实现手势识别与追踪功能，因此并未设计复杂的神经网络处理单元，仅采用一个卷积神经网络 (CNN) 完成手势检测与追踪。CNN 是一个并行处理网络，可以提取图像上物体的特征信息，然后利用这些特征信息训练模型，并将该模型应用于新的图像进行识别和追踪。

预处理之后，平台需要使用预先定义的算法对采集到的手势图像进行处理。本平台采用 KNN (核神经网络) 算法进行手势识别。KNN 算法是一种基于统计学习理论的分类算法。在识别手势时，首先将手势图像划分成若干个区域，然后利用这些区域内已有的手势特征 (如方向、位置、大小等) 训练一个分类器。当出现新的手势时，分类器将其与已经训练好的分类器进行比较，如果两个分类器之间存在显著偏差，则表明该新手势为新手势。将新出现的手势与已有分类器中已经训练好的分类器进行比较，判断是否为新手势。如果是，则表明该手势为新出现的手势；否则，表明该手势与已有分类器中已经训练好的分类器不一致。如果两个分类器都被识别为新手势，则将其从数据库中删除，并重新训练分类器；如果两个分类器都识别出旧的手势，它们将被保留以进行识别和跟踪。

4 实验结果与分析

本文在实验中采用了两种测试方法，一种是检测并跟踪手势，另一种是识别手势。检测使用 OpenCV，跟踪使用 Kinect2.0。OpenCV 是一个开源计算机视觉库，主要用于计算机视觉领域的图像处理、物体检测与识别等应用。

4.1 手势检测与跟踪

在手势检测与跟踪方面，本文主要采用基于背景差分法和帧差分法的手势检测与跟踪算法。为了验证本文提出的手势检测与跟踪算法的性能，首先采用帧差分法检测手势，然后采用光流法跟踪手势。将结果与其他三种检测方法的结果进行比较，本文提出的算法在手势检测和跟踪方面取得了优于其他三种检测方法的效果。在检测方面，本文采用了背景差分法和帧差分法两种基本算法；在跟踪方面，本文采用了基于 Kinect 2.0 传感器的目标跟踪方法。最后，通过实验验证了本

文提出算法的有效性。实验中, 首先将摄像头固定在某一位置并设置摄像头参数。然后将手势图像导入平台并进行图像预处理工作。在此过程中, 首先使用 Kinect 2.0 传感器采集图像并将其传输到云端。接下来, 我们使用 OpenCV 对采集到的图像进行处理和分析。最终获得手势图像。实验结果表明, 该算法是有效的^[5]。

4.2 手势识别

由于 Kinect 2.0 是一款优秀的触觉摄像头, 可以直接通过 Kinect 2.0 获取手势信息。获取手势信息后, 可以利用 OpenCV 进行手势识别。实验结果表明, 所提算法能够成功实现手势识别。当手势运动幅度较小时, 特征提取效果较为明显; 当手势运动幅度较大时, 特征提取效果相对较差。

4.3 实验环境

本文设计了一种高性能、低功耗、低成本的人机交互手势控制平台。该平台主要由两部分组成: 硬件部分主要包含 ARM Cortex-A8 处理器、三星 32 位 ARM7TDMI-SOC、OpenCV 三大核心组件。软件部分主要包含 Socket API、OpenCL、Kinect 三大核心库。该平台在满足功耗和性能要求的同时, 充分利用了 ARM Cortex-A8 处理器的性能, 以及三星 32 位 ARM7TDMI-SOC 的高性能、低功耗、低成本优势, 大大提升了人机交互手势控制平台的性能。该平台可以实现手势数据的实时采集和处理, 并将处理后的手势数据发送到 PC 机进行后续处理。

本文利用 OpenCV 对手势进行检测和跟踪, 提出了一种基于改进粒子滤波 (PSOF) 算法的手势识别算法。实验结果表明, 该方法能够有效地检测出手部区域并跟踪多个手部区域。此外, 本文还将基于深度卷积神经网络 (CNN) 的手势识别算法应用于手势识别。实验结果表明, 该方法能够有效地识别手势。

实验中我们采用高性能人工智能计算平台 (HPC) 作为硬件基础, 设计并实现了基于手势的人机交互控制平台, 该平台将深度学习技术引入到手势识别领域, 使得手势识别具有更高的准确率和鲁棒性。

5 结论

本文提出了一种基于红外摄像头的手势识别方法,

在不增加硬件成本的情况下实现了手势的识别与跟踪。通过在红外摄像头上安装 AI 处理单元, 实现了手势的识别与跟踪。该方法可应用于各种应用场景, 具有成本低廉、易于使用的特点。在嵌入式平台上, 可以利用 CPU、GPU、FPGA 等平台实现 AI 处理单元的硬件加速。该方法不依赖于复杂的硬件平台或特定的软件算法, 易于使用和扩展。

手势识别是人工智能和计算机视觉技术的重要应用领域, 也是人机交互中应用最广泛的方式之一。然而随着对手势识别精度和可靠性要求的提高, 传统方法存在的一些不足。本文提出一种基于深度学习的手势识别方法, 可以有效解决传统方法存在的问题。随着人工智能技术和计算机视觉技术的不断发展, 手势识别技术也必将取得更大的进步。

参考文献

- [1] Yang Tao, Yang Boxiong, Yin Ping, et al. Research on human-machine interaction gesture control recognition based on high-performance embedded AI computing platform [J]. Information Record Materials, 2019, 20 (11): 175-177.
- [2] Yang Yiping, Min Xiao. Human computer interaction technology of Gesture recognition based on computer vision [J]. Electronic Technology and Software Engineering, 2018 (12): 138-139.
- [3] Zhang Yu, Wang Shidong, Ma Yishan, Huang Ai. Design of Gesture recognition based on embedded system. Computer and Network, 2021, 47 (17): 44-45.
- [4] Wang Fei. Design and implementation of Gesture recognition based on embedded system. Radio Communication Technology, 2020, 46 (06): 717-721.
- [5] Liu Jinkui. Research on visual Gesture recognition method based on embedded system. Journal of Jiaozuo University, 2020, 34 (02) 94-96.

版权声明: ©2025 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。



OPEN ACCESS