

Non-Contact Virtual Mouse Based on Hand Gesture

Yanli Li, Guanqun Wang, Yucheng Hong, Hongmei Wei, Chenhao Gao, Dan Yue*

School of Science, Changchun University of Science and Technology, Changchun Jilin
Email: 2571062645@qq.com, *danzik3@126.com

Received: May 1st, 2018; accepted: May 18th, 2018; published: May 25th, 2018

Abstract

Gesture Recognition, which is efficient, direct and natural, has become one of the key techniques of human-compute interaction. Meanwhile, the high degree of freedom of hand results in the complexity of gesture, diversity of variation and rapidity of movement, which makes gesture recognition full of challenges. The author uses the HSV and YCgCr color space to recognize different gestures, and depending on the OPENCV2.3.1 visual library and Microsoft Visual Studio 2010, the detection and feature extraction based on C++ language was realized. The main works are listed as following: Acquisition and preprocessing of the hand gesture image, segmentation and feature extraction of the image, and hand gesture modeling, etc. Finally based on the research on the technologies of the vision-based hand gesture recognition, the author gives an optimal solution of virtual mouse and keyboard system.

Keywords

Skin Detection, Hand Gesture Recognition, Hand Gesture Segmentation, OPENCV Visual Library, Virtual Mouse and Keyboard

基于人手的非接触式 虚拟鼠标

李彦黎, 王冠群, 洪雨城, 魏红梅, 高晨皓, 岳丹*

长春理工大学理学院, 吉林 长春
Email: 2571062645@qq.com, *danzik3@126.com

收稿日期: 2018年5月1日; 录用日期: 2018年5月18日; 发布日期: 2018年5月25日

*通讯作者。

文章引用: 李彦黎, 王冠群, 洪雨城, 高晨皓, 魏红梅, 岳丹. 基于人手的非接触式虚拟鼠标[J]. 人工智能与机器人研究, 2018, 7(2): 68-77. DOI: 10.12677/airr.2018.72008

摘要

手势识别技术,是一种高效、直接、自然的人机交互方式,属于计算机视觉领域中较为活跃的研究主题。同时,手的高自由度导致手势复杂、变化多样、移动高速等特性,使手势识别技术成为极具挑战的课题之一。本文在HSV和YCgCr颜色空间对不同的手势进行识别,借助OPENCV2.3.1视觉库和Microsoft Visual Studio 2010实现了基于C++语言的检测和特征提取。本文对基于计算机视觉的自然手势识别的相关技术:手势图像采集和预处理、手势图像的分割和特征提取,以及手势建模等,并且提出了适合虚拟鼠标、键盘系统的解决方案。

关键词

肤色检测, 手势识别, 手势分割, OPENCV库, 虚拟鼠标, 键盘

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目前键盘和鼠标是我们和计算机交互的接口方式,我们希望有更简洁的、智能化和人性化的人机交互方式。计算机无接触式地收集人类在计算机前的视频信号,利用计算机视觉的相关理论分析视频信号,做到分辨人类的动作,明白人类的意图,即我们希望计算机能尽可能地“理解”我们。

手势是一种自然、直观、易于学习的人机交互手段。手势研究分为手势合成和手势识别,手势识别技术一直是人机交互的一个重要研究课题。以人手直接作为输入手段与其它输入方法相比较,具有自然性、简洁性和丰富性、直接性的特点。这说明用计算机来识别手势为人们提供了一个更自然的人机接口。但是由于其难度较大,目前的研究结果实用化程度很低。

经调查,国内多数采用类似数据手套或通过传感节点和汇聚控制节点构成的系统并置入速度计,陀螺仪和磁力计等惯性传感器设备等,进行对手部移动轨迹的跟踪和模仿,这是一套一体的动作捕捉系统,要求多个元件共同工作。少部分采用激光定位技术,其基本原理就是在空间内安装数个可发射激光的装置通过计算角度差,得到三维坐标完成动作的捕捉。但是此类产品因为内置多种传感器和器件导致系统繁杂昂贵,不适合大众使用。而较为成熟的技术多为国外引进,此类产品中大部分配件,系统都是集成形式,对国内不开放,具有技术壁垒,导致了国内的技术对国外的技术依赖性很大。这使本文论述的方法在国内的交互设计中具有相当的可行性。

本文着重研究基于单目视觉的手势识别,从手势分割及手势表示两方面着手,提出了一种单目视觉下的手势识别方法,并且结合鼠标、键盘操作,实现简单的鼠标、键盘事件。以此为基础,在之后的研究中仍然可以深化改良,使得本文将要论述的方式可以在人机交互中给出更加精确且可信任的数据回馈,使人们在与计算机的动作跟踪模拟的交互中取得直观有效的良好体验。

2. 研究方案概述

通过设计,开发出一款人机手势交互系统,该系统功能为通过简单手势实现对计算机的控制,譬如鼠标的一些基本功能——移动,点击等,以及PPT、图片翻页等功能,具体流程如下(图1)。

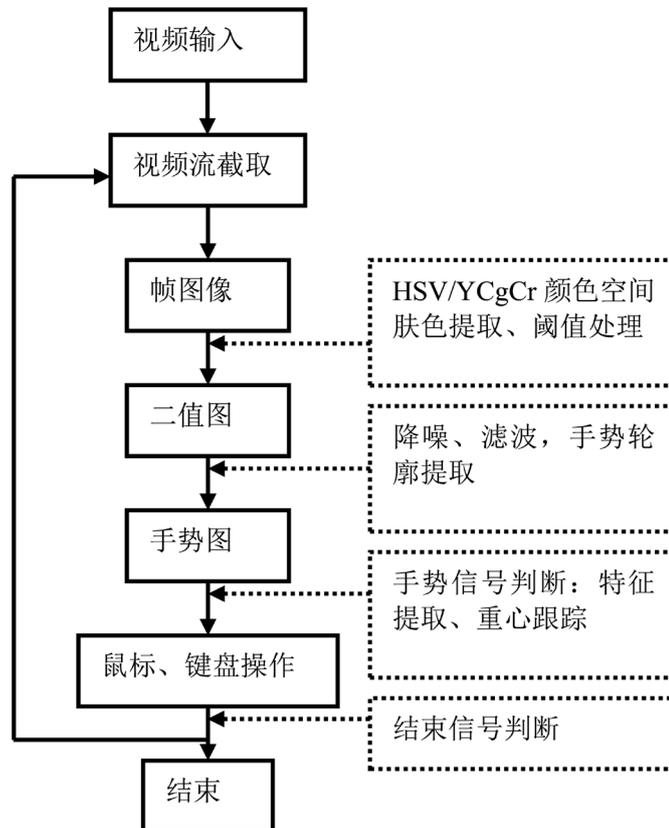


Figure 1. Flow chart
图 1. 流程图

结合 OPENCV 视频处理库实现视频流的输入, 通过捕捉视频流中的人手的动作来模拟鼠标的运动。对视频流中的信息进行提取时, 由于生活中常用的是 RGB 色彩空间的三个颜色分量都包含亮度信息, 存在相关性, 常常会影响手势检测的效果。所以我们考虑到使用 YCgCr 和 YCbCr 颜色空间进行肤色检测, 而肤色在 YCgCr 色彩空间比在 YCbCr 色彩空间具有更好的聚集性和辨识性, 故使用 YCgCr 色彩空间[1]。根据流程图, 可以看出我们的研究方案最终决定在 YCgCr、HSV 颜色空间进行肤色检测, 再利用 MeanShift 算法进行重心跟踪、特征提取, 为接下来的鼠标、键盘动作提供手势判断信号的判断条件, 最后结合 OPENCV 库在 VS2010 上来实现虚拟简单鼠标动作与键盘动作的目的, 各自通过画图工具、操作 PPT 上一页、下一页、全屏, 退出得以验证。利用人手的轮廓与凸缺陷检测指尖, 再利用不同手势的面积特点表示手势[2], 根据已定义手势的先验知识可以快速准确地进行识别输出。因此本文所用到的面积检测方法简单而又高效。

3. 研究方案分析

在本设计中, 我们采用基于计算机视觉的手势输入技术。故此, 我们将通过摄像头获取视频数据流。接着, 系统对每一帧图像检测到交互手势, 对满足条件的关键帧进行提取, 将该手势从视频信号中分割出来, 然后, 选择合适的手势特征进行手势分析, 分析过程需要提取模型特征参数。最后, 根据模型参数对手势进行分类并依据具体需要生成手势描述, 用以驱动相关应用。下面针对各个功能模块逐一加以分析。The design of Human-based contactless virtual mouse 将从三个方面进行描述: 手势分割、手势识别、鼠标动作。

3.1. 手势分割

手势识别最重要，也是最基础的一步就是手势分割了，手势分割是指将有意义的手势区域从包含手势的图像中提取出来的过程，手势分割的好坏直接影响后续的手势识别效果。本设计中利用 YCgCr、HSV 颜色空间转换进行肤色提取、辅以降噪、滤波等过程，实现手势图二值图的获得。

3.1.1. HSV、YCgCr 颜色空间

目前有许多颜色空间，应用于皮肤检测的颜色空间也有许多，如归一化 RGB 颜色空间，YIQ, CIE-Lab 和 CIE-Luv 等颜色空间。在 RGB 颜色空间中，三个颜色分量都包含亮度信息，存在相关性，因此用于肤色检测是算法的亮度适应性不好。而通过调查发下，应用广泛的是 HSV、YCgCr 颜色空间[3]。

HSV 色彩空间是一个圆锥形的模型，其中色相(H)是色彩的基本属性，取 0~360 度得数值。饱和度(S)是色彩的纯度，越高色彩越纯，低则变灰。取值为 0~100%。明度(V)也叫亮度，取值 0~100。相较于其他颜色空间而言，HSV 空间中肤色的聚类特征更为明显，通过调查、实验我们的结论是 H 取值范围为 7~29 时，定义为肤色范围。

虽然 HSV 空间有较高的肤色聚类特征，但对于与肤色相似的背景颜色，HSV 空间中还是不能很好地进行分辨，所以需要进一步的进行肤色提取，本设计采用双重颜色空间的肤色提取，主要是为了提高肤色确认的精确度，而应用到的第二个空间即 YCgCr 颜色空间。

YCgCr 颜色空间，是在 YCbCr 空间的基础上，由 De Dios 提出的。首先肤色在 Cg-Cr 颜色空间进行投影，对像素点区域进行拟合，初步去除图像中的非肤色，得到肤色第一次分割的图像，实验证明这种拟合可以起到好的去除非肤色像素的作用，该算法提高了手势检测的速度，并且有较高的检测率和较低的漏检率。其中，Y 仍分别代表颜色信息中的亮度，Cr 仍代表红色分量 R 与亮度 Y 的差，Cg 则相应代表绿色分量 G 与亮度 Y 的差值。RGB 到 YCgCr 的转换如下：

$$\begin{bmatrix} Y \\ Cg \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -81.085 & 112.000 & -30.915 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

调查、实验得到的肤色在 YCgCr 空间的范围满足

$$\begin{cases} Y \in [30 & 235] \\ Cg \in [90 & 127] \\ Cr \in [129 & 173] \end{cases}$$

实验得到应用 HSV 与 YCgCr 双重颜色空间的肤色提取，可以较好的剔除背景非肤色区域，三者二值图比较结果如下(图 2)。

根据对比图可以看出双重空间肤色检测相对于独立的 HSV、YCgCr 空间检测，可以更好地去除环境背景干扰，尤其是对类肤色背景的去留有比较好的效果，所以本设计选用双层颜色空间检测[4]，为手势识别提供更完善的输入。

3.1.2. HSV、YCgCr 颜色空间

图像处理是手势识别中的一个重要环节[5]，好的图像处理过程可以为识别过程提供较好的识别环境，减少干扰、错误。本设计的图像处理过程包括：腐蚀、膨胀，滤波、多边形逼近等。其中滤波包括中值滤波、去除小面积、多边形逼近[6]。去除小面积的过程采用遍历的方法查询最大面积并保留，其余删除，此操作目的是去除面积相对人手面积小的人脸背景噪声(原于人脸相对与摄像头较远，面积可以小于人手面

积), 具体实现过程展示如下(图 3~图 8)。

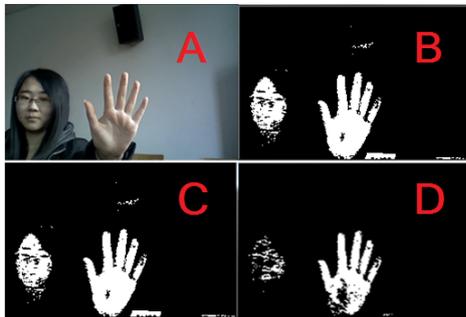


Figure 2. (A) Original; (B) HSV test results; (C) YCgCr test results; (D) Double space test results
图 2. (A) 原始图片; (B) HSV 检测结果; (C) YCgCr 检测结果; (D) 双重空间检测结果



Figure 3. Double skin extraction results
图 3. 双重肤色提取结果



Figure 4. Corrosion results
图 4. 腐蚀结果



Figure 5. Expansion result
图 5. 膨胀结果



Figure 6. Median filter results
图 6. 中值滤波结果

从以上图示，可以清楚的感受到本设计图像处理过程，基本可以从视频流中的到比较完整的手势图像。

3.2. 手势分割

手势识别的方法有很多种，与所用的手势表示(建模)息息相关[7]。本文用手势二值图像的重心表示手掌，应用凸缺陷检测指尖，通过不同手势凸包轮廓面积与真实轮廓面积比值 Q 的不同区分不同的手势。如图所示，通过不断测试，可以得到拳头的 Q 值为 $1\sim 1.2$ ，五指伸开时的 Q 值则应大于 1.2 ，单指、双指等 Q 值可以通过测试进一步细分(图 9~图 15)。



Figure 7. Remove small area results

图 7. 去除小面积结果



Figure 8. Polygonal approximation results

图 8. 多边形逼近结果

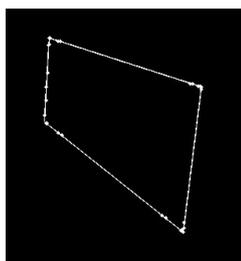


Figure 9. Fist Q value less than 1.2 (left)

图 9. 拳头 Q 值小于 1.2 (左)

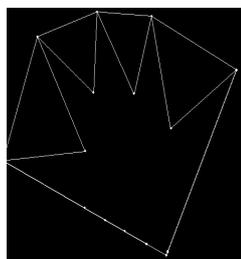


Figure 10. Five fingers spread Q value greater than 1.2 (right)

图 10. 五指伸开 Q 值大于 1.2 (右)

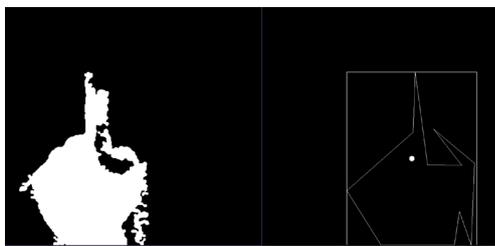


Figure 11. An extended Q value of about 1.05 (left)
图 11. 一指伸开 Q 值约 1.05 (左)



Figure 12. Two refers to the extended q value of about 1.09 (left)
图 12. 二指伸开 Q 值约 1.09 (左)

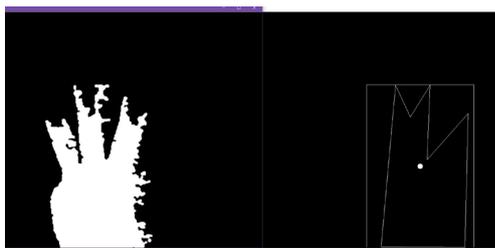


Figure 13. Three refers to the extended q value of about 1.13 (left)
图 13. 三指伸开 Q 值约 1.13 (左)

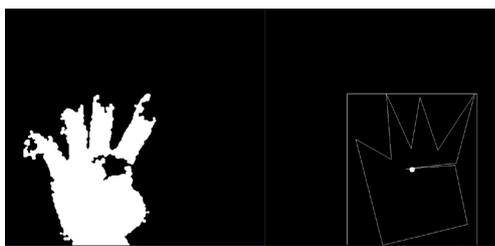


Figure 14. Four fingers extended q value of about 1.18 (left)
图 14. 四指伸开 Q 值约 1.18 (左)

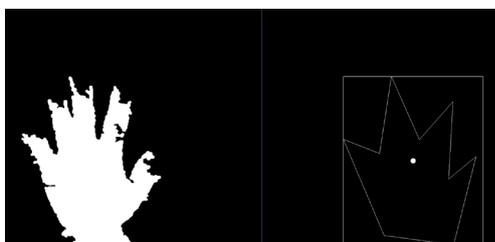


Figure 15. Five fingers extended q value of 1.25 (left)
图 15. 五指伸开 Q 值 1.25 (左)

同时, 我们应用 MeanShift 算法[8], 对重心进行跟踪, 为下一步鼠标动作提供可能。其算法思想可以简单理解为: 利用概率密度的梯度爬升来寻找局部最优。它要做的就是输入一个在图像的范围, 然后一直迭代(朝着重心迭代)直到满足要求为止[9]。

通过以上算法, 我们可以为接下来的鼠标、键盘动作提供了手势判断信号的判断条件, 为下一步做足准备。

3.3. 鼠标、键盘动作

鼠标、键盘动作是本设计的最后环节。本设计编辑两程序分别实现鼠标动作与键盘动作。鼠标程序, 实现了左键单击的鼠标动作、鼠标跟踪过程模拟鼠标的功能, 通过画图工具得以验证。键盘程序, 实现了方向上键、方向下键, F5 键、Esc 键的手势操作, 通过操作 PPT 上一页、下一页、全屏, 退出得以验证。

虚拟鼠标程序, 可以通过与 mouse event (MOUSEEVENTF_LEFTUP/DOWN, 0, 0, 0, 0)与 mouse event (MOUSEEVENTF_MOVE, x, y, 0, 0)的连用来实现, x、y 分别代表相对位移, 可以通过手在图像上的离上一位置的距离求得, 产生鼠标移动的增量来实现。但由于摄像头的分辨率往往小于屏幕的分辨率, 造成精确度不高。而为了交互的自然性, 鼠标是否按下用手势来决定, 五指伸张意味着左键未按下, 拳头意味着左键按下。同时虚拟鼠标程序, 在通过使用 mouse_event ()的基础上还可以得出对手势的跟踪和模仿实现相当的表征结果, 具体表征结果如图 16 和图 17 显示:

虚拟键盘程序, 可以通过与 keybd_event ()与跟踪实现, 具体对应如下(图 18)。

在大量的实验数据中, 选取部分数据进行表征本文结果的实现阶段, 下列图片数据具有典型的特征结果表现响应时间, 得出肤色检测在手势跟踪模仿中的反响时间可以较好的在一定时间内呈现稳定的特征结果的结论。可以明确的看出通过肤色检测技术可以达到一定程度上的对鼠标的模拟实现, 并且数据反馈表征结果良好(图 19)。

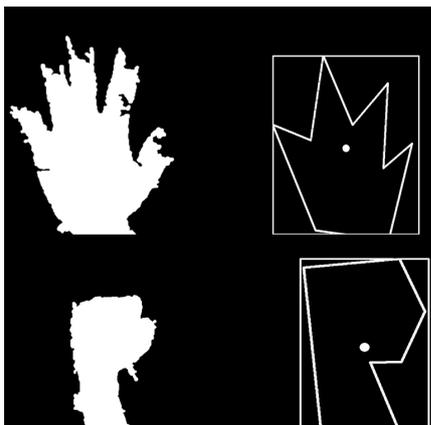


Figure 16. Five fingers open to control movement and clench fist to control left key stroke
图 16. 五指张开控制移动, 握拳控制左键击



Figure 17. Illustrates the drawing by mouse simulation
图 17. 通过鼠标模拟实现绘图

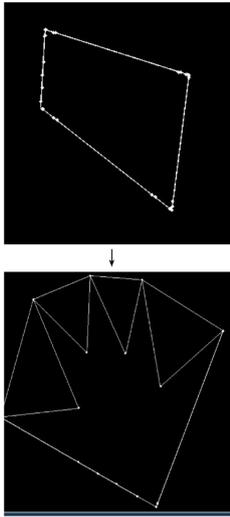
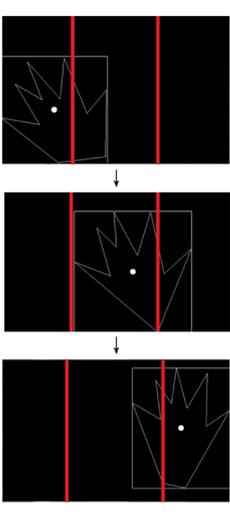
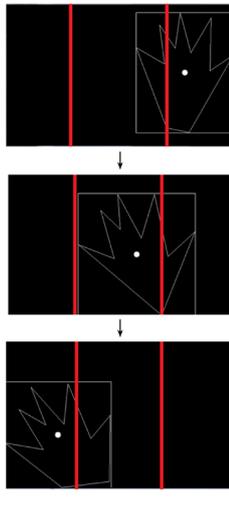
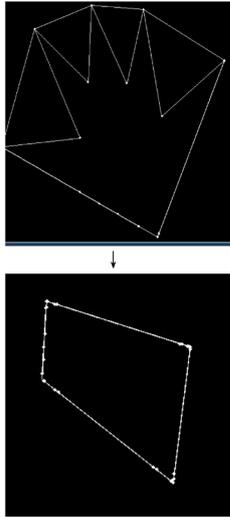
动作	拳头→五指伸张	五指伸张: 左移动至右, 过中间判定区域	五指伸张: 右移动至左, 过中间判定区域	五指伸张→拳头
功能	F5(ppt 全屏)	方向下键(下一页)	方向上键(上一页)	Esc(退出全屏)
动作图片展示				

Figure 18. Virtual keyboard program
图 18. 虚拟键盘程序

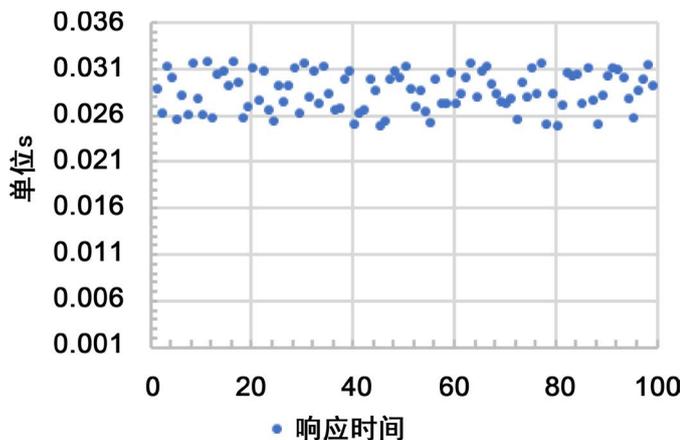


Figure 19. Response time
图 19. 响应时间

4. 研究结果

本设计方案的实验环境为: Intel & Pentium CPU, 2GB 内存, Windows7 操作系统, VS2010 + OPENCV2.3.1, 所用视频分辨率 640 × 480。

本设计方案在 HSV 和 YCgCr 颜色空间进行肤色检测, 借助 vs2010 软件, 利用 MeanShift 算法对重心进行跟踪, 为接下来的鼠标、键盘动作提供手势判断信号的判断条件, 进而对不同的手势进行识别, 再结合 opencv2.3.1 视觉库, 实现了对人手的识别过程。实验具体验证结果显示为: 1) 虚拟鼠标画图, 通过画图工具得以验证, 即实现鼠标跟踪与单击功能; 2) 实现虚拟键盘功能, 通过操作 PPT 上一页、下一页、全屏, 退出得以验证, 完成了 PowerPoint 程序的手势控制, 初步实现了人机交互的预期目标。

5. 结论

本文在 HSV 与 YCgCr 颜色空间上采用固定阈值方法分割人手, 利用轮廓及凸缺陷检测指尖, 根据手势面积特征识别手势。利用凸缺陷检测指尖的方法简单快捷。实验结果表明, 本文的手势识别方法对手势分割中间结果的精度要求不高, 因此能适应环境的变化。通过大量具有重复性且精度良好的结果印证, 本文的肤色检测技术显然可以初步实现在对手势进行跟踪和模仿的结果上拥有更优良的体验, 可以得出本文方法具有一定程度上的实现的可行性。这表明肤色检测技术在基于 HSV 和 YCgCr 空间检测下对单目视觉下的单手手势识别有着显然性的识别性, 而且对平移、平面旋转, 具有相当良好的识别能力, 但还有不足之处, 如获取的手势种类样本不够多、研究时间有限限制了程序优化, 使时间运行方面仍然有限制不能做到 24 ms 内。这里面既有优化程序以及实现交互水平方面的原因——凸缺陷的检测要求手势与屏幕保持在一定程度上的水平, 以此获得更加良好的输入, 同时也有硬件设备的原因——单目视觉摄像机无法得到深度信息, 计算机内存有限等问题。因此, 我们下一步的工作是继续完善该方法, 适当地增加手势的种类样本以补充本实验的丰富性和完整性, 考虑加入多手势识别并准确的跟踪和模仿手势得出轨迹数据, 提高程序反应速度等。目前可以得出的结论是, 通过大量的数据可以得到通过本文所述的肤色检测在实现虚拟鼠标设计中具有显然的相关性和合理性, 肤色检测可以使数据跟踪模拟在其他相关的检测模式中具有良好的反响, 获得更加直观且具有可信任的数据反馈。这是在使用其他方式的交互模式中无法达到的, 可以说肤色检测在交互技术中也是有着相当的技术基础和应用前景的, 本文所做的实验也只是基础, 之后也还会有更多的开发上升空间。

研究肤色检测是一项既有趣又充满挑战的工作, 人与计算机的自由交互是计算机发展的最终趋势, 因此手势识别技术的发展前景广阔。随着人们对手势识别技术研究热情的不断高涨, 科学技术不断进步, “手势识别时代” 已经到来, 相信在不远的将来, 出现在科幻电影中的情景就会在现实生活中上演。若想要加深在肤色检测基础上的人机交互技术, 仍然需要在未来投入更多的时间去开展更为深入的研究。希望这一具有创新性的工作可以给业内或者本领域内更多的研究人员在本方向上开展的研究起到些许的作用。因此本文的研究工作有一定的现实意义, 有好的实用性。

参考文献

- [1] 张争珍, 石跃祥. YCgCr 与 YCgCb 颜色空间的肤色检测[J]. 计算机工程与应用, 2010, 46(34): 167-170.
- [2] 翁汉良, 战荫伟. 基于视觉的多特征手势识别[J]. 计算机工程与科学, 2010, 34(2): 123-127.
- [3] de Dios, J.J. and Garcia, N. (2003) Face Detection Based on a New Color Space YCgCr. *Proceedings of International Conference on Image Processing*, 2, 909-912. <https://doi.org/10.1109/ICIP.2003.1247393>
- [4] 胡友树. 手势识别技术综述[J]. 中国科技信息, 2005(2): 42.
- [5] Tsapatsoulis, N. and Kollias, A.Y. (2000) Efficient Face Detection for Multimedia Applications. *IEEE International Conference on Image Processing*, 2, 247-250. <https://doi.org/10.1109/ICIP.2000.899289>
- [6] 崔育礼. 手势识别若干关键技术研究与应用[D]: [硕士学位论文]. 上海: 同济大学, 2008.
- [7] 戴丹. 基于图像的静态手势识别及在服务机器人中的应用[D]: [硕士学位论文]. 浙江: 浙江大学, 2007.
- [8] 郭秀梅, 赵秀艳, 王玉亮, 等. 基于 YCgCb 新颜色空间的人脸检测算法的研究[J]. 电子测量技术, 2008, 31(4): 79-81.
- [9] 江超, 艾娇燕. 基于的摄像头动态手势轨迹识别及其应用[J]. 计算机应用, 2012, 32(S1): 128-133.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2326-3415，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：airr@hanspub.org