

基于手势识别的机器人控制系统^①

师 哲, 李 皓, 张明明

(武汉工程大学 邮电与信息工程学院, 武汉 430000)

摘 要: 研究了一种基于人体手势识别的机器人控制系统. 首先, 利用图像识别技术, 通过 YCrCb 皮肤颜色模型提取手掌并分析指尖和手心的相关信息; 其次, 利用帧差法对手掌运动趋势和简单的手势信息进行识别; 最后, 通过无线蓝牙串口将识别出来的手势信号发送给机器人, 以达到手势控制机器人的目的. 系统是在 VS2010 下利用 OpenCV 计算机视觉库进行编译完成的, 实现了通过简单的手势信息控制机器人的目的, 从而摆脱了人机交互时必须依靠物理接触的限制. 实验结果表明, 该系统可以实现对机器人前进、后退、左转、右转、停止、加速的实时控制, 对手势信息的识别率在 90% 以上. 对进一步探索机器学习、自主识别等相关领域有着较高的参考价值.
关键词: 手势识别; 机器人; 人机互动; 手势控制系统; 模式识别

Robot Control System Based on Gesture Recognition

SHI Zhe, LI Hao, ZHANG Ming-Ming

(College of Post and Telecommunication of WIT, Wuhan 430000, China)

Abstract: A robot control system based on the hand gesture recognition is researched in this paper. First of all, this system uses image recognition technology and YCrCb skin color model to find the palm. Then it analyzes the fingertip and palm center. Furthermore, using frame difference method to identify the movement tendency of the hand and some simple gesture information. Finally, the signals recognized by the system will be transmitted to the robot. This is accomplished through the wireless Bluetooth serial port, in order to control the robot by the hand gesture. This system, which is compiled by OpenCV computer vision library under VS2010, can control the robot through some simple gesture and get rid of the physical contact during the Human-Machine interaction. The experimental result shows that this system can control the robot move forward, move backward, turn left, turn right, stop, speed up in real-time and the recognition rate is above 90%. This technology also has a high reference value on researches such as machine learning and automatic recognition.

Key words: hand gesture recognition; robot; human-machine interaction; gesture control system; pattern recognition

随着科学技术的发展, 人们一直在探寻着一种全新的人机交互方式来突破鼠标、键盘、遥控器的约束. 手势识别技术^[1]由于其具有自然、方便的人际交流特性而被人们所广泛研究. 早在 1993 年 B. Thomas 等人就提出了借助数据手套或在人手粘贴特殊颜色的辅助标记来进行手势动作的识别, 由此开启了人们对手势识别领域的探索. 通过十多年的发展, 尤其是计算机视觉的快速发展, 人们从粘贴特殊颜色的辅助标记逐步向裸手识别技术迈进.

微软公司就在 2010 年推出了一款名为 Kinect 的设

备, 用户只需要完成一些预定的手势动作就可以实现控制电脑玩体感游戏的目的. 目前, 互联网巨头谷歌及汽车厂商福特均在研发支持手势控制功能的智能汽车, 此类汽车通过车载摄像头识别特定手势, 以此来替代汽车仪表盘上的各种旋钮和按钮. 或许很快, 司机师傅们只需要挥一挥手, 就可以随意打开窗户或者空调. 本文研究的就是这种手势识别控制技术.

文章建立了一个动态手势识别的模型, 通过手势分割、手势识别和系统交互^[2]来实现手势控制机器人系统的设计. 利用 YCrCb 颜色空间提取手掌区域并识

^① 收稿时间:2014-06-01;收到修改稿时间:2014-07-08

别出手心和手指指尖的坐标, 对手势区域进行实时跟踪和分析, 得出相应的手势信息. 将分析出来的手势信息通过无线蓝牙串口发送给机器人, 进而控制机器人的运动.

1 肤色提取及手势分割

1.1 基于 YCrCb 的肤色提取

YCrCb 颜色空间^[3]具有将色度和亮度分离的特点, 对肤色的聚类特性比较好, 受亮度变化的影响小, 可以很好的区分肤色区域. 根据文献[4]中的数据可以知道, 人体皮肤颜色在 YCrCb 色度空间的分布范围大致为: $77 \leq Cb \leq 127, 133 \leq Cr \leq 173$ 选取这个范围作为肤色分割的阈值.

以下是 RGB 颜色空间与 YCrCb 颜色空间的转换关系:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.169 & 0.331 & 0.500 \\ 0.500 & -0.419 & 0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

1.2 复杂背景下的手势分割

考虑到实际图像采集过程中背景噪声可能会对手部提取产生干扰, 系统设计时首先需要排除类似于皮肤颜色背景噪声的影响. 实验选择了和皮肤颜色相近似的相框和衣服作为背景如图 1-1, 通过以上的皮肤肤色模型对图像进行分割得到图 1-2 所示的二值化图片.

OpenCV 提供了查找连通区域的函数, 可以返回各个连通区域的标号和像素点的面积. 由于手部区域在图像中占主导地位, 取像素点面积最大的连通区域为候选区域, 用矩形标记出该连通区域和中心点的坐标, 对其他不符合条件的区域进行填充即可得到如图 1-3 所示的手势分割区域.

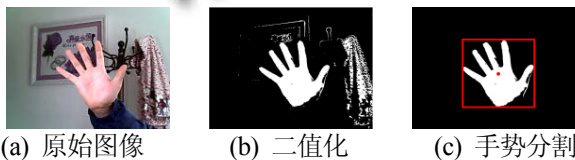


图 1 手势分割对比图

设计时还需考虑如果手不在图像捕捉范围时, 不能误将和肤色类似的背景区域理解为手势区域, 为了防止这样情况的发生, 规定当最大连通区域小于 5000

像素点时, 不予标记. 系统选择了许多较为苛刻的环境进行测试, 实验效果良好, 可以正确分割出手势区域.

2 手势信息分析及识别

2.1 指尖提取和手指个数识别

为了确定指尖的位置, 还需引入凸包^[5]这一概念. 提取出的手掌区域的二值化图片在计算机中是以二维平面的点集构成的. 凸包就是将最外层的点连接起来构成的凸多边形, 它能包含点集中所有的点, OpenCV 提供了 convexHull 函数来寻找凸多边形的各个顶点. 图 2 就是对手势图片的凸包顶点进行标记后的结果.

从图 2 中可以发现, 有些指尖被重复标记而且一些无关区域也被标记出来, 这就给精确识别指尖个数带来了麻烦, 需要去除这些不符合要求的点. 规定当凸包顶点间距离小于 500 像素点时, 只标记其中一个; 低于手心坐标的区域不予标记, 其中手心坐标提取公式^[6]如下:

$$C(x, y) = \frac{\sum_{m=0}^N P_m(x, y)}{N} \quad (2)$$

公式中, $P_m(x, y)$ 为手势区域中第 m 个像素点的坐标值, N 为手势区域中像素点的总个数, 得到的 $C(x, y)$ 为手心的坐标. 通过以上步骤就可以比较准确的识别出手心和指尖的位置, 统计蓝色圆圈的数目就可以得到手指个数, 如图 3.

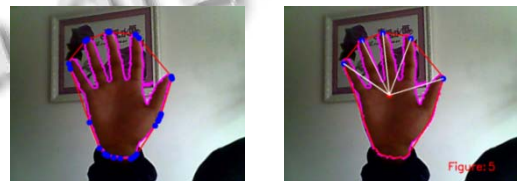


图 2 凸包顶点标记 图 3 指尖、手心标记

2.2 动态手势识别流程

上文识别出了手部区域, 手心、指尖位置和手指个数, 接下来通过帧差法^[7]对运动的手势信息进行识别处理. 程序设置 Hand 和 Count 两个变量分别标记是否捕捉到手和手进入的帧数, 程序流程图如图 4 所示.

情况一(Count=0, Hand=0): 当前图像和上一帧图像中都没有手出现, 表示这段时间内无手进出摄像头捕捉区域, 不做任何处理.

情况二(Count=0, Hand=1): 当前图像无手出现,

上一帧有手出现, 表示手刚刚离开摄像头捕捉区域, 保存该图像信息.

情况三(Count=1, Hand=1): 当前图像有手出现, 且为第一帧图像, 表示手刚刚进入图像捕捉区域, 保存该图像信息.

情况四(Count=K, Hand=1): 表示手一直在图像捕捉区域中.

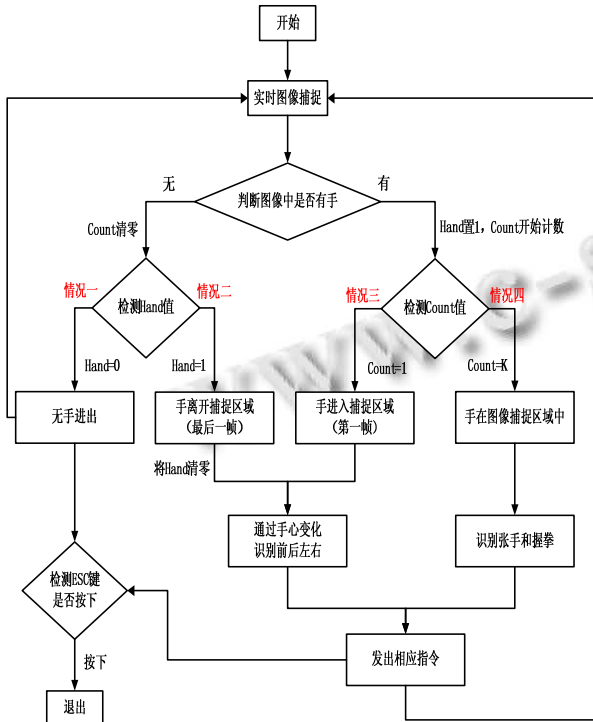


图 4 手势识别程序流程图

2.3 动态手势识别

2.3.1 上下左右的识别

结合情况二和情况三, 通过判断第一帧和最后一帧手势中心坐标的变化情况就可以识别出上、下、左、右这四种手势运动方向, 其识别过程可由如下公式表示:

$$(\Delta x, \Delta y) = C_{count-1}(x, y) - C_1(x, y) \quad (3)$$

$$\theta = \arctan\left(\frac{\Delta y}{\Delta x}\right) \quad (4)$$

$C_1(x, y)$ 和 $C_{count-1}(x, y)$ 分别表示第一帧和最后一帧手势中心的坐标, θ 为两点间的夹角, 通过判断 θ 值就可以知道手势运动的大体方向.

选取第一帧和最后一帧来代替手从进入捕捉区域到离开时系列图像帧, 即具有代表性也简化了程序设计的复杂度. 通过计算这两帧图像手心坐标的夹角的

正切值, 最多可识别出八种运动方向, 考虑到系统的精确度和实际设计的需求, 只选取了上、下、左、右这四种判别方向, 识别效果较为理想.

2.3.2 张手握拳的识别

对于情况四, 通过帧间差分的方法识别张手和握拳这两种动态手势信息. 张手可定义为连续 M 帧手指个数为 5 且接下来的连续 M 帧手指个数保存不变, 握拳可定义为连续 N 帧手指个数为 5 且接下来 N 帧内减少 T 个像素点. M、N 和 T 的值都需要根据系统的具体情况进行设置, 我们设置的阈值为 $M=N=20$, $T=1000$. 在此期间任何一帧不符合条件时 M、N 和 T 的值立即清零, 重新计数.

通过以上步骤就可以识别 6 种手势信息, 进而可产生 6 种不同的控制信号.

3 手势控制机器人系统设计

图 5 为手势控制机器人系统的原理图^[8]. 实验者对摄像头做出动态手势, 系统通过上文的手势识别流程识别出相应的手势动作, 当符合条件时给无线蓝牙串口写一个数据发送给机器人, 机器人接收该串口数据后执行相应的动作指令.



图 5 手势控制机器人系统的原理图

在实际设计过程中, 机器人若要调整运动方向, 必须先停止当前的运动再执行下一次运动, 指令如果发错要有纠错的能力, 发送指令之间要有一定的时间间隔, 以防止控制信息重复多次发送而导致机器人无法正确处理相应信息.

规定当发送前、后、左、右这四种指令时, 下一次指令只能发送加速或停止, 否则系统自动屏蔽不予发送, Hand 和 Count 值自动清零, 重新捕获图像. 每个

指令发送的时间间隔为三秒钟,三秒以内系统不对捕捉图像进行处理.如此以来,就可保证即使操作者错误发送一些指令信号,系统在识别的过程中也可以予以屏蔽,避免影响机器人正常工作.

表 1 手势动作对应串口数据

手势动作	指令	串口发送
上	前	W
下	后	S
左	左	A
右	右	D
握拳	加速	Q
张手	停止	E

表 2 动态手势识别结果

手势动作		上	下	左	右	握拳	张手
实验次数		20	20	20	20	20	20
快	识别个数	18	17	18	18	14	18
	识别率(%)	90	85	90	90	70	90
正常	识别个数	19	19	18	19	17	19
	识别率(%)	95	95	90	95	85	95
慢	识别个数	19	19	19	20	15	20
	识别率(%)	95	95	95	100	75	100
平均识别率(%)		93.3	91.7	91.7	95	76.7	95

5 总结与展望

基于计算机视觉的手势识别系统对环境较为敏感,手势识别易受到光照亮度、周围环境颜色和人体肤色的影响.从实验结果来看只要在正常光照条件下,周围没有大面积整块的与肤色接近的物体时系统是不会受到太大影响的.当环境背景较为复杂,特别是当手势区域和近似于肤色颜色区域重合时,系统会将手势部分和噪声区域整块的视为一个连通区域而影响手势的识别,如何消除这一点的影响,取得更好的鲁棒性将是我们下一步重点解决的问题.

参考文献

- 任海兵,祝远新,徐光佑.基于视觉手势识别的研究—综述.自动化学报,2000,29(6):900-912.
- Stefanov N, Galata A, Hubbard R. A real-time hand tracker

4 实验结果分析

实验者按照较快、正常、较慢这三种速度对每种动作进行了20次测试,实验结果和识别率如表2所示.可以发现手势运动速度的快慢对实验结果并没有太大的影响,除握拳这一手势动作以外,其他动作的平均识别率都能达到90%以上,如果操作者动作更加规范,其识别准确度更高.通过计算其时间复杂度,系统对不同手势动作的平均响应时间小于50ms,具有较高的实时性,可以满足对机器人进行实时控制的要求.

using variable-length Markov models of behavior. Computer Vision and Image Understanding, 2007, 108(1-2): 98-115.

- 邱迪.基于HSV和YCrCb颜色空间进行肤色检测研究.电脑编程技巧与维护,2012,10:74-75.
- Choi KM, Na YG, Chae Sb, Jung KH. A hand gesture-based remote control of robot. Proc. of KSBE Fall Conference. 2010. 196-199.
- Bradski G, Kaebler A.学习OpenCV中文版.于仕琪,刘瑞祯,译.北京:清华大学出版社,2008.
- 李平,李允俊.基于手势识别算法的鼠标终端.计算机系统应用,2013,22(8):83-71.
- 张晓宇,彭四伟.基于OpenCV的运动目标识别算法与实现.现代电子技术,2009,22:99-101.
- 陈一民,张云华.基于手势识别的机器人人机交互技术研究.机器人,2009,31(4): 351-356.