

智能空间下基于手势识别的人机交互

刘军¹, 田国会¹, 李荣宽¹, 刘贤锴²

(1. 山东大学控制科学与工程学院, 济南 250061; 2. 山东政法学院信息科学技术系, 济南 250014)

[摘要] 在智能空间框架下,提出了基于手势识别的人机交互。在基于手势识别的人机交互设计中,采用掩膜平滑滤波方法去除噪声。在 HIS 色彩空间中,通过非参数化的颜色直方图取得肤色范围,然后应用肤色阈值和面积几何特征对目标手势进行提取,最后通过最小欧氏距离分类方法进行模式识别得到用户的命令。经实验验证取得了良好的效果。

[关键词] 智能空间; 人机交互; 手势识别

[中图分类号] TP 242.6 **[文献标识码]** A **[文章编号]** 1005-0310(2010)02-0014-04

Human Machine Interface Based on Hand Gesture Recognition in Intelligent Space

LIU Jun¹, TIAN Guo-hui¹, LI Rong-kuan¹, LIU Xian-kai²

(1. School of Control Science and Engineering, Shandong University, Jinan 250061, China;

2. Department of Information Science and Technology, Shandong University of Political Science and Law, Jinan 250014, China)

Abstract: A gesture recognition-based human-machine interface (HMI) for home service robot is presented in this paper within the framework of intelligent space. In the gesture recognition-based HMI design, after mask filtering the target gesture was segmented by the complexion and geometry characters of the gestures. The complexion threshold by the non-parametric color histogram in the HIS color space was obtained. Intelligent space will receive the user's order from gestures which are classified using minimum Euclidean distance method. This HMI design in the intelligent space has got good results in the experiments.

Key words: intelligent space; human-machine interface (HMI); gesture recognition

0 引言

家庭服务机器人将会在不远的将来走入人们的日常生活中,它将给人们带来更准确、更快捷、更便利的服务。智能空间的概念最早是由日本东京大学的 Hashimoto 实验室于 1996 年提出的。智能空间(Intelligent Space)是为了使机器人在家庭或办公室等非确定性环境下更加高效地工作而建立的工作环境,在该环境中,装备有传感器、通信工具、服务器、局域网以及其他辅助环境识别工具。智能空间的核心思想就是将具备计算能力、感知能

力、执行能力的智能设备安装在室内的多个位置,使得普通的室内环境变成一个智能的、知识的、能够为用户提供支持的智能空间^[1-2]。

随着家庭服务机器人研究的深入,当机器人走入家庭之中为人类服务时,人们就会处于一种人机并存的智能环境中。在这种人机并存环境中,应用视觉手势识别的人机交互能够为人们提供有效的服务。视觉手势识别是指对视频采集设备拍摄到的包含手势的图像序列,采用计算机视觉技术来处理,最终识别手势。视觉手势识别在 1992 年前后才出现,在这之后,研究人员对视觉手势识别做了

[收稿日期] 2010-01-15

[基金项目] 国家“863”计划家庭服务机器人重点项目(2006AA040206);国家“863”计划项目(2009AA04Z220)。

[作者简介] 刘军(1986-),男,山东济宁人,山东大学控制科学与工程学院硕士研究生,研究方向为智能机器人。

深入的研究^[3]。研究者们一开始通过在人手上放置特定颜色的标记 (Marker) 来简化问题,如 Brian S. 等人应用感知手套,开发了一种手势交互系统^[4]。但是,这种方法不方便用于实际系统中,因此,现在主要研究的是不使用特定标志 (Marker-free) 的视觉手势识别。视觉手势识别一般分为静态手势识别和动态手势识别。动态手势识别包括运动序列的检测以及对每一帧图像中手势的识别理解,即包括运动检测和静态手势识别。因此,静态手势识别是手势识别的基础。在静态手势识别中,手的表示可以采用几何特征,如轮廓、指尖和手指的方向^[5]等以及非几何特征,如颜色、纹理等来表示^[6]。肤色是手很明显的特征,因此在手势识别中也最常用。Yingyos Sriboonruang 等利用肤色特征成功提取手势,然后检测拇指与食指指尖的距离,判断人的意图,用于实现操作电脑游戏^[7]。王西颖等对手势交互采用层次化建模方法,实现了在虚拟家居系统中的应用^[8]。李瑞峰等应用肤色特征,成功提取了手势,但光线干扰情况下的实验效果不佳^[9]。

在基于视觉的手势识别中,首先要对图像进行分割,检测到手势,然后提取手势特征,对手势进行分类。其中,如何从复杂图像中提取出手势图像,是手势识别的关键,也是研究的难点。在本文中,采用了肤色和几何特征提取出目标手势,然后通过线性分类方法进行模式识别。在图像预处理时,采用掩膜平滑滤波,在去除噪声的同时保留了区域边界的信息,然后通过肤色特征提取图像中的有用信息,再根据面积阈值去除其他干扰,从而得到最终的手势,最后在模式识别中采用线性分类方法中的最小欧式距离,将手势分类。

1 基于手势识别的人机交互技术

以智能环境下开关窗帘为例,当人做出某些手势时,分布于空间中的摄像头会检测到人的一些行为动作,然后对采集的图像进行处理,通过模式识别,明确人的意图,进而通过智能空间控制执行器开关窗帘。在手势识别中,对采集的图像首先进行预处理,去除噪声,然后根据肤色阈值以及面积几何特征提取目标手势,最后通过模式识别中的线性分类方法得到识别结果。

1.1 图像获取以及预处理

实验处理的图像是在实验室搭建的、在智能空间环境下采集的 24 位 bmp 图像,如图 1 所示。



图 1 原始图像

由于原始图像存在干扰噪声,所以要对图像进行目标提取之前的预处理。对于数字图像,常见的平滑处理方法有消除孤立点噪声、邻域平滑法、均值滤波以及自适应的局部平滑滤波方法等。在以上方法中,孤立点噪声、邻域平滑法以及均值滤波在消除噪声的同时,都不可避免地带来了平均化的缺憾,致使尖锐变化的边缘或线变得模糊。考虑到图像中目标物体和背景一般都具有不同的统计特征,即不同的均值和方差,为保留一定的边缘信息,可采用自适应的局部平滑滤波方法,又称为选择式掩膜平滑法。局部平滑滤波方法取 5×5 窗口,如图 2 所示。

在窗口内,以中心像素 $f(i, j)$ 为基点,制作 9 种形状的屏蔽窗口,分别计算每个窗口内的平均值及方差。由于含有尖锐边缘的区域,方差必定较平缓区域大,因此采用方差最小的屏蔽窗口进行平均化,这种方法在完成滤波操作的同时,又不破坏区域边界的细节。^[10]

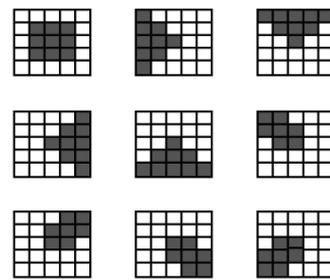


图 2 9 种屏蔽窗

计算各个掩膜窗口的均值 (a_i) 及方差 (k_i)

$$a_i = \frac{\sum_{m=1}^N f(i, j)}{N} \quad (1)$$

$$k_i = \sum_{m=1}^N (f^2(i, j) - a_i^2) \quad (2)$$

式中 $m=1, 2, 3, \dots, N$; N 为各个掩膜对应的像素个数。

算出各个掩膜对应的 a_i, k_i 后,取最小的 $k_{i_{\min}}$

对应的掩膜均值 $f(i, j)$ 作为平滑输出。

1.2 目标的分割提取

图像采集预处理之后,对其进行色度空间转换,将 RGB 颜色模型表示的原始图像转换为 HSI 颜色模型。

HSI 颜色模型是从人的视觉系统出发,用色调(Hue)、色饱和度(Saturation 或 Chroma)和亮度(Intensity 或 Brightness)来描述色彩。对于手势识别,要将上面采集的彩色图像从 RGB 转换为 HIS 坐标,以便实现基于肤色的图像分割提取。RGB 转换为 HIS 坐标变换关系如公式(3)所示:

$$\begin{cases} I = \frac{R + G + B}{3} \\ H = \frac{[90 - \arctan(F/\sqrt{3}) + \{0 \text{ } G > B; 180 \text{ } G < B\}]}{360} \\ S = 1 - \frac{\min(R, G, B)}{I} \\ \text{其中 } F = \frac{2R - G - B}{G - B} \end{cases} \quad (3)$$

试验中的肤色阈值由实际样本测得,将待测肤色图像样本转换为 HIS 空间表示,然后通过非参数化的颜色直方图的形式得到人体肤色的聚类特征。通过多次实验测算,得出肤色的 HIS 阈值范围是: $10 \leq H \leq 40$, $S \geq 40$, $I \geq 100$,其中,转换后图像 HIS 各数值均为 8 位数据存储。

图像进行色彩模型转换后,根据肤色的 HSI 色彩阈值范围,对原始图像进行提取,提取后的结果如图 3 所示。

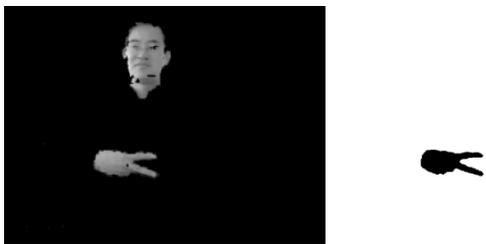


图 3 按肤色分割后的图像与最终手势图像

提取出肤色区域之后,再对图像进行灰度变换并进行二次滤除噪声操作。灰度变换先进行二值化操作并反色变换。在此之后,将获取的 24 位 bmp 图像转换为 256 色 bmp 图像,这样做可以提高图像的处理速度,而且不会影响后期图像识别。

下一步通过面积阈值提取出手势图像。在处理过程中,首先对二值图像中的连通域进行标记,然后对于同一标号的区域进行像素统计,得出每个连通区域的面积。实验测得,上图中,手势像素面

积为 4 047,通过实验,验证手势区域的肤色面积具有很好的聚类特性。经过对其他手势做相同处理,综合各次实验结果,设定手势面积阈值为 [3 000, 5 000] 像素点,最终获得的手势图像如图 3 所示。

1.3 图像识别理解

经过以上处理,已经得出了所期望的手势目标,下面将要对其进行识别分类。在图像识别中,有很多方法可用,在这里,采用模板匹配方法,取得了十分理想的效果。

设有 M 种类别: $\omega_1, \omega_2, \omega_3, \dots, \omega_M$, 每类的特征由向量表示,如 ω_i 类由 X_i 表示,其中 $X_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$, 任意被识别手势 $X = [x_1, x_2, \dots, x_n]^T$,

判别待测手势图像时,有如下公式:

$$d(X, X_i) = |X - X_i|^2 = (X - X_i)^T (X - X_i) = X^T X - (X^T X_i + X_i^T X - X_i^T X_i) \quad (4)$$

对待测样品与各已知类别计算二者之间的欧式距离 $d(X, X_i)$, 当取得最小值时,则 X 属于相对应的 X_i 类,即 $X_i \in \omega_i$ 。

手势识别算法步骤如下:

1) 对待测图像进行特征提取,计算出待测样品的高度和宽度,将每个样品的长和宽 N 等份,构成 $N \times N$ 均匀区域,如图 4 所示。其中, N 越大,模板也越大,特征越多,区分不同物体的能力越强,但同时计算量增大,所需样品库也增加。

2) 对每一个小区域的黑像素个数进行统计,再除以该小区域的面积总数,即得特征值。这样做的好处在于针对同一形状,不同大小的样品得到的特征值相差不大,能够将同一形状、不同大小的手势样品归为同一类。

3) 将待测手势样品 X 与训练集里的每个样品 X_i 按照上面公式进行模板匹配,找出与待测样品最接近的已知样品,则待测样品属于已知样品。

对图像提取后,通过模板匹配得到结果,如图 4 所示。

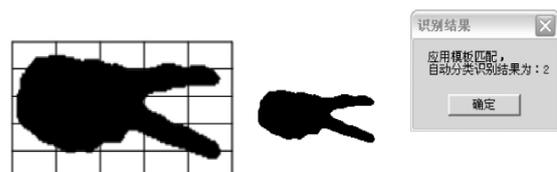


图 4 手势模板与识别结果

2 实验结果及分析

在实验室搭建的智能空间环境下,采集了 8 组

手势图像,并对其进行了处理识别,结果如图 5 所示。实验表明,在不同的复杂背景以及强光照条件下,基于肤色的手势分割都取得了很好的效果。在静态手势识别中,识别率可以达到 90% 以上。单个手势包括左手和右手两种不同情况,再配以运动。检测与识别人手的向上、向下、向左和向右等运动状态可以搭配出多种手势,给不同的手势赋予不同的含义,这便可以让机器人理解人的意图,进而执行相应的动作实现与人的交互。

3 结论

本文设计了基于手势识别的结合智能空间技

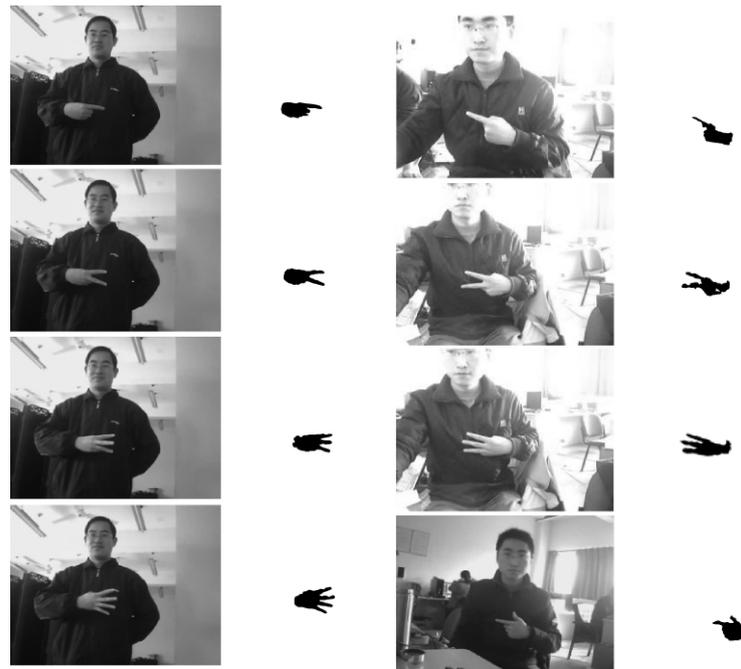


图 5 原始图像及实验结果

[参考文献]

- [1] 田国会. 家庭服务机器人研究前景广阔[J]. 国际学术动态, 2007(1): 28 - 29.
- [2] 田国会, 李晓磊, 赵守鹏, 等. 家庭服务机器人智能空间技术研究与进展[J]. 山东大学学报:工学版, 2007, 37(5): 53 - 59.
- [3] Pavlovic V, Sharma R, Huang T. Visual interpretation of hand gestures for human-computer interaction: A review [J]. IEEE Trans Pattern Anal. Machine Intell., 1997, 19(7): 677 - 695.
- [4] Tani B S, Maia R S, Wangenheim A. A Gesture Interface for Radiological Workstations [C]// Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07) 2007: 27 - 32; 20 - 22.
- [5] Ng C, Ranganath S. Real-time gesture recognition system and application [J]. Image and Vision Computing, 2002, 20(13/14): 993 - 1007.
- [6] Triesch J, Malsburg C. A gesture interface for human-robot interaction [C]// IEEE Int'l Conf on Automated Face and Gesture Recognition (FG'98), 1998: 546 - 551.
- [7] Sriboonruang Y, Kumhom P, Chamnongthai K. Visual Hand Gesture Interface for Computer Board Game Control [C]// Tenth IEEE International Symposium on Consumer Electronics (ISCE '06), 2006: 1 - 5.

(下转第 53 页)

术的人机交互系统,使传统的人机交互变得更加简便易用,符合智能敏捷型家庭助理机器人的项目要求。基于手势识别的人机交互,人们在人机并存环境下更加舒适地向智能空间传递意图,并接受反馈信息。这种交互方式都是面向被服务对象的人性化设计,能够理解人的手势,具有较高的智能性。在硬件方面,采用摄像头,成本较低,而且处理结果的传输通过无线局域网络,也满足了人机交互实时性的要求。所以,本文所提出的基于手势识别的人机交互具有良好的识别效果,能够满足家庭中复杂的人机并存环境的需要。

