

结合归一化中心矩与 DAGSVM 的机器人交互手势识别

蔡 军 李晓娟 张 毅 罗 元
(重庆邮电大学智能系统及机器人研究所 重庆 400065)

摘 要 将归一化中心矩和 DAGSVM 相结合应用于机器人人机交互中的手势识别。归一化中心矩作为手势特征具有平移和比例不变性,同时对方向变化比较敏感,这正是机器人识别不同指向的手势时特征所需具有的特性,然后,将这个手势特征向量输入到 DAGSVM 分类器进行分类识别。实验结果表明,该方法能够有效地识别手势,且用于控制机器人运动效果良好。

关键词 手势识别 归一化中心矩 有向无环图支持向量机

中图分类号 TP391 文献标识码 A DOI:10.3969/j.issn.1000-386x.2013.07.036

ROBOT INTERACTIVE HAND GESTURE RECOGNITION COMBINING
NORMALISED CENTRAL MOMENTS AND DAGSVM

Cai Jun Li Xiaojuan Zhang Yi Luo Yuan
(Research Center of Intelligent System and Robot, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract We combine the normalised central moments and the DAGSVM (directed acyclic graph support vector machine) together to apply them to hand gesture recognition in robot-human interaction. The normalised central moments is selected as the feature of hand gestures, which is robust to translation and scale, at the same time it is also sensitive to direction, and these are the necessary characters for robot to recognise different pointing hand gestures. Afterwards, we input the gesture feature vectors to DAGSVM classifier for classifying and recognising. Experimental results show that this method can effectively recognise the hand gestures, and has good effect when to be used to control the robot motion.

Keywords Hand gesture recognition Normalised central moments DAGSVM

0 引 言

随着机器人技术的发展,人们对机器人人机交互过程中手势的直观性有越来越高的要求。目前静态手势识别大部分都是基于手形的手势识别^[1-3],如文献[4,5]通过提取手势的形状特征识别手形来控制机器人的运动,该方法简单有效但由于控制指令不够直观明了,用户必需记住相应的指令才可以操作机器人,使用不方便。为此有的学者和研究机构采用信息融合的方法,如华南理工大学的肖志勇采用人眼和指尖的连线确定用户指向屏幕的位置^[6],从而达到人机交互的目的;东南大学的钱堃采用结合人脸和人手的姿态作为输入^[7],根据人脸和人手的相对位置的不同来控制机器人,实现机器人的简单运动。以上采用信息融合的两种方法中控制指令简单易学且比较直观,但是操作过程中用户的身体比较受限制,有时操作不方便。同时由于手势分割比较容易受光照和类肤色物体影响^[7],因此将手势区域和背景区域分割开也是本文手势识别中一个需要解决的关键问题。

基于上述状况,本文提出了一种识别不同指向的手势的方法。该方法通过提取对方向变化敏感的归一化中心矩作为手势特征,采用有向无环图支持向量机(DAGSVM)来识别不同指向的手势。同时为了能够更好地分割手势,本文利用微软推出的

Kinect 传感器采集场景的深度信息。它是一种 3D 体感摄影机,同时具有即时动态捕捉、图像识别、麦克风输入、语音识别、社群互动等功能。

1 系统构成与基本原理

机器人交互手势识别系统流程如图 1 所示。



图 1 机器人交互手势识别技术思路

收稿日期:2012-07-05。重庆市科技攻关项目(CSTC,2010AA2055);科技部国际合作项目(2010DFA12160)。蔡军,副教授,主研领域:模式识别,智能控制。李晓娟,硕士生。张毅,教授。罗元,教授。

采用 Kinect 传感器获取场景的深度图像,使用灰度直方图方法确定分割手势区域和背景区域的深度阈值,从而分割出手势区域,提取手势的归一化中心矩特征并将其输入到 DAGSVM 分类器中进行识别,DAGSVM 分类器是由多个支持向量机按有向无环图(DAG)拓扑结构构成的,其中支持向量机是根据采集的样本训练得到的,最后将识别结果转换为相应的控制指令控制机器人的运动。

2 结合归一化中心矩和 DAGSVM 的手势识别

2.1 手势分割

在人机交互中,交互手势往往位于摄像头所拍场景的最前面,因此可以根据手势区域和背景区域具有不同的深度值进行分割手势。本文使用 Kinect 采集场景的深度图像,尽管具有同一深度的像素点在深度图像中灰度值相等,但是每次手和 Kinect 之间的距离都不完全相同,所以无法用固定的深度阈值实现手势区域的分割。

本文采用灰度直方图方法来确定分割手势区域和背景区域的深度阈值。灰度直方图是用来统计图像中具有每种灰度级的像素个数分布,反映了图像中每种灰度出现的频率。首先由 Kinect 传感器采集深度图像,其中采集到的深度图像为灰度图如图 2(a)所示,然后计算深度图像所对应的灰度直方图如图 2(c)所示。在灰度直方图中由大到小寻找像素数目变化较快的灰度值作为深度分割阈值,同时对分割得到的手势二值图平滑去噪,手势分割效果如图 2(b)所示。

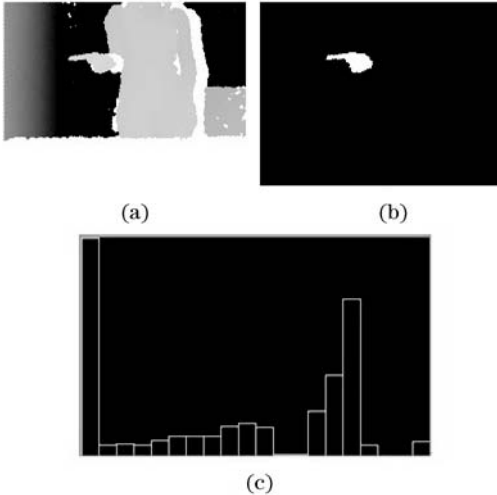


图2 手势区域分割

2.2 采用归一化中心矩的手势特征提取

M. K. Hu 已于 1962 年提出了矩的定义和关于矩的基本性质^[8],并具体给出了具有平移和比例不变性的同时对方向变化敏感的归一化中心矩的表达式。为了使机器人交互手势更直观,本文采用归一化中心矩作为手势特征。

对于手势二值图像 $f(x, y)$,其 $(p+q)$ 阶原点矩表示为:

$$m_{pq} = \sum_y \sum_x f(x, y) x^p y^q \quad p, q = 0, 1, 2 \quad (1)$$

其中心矩表示为:

$$\mu_{pq} = \sum_y \sum_x (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad p, q = 0, 1, 2, \dots \quad (2)$$

$$\text{其中 } \bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}.$$

式(2)中的中心矩用原点矩来表示,前四阶中心矩为:

$$\mu_{00} = m_{00} \equiv \mu \quad (3)$$

$$\mu_{20} = m_{20} - \mu \bar{x}^2 \quad (4)$$

$$\mu_{02} = m_{02} - \mu \bar{y}^2 \quad (5)$$

$$\mu_{11} = m_{11} - \mu \bar{x} \bar{y} \quad (6)$$

$$\mu_{12} = m_{12} - m_{02} \bar{x} - 2m_{11} \bar{y} + 2\mu \bar{x} \bar{y}^2 \quad (7)$$

$$\mu_{21} = m_{21} - m_{10} \bar{y} - 2m_{11} \bar{x} + 2\mu \bar{x}^2 \bar{y} \quad (8)$$

$$\mu_{30} = m_{30} - 3m_{20} \bar{x} + 2\mu \bar{x}^3 \quad (9)$$

$$\mu_{03} = m_{03} - 3m_{02} \bar{y} + 2\mu \bar{y}^3 \quad (10)$$

中心矩属于一种平移不变量,使之具有比例不变性,应进行如下归一化处理:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}} \quad p+q=2, 3, 4, \dots \quad (11)$$

本文手势特征向量选定为: $h = [\eta_{20}, \eta_{02}, \eta_{11}, \eta_{12}, \eta_{21}, \eta_{30}, \eta_{03}]$,该手势特征具有平移不变性和比例不变性,同时对方向变化敏感。

2.3 基于 DAGSVM 的手势识别

DAGSVM 即基于 DAG^[9]的支持向量机。支持向量机的基本思想是:通过非线性映射将输入向量映射到高维空间,并利用输入空间核函数取代高维空间中的点积运算,从而降低计算的复杂性。在高维空间中,依据结构风险最小化原则构造最优决策函数。具体算法思想介绍如下:

设样本集为 (x_i, y_i) , $x_i \in R^d$, $y_i \in \{-1, +1\}$ 。通过非线性映射 $\Phi(x)$ 将输入的特征向量映射到高维空间,并在高维空间构造出最优决策函数 $f(x) = \text{sgn}[\omega^T \cdot \Phi(x) + b]$,其中 ω 为权向量($\omega \in R$), b 为实常数。

对线性不可分情况,SVM 引入松弛因子 ξ_i ($\xi_i \geq 0$)和惩罚因子 c ($c \geq 0$),根据结构风险最小化原则,分类问题可以转化为以下约束优化问题:

$$\min \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n \xi_i \quad (12)$$

s. t

$$y_i [\omega^T \cdot \Phi(x) + b] \geq 1 - \xi_i \quad (13)$$

总可以在输入空间中找到一个满足 Merce 条件的核函数 $K(x_i, x_j) = (\Phi(x_i) \cdot \Phi(x_j))$ 取代高维空间中的点积运算。引入拉格朗日函数并根据 KKT 优化条件,求解其对偶二次规划问题,得到最优决策函数,如式(14)所示,其中 α_i 为拉个朗日乘子。最优决策函数其实是一个支持向量机二分类器。

$$f(x) = \text{sgn} \left[\sum_{i=1}^n \alpha_i y_i K(x, x_i) + b \right] \quad (14)$$

支持向量机本身只能解决二分类问题,当解决多分类问题时,需由多个支持向量机二分类器采用 DAG 拓扑结构构成多值分类器 DAGSVM。DAGSVM 需要训练 $k(k-1)/2$ 个支持向量机二分类器。本文定义了 5 个类,所以需要训练 10 个支持向量机二分类器。如图 3 所示,有 10 个节点,每个节点都是一个由类 i 和类 j 训练得到的支持向量机二分类器,如根节点 1V5 表示由类 1 和类 5 训练得到的支持向量机二分类器。将特征向量 $h = [\eta_{20}, \eta_{02}, \eta_{11}, \eta_{12}, \eta_{21}, \eta_{30}, \eta_{03}]$ 输入到 DAGSVM 分类器中,从根节点开始,经过 4 次支持向量机二分类器判断即可得到手势的分类结果。为了最大限度地减少误识率,将识别率高的分类器置于拓扑结构的较高层。

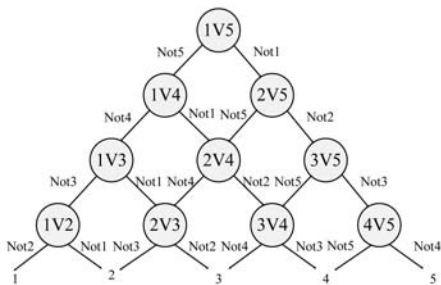


图3 基于 DAGSVM 的五类分类器结构

3 实验结果及分析

3.1 手势识别实验

实验中定义了5个手势,4个不同的手势者每个手势采集10个手势样本,共计200个样本。部分手势样本如图4所示,从左到右依次代表控制机器人的前进、后退、左转、右转和停止指令。对采集到的手势样本采用前述方法提取手势特征向量,训练支持向量机二分类器并组织多个支持向量机二分类器构成DAGSVM分类器,从而在手势识别时使用DAGSVM分类器对分割出的手势进行识别。



图4 部分手势样本

用定义的5个手势分别做了50次实验,实验结果如表1所示。实验表明,该方法能够有效地识别不同指向的手势且识别效果稳定。

表1 手势识别结果统计

手势分类	正确次数	识别率(%)
前进	46	92
后退	45	90
左转	45	90
右转	46	92
停止	47	94

本文也对在平面内旋转一定角度后的手势进行了识别,部分旋转后的手势如图5所示,其中除手势1和手势6以外都能够得到正确的识别,结果表明,只要该手势在平面内旋转角度不大于20度就能够正确识别手势。其他手势的识别情况与此大致相同。因此,归一化中心矩的方向敏感性在满足本文定义的手势被正确识别,同时也能够保证手势在发生小范围旋转时手势的正确识别。

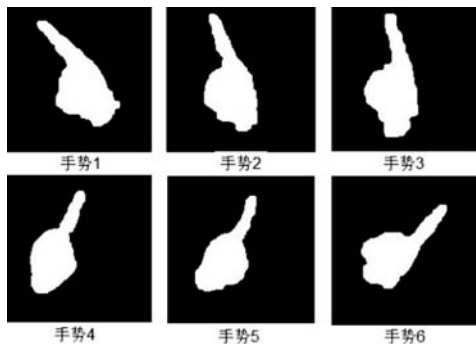


图5 旋转角度不同的“前进”手势

3.2 机器人手势交互实验

在实验中,对机器人在运动过程中采集到的手势进行识别。机器人由 ARM9 上位机、底层驱动模块和传感器模块共同完成驱动的智能轮椅服务机器人。将手势识别结果转换为相应的控制指令,然后通过 TCP/IP 无线网络将控制指令发送给机器人,从而达到控制机器人运动的目的。

该实验由4个不同的实验者(其中包括1个对机器人操作熟练的实验者和3个不熟练的实验者)在长800cm宽600cm的实验场地中进行,按照图6所设定的路线控制机器人运动,机器人一次完整的运动过程是从A点按箭头所示方向运动到B点,实验场地中两个长方形物体为障碍物。4个实验者实际使用该手势识别方法控制机器人的运动轨迹如图7所示。实验结果表明,该手势识别方法不论对机器人操作熟练的人还是不熟练的人,都能够很好地控制机器人的运动,且各手势指令直观易学,实验者在交互过程中感受良好。

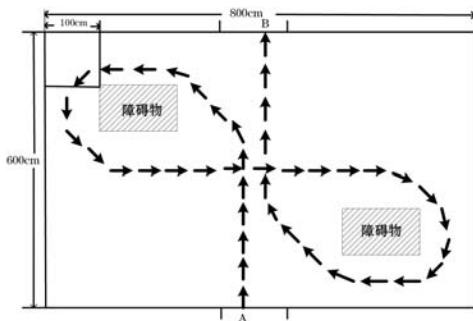


图6 设定的机器人运动路线

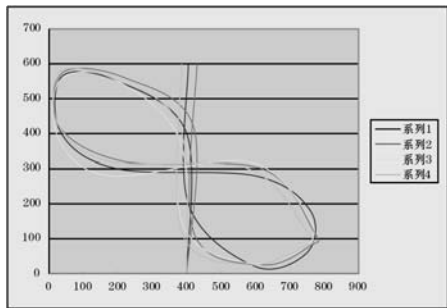


图7 机器人实际运动轨迹

4 结 语

文中通过采集场景的深度信息有效地将手势区域和背景区域分割开。采用具有平移和比例不变性的同时对方向变化比较敏感的归一化中心矩提取手势特征,并结合 DAGSVM 实现了一种简单直观的机器人手势交互方式。同时实现进一步精确控制机器人的运动是本文下一步的工作目标。

参 考 文 献

[1] Huang D Y, Hu W C, Chang S H. Vision-based hand gesture recognition using PCA + Gabor filters and SVM[C]//5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE, 2009:1-4.

[2] Ren Z, Meng J J, Yuan J S, et al. Robust hand gesture recognition with kinect sensor[C]//19th ACM International Conference on Multimedia ACM Multimedia, Association for Computing Machinery, 2011:

759-760.

[3] Choi J, Park H, Park J. Hand gesture recognition using distance transform and shape decomposition[C]//18th IEEE International Conference on Image Processing, IEEE, 2011:3605-3608.

[4] Murthy G R S, Jadon R S. Hand gesture recognition using neural networks[C]//2010 IEEE 2nd International Advance Computing Conference, IEEE, 2010: 134-138.

[5] 罗元, 谢彧, 张毅. 基于 Kinect 传感器的智能轮椅手势控制系统的设计与实现[J]. 机器人, 2012, 34(1):110-113,119.

[6] 肖志勇, 秦华标. 基于视线跟踪和手势识别的人机交互[J]. 计算机工程, 2009, 35(15):198-200.

[7] 钱望, 马旭东, 戴先中, 等. 基于最优 DAGSVM 的服务机器人的交互手势识别[J]. 中国图象图形学报, 2009, 14(1):118-124.

[8] Hu M K. Visual patter recognition by moment invariants[J]. IRE Transactions on Information Theory, 1962, 2(8):179-187.

[9] Platt J C, Cristianini N. Large margin DAGs for multiclass classification[J]. Advances in Neural Information Processing Systems, 2000, 12(3):547-553.

(上接第 13 页)

由图 6 可以看出,一次谱聚类算法(OSC)和基于图划分算法 CBG 二者比另外两种算法要好得多。而基于图划分算法 CBG 又比一次谱聚类算法(OSC)要好,这是在程序中通过不停的迭代优化的结果。可见,基于图划分算法 CBG 优于 OSC,是一次谱聚类算法(OSC)的优化补充。

2.2 数据集 the German credit dataset 上的实验结果

该数据集取自 UCI Machine Learning Repository^[12],它由 7 个数值属性和 13 个符号属性数据组成的,共有两个类,共有 1000 个样本数据。为了便于计算,实验中随机取了 300 个样本作为数据集。采用错误率作为评判算法的性能。实验效果如图 7 所示。

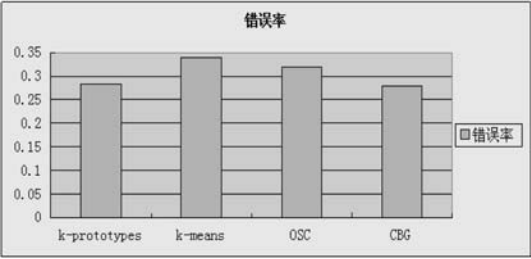


图 7 各种算法在 German 数据集上的结果

从图 7 中可以看出,对于 German credit dataset 这个数据集,各种算法的效果都不理想。但是,基于图划分算法 CBG 性能好于另外两种算法。另外,一次谱聚类算法(OSC)比 k-means 算法表现要好,可见谱聚类算法对于该数据集表现要比 k-means 好,也证明了我们选取谱聚类算法是正确的选择。

2.3 数据集 the heart-disease/cleve 上的实验结果

该数据集同样来自于 UCI Machine Learning Repository^[12],此数据集包含 7 个符号属性和 6 个数值属性,共有两个类。去掉一些丢失数据的数据点后,一共得到一个 296 个对象点的数据集。实验效果如图 8 所示。

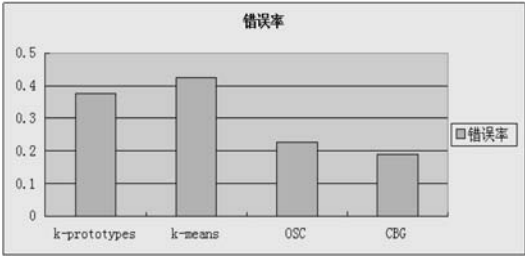


图 8 各种算法在 heart-disease/cleve 数据集上的结果

由图 8 可以看出,一次谱聚类算法(OSC)和基于图划分算法 CBG 二者比另外两种算法要好得多。而基于图划分算法 CBG 又比一次谱聚类算法(OSC)的表现要好一些,这是因为在程序中通过不停迭代优化的结果。可见,基于图划分算法 CBG 是一次谱聚类算法(OSC)的优化,迭代循环寻优是有道理的,是符合客观规律的。

3 结 语

本文提出了一个有效的混合属性聚类算法,算法统一计算两类属性,采取循环迭代的方法对聚类结果进行优化,使结果达到最优。实验结果表明,基于图划分的混合数据聚类算法(CBG)具有理想的聚类效果,性能优于其他算法。算法的不足是采取了谱聚类算法和循环迭代,当数据量很大时,算法的计算复杂度较高。下一步研究工作包括如何降低算法复杂度和自动确定类别数,实现自动聚类。

参 考 文 献

[1] Cheeseman P, Stutz J. Bayesian classification (AutoClass); Theory and results Advances in Knowledge Discovery and Data Mining [M]. AAAI Press/The MIT Press, 1996:153-180.

[2] Li C, Biswas G. Unsupervised Learning with Mixed Numeric and Nominal Data [J]. IEEE Trans. Knowl. Data Eng. ,2002,14(4):673-690.

[3] Goodall D W. A New Similarity Index Based On Probability [J]. Biometrics ,1966,22:882-907.

[4] He Z, Xu X, Deng S. Clustering Mixed Numeric and Categorical Data: A Cluster Ensemble Approach[OL]. eprint arXiv:cs/0509011, 2005.

[5] Ng A, Jordan M, Weiss Y. On Spectral Clustering: Analysis and an algorithm [C]// NIPS. Vancouver, British Columbia, Canada: MIT Press, 2001:849-856.

[6] 陈犇, 王雷. 基于 K-prototypes 的混合属性数据聚类算法[J]. 计算机应用, 2010, 30(8):2003-2010.

[7] 赵宇, 李兵. 混合属性数据聚类融合算法[J]. 清华大学学报, 2006, 46(10): 1673-1676.

[8] 蔡晓研, 戴冠中. 谱聚类算法综述[J]. 计算机科学, 2008:35(7): 14-18.

[9] Meila M, Xu L. Multiway cuts and spectral clustering[R]. U. Washington Tech Report, 2003.

[10] Huang Z. Clustering Large Data Sets with Mixed Numeric and Categorical Values [C]// Proceedings of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining, (PAKDD). Singapore, 1997:21-34.

[11] 吴东磊, 于洪, 杨大春. 一种处理混合型数据的层次聚类算法[J]. 计算机应用研究, 2009, 26(8):2885-2887.

[12] <http://archive.ics.uci.edu/ml/datasets/>.