

基于圆弧扫描线的手势特征提取和实时手势识别

刘砚秋 王修晖

(中国计量学院信息工程学院, 杭州, 310018)

摘要: 基于手势的人机交互是当前备受关注的自然人机交互模式之一, 实时手势识别是其中最重要的步骤。本文提出了一种基于圆弧扫描线的手势特征提取和实时手势识别方法。首先, 基于一种抽象描述手掌和五指关系的简洁人手海龟模型, 结合肤色特征和腕部标记分割出人手部图像, 并进行二值化处理和统一尺寸来建立手势训练集。然后, 以手掌中心为圆心构造同心圆来提取训练集中不同手势样本的特征, 并使用线性判别分析(Linear discriminant analysis, LDA)算法对手势特征向量进行离线预处理。最后, 使用改进的加权K近邻(Weighted-K-nearest neighbor, W-KNN)算法进行实时手势分类和识别。为了验证本文方法的有效性, 在自建小型手势数据库上进行了算法分析和比较, 并在多投影系统下进行实时交互测试。实验结果表明本文算法具有较高的识别效率。

关键词: 人手模型; 手势识别; 线性判别分析; 人机交互; K近邻法

中图分类号: TP391 **文献标志码:** A

Gesture Feature Extraction and Recognition Based on Circular Scan Lines

Liu Yanqiu, Wang Xiuhui

(College of Information Engineering, China Jiliang University, Hangzhou, 310018, China)

Abstract: Human computer interaction based on hand gesture is one of the most popular natural interactive modes, which severely depends on the methods for real-time gesture recognition. Here, an effective hand feature extraction method is described, and the corresponding hand gesture recognition method is proposed. First, based on a simple tortoise model, one segments the human hand images by skin color features and tags on the wrist, and normalizes them to create the train set. Then feature vectors are computed by drawing concentric circles according to the center of the palm, and linear discriminant analysis (LDA) algorithm is used to deal with those vectors. Finally, an improved K-nearest neighbor (KNN) algorithm is presented to achieve real-time hand gesture classification and recognition. Experimental results with a self-defined hand gesture data set and multi-projector display systems prove the efficiency of the new approach.

Key words: hand gesture model; hand gesture recognition; linear discriminant analysis; human computer interaction; K-nearest neighbor method

引 言

手势是一种符合人类日常交流习惯的交互媒介,将人的手部形态直接作为计算机系统的输入信息,可以通过预先定义的特定手势来对目标物体或场景进行控制。手势是人手或者手和手臂相结合产生的各种姿势或动作,包括静态手势(指交互姿态,单个手形)和动态手势(指交互动作,由一系列姿态组成)^[1]。随着计算机视觉技术在各个领域的应用日益广泛,如何实现自然和谐的人机交互成为备受关注的研究热点,相关研究包括面部表情识别、手势识别、头部运动跟踪以及背影识别等^[2]。其中,手势识别作为高级人机交互系统的关键技术,在计算机视觉、视频会议和机器人等领域有着广阔的应用前景。根据感知设备的不同,关于手势识别的研究成果可以分为基于硬件传感器(如数据手套)和基于计算机视觉的方案两大类。基于硬件传感器的手势识别方案主要有使用数据手套来感应手形变化的方法^[3]和通过在指关节设置感应环来识别手指形态的方法^[4]。由于此类方法通常需要给交互者的手部佩戴某种物理感应装置,这在一定程度上影响了交互界面的友好性。另外,数据手套等物理感应设备的灵敏度对手的尺寸和形状变化过于敏感,而且其内部传感部件容易磨损老化,这些都进一步限制了此类辨识方法的应用场合。

基于视觉的手势识别主要通过对手部肤色、手形轮廓或者预设标识的分析识别来提取手部特征,并通过进一步的手形匹配来理解当前手势的交互语义,相关研究又可以分为基于模型的和无模型的手势识别方法。文献[5]设计了一种简单的 Tortoise 模型来表征人手部的基本特征,并结合遗传算法中优胜劣汰的规则,在几何与纹理混合特征空间实现目标手形与模式库中手形匹配和识别手势,提高了实时性,但其对环境光照稳定性要求较高。文献[6]提出了一种基于手势空间分布特征的静态手势识别算法,具有较强的实用性,但是该方法未能处理人手区域和人脸区域部分重叠或完全重叠的情况,而且当手势区分度较小时,识别率会显著降低。文献[7]提出了一种基于区域生长与 Meanshift 算法相结合的动态变形手势跟踪算法,该方法无需对手势建立模型,但是依赖于前期手势分割的精度,当手势变化剧烈时,会对跟踪和识别效果产生较大影响。文献[8]介绍了一种基于视觉的手势交互桌面应用开发框架,该框架采用 RGB-D 信息作为数据输入,并利用 3D 手形轮廓降低手势匹配的复杂度,能够支持复杂手势的实时识别,但是该方法尚无法处理身体与手臂一起出现在视场中时的手部分割与手势识别。文献[9]提出了一种基于 LDA 主题模型来识别特定生物条件下 miRNAs 和靶标 mRNAs 之间的调控模块。此类方法相对而言,基于计算机视觉的手势识别更加符合人类的日常交互习惯,而且识别的准确性和系统的耐用性也有较大幅度地提高。但是,由于人手复杂的非刚体变形和视频信息的不稳定性,给手势识别带来了一定的困难。为了克服当前基于视觉的手势识别方法中实时性不足的问题,本文在文献[5]的基础上,提出一种更加简介有效的手势特征表示和实时识别方法。

1 手势特征的提取

1.1 简单海龟模型

在进行手势图像分割和手势特征提取时,可以将手部分解成椭圆形的手掌和 $n(n \in [0, 5])$ 根被看作刚体的手指,本文使用一种特征显著且便于检测和描述的人手海龟模型^[5]来描述人手的基本特征(如图 1 所示)。人手海龟模型的参数形式定义如下

$$y = f(r_1, r_2, n, L_1, \dots, L_n, W_1, \dots, W_n, \theta_1, \dots, \theta_n, R, G, B) \quad (1)$$

$$\text{s. t.} \begin{cases} r_1 \geq 1.5r_2 \\ n \in [0, 5] \\ 1.2r_1 \geq L_i \geq 0.3r_1, i = 1, \dots, n \\ r_1 \geq 3.0W_i, i = 1, \dots, n \\ \theta_i \in [0, 90], i = 1, \dots, n \\ R \in [0, 255], G \in [0, 255], B \in [0, 255] \end{cases} \quad (2)$$

式中: r_1 为手掌部半径; r_2 为手腕部半径; n 为手指个数; L_1, \dots, L_n 为每根手指的长度; W_1, \dots, W_n 为每根手指的宽度; $\theta_1, \dots, \theta_n$ 为每根手指与手腕间夹角 ($\theta_i \in [0, 90), i = 1, \dots, n$); R, G, B 为皮肤颜色。

人手海龟模型具有以下优势:(1) 采用近似对称的椭圆形手掌模型, 简洁的模型表达提高了手形分割和特征提取速度, 同时, 模型的对称性也减少了由于手部旋转造成的误判率;(2) 分布比较随意的矩形手指模型, 使得在具体应用中可以通过增加摄像机数量, 从不同二维图像中计算总的手指数来判断手形, 可以进一步降低由于遮挡带来的误判率;(3) 使用手掌和手指的相对尺寸作为区分不同手势的依据, 在一定程度上可以减少手臂相对镜头运动造成的手形缩放所带来的误判率。

1.2 基于圆弧扫描线的手势特征提取

在人手海龟模型基础上, 使用圆弧扫描线处理二值化的手型图像, 如图 2(a) 所示。其基本思想是: 获得手掌的中心点位置 G , 然后以 G 为圆心, R_i 为半径画一簇同心圆, i 为同心圆编号, $i \in [R_{\min}, R_{\max}]$ 。 R_{\min} 和 R_{\max} 分别为最内层同心圆和最外层同心圆的半径。 R_{\min} 的值取决于画圆起始条件: 初始化半径 R 为 1, 当圆弧完全落在手掌区域内时, 增加半径 $R = R + S$, 直到有部分圆弧落到掌心外部为止。而 R_{\max} 的值取决于两个画圆终止条件之一:(1) 当前圆弧与手部没有交叉;(2) 当前圆弧完全落在二值图像外部。

设训练样本图像集为 $T = \{t_{i,j} \mid i \in [1, C], j \in [1, N_i]\}$, 其中 C 为手势类型数, N_i 为第 i 类手势的样本总数。设 W 为使用 LDA 方法求取的投影矩阵, 对训练样本图像集中的每个样本 $t_{i,j}$, $v_{i,j}$ 和 $v'_{i,j}$ 分别为使用 W 投影变换前后的特征向量, 则本文的手势特征提取算法描述如下。

算法 1 基于圆弧扫描线的手势特征提取

输入 包含训练手型的图像序列 $T = \{t_{i,j}\}$

输出 投影矩阵 W , 投影变换前的特征矩阵维数 D 以及投影变换后的特征向量集 $V' = \{v'_{i,j}\}$

第 1 步 手部图像分割。在人手海龟模型的基础上, 根据肤色特征和手腕部的特征标记分隔出手部图像 $T' = \{t'_{i,j} \mid i \in [1, C], j \in [1, N_i]\}$ 。

第 2 步 二值化及尺寸归一化处理。对分割后的图像集 T' 中的各图像进行二值化处理, 并缩放到统一的大小, 得到图像集 $T'' = \{t''_{i,j} \mid i \in [1, C], j \in [1, N_i]\}$ 。

第 3 步 手势特征提取。使用霍夫变换进行椭圆检测识别 T'' 中各图像的掌心位置, 并以此为圆心构造一簇同心圆来提取手势特征向量。

第 4 步 维数统一化处理。因为提取的手势特征为各圆弧扫描到手部边缘的次数, 这种方法可能会造成提取的特征向量维数不一致的问题, 因此需要进行维数一致化处理。本文将每个样本提取的特征向量统一为最小维数(设最小维数为 n_1), 对于维数多的向量 P (设 P 的维数为 n_2), 将向量 P 中取自内侧圆弧的 $(n_2 - n_1 + 1)$ 个分量进行加权合并, 越靠近内侧, 权重越小。处理后得到标准化特征向量集记作 $V = \{v_{i,j}\}$, 其中各向量维数均为 $D = n_1$ 。

第 5 步 在第 4 步的基础上, 使用 LDA 方法^[10,11] 求取最优的投影矩阵 W , 并使用 W 对 $V = \{v_{i,j}\}$ 中

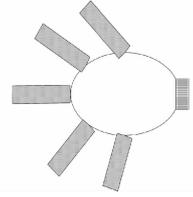
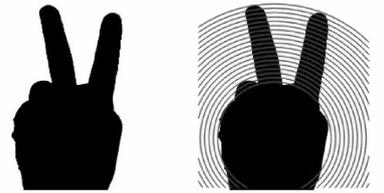


图 1 人手海龟模型

Fig. 1 Tortoise model for human hand



(a) 原始图像 (b) 同心圆
(a) Original image (b) Concentric circles

图 2 特征提取示意图

Fig. 2 Schematic diagram for extracting gesture features

的向量进行投影变换,得到对应的投影向量集 $\mathbf{V}' = \{v'_{i,j}\}$ 。

2 基于改进 KNN 的手势识别

本文采用 K 近邻(K-nearest neighbor, KNN)算法对输入的手势样本进行在线识别。KNN 算法的基本思想是:选择训练样本集中与待识别样本最接近的 K 个样本,如果这些样本中大多数属于某一类型,则待分类样本就判定为该类型。然而,传统 KNN 算法中,只利用了待分类样本中 K 个最近邻的样本信息,而没有考虑不同样本类别子集的内聚性对分类结果的影响。针对上述 KNN 算法存在的不足,本文结合小样本库手势识别的特殊性,提出了加权 KNN 算法(Weighted KNN, W-KNN)来降低手势识别的误判率。在算法 1 输出结果的基础上,本文采用的手势识别算法表述如下。

算法 2 基于 KNN 法的手势识别

输入 待识别手势图像 I_x , 投影变换前的特征矩阵维数 D , 投影矩阵 \mathbf{W} , 投影变换后的训练集特征向量集 $\mathbf{V}' = \{v'_{i,j}\}$

输出 I_x 中手势所属类别 $C_x \in [1, C]$

第 1 步 生成待识别手势的初始特征向量。采用类似算法 1 中第 1~3 步的方法,从 I_x 中将手部图像分割出来,并对其进行二值化和尺寸归一化处理,然后借助于圆弧扫描线方法得到维数为 D_x 的特征向量 \mathbf{P}_x 。

第 2 步 维数统一化处理。与算法 1 有所不同,这里的待处理向量 \mathbf{P}_x 的维数 D_x 可能比训练集的统一维数 D 大,也可能比 D 小。当 $D_x > D$ 时,采用类似于算法 1 的处理办法;当 $D_x < D$ 时,则在 \mathbf{P}_x 的前部增加 $D - D_x$ 维,并设其值为 1(相当于在手掌内部画圆,因此只有一个交点)。处理后得到的特征向量记作 \mathbf{V}_x 。

第 3 步 投影变换。使用投影矩阵 \mathbf{W} 处理 \mathbf{V}_x , 得到 LDA 方法优化后的特征向量 \mathbf{V}'_x 。

第 4 步 求取最近邻。计算 \mathbf{V}' 中各向量与 \mathbf{V}'_x 的距离并进行排序,得到 K 个距离 \mathbf{V}'_x 最近的向量,记为 $\mathbf{V}_N = \{v'_{i,j,k} \mid i \in [1, C], j \in [1, N_i], k = 1, 2, \dots, K\}$ (3)

式中: C 为手势类型数, N_i 为第 i 类手势样本总数。

第 5 步 计算权重因子。设 $v'_{i,j,p}$ 为 \mathbf{V}_N 中的任一向量,则该向量的权重因子为

$$W_p = \sigma_1 / \sigma_2 \quad (4)$$

式中: σ_1 和 σ_2 分别为将待识别手势对应的特征变量 \mathbf{P}_x 并入 \mathbf{V}_N 前后数据集的标准差。

第 6 步 加权投票。根据 \mathbf{V}_N 中 K 个最近邻的加权投票结果,确定待识别手势的类型。

3 实验结果与分析

3.1 实时手势识别系统

为了测试本文方法的有效性,本文建立了一个针对多投影显示平台的手势识别系统。该系统的结构如图 3 所示。其中手部跟踪和手型图像分割是该系统的重要组成部分,由于篇幅所限,在此不做详细的描述,系统的这方面细节以及通过状态机实现高层操作语义的内容详见文献[5]。该系统完成手势识别分为离线预处理阶段和实时识别阶段。

离线预处理阶段处理流程如下:(1) 数据输入。输入目标交互者的不同手型图像序列;(2) 手型图像分割。采用高斯背景建模法和肤色特征等检测手部图像,然后进行图像分割,并缩放到统一尺寸,建立训练集;(3) 手势特征提取。使用一组以手掌中心为圆心的同心圆来提取手势特征的训练数据,并用线性判别分析对训练数据进行预处理和压缩,在离线状态完

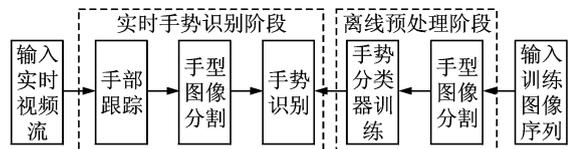


图 3 手势识别系统

Fig. 3 Hand gesture recognition system

成手势分类器的训练。

实时识别阶段处理流程如下:(1) 数据输入。输入为来自监控摄像头的实时视频流;(2) 手型图像分割。与离线预处理阶段不同的是,这个阶段的输入数据中,手部位置变动较大,因此,需要在考虑运动连贯性的基础上,采用高斯背景建模法和肤色特征等进行手部检测和分割,并缩放到统一尺寸;(3) 手势实时识别。使用改进的 W-KNN 算法,根据离线训练好的手势分类器对上一步分隔出的手势数据进行分类。

在该系统中,将手势分类器的训练等计算量较大的部分放到预处理阶段,有利于提高手势识别的实时性,而采用改进的 W-KNN 算法进行手势分类和识别,显著提高了手势识别率。

3.2 算法比较和分析

本节分别将本文提出的手势识别方案和数据手套硬件实现的手势识别方案,以及序贯相似性检测算法(Sequential similarity detection algorithm,SSDA)和文献[5]中的手势识别算法进行了比较和分析。测试用例选择了场景物体的“选中”、“平移”和“旋转”等 3 种。数据手套采用了 5DT Data Glove Ultra 系列中的右手 5 个传感器类型。摄像头采用了 HISUNG 的 IPC-EH5110PL-IR3 网络摄像头,焦距 6 mm,感光面积 1/4 英寸。程序运行环境为 HP 商用 PC 机(HP Pro 3380 MT,内存 4 GB,CPU i5-3470)。

表 1 所示为本文方案和数据手套方案的比较结果。可以看出,数据手套方案的误判率比本文方案高得多,尤其是对平移操作中一些区分度较小的手势。这主要是由数据手套中的感应器件对手部尺寸的严重依赖性和硬件传感器本身固有的灵敏度限制造成的。需要指出的是,虽然本文方案的识别速度已经能够满足实时交互要求,但是仍然和数据手套方案的识别速度不可比拟,因此这里只从误判率方面进行比较。

表 2 为本文算法与 SSDA 算法以及文献[5]算法的比较结果。从误判率来看,本文算法的误判率远低于 SSDA 算法和文献[5]算法,尤其是对平移操作中一些区分度较小的手势。这主要是由数据手套中的感应器件对手部尺寸的严重依赖性和硬件传感器本身固有的灵敏度限制造成的。另一方面,平均辨识速度^[5]方面,本文算法略低于 SSDA 算法,但是较文献[5]算法有显著提高。

表 1 本文方案与数据手套方案的比较

Tab. 1 Comparison between the proposed algorithm and data glove scheme

方案	测试用例	测试次数	误判率/%
本文方案	选中	200	0.9
	平移	200	1.4
	旋转	200	1.6
数据手套方案	选中	200	4.4
	平移	200	9.2
	旋转	200	9.7

表 2 本文算法与其他算法比较

Tab. 2 Comparison between the proposed algorithm and other algorithms

算法	测试用例	测试次数	误判率/%	平均辨识速度(次/s)
本文算法	选中	200	0.9	10.1
	平移	200	1.4	9.3
	旋转	200	1.6	10.5
SSDA	选中	200	8.1	9.3
	平移	200	11.3	8.9
	旋转	200	12.0	9.1
文献[5]算法	选中	200	3.1	16.3
	平移	200	5.7	14.7
	旋转	200	5.9	15.4

4 结束语

手势识别的识别率和实时性是评价手势识别算法优劣的关键指标,而如何在识别率和实时性之间

实现折衷是算法实用性好坏的重要因素。本文提出了一种基于小样本库的手势特征表示和实时识别方法,该方法在离线阶段基于特定的人手模型,并结合肤色特征和手腕上的特征标记完成手势图像的分割;然后借助于以掌心为中心的同心圆来提取训练数据的特征;最后,在实时交互时使用改进的加权KNN算法进行手势分类和识别。实验数据表明,本文方法在保持较高识别率的同时,显著地提高了手势识别的实时性。另外,该方法在实际应用中还有若干问题有待解决,如不同人种带来的肤色显著差异,以及在实际交互中手掌和摄像头角度问题等,都有待进一步研究。

参考文献:

- [1] 任海兵,徐光祐,林学闯. 基于特征线条的手势识别[J]. 软件学报,2002,13(5): 987-994.
Ren Haibin, Xu Guangyou, Lin Xueyan. Hand gesture recognition based on characteristic curves [J]. Journal of Software, 2002, 13(5): 987-994.
- [2] 庞海波,李占波,丁友东. 基于时间序列手势轮廓模型的动态手势识别[J]. 华南理工大学学报:自然科学版,2015,43(1): 140-146.
Pang Haibo, Li Zhanbo, Ding Youdong. Dynamic gesture recognition based on gesture contour model of time series [J]. Journal of South China University of Technology: Natural Science Edition, 2015, 43(1): 140-146.
- [3] 蔡兴泉,郭天航,臧坤等. 基于无线数据手套的手势识别方法研究[J]. 系统仿真学报,2014,26(1):72-77.
Cai Xingquan, Guo Tianhang, Zang Kun, et al. Gesture recognition method based on wireless data glove [J]. Journal of System Simulation, 2014, 26(1): 72-77.
- [4] Haritaoglu I, Harwood D, Davis L S. W4: Real-time surveillance of people and their activities [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 809-830.
- [5] 王修晖,鲍虎军. 基于自适应遗传算法的手势识别[J]. 计算机辅助设计与图形学学报, 2003, 19(8): 1056-1063.
Wang Xiuhui, Bao Hujun. Gesture recognition based on adaptive genetic algorithm [J]. Journal of Computer-Aided Design & Computer Graphics, 2003, 19(8): 1056-1063.
- [6] 杨波,宋晓娜,冯志全,等. 复杂背景下基于空间分布特征的手势识别算法[J]. 计算机辅助设计与图形学学报, 2010, 22(10): 1841-1848.
Yang Bo, Song Xiaona, Feng Zhiquan, et al. Gesture recognition in complex background based on distribution features of hand [J]. Journal of Computer-Aided Design & Computer Graphics, 2010, 22(10): 1841-1848.
- [7] 张秋余,胡建强,张墨逸. 基于区域生长的 Meanshift 动态变形手势跟踪算法[J]. 模式识别与人工智能, 2010, 23(4): 580-585.
Zhang Qiuyu, Hu Jianqiang, Zhang Moyi. Mean shift dynamic deforming hand gesture tracking algorithm based on region growth [J]. PR&AI, 2010, 23(4): 580-585.
- [8] 姚远,张林剑,乔文豹. RGB-D 图像中手部样本标记与手势识别[J]. 计算机辅助设计与图形学学报, 2013, 25(12): 1810-1817.
Yao Yuan, Zhang Linjian, Qiao Wenbao. Hand part labeling and gesture recognition from RGB-D data [J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 25(12): 1810-1817.
- [9] 张俊鹏,贺建峰. 基于 LDA 主题模型的功能性 miRNA-mRNA 调控模块识别[J]. 数据采集与处理, 2015, 30(1): 155-163.
Zhang Junpeng, He Jianfeng. Identifying of functional miRNA-mRNA regulatory modules based on LDA topic model [J]. Journal of Data Acquisition and Processing, 2015, 30(1): 155-163.
- [10] Huang Rui, Liu Qingshan, Lu Hanqing, et al. Solving the small sample size problem of LDA [C] // 16th International Conference on Pattern Recognition. Quebec, Canada: IEEE Computer Society, 2002: 29-32.
- [11] Kyperountas M, Tefas A, Pitas I. Weighted piecewise LDA for solving the small sample size problem in face verification [J]. IEEE Transactions on Neural Networks, 2007, 18(2): 506-519.

作者简介:



刘砚秋(1977-),女,讲师,研究方向:数字图像处理、人机交互, E-mail: lyq@cjlu.edu.cn。



王修晖(1978-),男,博士,副教授,研究方向:计算机图形学、计算机视觉和模式识别, E-mail: wangxiuhui@cjlu.edu.cn。

