

基于 Kinect 的动态孤立手势识别研究

田喜平¹, 赵红丹²

(1 安阳师范学院 软件学院, 河南 安阳 455000; 2 安阳师范学院 计算机与信息工程学院, 河南 安阳 455000)

摘 要: 当前方法对动态孤立手势的识别, 过程复杂、成本高, 动态手势的移动轨迹易受到外界环境的干扰、识别准确率低, 提出了基于 Kinect 的动态孤立手势识别方法, 利用 Kinect 传感器获取动态手势信息, 对人体手部进行实时、准确的定位跟踪, 并对手部图像进行平滑去噪处理, 提取动态手势轨迹的特征; 引入隐马尔可夫模型(HMM)对动态孤立手势及手部运动轨迹的样本集进行有效训练, 最终实现动态孤立手势的精确识别. 实验证明提出的方法在噪声干扰和光线缺失的环境下, 对动态孤立手势仍具有较高的识别率, 鲁棒性强.

关键词: Kinect; 手势识别; HMM

中图分类号: TP391

文献标识码: A

文章编号: 1000—7180(2017)02—0115—04

Research on Dynamic Isolated Gesture Recognition Based on Kinect

TIAN Xi-ping¹, ZHAO Hong-dan²

(1 School of Software Engineering, Anyang Normal University, Anyang 455000, China;

2 School of Computer and Information Engineering, Anyang Normal University, 455000, China)

Abstract: Traditional methods for the extraction and identification of the data of dynamic isolated gestures, which constitute a complex and high cost, which is easy to be disturbed by the external environment, and the recognition accuracy is low. The study of dynamic hand gesture recognition method based on Kinect, obtaining the depth information of dynamic gesture using Kinect sensor, the human hand positioning real-time and accurate tracking, feature and image denoising processing, the extraction of hand movement; finally introduce the hidden Markov model (HMM) for effective training on dynamic gesture and moving track sample set, to achieve dynamic hand gesture recognition. Experiments show that the proposed method has high recognition rate and high robustness in the environment of noise interference and light loss.

Key words: Kinect; gesture recognition; HMM

1 引言

近年来,随着计算机图像识别技术^[1-2]的快速发展,手势识别已成为一种重要的人机交互手段^[3-4]. 动态手势识别通过计算机系统对人体手势做出精准的判断和解释,相对于传统的输入方式,具有自然性和直观性的优点^[5-6]. 动态手势轨迹信息具有很强的扩展性和移植性,符合人类的习惯,在手语理解、医学研究^[7-8]、图像拼接、视觉导航、三维设计等诸多领域有着广泛的用途^[9-10]. 动态手势交互手段是未来

人机交互的发展趋势,具有很高的研究价值和推广意义. 传统的动态手势识别方法,手势信息的提取多基于数据手套传感器^[11-12],采用 BP 神经网络等模糊^[13-14]分析方法对手势移动轨迹和数据信息进行分析处理,容易受到摄像头、背景、光照及外界噪声的干扰,对动态手势的识别率低、精确度差、计算复杂、成本高.

本文提出一种基于 Kinect 的动态孤立手势识别方法,通过引入隐马尔可夫模型(HMM)对动态孤立手势及手部运动轨迹的样本集进行有效的训

练,最终实现动态孤立手势的精确识别。

2 一种基于 Kinect 的动态孤立手势识别方法研究

2.1 基于 Kinect 传感器的动态手势深度信息获取

动态孤立手势包含着丰富的信息,准确、清晰地获取这些深度信息是对孤立手势进行有效识别的关键步骤。Kinect 传感器作为一种 3D 体感摄像设备^[15-16],可以同时获取人体手势深度图像数据信息和 RGB 信息,并能够跟踪和识别一系列动态孤立的手部姿态。

获取动态手势的深度信息,需要将人体手部从复杂的背景环境中抽离出来。对动态手势的检测,首先要通过辨识预先定义好的手部姿态来实现,在实验过程中保持正确的手部姿态、距离和运动频率,直到 Kinect 检测到运动中的手部。将手部姿态测试的阈值设定为 2 mm,保证调试的精度,避免设备的微抖动。滤除无效的手势帧序列,可以去除冗余的运算量,提高系统的识别率。

2.2 动态手势信息的预处理及手势轨迹的特征提取

基于 Kinect 传感器获取的动态手势深度信息,通常包含较大的噪声干扰,影响图像的清晰度及后续对动态孤立手势的识别,因此需要对动态手势信息进行预处理,即对包含深度信息的图像进行滤波操作。

设动态手部信息的图像为 $v = (v(i), i \in A)$, 其中 A 为像素的集合,令 $i, j \in A$, 其中 i 为图像中心像素点, j 为图像的其他像素点,那么该算法的降噪灰度值公式为

$$T(i) = \sum_{j=1}^i w(i, j) v(j) \quad (1)$$

式中, $w(i, j)$ 为像素点的权值,在对动态手部图像平滑区域降噪中,引入归一化变量 z 和滤波参数 h ,能够提高图像平滑区域的降噪效果,并且保证图像的细节特征不丢失,其计算公式为

$$z = \sum \exp\left(-\frac{\|v(i) - v(j)\|_2^2}{h^2}\right) \quad (2)$$

$$\|v(i) - v(j)\|_2^2 = \sum_{i=1}^n |v(x+i) - v(y+i)|^2 \quad (3)$$

$$w(i, j) = \frac{1}{z} \exp\left(-\frac{\|v(i) - v(j)\|_2^2}{h^2}\right) \quad (4)$$

通过式(1)~(4)的求解,调整中心像素点和其他区域像素的灰度关系及像素点的权值,能够实现

对动态手部图像的降噪处理。

图像平滑去噪后对动态孤立手势实行分割处理,是手势辨别里最关键的根本步骤,总用的手势分割法为基于手部轮廓的分割,方便易行但辨别率很低,在复杂背景干扰的情况下,有失效的风险。本文采用基于阈值的分割方法来提取手部的运动轨迹,分割结果好,在一定水平上,还能够达成在不同光照及繁杂状况下对手部特征的急速获取。

动态孤立手势不同于静态手势,动态手势既涉及到时间上的变化,又涉及到空间上的变化,手势轨迹包含位置、速度和角度 3 个要素。手势轨迹曲线上点坐标位置,能够作为手势轨迹辨别的一项关键特征,但就算让一样的手势轨迹也会有不一样,坐标点会产生改变,所以需要让手势轨迹在不一样的时刻切线角度发生转变,来当作手部运动轨迹的特征。

假设 Kinect 传感器得到第 t 时刻及第 $t+1$ 时刻的手部中心位置的坐标分别为 $P_1(x_t, y_t)$ 和 $P_2(x_{t+1}, y_{t+1})$,如图 1 所示。

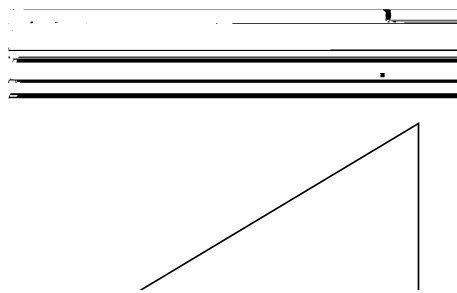


图 1 手势轨迹角度示意图

设手势移动轨迹的角度为 θ :

$$\theta = \arctan\left(\frac{y_{t+1} - y_t}{x_{t+1} - x_t}\right) \quad (5)$$

根据 θ 选取的不同角度,对不同的链码分别进行离散化处理,完成手部移动轨迹特征的提取,并形成大量的训练样本。

2.3 基于 HMM 的动态孤立手势识别研究的实现

引入 HMM 方法对动态手势移动轨迹样本集进行有效地训练,并实现动态孤立手势的有效识别。

HMM 是一种统计分析模型,其系统模型如图 2 所示,在动态孤立手势识别系统中,设 HMM 为一个五元组 $\{Q, V, A, B, \pi\}$, 其中 $Q = \{q_1, q_2, \dots, q_n\}$ 为状态的有限集合, n 为状态的个数,第 i 时刻的状态为 q_i . $V = \{v_1, v_2, \dots, v_m\}$ 为状态集的观测集合,观测值的个数为 m ,状态分布矩阵 $A = \{a_{ij}\}$ 可以表示为

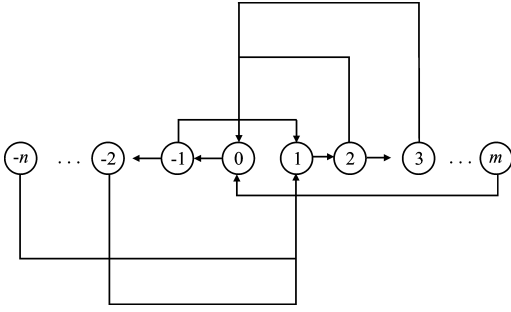


图2 HMM 系统模型图

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} \\ a_{21} & a_{22} & \cdots & a_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} \end{bmatrix} \quad (6)$$

$$a_{ij} = C(x_{t+1} = q_j | x_t = q_i), i \geq 1, j \leq n \quad (7)$$

$B = \{b_j\}$ 为 j 状态下的观测概率分布:

$$b_j = C(x_t = v_k | x_t = q_j), 1 \leq j \leq n \quad (8)$$

状态集的初始分布 $\pi = \{\pi_i\}$, 其中,

$$\pi_i = C(x_i = q_i), 1 \leq i \leq n \quad (9)$$

设观测序列值 $K = \{k_1, k_2, \dots, k_i\}$, $\lambda = \{A, B, \pi\}$, 基于 HMM 的样本训练过程如下.

对动态手势移动轨迹样本集进行有效地训练, 可以通过对 HMM 参数的优化和更新来完成, 调节参数为 $\lambda = \{A, B, \pi\}$, 对以下公式进行概率最大化处理:

$$C(K|\lambda) = \prod_{i=1}^k C(k_i|\lambda) = \prod_{i=1}^k C_i \quad (10)$$

定义一个变量 $\xi(i, j)$, 与序列 K 相对应, 假设在 $t-1$ 时刻的状态为 S_i , 在 t 时刻的状态为 S_j 的概率为

$$\begin{aligned} \xi(i, j) &= \frac{C\lambda(q_{t-1} = S_i, q_t = S_j, k_1, k_2, \dots, k_i)}{K(k_1, k_2, \dots, k_i)} \\ &= \frac{\lambda\alpha_{t-1}(i)a_{ij}b_j(k_t)\beta_t(j)}{\sum_i \sum_j \lambda\alpha_{t-1}(i)a_{ij}b_j(k_t)\beta_t(j)} \end{aligned} \quad (11)$$

再定义一个变量 $\kappa(i)$, 在 t 时刻的状态为 S_i 的概率为

$$\begin{aligned} \kappa(i) &= \frac{C\lambda(q_t = S_i, k_1, k_2, \dots, k_i)}{K(k_1, k_2, \dots, k_i)} \\ &= \frac{\lambda\alpha_t(i)\beta_t(i)}{\sum_i \alpha_t(i)\beta_t(i)} \end{aligned} \quad (12)$$

通过对参数变量在不同时刻状态的概率的统计, 来完成手部姿态轨迹的精确计算和动态孤立手势的有效识别.

本文提出基于 Kinect 传感器的动态孤立手势

识别方法研究, 采用 Kinect 传感器得到人体动态手势的深度讯息, 对人体手部实行定位追踪, 并对图像实行平滑去噪治理, 得到人体手部运动轨迹的特征. 引入隐马尔可夫模型(HMM)对动态手势及其移动轨迹样本集进行有效地训练, 能够有效地完成对动态孤立手势的识别.

3 实验结果与分析

为验证本文提出的基于 Kinect 的动态孤立手势识别方法研究的有效性, 本文针对 Kinect 提取动态孤立手势姿态图像效果、去噪表现及在不同光线和背景的条件下方方法的鲁棒性进行了仿真试验.

本文方法提取动态孤立手势图像后, 对图像进行了平滑去噪处理, 图像的清晰度更高.

本文定义了 4 种基本手势, 前进、转弯、停止、点击, 动态手势特征在考虑手形特征之外, 还要对手势在时空序列的转变进行考虑.

在 HMM 建模进程里, 每条有效轨迹对应一个 HMM, 训练模型状态数越多越好, 所以跟着情况数的提高, 训练的模型越能描述轨迹. 各提取以上 4 种不同手势下的 200 个测试数据样本.

为了验证本文方法对比传统基于 BP 神经网络方法具有更强的鲁棒性, 分别在正常光线情况下、黑暗光线情况下及复杂背景情况, 对比两种方法的动态孤立手势识别率, 从实验结果可以看出, 在正常的光线条件下传统方法和本文方法都具有较高的识别率, 但在光线条件差和复杂背景的条件下, 传统的 BP 神经网络方法的动态手势识别率较差, 并且不稳定; 而本文方法在黑暗环境下及复杂的背景下, 仍能保持较高的动态孤立手势识别率.

综合考虑以上 3 种不同的情况, 将两种动态孤立手势识别方法的综合识别率拟合合成曲线, 如图 3 所示, 综合识别率曲线表明本文基于 Kinect 的动态孤立手势识别方法具有更高的识别率, 表现更稳定.

仿真实验证明了本文提出的基于 Kinect 的动态孤立手势识别方法的有效性, 在不同光线和复杂背景条件下仍具有较强的鲁棒性.

4 结束语

本文提出了一种基于 Kinect 的动态孤立手势识别方法研究, 利用 Kinect 传感器获取动态手势信息, 对人体手部进行实时、准确的定位跟踪和图像处理, 提取动态手势轨迹的特征, 基于隐马尔可夫模型(HMM)对动态孤立手势及手部运动轨迹的样本集

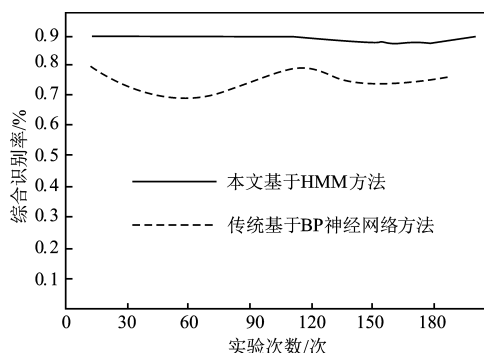


图3 两种方法综合识别率曲线对比

进行有效地训练,最终实现动态孤立手势的精确识别,实验证明了本文方法的有效性。

参考文献:

- [1] 温玉春. 计算机图像处理技术应用研究[J]. 现代商贸工业, 2011, 23(2): 230-231.
- [2] 孙宗保, 蔡健荣, 姜松. 数学形态学在计算机图像处理技术检测粒度中的应用[J]. 计算机测量与控制, 2003, 11(9): 665-667.
- [3] 孙正兴, 冯桂焕, 周若鸿. 基于草图的人机交互技术研究进展[J]. 计算机辅助设计与图形学学报, 2005, 17(9): 1889-1899.
- [4] 陈一民, 张云华. 基于手势识别的机器人人机交互技术研究[J]. 机器人, 2009, 31(4): 351-356.
- [5] 包加桐, 宋爱国, 郭晏, 等. 基于 SURF 特征跟踪的动态手势识别算法[J]. 机器人, 2011, 33(4): 482-489.

- [6] 孔晓明, 陈一民, 陈养彬, 等. 基于视觉的动态手势识别[J]. 计算机工程与设计, 2005, 26(11): 2934-2936.
- [7] 王兆丽, 王力. 一个简单高效的动态手势识别方法[J]. 计算机工程与应用, 2002, 38(19): 73-75.
- [8] 刘江华, 程君实, 陈佳品. 基于光流的动态手势识别[J]. 计算机工程, 2002, 28(4): 104-105.
- [9] 郑轶, 沈旭昆. 基于连续数据流的动态手势识别算法[J]. 北京航空航天大学学报, 2012, 38(2): 273-279.
- [10] 于美娟, 马希荣. 基于 HMM 方法的动态手势识别技术的改进[J]. 计算机科学, 2011, 38(1): 251-252.
- [11] 刘江省, 姚英学, 李建广, 等. 虚拟装配中基于数据手套的虚拟操作研究[J]. 系统仿真学报, 2004, 16(8): 1744-1747.
- [12] 王家顺, 王田苗, 魏军, 等. 一种面向遥操作的新型数据手套研制[J]. 机器人, 2000, 22(3): 201-206.
- [13] 江立, 阮秋琦. 基于神经网络的手势识别技术研究[J]. 北京交通大学学报, 2006, 30(5): 32-36.
- [14] 曾芬芳, 王建华. 基于神经网络的手势识别[J]. 机器人, 1999, 20(1): 40-45.
- [15] 淦创. 基于 3D 体感技术的动态手势识别[J]. 光电技术应用, 2012, 227(4): 55-58.
- [16] 陈义, 马云林. 基于视觉的手势识别技术在车载主机上的应用[J]. 电子设计工程, 2016, 24(8): 141-144.

作者简介:

田喜平 女, (1980-), 硕士研究生, 讲师. 研究方向为智能信息处理、Agent 与电子商务。

赵红丹(通讯作者) 女, (1982-), 硕士研究生, 讲师. 研究方向为自然语言处理. E-mail: zhd. cn@126. com.

(上接第 114 页)

- [3] Seyed Saeid Mirkamali, Nagabhushan P. Object removal by depth-wise image inpainting[J]. Signal, Image and Video Processing, 2015, 8(9): 1785-1794.
- [4] 李俊柏, 黄文清, 黄牧. 改进的 Criminisi 图像修复算法[J]. 工业控制计算机, 2014, 7(27): 127-131.
- [5] Anamandra Sai Harees. Exemplar-based color image inpainting: A fractional gradient function approach[J]. Pattern Analysis and Applications, 2014, 2(17): 389-399.
- [6] Muhammad S, Irfan M, Baik Sung Wook. Sparse coded image super-resolution using K-SVD trained dictionary based on regularized orthogonal matching pursuit[J]. Bio-Medical Materials and Engineering, 2015, 1(26): 1399-1407.
- [7] Kazi J N, Patil Y M. Comparative analysis of fragment based and exemplar based inpainting techniques[J]. International Journal of Advanced Research in E-

lectrical, 2013, 3(2): 1020-1026.

- [8] 曾令军. 基于压缩感知的射电天文图像重构[D]. 昆明: 昆明理工大学, 2014: 23-42.
- [9] Many V Afonso, João M R Sanches. A total variation recursive space-variant filter for image denoising[J]. Elsevier Digital Signal Processing, 2015, 11(40): 101-116.
- [10] Papafitsoros K, Schönlieb C B. A combined first and second order variational approach for image Reconstruction[J]. Journal of Mathematical Imaging and Vision, 2014, 2(48): 308-338.
- [11] Berntsson F, Baravdish G. Coefficient Identification in PDEs applied to image inpainting[J]. Applied Mathematics and Computation, 2014, 242(11): 227-235.

作者简介:

徐俊 男, (1981-), 硕士, 工程师. 研究方向为图像处理、计算机技术. E-mail: bxujunx1981jd@sina. com.