

基于 Leap Motion 的动态手势识别方法^①



高宇¹, 何小海¹, 吴晓红¹, 王正勇¹, 张豫堃²

¹(四川大学 电子信息学院 图像信息研究所, 成都 610065)

²(成都西图科技有限公司, 成都 610065)

通讯作者: 何小海, E-mail: hxx@scu.edu.cn

摘要: 伴随虚拟现实 (Virtual Reality, VR) 技术的发展, 以及人们对人机交互性能和体验感的要求提高, 手势识别作为影响虚拟现实中交互操作的重要技术之一, 其精确度急需提升^[1]. 针对当前手势识别方法在一些动作类似的手势识别中表现欠佳的问题, 提出了一种多特征动态手势识别方法. 该方法首先使用体感控制器 Leap Motion 追踪动态手势获取数据, 然后在特征提取过程中增加对位移向量角度和拐点判定计数的提取, 接着进行动态手势隐马尔科夫模型 (Hidden Markov Model, HMM) 的训练, 最后根据待测手势与模型的匹配率进行识别. 从实验结果中得出, 该多特征识别方法能够提升相似手势的识别率.

关键词: 手势识别; 虚拟现实; Leap Motion; 特征提取; 隐马尔科夫模型

引用格式: 高宇, 何小海, 吴晓红, 王正勇, 张豫堃. 基于 Leap Motion 的动态手势识别方法. 计算机系统应用, 2019, 28(11): 208-212. <http://www.c-s-a.org.cn/1003-3254/7172.html>

Dynamic Gesture Recognition Method Based on Leap Motion

GAO Yu¹, HE Xiao-Hai¹, WU Xiao-Hong¹, WANG Zheng-Yong¹, ZHANG Yu-Kun²

¹(Image Information Institute, College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

²(Chengdu Xitu Technology Co. Ltd., Chengdu 610065, China)

Abstract: With the development of Virtual Reality (VR) technology and the increasing demand for human-computer interaction performance and experience, gesture recognition is one of the important technologies affecting the interaction in VR, and its accuracy needs to be improved. Aiming at the problem that the current gesture recognition method performs poorly in some similar gesture recognition, a multi-feature dynamic gesture recognition method is proposed. Firstly, this method uses Leap Motion to track the dynamic gestures to acquire data, then adds the displacement vector angle and the inflection point judgment in the feature extraction process, after that performs the training of the dynamic gesture Hidden Markov Model (HMM). Finally, the recognition is carried out according to the matching ratio of the gesture to be tested and the model. It is concluded from the experimental results that the multi-feature recognition method can improve the recognition rate of similar gestures.

Key words: Virtual Reality (VR); gesture recognition; Leap Motion; feature extraction; Hidden Markov Model (HMM)

随着计算机、虚拟现实等各方面技术的显著进步, 人机交互技术在虚拟现实中扮演的角色地位逐渐提升^[1]. 人们不仅对人机交互的实时性、识别率等标准的要求愈发严苛, 而且提出人机交互方式更简单、准确

率更高和参与性更强的发展需求^[2]. 手势作为一种变化方式和表达含义都很丰富的肢体语言, 自然成为人机交互的一种重点媒介.

手势识别有动态手势识别和静态手势识别, 相较

① 基金项目: 四川省科技计划项目 (2018HH0143); 四川省教育厅项目 (18ZB0355)

Foundation item: Science and Technology Program of Sichuan Province (2018HH0143); Program of Education Bureau, Sichuan Province (18ZB0355)

收稿时间: 2019-04-25; 修改时间: 2019-05-21; 采用时间: 2019-05-27; csa 在线出版时间: 2019-11-06

于静态手势识别而言,动态手势识别特点在于跟踪目标伴随时间和空间的改变而产生形态和运动轨迹的变化.动态手势识别一般分为特征数据采集、手势模型训练和手势分类识别3个步骤^[3].其中,手势模型的训练和建立是至关重要的一步.选择不同手势之间区分度较大的特征参与手势模型的训练,才能准确建立出相似手势的不同手势模型.张琪祥利用掌心位移、掌心法向量、掌心球半径和五指伸展程度作为特征训练HMM手势模型^[4],但是一些相似手势的识别率差强人意;Lu W等人选取相邻指尖距离、指尖梯度和五指分开度作为特征进行手势模型训练^[5],然而对相似单指动态手势的识别不够理想.针对以上识别率不足的问题,提出在特征数据的提取中不仅提取掌心位移、掌心法向量、掌心球半径和手指伸展程度,而且加入掌心位移向量角度的提取,并且在进行HMM手势模型训练的过程中将拐点判定计数也作为手势的特征参与训练,建立更准确的手势模型,提升手势识别的准确率.

1 Leap Motion 简介

动态手势识别中手部会根据时间和空间的改变而产生手势和运动轨迹的变化.提取用于手势模型训练的特征量需要能以高精度追踪到掌心、关节、指尖等的位置信息和手指手掌的方向向量及法向量信息,得到针对手部特征信息的描述.Leap Motion作为一种新型传感器,能够检测出手部更细微的动作变化、提取出更详细的手部信息,动态手势在Leap Motion的跟踪下具有清楚的运动特征.

如图1所示,Leap Motion主要由2个摄像头和3个红外LED构成,鉴于使用红外检测目标手部的特点,不易受外界条件的干扰,能够采集到超低噪声的帧图像便于后续的数据处理^[6].Leap Motion对手部运动的追踪频率可达到每秒200帧,能够实现精度高达0.01 mm的追踪,可追踪空间呈一个塔尖在设备中心的倒四棱锥型体,范围大约在设备前方25~600 mm,并且拥有150°的圆形空间视场.Leap Motion将能够采集到的数据封装在每一帧当中,每帧数据均有一个跟踪数据列表,包含目标手及其组成部分等,也包含识别出的手势和发生在跟踪范围内运动因素的描述^[7].本文从帧中提取深度信息,然后利用预处理后的特征数据进行手势模型的训练和未知手势的识别.



图1 Leap Motion 结构示意图

2 动态手势特征的提取和预处理

Leap Motion跟踪目标手得到的帧数据中包含PLAM、HAND等对象,其中PLAM对象中包含掌心法向量、手指方向向量和掌心坐标等数据,HAND对象中包含手指伸展程度等数据.经典手势识别方法中通过从帧数据中提取或计算的特征有:掌心位移、掌心法向量、掌心球半径、手势速度、手指伸展程度、手指分开度和运动方向变化率等.由于一些动作类似手势的以上特征也非常相似,所以导致识别率差强人意.经过对相似手势分析和调研,手势位移向量角度和拐点的判定计数在相似手势中区分度较大,故本文动态手势识别方法中添加提取以上两种特征.其中掌心球半径、掌心法向量和手指伸展程度均可从Leap Motion采集到的帧数据中直接得出.而剩余的掌心位移、手势位移向量角度、手势速度、拐点判定计数和运动方向变化率则需要通过一定的计算处理之后才能得出.

2.1 经典特征提取

动态手势识别方法中提取的经典手势特征包含掌心位移、手势速度和运动方向变化率等.

(1) 掌心位移数据的获取

从一个时刻起,到另一个时刻止,两个时刻追踪目标手的掌心坐标之间的欧氏距离即为掌心位移.计算过程如下:

$$P = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2 + (z_i - z_c)^2} \quad (1)$$

其中, P 为掌心位移, (x_i, y_i, z_i) 为掌心在 i 时刻的空间坐标, (x_c, y_c, z_c) 为掌心在 c 时刻的空间坐标.

(2) 手势速度的获取

速度特征能够体现出手势运动的快慢,且在本文中速度特征还作为判定手势运动开始、结束以及运动拐点的标准之一.因为Leap Motion的追踪频率能够达

到 200 帧/秒, 所以本文中把每两帧的帧间速率近似为瞬时速率, 手势在第*i*帧的速率计算如下:

$$V_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2} / T \quad (2)$$

其中, *T* 为帧间时间差. 当手势速度大于阈值 σ 时满足开始采集手势的速度条件, 小于阈值 σ 时则满足停止手势采集的速度条件.

(3) 运动方向变化率的获取

获取运动方向变化率是为能够区分出由于拐点造成的速度低于阈值和手势停止时速度低于阈值这两种情况, 避免发生误判. 运动方向变化率需要从帧数据中提取出掌心的空间坐标 (x_t, y_t, z_t) , 根据如下公式:

$$\theta_{xy} = \left\{ \begin{array}{l} \arctan\left(\frac{x_t - x_{t-1}}{y_t - y_{t-1}}\right) + \pi, (y_t - y_{t-1} < 0) \\ \arctan\left(\frac{x_t - x_{t-1}}{y_t - y_{t-1}}\right) + 2\pi, (x_t - x_{t-1} < 0) \\ \arctan\left(\frac{x_t - x_{t-1}}{y_t - y_{t-1}}\right), (other) \end{array} \right. \quad (3)$$

可求出 *x* 在单位时间 *t* 内于 *y* 轴上的手势运动方向 θ_{xy} , 用同等方法也可求出 θ_{xz} 和 θ_{zy} .

2.2 新特征提取

本文方法中, 为提升动作类似的手势识别率, 加入新特征拐点的判定计数和手势位移向量角度的提取, 提升相似手势的区分度, 建立更准确的手势模型.

(1) 拐点的判定及计数的获取

由于手势位于拐点处附近时速度是先降低后升高, 所以得出当瞬时速率为 0 的某一时刻 T_i , 通过判断该时间段内手势序列中前后两端速率满足:

$$\begin{cases} V_{T_i} < V_{T_{i+1}} \cdot \tau \\ V_{T_i} < V_{T_{i-1}} \cdot \tau \end{cases} \quad (4)$$

式中, τ 取 0.5, 判定该点为手势运动的拐点位置.

单指手势的传统特征掌心位移、掌心法向量等由于手势的随机性容易发生局部相似或者整体相似的情况.

但是如图 2 所示, 它们的拐点存在情况和数目有明确的差别. 单指画叉和单指画三角形虽然有可能拐点数目相同, 但是结合掌心位移等特征也能明确的区分. 故在判定出一个拐点后并对拐点计数作加 1 操作, 以此为特征提升相似手势的区分度.

(2) 手势位移向量角度的获取

通过获取掌心在手势运动的开始位置和停止位置的空间坐标 (x_o, y_o, z_o) 和 (x_s, y_s, z_s) , 则手势位移向量 \vec{k} 计算如下:

$$\vec{k} = (x_s - x_o, y_s - y_o, z_s - z_o) \quad (5)$$

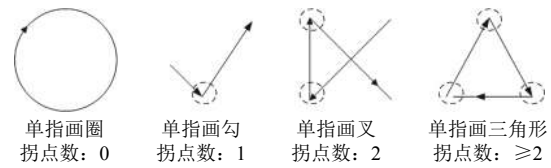


图 2 单指手势指尖轨迹示意图

定义手势位移向量角度为 \vec{k} 与 *y* 轴负方向上的单位向量 \vec{n} 之间的夹角 θ_s , θ_s 计算过程如下:

$$\begin{aligned} \theta_s &= \arccos\left(\frac{\vec{k} \cdot \vec{n}}{|\vec{k}| |\vec{n}|}\right) \\ &= \frac{(y_s - y_o) \times (-1)}{\sqrt{(x_s - x_o)^2 + (y_s - y_o)^2 + (z_s - z_o)^2} + \sqrt{(-1)^2}} \end{aligned} \quad (6)$$

如图 3 中所示, 向上划和向后划、向下划和向前划分别为两对相似手势, 每对相似手势由于手部运动的不规则性均有可能出现掌心位移、掌心球半径、五指伸展程度、五指分开度和经典特征等相似的情况, 造成手势识别混乱的结果. 但是手势位移向量角度在这些相似手势中有明显的差别, 经过分析实验得出结果如表 1 所示. 故添加此特征的提取有助于提升手势识别率.

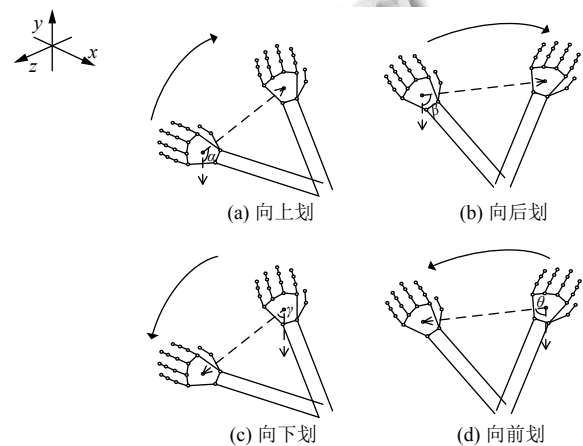


图 3 手势位移向量角度

3 基于 HMM 模型的动态手势识别

3.1 基于 HMM 模型的动态手势识别的原理

基于 HMM 的动态手势识别流程如图 4 所示, 首先用 Leap Motion 采集训练数据集并从中提取出需要的特征数据, 然后对提取数据进行预处理, 再将处理好

的数据用于训练动态手势 HMM 模型, 模型训练完成后即可对未知手势进行识别。

隐马尔科夫模型 (HMM) 是被 Baum LE 等人在 20 世纪 60 年代提出的一种具有马尔科夫性质的统计模型。一个完整的 HMM 由 (π, A, B, S, V) 五种参数构成, 其中 π 表示隐含状态的初始分布概率, A 是隐含状态转移矩阵, B 则代表观测状态发生的概率矩阵, S 是隐含状态集合, V 是观测状态集合。其中 B, V 组成的观

测状态随机输出过程和 π, A, S 组成的隐含状态随机转移过程构成 HMM 的双重随机性^[8]。

表 1 不同手势位移向量角度的范围

手势	手势位移向量角度范围
向上划	[155°, 180°]
向后划	[80°, 120°]
向下划	[0°, 30°]
向前划	[70°, 90°]

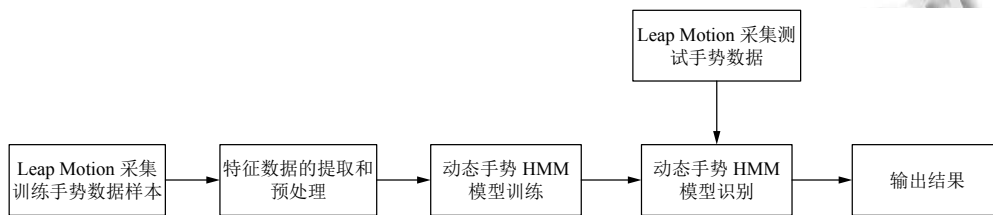


图 4 动态手势识别流程图

HMM 模型的应用中主要需解决的 3 类问题有:

(1) 评价问题: 在已知参数 $\lambda = (\pi, A, B, S, V)$ 和观测状态序列 O 的条件下, 计算出全部模型的输出概率 $P(O|\lambda)$ 。

(2) 解码问题: 在已知参数 $\lambda = (\pi, A, B, S, V)$ 和观测状态序列 O 的条件下, 求出与观测序列 O 匹配度最高的隐藏状态序列 Q 。

(3) 学习问题: 观测状态序列 O 已知的条件下, 通过反复训练和迭代得出使 $P(O|\lambda)$ 最大的模型参数 λ 。

HMM 模型在时序信号的处理上表现突出, 已经广泛被运用于语音识别等范畴中^[9]。由于语音序列和动态手势序列类似程度颇高, 因此动态手势识别领域也愈加频繁的使用 HMM 模型参与研究。基于 HMM 的动态手势识别中需解决的问题主要是学习问题和评价问题, 动态手势识别中通过训练而建立每一个手势模型的过程即为解决 HMM 中的学习问题的过程, 评价问题的解决即通过得到的手势观测状态序列而判断出该手势属于哪一个模型。

本文中解决 HMM 学习问题时选择 Baum-Welch 算法^[10]来求解最优模型参数 λ , λ 的求取过程即动态手势模型的训练过程。在得到观测序列即得到特征向量时, 通过反复迭代计算并更新模型参数 $\bar{\lambda} = (\bar{\pi}, \bar{A}, \bar{B})$, 直到满足式 (7) 为止得到最优模型参数, 其中 $\bar{\lambda}_a$ 和 $\bar{\lambda}_b$ 为相邻两次迭代的结果。

$$|\log P(O|\bar{\lambda}_a) - \log P(O|\bar{\lambda}_b)| < \varepsilon \quad (7)$$

本文中解决 HMM 评价问题选择前向算法^[11], 评价问题的处理过程和结果即为对未知动态手势的识别过程和手势识别结果。在经过手势模型的训练阶段, 已经得出每个手势模型的最优模型参数 λ 。前向算法从未知手势观测序列 O 中提取特征量, 然后计算出特征向量与每个 HMM 的匹配概率, 选择匹配概率 $P(O|\lambda)$ 最大的手势模型作为识别结果。

3.2 结合新特征的实验分析

为检验出加入新特征后的手势识别方法的效果, 实验设计单指画圈、单指画叉、单指画三角形三种单指手势运动, 此类手势为食指向前舒展, 其余手指握紧; 向左划、向右划、向前划、向后划、向上划、向下划六种手势运动选自文献[4], 此类手势运动均为五指伸展, 且运动时掌心方向与运动方向相同进行划动。文献[4]方法对每种手势采集的样本数据包括掌心位移、掌心球半径、掌心法向量、五指伸展程度; 文献[5]采集的特征数据包括相邻指尖距离、指尖梯度和五指分开度; 而加入新特征后的方法中的样本数据包括掌心位移、掌心球半径、掌心法向量、拐点的判定及计数、位移向量角度和运动方向变化率。实验分别对每种手势采集特征数据样本, 通过加载同一手势下特征数据样本训练得到 9 种 HMM 手势模型, 最后通过各个模型对未知样本的评价来判断对每个模型的匹配率, 选择与待测样本具有最高匹配率的模型作为识别结果。

实验数据的录制采样由 4 人完成, 每个人对每种

手势进行40组采样. 每组手势的160组样本数据按照3:1比例随机分成120组数据作为HMM模型的训练数据和40组数据作为HMM模型的测试数据. 以上述方式进行5次交叉实验, 取识别率的平均值作为最终结果. 实验主要对比本文方法和文献[4]、文献[5]的方法, 得出实验结果如表2所示.

表2 HMM手势模型识别结果

设计手势	文献[4]方法 识别率(%)	文献[5]方法 识别率(%)	本文方法 识别率(%)
单指画圈	95.02	93.21	94.12
单指画叉	81.87	84.26	89.85
单指画三角形	84.63	85.83	89.85
向上划	93.73	93.22	92.99
向后划	87.11	88.41	91.24
向下划	94.21	93.17	93.96
向前划	87.66	87.43	92.67
向左划	93.91	92.78	93.58
向右划	89.21	90.39	90.18
平均识别率	89.48	89.85	92.30
单个模型平均 训练耗时(s)	0.69	0.71	0.79

由表2可以得出, 对于参与实验的9组手势, 文献[4]方法中近似的手势之中有一组识别率较高而剩余的一组或两组识别率较低. 如单指画圈、单指画叉和单指画三角形为近似手势, 单指画圈的识别率较高而其余两组的识别率略显不足. 分析可得就文献[4]的方法而言, 单指画叉和单指画三角形本身就与单指画圈整体类似或者局部类似, 导致一些单指画叉和单指画三角形的手势被误判成单指画圈的手势. 文献[5]方法在单指近似手势中表现略比文献[4]的方法更好, 但是其余的相似手势识别率依旧不够理想. 分析可得文献[5]主要提取的特征量基本代表手指之间的相对特征, 而后6组实验手势手掌姿态均保持高度相似所以导致此类相似手势识别率较低. 在本文方法中加入拐点的判定计数使单指相似手势能够区分开, 如单指画圈就没有拐点, 而单指画叉和单指画三角形的拐点均在2个及以上. 同一类型中相似手势的位移向量角度均有较大的差异, 所以加入位移向量角度作为特征量, 突出相似手势的不同点, 使识别率提升. 从实验结果中能够总结出: 由于录制实验数据时极少的误差对于本文方法提出的新特征判定的影响大于对文献[4]和文献[5]的影响, 故一些手势的识别率略低于文献[4]、文献[5]的方法, 但差异范围低于1%且仍旧表现出较高的识别

率. 而对于文献[4]、文献[5]方法中识别率较低的手势, 本文方法识别率均有3%-8%的提升. 总体而言, 平均识别率由89.48%和89.85%提升至92.30%, 且本文方法的一组数据平均训练时间为0.79s, 相较于文献[4]和文献[5]的0.69s和0.71s并未增加过多, 即本文方法在保障工作效率的前提下使识别率提升, 可见本文方法的有效性.

4 结束语

Leap Motion的出现使HMM模型在基于时序序列的模式识别领域内具有更好的应用前景, 本文讨论了基于HMM模型的动态手势识别的研究现状, 根据对现存且具有代表性的基于HMM的动态手势识别方法进行优劣分析, 通过在模型训练的过程中添加拐点的判定计数和位移向量角度这两个特征, 使相似手势的区分识别率提升, 进而提升手势识别的整体识别率.

参考文献

- Sagayam KM, Hemanth DJ. Hand posture and gesture recognition techniques for virtual reality applications: A survey. *Virtual Reality*, 2017, 21(2): 91-107. [doi: 10.1007/s10055-016-0301-0]
- 张凤军, 戴国忠, 彭晓兰. 虚拟现实的人机交互综述. *中国科学: 信息科学*, 2016, 46(12): 1711-1736.
- 田喜平, 赵红丹. 基于Kinect的动态孤立手势识别研究. *微电子学与计算机*, 2017, 34(2): 115-118.
- 张琪祥. 基于Leap Motion的手势识别研究及应用[硕士学位论文]. 北京: 北京邮电大学, 2018.
- Lu W, Tong Z, Chu JH. Dynamic hand gesture recognition with Leap Motion controller. *IEEE Signal Processing Letters*, 2016, 23(9): 1188-1192. [doi: 10.1109/LSP.2016.2590470]
- 黎珍. 基于Leap Motion交互技术的研究. *通讯世界*, 2017, (7): 297. [doi: 10.3969/j.issn.1006-4222.2017.07.213]
- 黄俊. 基于Leap Motion的手势交互在虚拟场景中的研究[硕士学位论文]. 成都: 西南交通大学, 2016.
- Vennila G, Manikandan MSK, Suresh MN. Dynamic voice spammers detection using hidden Markov model for voice over internet protocol network. *Computers & Security*, 2018, 73: 1-16.
- 林巧民, 齐柱柱. 基于HMM和ANN混合模型的语音情感识别研究. *计算机技术与发展*, 2018, 28(10): 74-78. [doi: 10.3969/j.issn.1673-629X.2018.10.015]
- 郑敏嘉, 卢询, 程鑫. 基于HMM模型的电力负荷预测模型研究. *机电工程技术*, 2018, 47(11): 170-173. [doi: 10.3969/j.issn.1009-9492.2018.11.049]
- 曹凯, 于善义, 于少伟. 基于多隐马尔可夫模型的车辆机动行为识别与预测. *信息与控制*, 2014, 43(4): 506-512.