

多种手势对应同一语义的 柔性映射交互算法的研究

冯志全¹, 乔宇¹, 冯仕昌², 周晓燕¹, 徐涛¹, 杨晓晖¹, 范雪¹, 毕建平³

(1. 济南大学信息科学与工程学院, 山东济南 250022; 2. 厦门大学信息学院智能科学与技术系, 福建厦门 361005;

3. 山东中医药大学附属医院, 山东济南 250011)

摘 要: 针对智能交互界面中手势识别错误导致交互界面变化错误和手势不识别两个基本问题, 本文设计并实现了基于手势交互的智能教学界面, 该系统可以通过获取教师的手势信息与教师进行交互. 主要创新点在于提出了多种手势对应同一语义的柔性映射交互算法. 本文选取了 14 种自然交互手势, 分析了对应同一语义的多种手势之间的共同特征. 实验结果显示, 该算法能够有效降低用户负荷. 该算法已经用于一个基于手势交互的智能教学系统界面中.

关键词: 交互模型; 自然用户界面; 认知负担; 基于手势的交互; 人机交互

中图分类号: TP391 文献标识码: A 文章编号: 0372-2112(2019)08-1612-06

电子学报 URL: <http://www.ejournal.org.cn> DOI: 10.3969/j.issn.0372-2112.2019.08.002

Research on Flexible Mapping Algorithm of Multi-gestures to One Semantic

FENG Zhi-quan¹, QIAO Yu¹, FENG Shi-chang², ZHOU Xiao-yan¹,
XU Tao¹, YANG Xiao-hui¹, FAN Xue¹, BI Jian-ping³

(1. College of Information Science and Engineering, University of Jinan, Jinan, Shandong 250022, China;

2. School of Information Science and Engineering, Xiamen University, Xiamen, Fujian 361005, China;

3. Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, Shandong 250011, China)

Abstract: Aiming at the two basic problems, the error of interface interaction caused by hand gesture recognition errors and gesture non-recognition error, the intelligent teaching interface based on gesture interaction is designed and implemented. And the system can interact with the teacher by obtaining the gesture information of the teacher. The main innovation of this paper is to propose a flexible mapping interaction algorithm with multiple gestures mapping to the one semantic. The paper selects fourteen kinds of natural interactive gesture, and analyzes the common characteristics of various gestures mapping to the one semantic. The experimental results show that the proposed algorithm can effectively reduce the users' burden. The algorithm has been used in an intelligent teaching system interface based on gesture interaction.

Key words: interactive mode; natural user interface; cognitive load; gesture interaction; human computer interaction

1 引言

手势是人与人之间非语言交流的最重要方式^[1], 手势交互是继鼠标、键盘和触摸屏之后一种新型的非接触式的人机交互^[2], 以其直观、自然、人机和谐的方式在各个领域取得了广泛应用. 如: 基于手势交互的车

辆控制^[3]、基于手势的机器人控制^[4]、基于手势交互的智能电视控制^[5-6]和基于手势交互的 3D 模型控制等^[7], 基于手势交互的电子沙盘控制^[8], 有力的推动了人机交互的友好发展. 与此同时, 我们注意到在手势与交互界面进行交互的过程中往往遇到两个问题, 一是, 手势识别错误导致交互界面变化错误; 二是, 手势不识别

别. 本文主要从如何解决上述两个问题并评价该解决方案展开的.

随着交互技术的发展, 传统的接触式交互设备如: 键盘、鼠标和游戏手柄等已经不能满足目前人机交互的需求. 新的交互设备如: Kinect^[9,10]、Leap Motion^[7] 和数据手套^[11] 等的出现使得基于手势的交互取得了长足的发展. 目前, 基于手势的交互主要涉及三种技术: 手势跟踪、手势分割、和手势识别. 对于手势跟踪, 目前有如下方法: 基于 Kinect 的手势跟踪方法^[9,12]、基于 Camshift 的手势跟踪方法^[13] 和基于深度学习的手势跟踪方法^[14]; 对于手势分割, 目前有如下方法: 结合 RGB-D 颜色空间的最大期望手势分割方法^[15]、基于神经网络的手势分割方法^[16] 和结合肤色空间的模糊分类手势分割方法^[17]; 对于手势识别, 目前有如下方法: 基于手部分区的手势识别方法^[10] 和同时追踪躯体和对手势的动态手势识别方法^[18] 以及基于模板匹配的三维手势识别算法^[19]. 在基于视觉的手势交互过程中, 我们通过 Kinect 和 Leap Motion 等捕捉设备捕捉手势信息, 经过一系列处理后即可获得手势识别结果, 然后根据得到的识别结果对三维模型^[20]、机器人^[21,22] 和虚拟物体^[7] 等进行操控. 穿戴设备获取的数据非常精确, 但穿戴设备较为昂贵, 在普及程度上有一定的限制. Kumar 提出了一种方法通过数据手套获取手势信息用来做空中输入^[23]. 李东洁等人针对数据手套的精度和实时性问题进行了一定优化^[24].

在交互界面的设计与实现过程中我们遇到了两个问题即在智能交互界面中手势识别错误导致交互界面变化错误和手势不识别两个问题. 针对这两个问题, 本文提出了多种手势对应同一语义的柔性映射交互模型.

2 对应同一语义的多种手势

对于含有一个轨迹跟踪点的手势本文选择弗雷歇距离判断手势轨迹是否相似, 是否可以归为一类; 对于含有多个轨迹跟踪点的手势本文采用多个轨迹跟踪点与开始状态下多个跟踪点的中心点的运动趋势判断多个轨迹手势是否可以归为一类.

2.1 弗雷歇距离

设 A 和 B 是 S 上的两条连续曲线, 即 $A: [0, 1] \rightarrow S$, $B: [0, 1] \rightarrow S$; 又设 α 和 β 是单位区间的两个重参数化函数, 即 $\alpha: [0, 1] \rightarrow [0, 1]$, $\beta: [0, 1] \rightarrow [0, 1]$; 则曲线 A 与曲线 B 的弗雷歇距离 $F(A, B)$ 定义为:

$$F(A, B) = \inf_{\alpha, \beta \in [0, 1]} \max \{ d(A(\alpha(t)), B(\beta(t))) \} \quad (1)$$

其中 d 是 S 上的度量函数.

在 $F(A, B)$ 的计算式中, 先固定最外层的 α 和 β , 也就是对每一个选定的 α 与 β 组合计算式 (2), 式 (1) 中 $d(A, \alpha, B, \beta)$ 均视为被固定住的已知函数, 只将

t 当作变量. 此时, 由于变量 t 将在单位区间 $[0, 1]$ 内遍历所有的连续值 (无穷多个), 为了便于直观理解, 我们将该区间做离散化处理, 即在该区间采样若干点来做分析, 然后通过逐渐增加采样点的个数来提高精度, 最后通过求极限的思想来理解两条曲线的弗雷歇距离.

$$F_{\alpha, \beta}(A, B) = \max_{t \in [0, 1]} \{ d(A(\alpha(t)), B(\beta(t))) \} \quad (2)$$

$F_{\alpha, \beta}(A, B)$ 的离散化计算式为:

$$\bar{F}_{\alpha, \beta}^n(A, B) = \max_{t \in \{t_k\}_{k=0}^{n+1}} \{ d(A(\alpha(t)), B(\beta(t))) \} \quad (3)$$

因此, $F(A, B)$ 的离散化计算式为:

$$\begin{aligned} \bar{F}^n(A, B) &= \inf_{\alpha, \beta} \{ \bar{F}_{\alpha, \beta}^n(A, B) \} \\ &= \inf_{\alpha, \beta} \left\{ \max_{t \in \{t_k\}_{k=0}^{n+1}} \{ d(A(\alpha(t)), B(\beta(t))) \} \right\} \quad (4) \end{aligned}$$

2.2 多个跟踪点的手势分类匹配

在本文中, 轨迹跟踪点与中心点的运动趋势为三种, 分别为相互靠近、相互远离和相互静止. 具体计算多个跟踪点的手势运动趋势步骤如下:

通过式 (5) 计算得到开始状态下多个轨迹跟踪点

$$\begin{cases} C_x = \frac{\sum_{i=1}^n x_i}{n} \\ C_y = \frac{\sum_{i=1}^n y_i}{n} \end{cases} \quad (5)$$

的中心点.

(C_x, C_y) 为计算得到的开始状态下多个轨迹跟踪点的中心点位置信息;

然后, 我们通过式 (6) 计算其各个跟踪点与中心点的相互运动趋势

$$\text{value} = \frac{(x_{i+1} - C_x)^2 + (y_{i+1} - C_y)^2}{(x_i - C_x)^2 + (y_i - C_y)^2} \quad (6)$$

(x_i, y_i) 为当前时刻轨迹点坐标; (x_{i+1}, y_{i+1}) 为下一时刻轨迹点坐标.

考虑到用户在做手势时会有抖动的问题, 因此我们采取式 (7) 作为前置限制条件, 当满足式 (7) 时, 轨迹跟踪点与中心点为相互静止状态.

$$\begin{aligned} &\| ((x_{i+1} - C_x)^2 + (y_{i+1} - C_y)^2) - \\ &((x_i - C_x)^2 + (y_i - C_y)^2) \| < \text{value}_{\text{threshold}} \quad (7) \end{aligned}$$

其中 $\text{value}_{\text{threshold}}$ 为手势抖动的阈值, 在该范围内手势的轨迹跟踪点与中心点为相互静止状态.

3 智能教学界面中的柔性映射交互算法

为了解决教师与智能教学界面交互过程中手势识别错误导致智能教学界面变化错误和手势不识别两个问题, 本文提出了多种手势对应同一语义的柔性映射交互算法. 智能教学界面的主流程如图 1 所示.

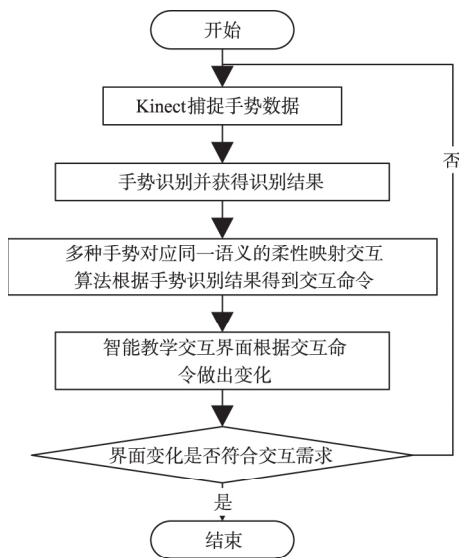


图1 智能教学界面的主流程图

3.1 获取手势信息和手势识别结果

本文采用 Kinect 同时捕捉手势区域的彩色信息和轨迹信息. 采用式 (8) 将手势区分为动态手势和轨迹手势. 其中动态手势符合该公式, 反之, 不符合该式的手势为轨迹手势.

$$\begin{cases} |handPosition.x - preHandPosition.x| < value \\ |hand4Position.y - preHandPosition.y| < value \end{cases} \quad (8)$$

handPosition 为手势结束时坐标的位置, preHandPosition 为手势开始时坐标的位置, value 为区分两种手势的阈值, 经过实验验证, 本文选取 20 作为 value 的值, 此时, 能够取得良好的效果.

本文采用结合主方向的类 Hausdorff 距离的静态手势识别方法^[25]对带有抓住状态的手势的抓住状态进行识别; 对于轨迹手势, 本文采用十二方向编码的隐马尔科夫模型方法^[26]进行识别; 本文采用基于形状上下文的密度分布特征动态手势识别方法^[27]对动态手势进行识别, 最终我们得到手势识别结果 X_{ij} .

3.2 多种手势对应同一语义柔性映射交互算法

本文中的智能教学界面包含 14 种交互手势和 5 种交互语义集合, 即 $X_{gesture}$ 和 $Y_{gesture}$, 其分别对应柔性映射交互算法的 14 种输入和 5 种输出. 柔性映射交互算法的内部映射规则建立流程如图 2 所示, 具体步骤如下:

首先, 我们将选取得到的 14 种手势按照其显著特征分为三组. 其次, 我们按照分组结果采用相应的衡量方法进行归类. 本文针对每一个手势采集了 100 个手势样本并对其进行特征统计. 对于带有抓住状态的手势, 考虑其抓住状态和抓住状态外的手势轨迹的相似性. 带有抓住状态的四组手势的手势轨迹的弗雷歇距离统计结果如表 1 所示.

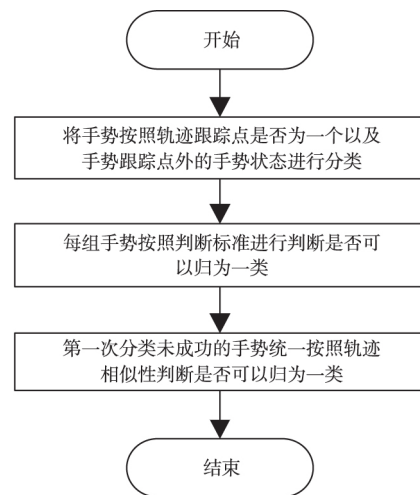


图2 多种手势对应同一语义柔性映射算法内部映射规则建立流程图

表 1 带有抓住状态手势的轨迹相似性(弗雷歇距离表示, 值越小代表越相近)

手势	弗雷歇距离
左手抓住不动 右手抓住远离手势 - 右手抓住远离手势	65.732787
左手抓住不动 右手抓住远离手势 - 左手抓住不动 右手靠近手势	396.225946
左手抓住不动 右手抓住远离手势 - 右手抓住靠近手势	395.786842
右手抓住远离手势 - 左手抓住不动 右手抓住靠近手势	397.180234
右手抓住远离手势 - 右手抓住靠近手势	396.223086
左手抓住不动 右手抓住靠近手势 - 右手抓住靠近手势	64.589072

对于具有多个轨迹跟踪点的手势, 本文考察其每个跟踪点与开始状态下多个跟踪点计算出的中心点的相互运动趋势. 具有多个轨迹跟踪点的 9 个手势的多个跟踪点的运动趋势统计结果如表 2 所示.

表 2 多个手势跟踪点与中心点的相互运动趋势

手势	跟踪点相互运动趋势
二指放开手势	相互远离
三指放开手势	相互远离
五指放开手势	相互远离
双手远离手势	相互远离
二指闭合手势	相互靠近
三指闭合手势	相互靠近
五指闭合手势	相互靠近
双手靠近手势	相互靠近
拧旋钮手势	相互静止

最后, 本文将单独分为一类的手势采用统一的轨迹匹配方式进行分类. 对于拧旋钮手势因其有两个轨迹跟踪点即两条轨迹曲线, 因此本文将这两条轨迹曲

线进行首尾连接成一条轨迹曲线再与其他手势的轨迹进行匹配. 拧旋钮手势和单手画圆手势的弗雷歇距离为 130.825195 小于阈值 150, 因此该两种手势可以分为一类. 至此, 本文将一个大类的手势集合按照匹配原则最终分为了 5 类, 此时本文的柔性映射的内部映射机制建立完毕, 即建立了满足式(9)的柔性映射交互模型.

$$F(X_{ij}) = y_i (i=1, 2, \dots, 5; j=1, 2, \dots, 14) \quad (9)$$

4 实验及分析

本文假设教师可以非常容易的学习并使用本文设计并实现的智能教学界面并且教师操作智能教学界面时的认知负荷较小; 本文同样假设教师可以适应采用手势控制教学界面完成教学任务. 下面, 本文将进行相关实验来论证上述两个假设.

4.1 实验

本文选取 10 名实验者, 在实验开始的时候, 依次告诉实验者表达这一语义的有哪些手势. 本文采用美国航空航天局任务负荷指数 (National Aeronautics and Space Administration Task Load Index, NASA-TLX) 认知负荷测评法来评估本文算法的认知负荷, 同时采用一种手势对应一种语义交互算法作为对比算法进行对比进而评估本文算法的认知负荷. 每项指标的最高分为 100, 最低分为 0. 分数越高表示完成整个交互过程实验者的认知负荷越大.

在实验者依次完成交互语义实验后, 实验者需参照 NASA-TLX 评分规则对其中所涉及到的六项指标依次进行打分并依次回答表 3 中的问题.

表 3 实验者在实验后需回答的问题

编号	问题
1	记住对应一种语义的多种手势是否是一个负担?
2	智能教学界面是否容易通过手势进行操控?
3	采用多种手势表达同一种语义这种方式在你看来是一种自然的交互方式?

4.2 实验结果及其分析

本文设计实现了证明“同底面积同高的三棱柱是三棱锥体积的三倍”这一数学定理的智能教学界面. 下面我们选取智能教学界面中的放大语义简述该智能教学界面所完成的功能.

如图 3 所示为三个棱锥拼合为一个三棱柱状态的智能教学界面, 此时想要将该三棱柱进行放大操作, 我们只需任选一种手势如五指放开手势对智能教学界面进行操作, 操作完成后如图 4 所示, 此时, 智能教学界面中的三棱柱变大; 若此时智能教学界面中的三棱柱没有变化, 只需更换对应放大语义的不同手势如二指放开手势操作智能教学界面, 直至智能教学界面做出正



图3 五指放大手势执行前三棱柱初始状态

确变化; 若此时智能教学界面中的三棱柱做错误变化, 我们只需要做出撤回手势使得智能教学界面回到变化之前的状态, 再选择对应放大语义的任一手势对智能教学界面进行操作直至智能教学界面的变化符合我们期望.



图4 五指放大手势执行后三棱柱状态

实验结果如图 5 所示, 与一种手势对应一种语义交互算法相比本文的柔性映射交互算法的认知负荷降低了 51.2%. 两种交互算法在脑力要求方面的评分基本持平. 其余的五个方面多种手势对应一种语义柔性映射交互算法都有着很好的表现, 因为该柔性映射交互算法解决了交互过程中遇到的手势不识别以及手势识别错误导致智能教学界面变化错误等两个问题.

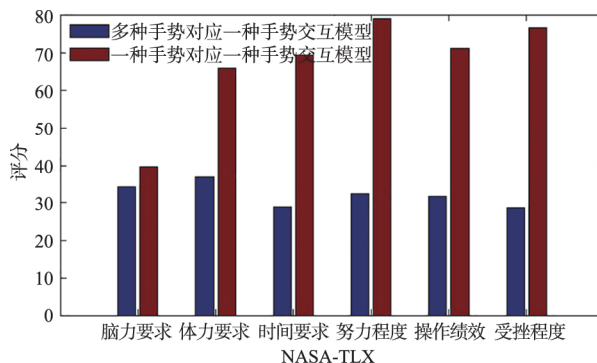


图5 两种交互算法的NASA-TLX评分

根据实验的反馈可知, 尽管多种手势对应一种语义柔性映射交互算法需要记住多种手势, 但是对应一

种语义的多种手势之间有着共同之处,一旦学会了,可以记忆很长一段时间,并且该智能教学界面通过多种手势对应同一语义柔性映射交互算法能够非常容易控制。

5 结语

本文提出了多种手势对应同一语义柔性映射交互算法能够很好的解决在与智能教学界面交互过程中遇到的手势识别错误导致智能教学界面变化错误和手势不识别两个问题,并且该柔性映射交互算法是一个符合人类日常交互习惯的交互算法,能够显著地减少用户的交互负荷。本文分析了该算法在智能教学界面中的表现,在我们未来的研究计划中,我们将该模型扩展到其他交互界面上如场景漫游和三维模型交互。

参考文献

- [1] Yin X, Xie M. Finger identification and hand posture recognition for human-robot interaction [J]. *Image and Vision Computing* 2007, 25(8): 1291 - 1300.
- [2] Nguyen, Viet T. Enhancing Touchless Interaction with the Leap Motion Using a Haptic Glove [D]. Joensuu, Kuopio, Finland: University of Eastern Finland 2014. 1 - 67.
- [3] Riener A, Ferscha A, Bachmair F, et al. Standardization of the in-car gesture interaction space [A]. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* [C]. New York, USA: ACM 2013. 14 - 21.
- [4] 齐静, 徐坤, 丁希伦. 机器人视觉手势交互技术研究进展 [J]. *机器人* 2017, 39(4): 565 - 584.
- [5] Ahn J, Kim K. Investigating smart TV gesture interaction based on gesture types and styles [J]. *Journal of the Ergonomics Society of Korea* 2017, 36(2): 109 - 121.
- [6] 徐治鹏, 冯志全, 刘慧, 等. 面向智能电视的隐式手势交互建模与算法 [J]. *计算机辅助设计与图形学学报* 2017, 29(2): 344 - 353.
Xu Z, Feng Z, Liu H, et al. Implicit gesture interaction modeling and algorithm for smart TV [J]. *Journal of Computer-Aided Design & Computer Graphics* 2017, 29(2): 344 - 353. (in Chinese)
- [7] Cui J, Kuijper A, Sourin A. Exploration of natural free-hand interaction for shape modeling using leap motion controller [A]. *Proceedings of the International Conference on Cyberworlds (CW)* [C]. Chongqing, China: IEEE, 2016. 41 - 48.
- [8] 张玉军, 孟晓军, 王刚. 基于手势交互的三维电子沙盘系统设计及实现 [J]. *指挥控制与仿真* 2016, 38(2): 110 - 114.
- [9] Carter M, Newn J, Velloso E, et al. Remote Gaze and Gesture Tracking on the microsoft kinect: investigating the role of feedback [A]. *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* [C]. New York, US: ACM 2015. 167 - 176.
- [10] Ren Z, Yuan J, Meng J, et al. Robust part-based hand gesture recognition using Kinect sensor [J]. *IEEE Transactions on Multimedia* 2013, 15(5): 1110 - 1120.
- [11] Camastra F, De Felice D. LVQ-based hand gesture recognition using a data glove [A]. *Neural Nets and Surroundings* [M]. Berlin Heidelberg: Springer 2013, 159 - 168.
- [12] 张莹莹, 郭星. 基于 Kinect 动态手势识别算法的研究与实现 [J]. *计算机技术与发展* 2017, 27(12): 11 - 15.
- [13] Hu Z F, Wang Y K, Luo Y, et al. An Improved camshift-based particle filter algorithm for real-time hand gesture tracking [J]. *Sensors & Transducers* 2014, 177(8): 307 - 312.
- [14] Sanchez-Riera J, Hsiao Y S, Lim T, et al. A robust tracking algorithm for 3d hand gesture with rapid hand motion through deep learning [A]. *Proceedings of the 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)* [C]. Cheng Du, China: IEEE, 2014. 1 - 6.
- [15] Ju Z, Wang Y, Zeng W, et al. A modified EM algorithm for hand gesture segmentation in RGB-D data [A]. *Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* [C]. Beijing, China: IEEE, 2014. 1736 - 1742.
- [16] Huang L P, Wu Q X, Zhang G R, et al. Gesture segmentation using spiking neural networks inspired by the visual system [A]. *Proceedings of 2015 International Conference on Artificial Intelligence and Industrial Engineering* [C]. Thailand, Phuket: Atlantis Press 2015. 184 - 187.
- [17] Zhang Q, Lu J, Zhang M, et al. Hand gesture segmentation based on mixed skin-color model and FCM algorithm [J]. *Journal of Information & Computational Science* 2015, 12(9): 3527 - 3536.
- [18] Song Y, Davis R. Continuous body and hand gesture recognition for natural human-computer interaction [A]. *Proceedings of the 24th International Conference on Artificial Intelligence* [C]. Buenos Aires, Argentina: AAAI Press, 2015. 4212 - 4216.
- [19] 刘杰, 黄进, 韩冬奇, 等. 模板匹配的三维手势识别算法 [J]. *计算机辅助设计与图形学学报* 2016, 28(8): 1365 - 1372.
Liu J, Huang J, Han D, et al. Templatematching algorithm for 3D gesture recognition [J]. *Journal of Computer-Aided Design & Computer Graphics* 2016, 28(8): 1365 - 1372. (in Chinese).

- [20] Wang K R ,Xiao B J ,Xia J Y ,et al. A real-time vision-based hand gesture interaction system for virtual EAST [J]. Fusion Engineering and Design ,2016 ,112: 829 -834.
- [21] Qian K ,Niu J ,Yang H. Developing a gesture based remote human-robot interaction system using Kinect [J]. International Journal of Smart Home 2013 ,7(4) : 203 -208.
- [22] 李文瀚,郑新,金宁,等. 基于手势识别的智能机器人系统设计[J]. 中国科技信息 2018 ,575(2) : 76 -77.
- [23] Kumar P ,Verma J ,Prasad S. Hand data glove: a wearable real-time device for human-computer interaction [J]. International Journal of Advanced Science and Technology , 2012 ,43: 15 -26.
- [24] 李东洁,李洋洋,杨柳. GL-RBF 优化的数据手套手势识别算法 [J]. 哈尔滨理工大学学报 2017 ,22(4) : 7 -12.
- [25] Yang X W ,Feng Z Q ,Huang Z Z ,et al. Gesture recognition based on combining main direction of gesture and hausdorff-like distance [J]. Journal of Computer-Aided Design & Computer Graphics. 2016 ,28(1) : 75 -81.
- [26] Zhang Y ,Wu S H ,Luo Y. Applications and recognition of gesture trajectory using HMM [J]. Semiconductor Optoelectronics 2015 ,36(4) : 650 -656.
- [27] Xiaoyan Z ,Zhiqian F ,Yu Q ,et al. The Man-machine finger-guessing game based on cooperation mechanism [J]. Theoretical Computer Science 2017 ,6: 92 -109.

作者简介



冯志全(通信作者) 男,1964年6月26日出生,现为济南大学教授,从事智能感知与自然交互研究。主持多项国家重点研发计划项目子课题、国家自然科学基金、山东省自然科学基金重点项目等项目,发表专业论文100余篇,授权发明专利30多项。

E-mail: ise_fengzq@ujn.edu.cn



乔宇 男,1993年出生于河北辛集人。济南大学硕士学位。主要从事人机交互相关技术的研究。