

doi: 10.3969/j.issn.1003-4271.2014.06.13

基于单目视觉和简单手势的空间交互技术

陈雅茜, 欧长坤, 郭墨阳

(西南民族大学计算机科学与技术学院, 四川 成都 610041)

摘 要: 和传统接触型交互设备相比, 新颖的非触摸型空间交互技术在空间自由度方面有着明显优势. 主要探讨基于单目摄像头和简单手势的双手空间交互技术. 首先利用肤色分割、滤波优化、运动轨迹预测等算法实现了双手的跟踪与识别. 由于需要用空间手势及其运动状态来替代鼠标, 提出了基于中心偏离控制和基于线性映射的两种鼠标映射模式, 进而实现了四种基本空间手势. 基于提出的空间交互技术, 实现了软件控制、游戏体验等多个应用. 最后讨论了基于单目视觉的双手空间交互技术在手部分割、学习算法、交互方式等方面的改进方向.

关键词: 单目视觉; 人机交互; 空间交互; 肤色分割; 鼠标控制

中图分类号: TP391.41

文献标识码: A

文章编号: 1003-4271(2014)06-0871-06

对于鼠标、触摸屏等设备, 用户必须接触到设备之后才能产生交互效果, 空间自由度存在一定的局限^[1]. 因此, 新颖的非触摸型空间交互技术的研究具有一定的实用意义^[2]. 基于单目摄像头的空间交互技术由于低成本、易推广等特性, 一直都是人机交互领域的一大研究热点^[3-4]. 文献[5]实现了基于单目摄像头的单手手势识别, 但并未涉及空间交互技术. 文献[3]设计实现了三个基于手势控制的人机交互原型系统, 但仅是针对特定的软件系统集成, 不具备通用性. 文献[4]实现了基于单目视觉的双手识别, 但空间交互的限制性较多, 自然手势交互还需改进.

本文探讨了基于单目摄像头和简单手势的双手空间交互技术, 利用 YCbCr 肤色分割、Kalman 滤波优化、GM(1,1)运动轨迹预测等算法实现了双手的跟踪识别, 基于以上算法深度优化了用户在鼠标操作上的体验. 提出了两种鼠标映射模式, 实现了四种基本空间手势. 最后还讨论了单目视觉的双手空间交互技术在手部分割、学习算法、交互方式等方面的改进方向.

1 系统框架

系统框架如图 1 所示.

- (1)核心运算模块: 封装膨胀、腐蚀和二值化等运算函数.
 - (2)手部提取与分割模块: 从图像中分割出多个目标, 计算各个目标的重心并更新目标坐标.
 - (3)手部跟踪模块: 利用运动轨迹预测算法对手部的下一个可能位置做预测.
 - (4)手势识别模块: 对手部提取结果中的手部轮廓进行凸性分析, 识别当前手势. 同时使用预测模块的结果对多目标进行标记.
 - (5)滤波优化模块: 对目标的坐标进行滤波处理, 达到平滑鼠标体验的目的.
 - (6)鼠标控制模块: 监控系统鼠标, 提供两种鼠标映射模式.
- 限于篇幅, 本文主要讨论手部提取与分割、手部跟踪以及鼠标控制模块中的鼠标映射模式.

收稿日期: 2014-09-12

通信作者: 陈雅茜(1981-), 女, 汉族, 四川雅安人, 副教授, 博士, 研究方向: 人机交互、信息可视化, Email: yaxichen@swun.cn.

基金项目: 2014 年国家外专项目; 西南民族大学中央高校基本科研业务费专项基金(青年教师基金项目)(13NZYQN19);西南民族大学 2013 年教育教学改革项目(2013ZC70);2014 年度西南民族大学国家级大学生创新创业训练计划项目(201410656005)

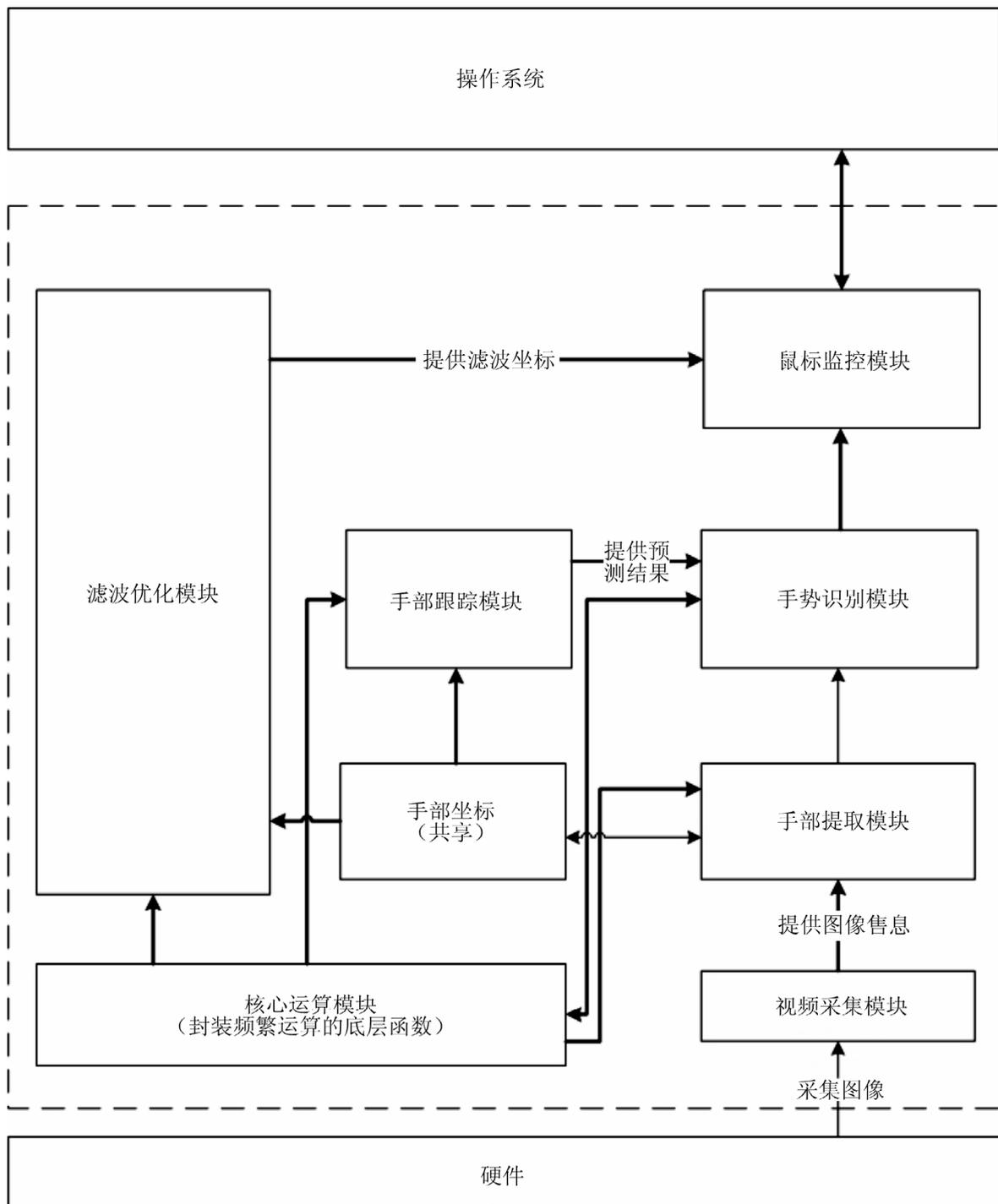


图 1 系统框架

Fig.1 System framework

2 手部跟踪识别

单手跟踪需要对干扰目标进行筛选和提出,而多手跟踪则需要对各个目标进行跟踪预测.在识别方面,如果采用的是简单手势,则只需要对手部进行凸性分析即可.

2.1 手部分割提取

本文利用椭圆肤色模型对画面中的手实现分割提取:将 RGB 颜色空间变换到 YCbCr 空间中,利用肤色在

YCbCr 颜色空间中的聚类性质提取肤色^[6-8]. 在 CbCr 平面上, 肤色分布符合椭圆分布, 满足:

$$\frac{(x-ecx)^2}{a^2} + \frac{(y-ecy)^2}{b^2} = 1 \quad (1)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} Cb-cb_0 \\ Cr-cr_0 \end{bmatrix} \quad (2)$$

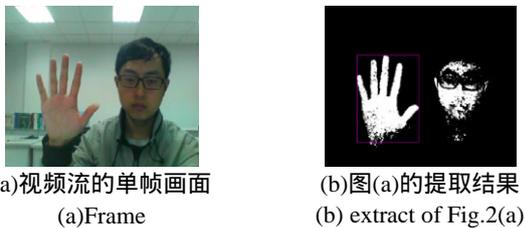
再根据判别函数 $D(Cb, Cr)$ (见式 3) 分割出手势.

$$D(Cb, Cr) = \begin{cases} 1 & \frac{(x-ecx)^2}{a^2} + \frac{(y-ecy)^2}{b^2} \leq 1 \\ 0 & \text{其他} \end{cases} \quad (3)$$

利用式(1)~(2)建模, 经实验得到:

$$77 \leq Cb \leq 127, 133 \leq Cr \leq 173 \quad (4)$$

如图 2 所示, 在光线明亮、画面环境稍复杂且有强光源条件下的手部提取效果良好. 图 3 是环境光源明暗不均时的手部提取效果. 由于没有提取到距离摄像头较远的脸部肤色, 其效果较图 2 更为良好.



(a)视频流的单帧画面 (a)Frame
(b)图(a)的提取结果 (b) extract of Fig.2(a)

图 2 复杂条件下的手部提取效果

Fig.2 Hand extraction effects in a complex condition



(a)视频流的单帧画面 (a)Frame
(b)该画面的提取结果 (b)extract of Fig.3(a)

图 3 环境光源明暗不均的手部提取效果

Fig.3 Hand extraction effects in nonuniform light condition

2.2 滤波优化

由于要用手部的空间位置替代屏幕上的鼠标位置, 则须利用平滑算法控制鼠标位置的抖动. 本文使用的 Kalman 滤波器^[9]假设系统是线性的, 它的系统行为是基于最小二乘的最佳估计, 不断把协方差递归, 估算出最优值, 系统的模型和观测估计拥有最小的方差. Kalman 滤波器的实时运行效率很高, 可以随不同时刻而改变它的值:

$$X(k|k) = X(k|k-1) + K_g(k)(Z(k) - HX(k|k-1)) \quad (5)$$

其中, H 是测量矩阵, $Z(k)$ 是测量值, $K_g(k)$ 是 Kalman 增益, $X(k|k)$ 是对 $X(k|k-1)$ 的改进.

2.3 基于运动轨迹预测的手部跟踪

本文将改进的灰色理论 GM(1,1)模型^[10-12]的预测算法应用于对手部运动轨迹的实时跟踪. 本文使用的待预测序列长度为 30.

记录每一帧中手部对应的坐标: $p^{(i)} = (x^{(i)}, y^{(i)})$, 初始轨迹坐标序列: $p_0 = (p_0^{(1)}, p_0^{(2)}, \dots, p_0^{(n)})$;

计算累加生成序列: $p_1 = (p_1^{(1)}, p_1^{(2)}, \dots, p_1^{(n)})$ 其中 $p_1^{(k)} = \sum_{i=1}^k p_0^{(i)} (k=1, 2, \dots, n)$, 及均值生成序列: $m_1 = (m_1(2), m_1(3), \dots, m_1(n))$,

其中

$$m_1(k) = \frac{1}{2}(p_1^{(k)} + p_1^{(k-1)}) (k=2, \dots, n)$$

$$\text{令 } u = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, Y = (p_0^{(2)}, p_0^{(3)}, \dots, p_0^{(n)})^T, B = \begin{bmatrix} -m_1^{(2)} & 1 \\ -m_1^{(3)} & 1 \\ \vdots & 1 \\ -m_1^{(n)} & 1 \end{bmatrix}, \text{ 计算 } (\hat{\alpha}, \hat{\beta}) = (B^T B)^{-1} B^T Y;$$

得到序列预测值为:

$$\hat{p}_1^{(n+1)} = \left(p_1^{(n+1)} - \frac{\hat{\beta}}{\hat{\alpha}} \right) e^{-\hat{\alpha}n} + \frac{\hat{\beta}}{\hat{\alpha}} \quad (6)$$

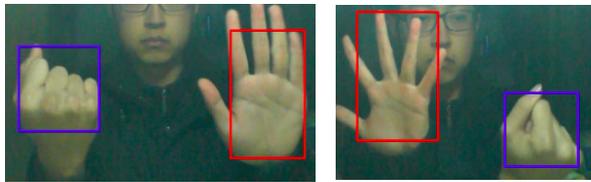
求出预测坐标:

$$\hat{p}_0^{(n+1)} = \hat{p}_1^{(n+1)} - p_1^{(n)} \quad (7)$$

在预测坐标附近进行小范围目标搜索, 如果找到, 则标记此目标为跟踪目标.

确定跟踪目标后, 通过图像分析来获得精确坐标, 并将下一次的预测基准序列更新为

$p_0 = (p_0^{(1)}, \dots, p_0^{(n)}, p_0^{(n+1)})$, 返回(1); 若根据预测没有搜索到目标, 则抛弃前有结果, 重新获取预测坐标序列 $p_0 = (p_0^{(1)}, p_0^{(2)}, \dots, p_0^{(n)})$, 返回(1).



(a) 目标移动前 (b) 目标移动后
(a) Before moving (b) Object moved

图 4 手部跟踪效果一
Fig.4 Hand tracking effects



(a) 目标移动前 (b) 目标移动后
(a) Before moving (b) Object moved

图 5 手部跟踪效果二
Fig.5 Hand tracking effects

实现结果见图 4、5 所示, 该预测算法不仅速度快、精度高, 系统开销小于 CamShift /MeanShift 等算法, 且跟踪效果与手空间姿势无关.

2.4 手势识别

在目标跟踪的基础上, 本文实现了对手掌和握拳两种基本手势的识别, 对[5]提出的静态手势识别思想进行了改进与优化: 以目标重心为圆心画圆, 并统计圆上不相交区域的个数即手指的个数, 对区域进行凸性分析(仅判断是否存在不相交区域), 从而确定当前手势, 具体算法如下:

先得出分割提取后手的凸包坐标^[13], 不妨设共 N 个, M 是指标集:

$$A = \{p_m(x, y) | m \in M\}, M = \{1, 2, \dots, N\} \quad (8)$$

则手的心为: $G(x, y) = \frac{1}{N} \sum_{m=1}^N p_m(x, y);$

得到最大半径: 存在唯一的 $s \in M$, 对任意的 $m \in M$, 有:

$$r_{\max} = \|p_s(x, y) - G(x, y)\| \geq \|p_m(x, y) - G(x, y)\| \quad (9)$$

对所有满足(10)和(11)的 $q(x, y)$, 有 $\varphi(q(x, y)) = 1$, 则认为手势为拳, 否则为掌. 其中 $\|\cdot\|$ 是 L_2 范数, $\langle \cdot, \cdot \rangle$ 是两向量的夹角, $\varphi(\cdot)$ 是向量所在像素的灰度值.

$$\|q(x, y) - G(x, y)\| - \frac{1}{2} r_{\max} \leq \Delta r \quad (10)$$

$$\langle (q(x, y) - G(x, y)), (p_s(x, y) - G(x, y)) \rangle \leq \frac{\pi}{8} \quad (11)$$

3 鼠标映射模式

本文通过 4 种基本手势替代鼠标操作: 手掌移动表示鼠标移动; 由手掌转变为握拳表示按下鼠标左键; 由握拳转变为手掌表示释放鼠标左键; 长时间的握拳表示单机鼠标右键. 本文设计了两种鼠标映射模式来实现对鼠标移动的控制:

模式 1 是基于中心偏离控制的映射模式, 见图 6(a): 画面中设定一个有效区域(长宽减小 ε), 当手位于区域内, 触发鼠标控制. 手离画面中心越远, 鼠标沿该方向的移动速度越快, 见式(12). 对给定的 $r \geq 2, k \in (0, 1)$, $W(x, y)$ 是图像中心坐标. 手位于稳定区域内则鼠标指针不移动.

$$\left((\Delta x)_{\frac{1}{r}}, (\Delta y)_{\frac{1}{r}} \right) = k(G(x, y) - W(x, y)) \quad (12)$$

模式 2 是基于线性映射的映射模式, 见图 6(b): 设定画面中的有效区域(长宽减小 ε), 将有效区域通过线性映射变换(式 13)到系统屏幕中.

$$\frac{x'}{x + \frac{img_{width} - \varepsilon}{2}} = \frac{y'}{\frac{img_{height} - \varepsilon}{2} - y} = \frac{screen_{width}}{screen_{height}} \quad (13)$$

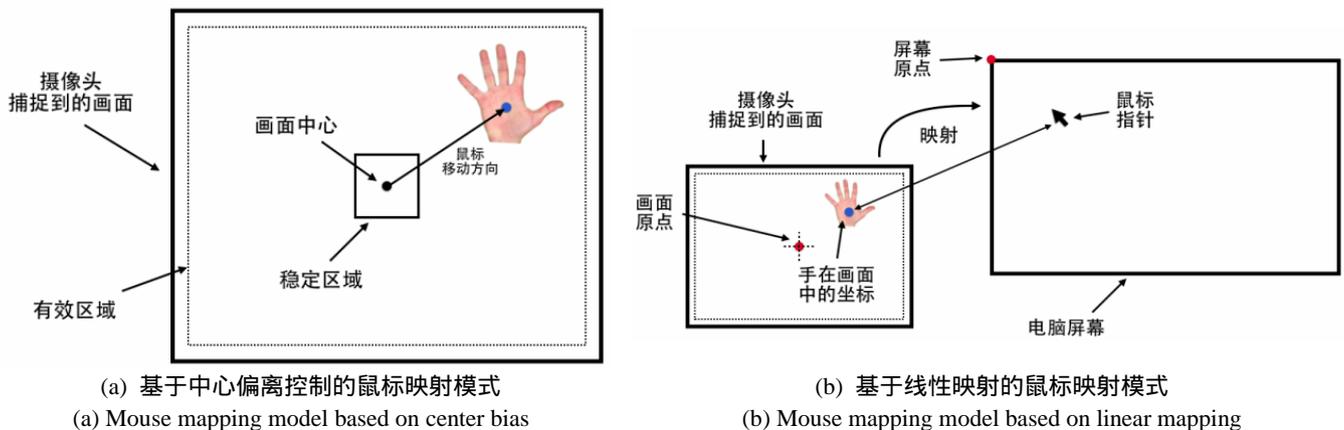


图 6 两种鼠标映射模式
Fig.6 Mouse mapping models

经用户测试^[14-15], 我们发现模式 2 适合普通情景下的鼠标控制, 学习成本较低; 而模式 1 则更适合 FPS 等游戏场景中的鼠标控制。

4 结束语

本文通过椭圆肤色模型、轨迹预测及 Kalman 滤波优化等算法实现了基于单目摄像头和简单手势的双手空间交互技术。未来需在识别算法、交互方式等方面做进一步改进:

1) 手部分割: 后续将重点研究如何在复杂光源及近似肤色干扰等情况^[16-18]下确保分割算法的正确性。另外, 可引入机器学习等算法对椭圆肤色模型的肤色范围参数等先验结果进行学习。

2) 学习算法的引入: 引入具有先验知识的学习算法能够大幅提升识别率。在推荐系统中, 经过巧妙设计特征, 使用推荐算法和先验知识能更有效的对录入的用户的行为信息进行管理和挖掘, 进而激发用户的参与性, 主动改进并完善识别系统, 大幅度提高识别率^[21]。再如, 虽然单目视觉深度信息的提取几乎不可行^[19-20], 但可以考虑通过监督学习算法(如 Markov 随机场)恢复一些粗糙的深度信息^[22-23]。

3) 交互方式的改进: 本文通过 4 种预设手势实现了基本鼠标操作, 但距用户自然多变的交互需求还有一定差距。因此不仅需要对基于自然手势的交互方式进行研究, 还可以尝试通过多通道输入信号筛选出更多有效的交互信息, 并将这些参数提供给游戏的物理引擎^[1-2], 从而带来更加真实、自然的交互体验。

参考文献:

- [1] ANDREW D. WILSON, SHAHRAMIZADI, OTMARHILLIGES. Bringing physics to the surface[C]//ACM UIST '08. MONTEREY, CA, October 2008.
- [2] OTMARHILLIGES, SHAHRAMIZADI, ANFREW D WILSON. Interactions in the Air: Adding Further Depth to Interactive Tabletops[C]//ACM UIST '09. VICTORIA, BC, October 2009.
- [3] 柴秀娟. 用于视觉交互系统的手势跟踪和识别研究[D]. 北京: 北京邮电大学, 2009.
- [4] 赵书兴. 基于单目视觉的双手手势识别技术研究[D]. 南京: 南京师范大学, 2013.
- [5] 李平, 李允俊. 基于手势识别算法的鼠标终端[J]. 计算机系统应用, 2013, 22(8):83-84.
- [6] YANG J, WAIBEL A. A Real-Time Face Tracker[J]. Proceedings 3rd IEEE Workshop on, 1996:142-147.
- [7] HSU RL, MOHAMED AM, JAIN AK. Face Detection in Color Images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 696-706.
- [8] SORENSON H W. Least-squares estimation: from Gauss to Kalman[J]. IEEE Spectrum, 1970: 7-12.
- [9] 彭丁聪. 卡尔曼滤波的基本原理及应用[J]. 软件导刊, 2009(11): 32-34.
- [10] 谭冠军. GM(1,1)模型的背景值构造方法和应用[J]. 系统工程理论与实践, 2000(4): 91-97.
- [11] 刘思峰, 邓聚龙. GM(1,1)模型的适用范围[J]. 系统工程理论与实践, 2000(5):98-103.
- [12] 张大海, 江世芳. 灰色预测公式的理论缺陷及改进[J]. 系统工程理论与实践, 2002(8):140-142.
- [13] K HOMMA, EI TAKENAKA. An image processing method for feature extraction of space-occupying lesions[J]. Journal of Nuclear Medicine, 1985, 26: 1472-1477.

- [14] 欧长坤. 基于单目摄像头和计算机视觉的空间交互技术的研究与实现[R]. 成都:西南民族大学, 2013.
- [15] 欧长坤. Demo 项目[EB/OL]. (2013-12-1)[2014.05.25]. <http://www.euryugasaki.com/works/demo>.
- [16] 尼璐璐. 基于物理的高光条件下人脸图像肤色检测技术的研究[D]. 天津: 天津大学, 2007.
- [17] 郑利华,张亚红. 一种基于 HSV 空间的人像高光区域自适应修正方法[J]. 桂林航天工业高等专科学校学报, 2012(03):243-246
- [18] 黄廷辉, 杨飞, 崔更申. 光照鲁棒性较强的肤色检测算法[J]. 计算机应用, 2014,34(04):1130-1133,1138.
- [19] D SCHARSTEIN, R SZELISKI. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. Int'l Journal of Computer Vision, 47:7-42, 2002.
- [20] DAVID A.FORSYTH AND JEAN PONCE. Computer Vision: A Modern Approach[M]. Prentice Hall, 2003.
- [21] 陈雅茜, 刘韬, 方诗虹. 推荐系统及其相关技术研究[J]. 西南民族大学学报: 自然科学版, 2014, 40(3): 439-442
- [22] SAXENA A, CHUNG S H, NG A Y. Learning depth from single monocular images[J]. In Neural information processing system (NIPS) 2005, 18: 119-121.
- [23] SAXENA A, CHUNG S H, NG A Y. 3-D Depth Reconstruction from a Single Still Image[J]. Int J Comput Vis, 2008(76): 53-69.

Space interactions based on monocular vision and simple gestures

CHEN Ya-xi, OU Chang-kun, GUO Zhao-yang

(Southwest University for Nationalities, Chengdu 610041, P.R.C.)

Abstract: Considering the issue of spacial freedom, novel non-touch space interaction precedes traditional interactive devices with touch screen. This paper mainly discusses space interactions based on monocular camera and simple gestures. Firstly, hands tracking and recognition are realized by using skin color segmentation, filter optimization and trajectory prediction. Since mouse movement should be replaced by spacial gestures, two mouse mapping modes (central-bias and linear-mapping) are proposed, based on which, four simple spacial gestures are designed. Based on space interaction technology, this paper implemented various applications such as software control and game experience. Future improvements are then discussed, including hand segmentation, learning algorithm and interaction modes.

Key words: monocular vision; human-computer interaction; space interaction; skin color segmentation; mouse control

本期重要作者介绍



马志霞
教授

马志霞,女,数学学科教授,硕士生导师。四川省数学学会理事,成都市应用数学学会常务理事。研究方向为常微分方程稳定性理论,在《数学年刊》、《数学杂志》、《Journal of Inequalities and Applications》、《Advances in Difference Equations》和《四川大学学报》(自然科学版)等国内外专业杂志发表学术论文20余篇,其中3篇被SCI收录。主持省部级科研项目1项、校级科研项目3项,主研省部级科研项目3项,主研省级及校级教改项目5项。主要担任《数学分析》、《常微分方程》、《数学模型》等课程的教学,主讲的《常微分方程》被列为校级精品课程。2012年获国家民委民族院校教学观摩课比赛一等奖,2007年获校第七届优秀教学质量二等奖。指导学生参加全国和美国大学生数学建模竞赛获一等奖5项,二等奖9项,省级奖若干项。



李清寒
副教授

李清寒,男,四川中江人,理学博士(博士后),副教授,硕士生导师,中国化学学会会员,国际纯粹与应用化学联合会(IUPAC)会员,中国环境科学学会高级会员,四川省化工学会理事,国际重要英文期刊Chem. Commun., Molecule杂志论文评审人,中文核心期刊《应用化学》杂志论文评审人。先后获得西南民族大学第二届优秀青年骨干教师(2011),西南民族大学校大学生创新创业训练优秀指导教师(2014)奖。

1995年6月毕业于四川师范大学化学系,2003年6月毕业于四川大学化学学院获理学硕士学位;同年9月进入四川大学教育部绿色化学与技术重点实验室攻读有机化学专业博士学位,师从冯小明院士,从事不对称催化反应研究,2006年6月获理学博士学位。2006年7月进入西南民族大学化学与环境保护工程学院。2009年8月-2011年7月进入台湾中兴大学化学系博士后流动站,师从中兴大学特聘教授高汉谋博士从事金属有机化学及不对称催化反应研究。现从事不对称催化、金属有机化学、绿色有机合成及生物有机化学的研究,主要研究领域涉及:不对称催化

及合成反应,金属催化的碳-碳/杂原子(O, S, N, etc.)键形成反应,具有生物活性化合物如抗菌、抗病毒及抗肿瘤化合物的合成研究。先后参与5项国家自然科学基金,5项省部级科研基金,主持1项省部级科研基金及6项校级基金,参与开发2个国家级新药。申报发明专利3项,授权一项。已在国内外核心期刊上发表学术科研论文40余篇,SCI收录28篇,其中影响IF>2.0的12篇,IF>3.0的6篇。



陈雅茜
副教授

陈雅茜,女,副教授。2007年9月获国家留学基金委(CSC)全额奖学金资助赴德国慕尼黑大学师从Andreas Butz教授攻读人机交互方向博士学位,于2010年7月获得德国慕尼黑大学自然科学博士学位。主要研究兴趣包括人机交互(Human-Computer Interaction, HCI)、信息可视化(Information Visualization, IV)、推荐系统(Recommender Systems, RecSys)。从事过多媒体库的探索式用户体验模式、基于多公共显示屏的信息推荐、多媒体库可视化等课题的研究。目前从事基于物理仿真和空间交互的Tabletop自然交互模型的研究。在高级别国内国际期刊和会议上发表10余篇科研论文,多篇EI、ISTP检索,出版译著一部(机械工业出版社),有着多次参加、组织和评审国际会议的经验,主持省部级科研项目1项、校级项目2项、国家外专项目3项、横向项目1项;主持校级教改项目2项,参研省级教改项目1项,指导大学生创新项目国家级1项、校级2项,指导学生参加四川省大学生计算机作品大赛获得省二等奖。