

一种基于视觉的手指与全息影像交互研究

于瀛洁,李雨浪,郑华东
(上海大学精密机械工程系,上海 200072)

摘要:利用全息技术进行真三维显示是显示技术研究领域中的一个亮点。为了探讨人与三维全息影像的交互问题,提出了基于视觉的手指与全息影像的交互方法。结合背景差分和色度差分检测图像中的人手区域,然后分析手部轮廓曲率定位指尖位置,并通过指尖位置引导人头三维全息实影像的显示,实现了人与全息实影像的动态交互。最后,通过对比指尖坐标的程序检测值与人工标定值,验证了指尖定位的准确性;通过分析指尖位置与人头影像的对应关系,验证了交互的正确性。

关键词:人手检测;指尖定位;三维全息实影像;交互

中图分类号:0438.1 **文献标识码:**A

Research of vision-based interaction method between finger and holographic images

YU Ying-jie, LI Yu-lang, ZHENG Hua-dong

(Department of Precision Mechanical Engineering, Shanghai University, Shanghai 200072, China)

Abstract: The application of holographic technology to display true 3 D images is a bright spot in the display field. In order to investigate the interaction between people and 3 D holographic images, a vision - based interaction method between finger and holographic images is proposed. Combining background difference with color difference method, detect hand area, and then locate fingertip by analyzing hand contour curvature, guide the display of 3D head holographic real images via tracking fingertip, then, the dynamic interaction between people and holographic images can be realized. Finally, validate the veracity of fingertip locating by comparing fingertip positions detected by program and by manual work, and validate the correctness of interaction by analyzing the corresponding relationship between fingertip positions and head images.

Key words: hand detecting; fingertip locating; 3 D holographic real image; interaction

1 引言

全息技术利用光的干涉和衍射记录和重建光波面,在全息图的再现过程中,全息技术不仅能精确地重建光波的幅值信息,也能重建其位相信息,因此全息技术本身就是一种三维显示技术。全息技术的这种特性,使其在三维立体显示中一直备受关注^[1-3]。利用全息技术进行真三维显示是三维立体显示领域中的一个亮点。然而,这项技术还存在一些问题,比如:三维全息实影像的承载介质问题;满足实时要求的计算全息方法问题;显示交互问题等。本文利用

计算全息获得全息图,然后对其进行光电再现,结合计算机视觉检测人的手指位置,引导数字全息光电再现实像的动态显示,在数字全息三维实影像的实时动态交互显示方面做了实验性的研究。

基金项目:国家自然科学基金(No. 60772124);上海市科委科国际合作项目(No. 09530708700);上海大学研究生创新基金(No. SHUCX092192)资助。

作者简介:于瀛洁(1969-),女,博士,研究员,博士生导师,研究领域为工程中的光学检测方法。E-mail:randy-2010@hotmail.com

收稿日期:2009-11-30

2 系统概述

基于视觉的手指与全息影像交互系统由三部分组成:数字全息光电再现系统、离子雾生成器和指尖的视觉检测系统,如图1所示。数字全息光电再现系统能够对数字全息图进行光电再现,并将再现实像投射出来。离子雾生成器能够生成离子雾柱,以此作为承载介质对投射出来的全息实像进行承载,使再现实的影像悬浮在空中,并保持全息像的三维效果,增强像的真实感。在交互上,视觉系统检测固定手型的指尖位置,将检测结果反馈给数字全息光电再现系统以引导再现出来的全息影像,从而实现数字全息实影像的交互显示。

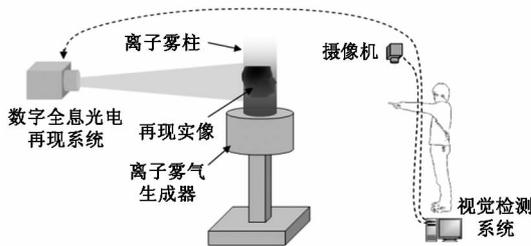


图1 基于视觉的手指与全息影像交互系统框图

3 人手检测与指尖定位

本文提出了基于视觉的手指与全息影像的交互方法,该方法通过手指的移动引导全息实影像的旋转显示,是一种直观友好的交互方式。

3.1 人手检测

为了检测到指尖位置,首先需要从摄像机图像中检测出人手区域。基于视觉的手部区域检测有很多种,比较常用的是基于标记物和标记手套的检测方法^[4-5]、基于红外图像的检测方法^[6]、基于肤色的检测方法^[7]以及基于图像差分的检测方法^[8-9]。在这几种方法中,基于标记物和标记手套的检测方法由于需要给人手佩戴标记物,使其操作自然性大大降低;基于红外图像的检测方法因红外像机昂贵的价格限制了该方法的使用;肤色检测是目前研究得比较多的一种方法,但单纯的肤色检测会受到背景色、环境光、肤色差异等多种因素的影响;图像差分法也只是对静态背景下的手部检测效果比较好。综合考虑各种方法的优势和不足,并注意在与全息影像的交互过程中,光电再现系统再现的是实际光束的影像,投射在离子雾柱上还有部分的透射和散射,会一定程度上改变人手的检测光场,本文提出采用图像差分法和肤色检测相结合的方法对图像中的人手进行检测。首先利用图像差分法分割出图像中

的运动前景,然后再采用肤色检测进一步确定手部区域。

图像差分法可分为背景差分法和帧间差分两种,本文采用图像的背景差分法。背景差分法最重要的是获取参考背景图像,比较简单的方法是取若干帧背景图像的平均值作为背景参考图像,并在适当的时候对其进行更新,本文也采用这种方法来获取背景参考图像。设摄像机获取的图像序列为:

$$f_i(x, y) = (r_i^f(x, y), g_i^f(x, y), b_i^f(x, y)), 1 \leq x \leq N, 1 \leq y \leq M, i = 0, 1, 2, \dots \quad (1)$$

其中, $r_i^f(x, y)$, $g_i^f(x, y)$, $b_i^f(x, y)$ 分别是 (x, y) 处像素点的 RGB 值; N 和 M 是图像的高和宽; i 为图像序列的帧号。

如式(2)所示,在交互开始之前,采集 10 帧图像进行平均:

$$B(x, y) = \frac{1}{10} \sum_{i=1}^{10} f_i(x, y) \quad (2)$$

这样就可以获得最初的背景参考图像。

传统的背景差分法对图像的灰度值进行差分,这种差分对背景参考图像要求很高,如果没有及时对其进行更新,会影响检测效果。为了适应光照因素的影响,本文在 YCbCr 空间中进行色度差分来分割前景^[10]。在 YCbCr 色彩空间中, Y 为亮度, Cb 和 Cr 分量分别表示红色和蓝色的色度。在进行差分运算时,去除 Y 分量,只采用图像的 Cb 和 Cr 分量,令:

$$D_i(x, y) = \alpha \times |Cb_i(x, y) - Cb_b(x, y)| + (1 - \alpha) \times |Cr_i(x, y) - Cr_b(x, y)|, \quad 0 \leq \alpha \leq 1 \quad (3)$$

其中, $Cb_i(x, y)$, $Cr_i(x, y)$, $Cb_b(x, y)$, $Cr_b(x, y)$ 分别表示第 i 帧和背景图像像素点 (x, y) 处的 Cb, Cr 分量, α 为加权因子。

利用式(3),可得到二值化的色度差分图像 $d_i(x, y)$:

$$d_i(x, y) = \begin{cases} 0, & D_i(x, y) < Th \\ 1, & D_i(x, y) \geq Th \end{cases} \quad (4)$$

式中, Th 为分割阈值。

图2所示为灰度差分效果与色度差分效果的比较,可以看到利用色度差分能够消除光影对差分检测的影响。另外,如图2(c)所示,当背景中物体的灰度值与皮肤灰度值比较接近时,检测到的手部区域会存在很大的孔洞,而色度差分也能一定程度上减弱这种影响。



(a)背景参考图像 (b)第495帧图像 (c)灰度差分效果 (d)色度差分效果

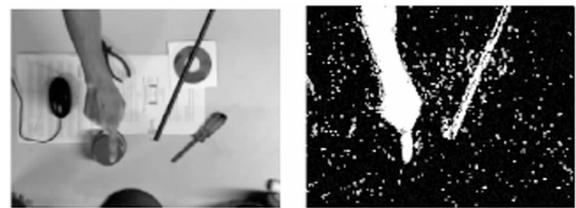
图2 色度差分与灰度差分比较

利用色度差分法已经能够在一定程度上检测出图像中的人手区域,然而如果图像中有非肤色的运动物体出现,如图3所示,会对检测造成影响。为了进一步确定图像中的人手,可以通过肤色检测进一步锁定图像中的人手区域。在YCbCr空间中,肤色在CbCr平面内具有良好的聚类特性,可以利用简单门限法对肤色区域进行提取,即:

$$s_i(x,y) = \begin{cases} 1, & 77 \leq Cb_i(x,y) \leq 127, 133 \leq Cr_i(x,y) \leq 173 \\ 0, & \text{其他} \end{cases} \quad (5)$$

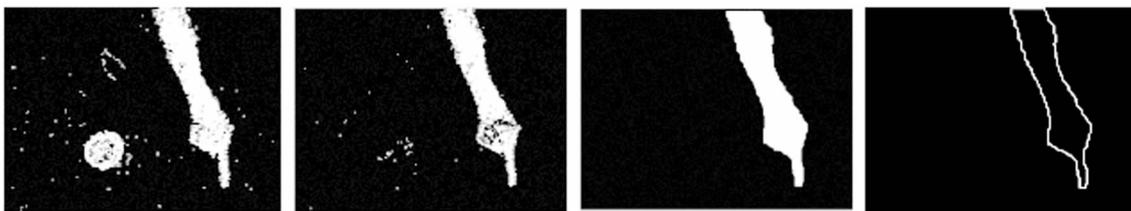
其中, $s_i(x,y)$ 表示肤色检测后得到的二值化图像;

$Cb_i(x,y)$ 和 $Cr_i(x,y)$ 表示第 i 帧图像的Cb, Cr分量。图4(a)为对图2(b)所示的图像进行肤色检测后的效果。



(a)有非肤色物体干扰时的图像 (b)色度差分效果

图3 有非肤色物体出现时的检测效果



(a)肤色检测效果 (b)与运算后结果 (c)滤波后的效果 (d)人手轮廓曲线

图4 结合肤色检测得到的效果

肤色检测结果 $s_i(x,y)$ 与色度差分结果 $d_i(x,y)$ 进行图像与运算,得到最终的检测结果 $r_i(x,y)$, 即:

$$r_i(x,y) = s_i(x,y) \cdot d_i(x,y) \quad (6)$$

图4(b)所示为与运算后的结果。可以看到,与运算后的图像中除了手部区域之外,已经没有了其他物体的影响。但图像中还存在噪点,而且手部区域也不完整。通过中值滤波,图像的膨胀和腐蚀操作可以去除图像中的噪点,并填补手部区域出现的孔洞。手部区域最终的检测结果为图像中最大的连通区域,如图4(c)所示。图4(d)为提取出的手部轮廓曲线。

为了进一步增强检测效果,需要对背景参考图像进行更新。背景图像的更新可在图像中最大的连通区域检测之后进行,如果图像中的最大连通区域的面积小于一定的阈值,就可以认为图像中没有带肤色的前景出现,此时对背景参考图像 $B(x,y)$ 进行更新^[11]:

$$B(x,y) = (1 - \beta)B(x,y) + \beta f_i(x,y), \quad 0 \leq \beta \leq 1 \quad (7)$$

其中, β 为更新速度系数。

3.2 指尖定位

提取出手部轮廓曲线后,可以根据此轮廓的曲率进行指尖位置搜索^[12],通过对手部轮廓按轮廓点顺序进行定长扫描,将满足公式(8)的点设为指尖候选点:

$$\text{ratio} = D_p / D_{ab} \geq \varepsilon \quad (8)$$

其中, D_{ab} 为扫描轮廓起始点 a 与终点 b 连线 AB 的长度; D_p 为扫描轮廓中点 p 到 AB 的垂直距离。参照图5(a), ε 为比值 ratio 的最小阈值(本文中取 $\varepsilon = 1.5$)。

比值 ratio 大于 ε 的情况下, p 点被设置为指尖候选点。对于图4(d)中的轮廓曲线,按顺序进行扫描,计算各轮廓点处的 ratio 值,如图5(b)曲线所示, X 轴为轮廓点索引值, Y 轴为各轮廓点处的 ratio 值,曲线中可以明显看到一个峰值,此峰值所在的轮廓点即为指尖点,图5(c)所示为指尖的定位结果。

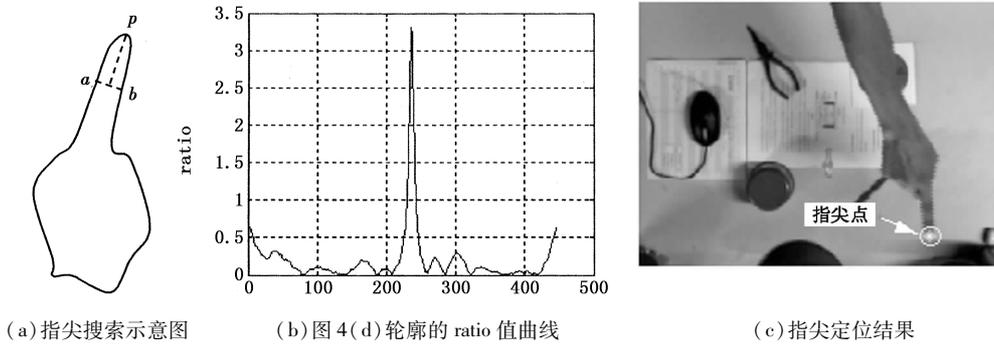


图5 指尖定位

当手指比较靠近图像边界,如图6所示的情况,由于图像边界对检测的影响,手部轮廓的ratio值曲线会出现两个峰值,这时候指尖的定位可以通过去



图6 手指靠近图像边界时的指尖定位

除图像边界附近轮廓点的ratio值实现。如此,指尖点定位的正确率可达96%以上。

4 与全息影像交互的实现

本文所设计的交互是手指与人头全息实影像的交互,通过手指在图像平面内的圆周运动带动人头全息影像的旋转,如图7所示。图中, θ 为指尖与图像坐标原点的连线和图像坐标系X轴的夹角; φ 为人头的旋转角度。

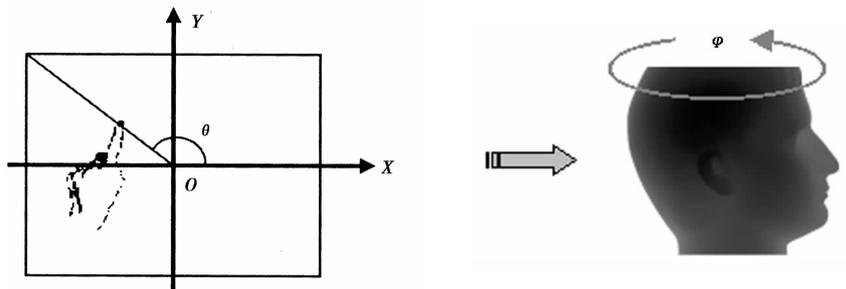


图7 手指引导人头全息影像旋转示意图

在交互的实现上,将人头旋转一周离散化为60个不同角度的序列图像,每相邻两幅图像中人头夹角为 6° 。然后利用计算全息的方法,计算好这60幅不同角度人头图像的数字全息图序列,在交互的时候,只要跟据指尖图像的 θ 值对相应的数字全息图进行光电再现即可:

$$\varphi = g(\theta) = \text{fix}(\theta/6^\circ) \times 6^\circ \quad (9)$$

其中,fix表示向零取整。

5 交互实验

5.1 实验环境

本文以 HOLOEYE Photonics AG 公司的一款 PLUTO-VIS 空间光调制器(分辨率 $1920 \text{ pixels} \times 1080 \text{ pixels}$,反射模式,最大刷新速率 60 Hz)为核心器件建立数字全息光电再现系统,如图8所示。并将光电再现出来的全息实影像投射到图9(a)所示的离子雾柱上,就可以得到悬浮于空中的人头全息实影像,图9(b)所示是当人头处在 0° 时再现出来的全息影像效果。

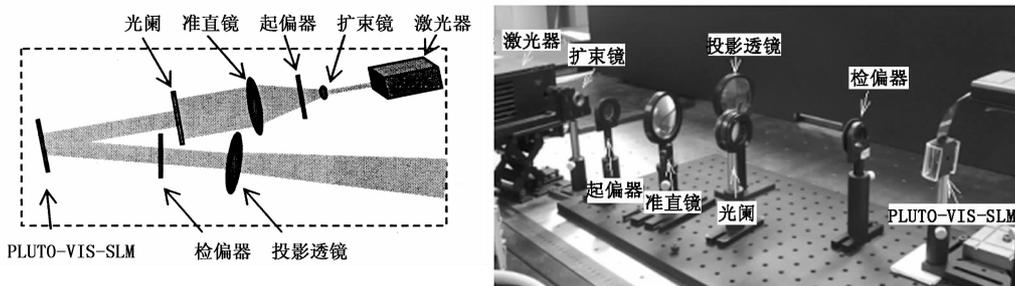


图8 数字全息光电再现系统光路图及实际照片



(a) 离子雾柱生成器



(b) 人头的全息影像

图9 离子雾柱生成器和人头的全息影像

5.2 与人头全息影像的交互

在交互过程中,如图 10 所示,当手指在水平面内做顺时针圆周运动时,离子雾柱中的人头全息影像跟随着手指做顺时针旋转;反之,人头全息影像则做逆时针旋转;当手指静止后,全息影像也停止旋转。

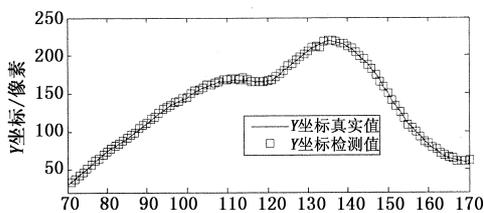
5.3 实验结果及分析

用 VC++6.0 实现了本文所采用的人手检测及指尖定位方法,并在一台 CPU 为 Pentium IV 2.80 GHz,内存 512 M 的普通 PC 机上进行了实验,图像采集设备为 JVC 公司的 TK-C1480BEC 摄像机和大恒公司的 CG410 图像采集卡,视频图像分辨率为 300×224 像素,处理速度约为 19 f/s 图像,基本上能满足实时交互的需求。另外,通过程序自动检测的指尖坐标与人工检测的指尖坐标进行比较,分析连续 100 帧图像的指尖定位效果,

得到的结果是: X, Y 坐标的平均误差分别为 0.94 和 1.22 像素,标准误差分别为 0.75 和 1.25 像素。图 11 是系统检测的指尖位置与人工标定真实值的对比结果(第 70 ~ 170 帧)。

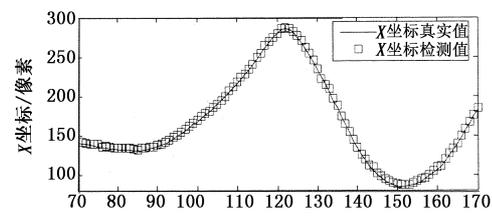


图 10 交互照片



图像序列的帧编号

(a) Y 坐标比较



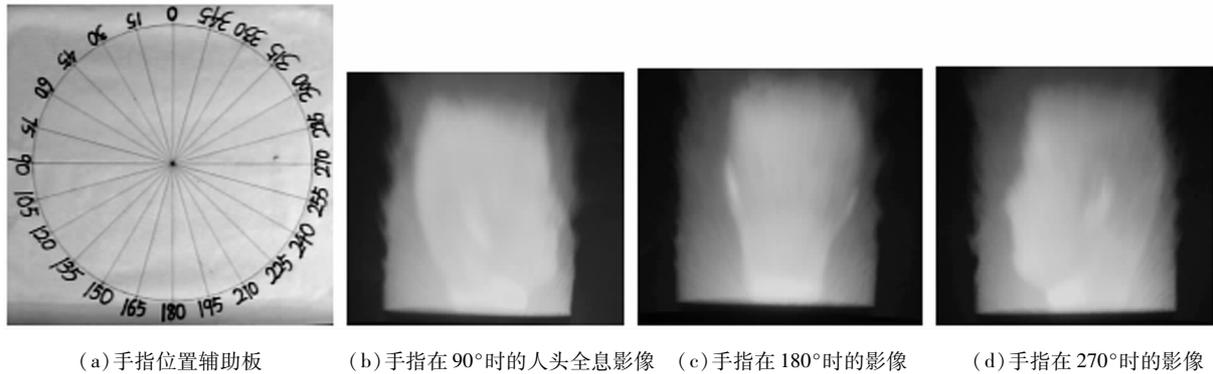
图像序列的帧编号

(b) X 坐标比较

图 11 检测值与真实值的对比结果

为了进行交互正确性的验证,本文采用一个手指位置板来辅助分析,如图 12(a)所示,在分析时,根据手指在位置板上的位置来判断全息影像显示的正确性。根据验证,手指可以比较稳定地引导全息

影像的显示,但有时会出现偶然的跳变,这主要是由于手指检测失败引起的。图 12 的其他几幅图分别是手指处在位置板 $90^\circ, 180^\circ$ 和 270° 时的人头全息影像。



(a) 手指位置辅助板 (b) 手指在 90° 时的人头全息影像 (c) 手指在 180° 时的影像 (d) 手指在 270° 时的影像

图 12 交互的正确性验证

6 结 论

虽然利用全息技术来实现真三维显示目前还存在层层困难,但随着新材料、新技术以及计算机技术的发展,这项技术一定会走向实用化。在显示技术领域,交互是一个重点问题。为了探讨人与全息影像的交互问题,本文提出了一种基于视觉的手指与全息影像的交互方法,增强了交互的友好性。结合图像差分法和肤色检测法检测图像中的人手区域,并对指尖进行定位,通过指尖位置引导全息实影像的显示,实现了人与全息实影像的动态交互功能。然而,本文所采用的人手检测和指尖定位方法,虽然在静止背景和光场有一定变化的条件下,能稳定地对指尖进行定位,并可以一定程度上抵抗非肤色移动前景的影响,但当背景有变化或出现类似肤色的移动前景时,检测就会失败。此问题的解决可以通过采用更为有效的背景更新算法和加强手部区域的跟踪来达到。在交互上,由于目前的计算全息方法得到数字全息图很耗时,所以不可能在交互过程中根据指尖位置实时计算全息图,只能预先计算好全息图,在交互过程中调用相应的全息图进行显示,限制了交互的多样性。因此,真正的时实交互有待采用更为有效的计算全息方法或采用硬件计算方法来实现。

参考文献:

- [1] M Lucente. Interactive three-dimensional holographic displays: seeing the future in depth[J]. Proc. Siggraph Computer Graphics, 1997, 31: 63 - 67.
- [2] Maurice Stanley, Mark AG Smith, Allan P Smith, et al. 3 D electronic holography display system using a 100 Megapixel spatial light modulator [J]. Proc. SPIE, SPIE-Int'l Soc. for Optical Eng. , 2003, 5249: 297 - 308.
- [3] Christopher Slinger, Colin Cameron, Stuart Coomber, et al. Recent development in computer-generated holography: toward a practical electroholography system for interactive 3D visualization [J]. Proc. SPIE. , SPIE-Int'l Soc. for Optical Eng. , 2004, 5290: 27 - 41.
- [4] K Dorfmueller-Ulhaas, D Schmalstieg. Finger tracking for interaction in augmented environments [J]. IEEE Computer Society, 2001: 55 - 64.
- [5] K Lanhans, D Bahr, D Bezecky, et al. FELIX 3 D display: An interactive tool for volumetric imaging [J]. Proceedings of SPIE, 2002, 4660: 176 - 190.
- [6] K Oka, Y Sato, H Koike. Real-time tracking of multiple fingertips and gesture recognition for augmented desk interface systems [J]. Proceeding of the IEEE International Conference on Automatic Face and Gesture Recognition, 2002: 429 - 434.
- [7] Miaolong Yuan, Farzam Farbiz, Corey Mason Manders, et al. Robust hand tracking using a simple color classification technique [C] // Proceedings of the 7th Acm Siggraph International Conference on Virtual-Reality Continuum and Its Application in Industry, Singapore, 2008.
- [8] Hardenberg Christian Von, Bérard Franois. Bare-hand human-computer interaction [J]. Proceedings of the ACM Workshop on Perceptive User Interfaces, 2001: 1 - 8.
- [9] Song Peng, Winkler Stefan, Gilani Syed Omer, et al. Vision-based projected tabletop interface for finger interactions [J]. Human-Computer Interaction, 2007, 4796: 49 - 58.
- [10] 毛晓波, 谢晓芳, 张晓林. 消除运动物体阴影的最大色度差分检测法 [J]. 电子技术应用, 2007, (1): 61 - 63.
- [11] 张继平, 刘直芳. 背景估计与运动目标检测跟踪 [J]. 计算机技术与自动化, 2004, 23 (4): 51 - 54.
- [12] 王西颖, 张习文, 戴国忠. 一种面向实时交互的变形手势跟踪方法 [J]. 软件学报, 2007, 18 (10): 2423 - 2433.