

基于 ARToolKit 手指交互系统的研究

盛 君, 王毅刚*

(杭州电子科技大学 计算机学院图形图像研究所, 浙江 杭州 310018)

摘要:针对传统交互方式依赖硬件交互设备问题,采用 ARToolKit 开发包进行增强现实系统二次开发,提出了一种更为方便、自然的交互方式—用指尖触碰虚拟菜单来实现对菜单命令的交互。即在图像空间,通过用户指尖识别,判断其是否落在虚拟菜单区域内来实现对其他虚拟物体的交互。实验结果表明,系统在光照适中的情况下具备良好的实时效果,并为今后进一步的研究提供了发展的平台。

关键词:增强现实;手指交互;指尖检测

中图分类号:TH7;TP391

文献标识码:A

文章编号:1001-4551(2010)06-0116-04

Research on finger interactive system based on ARToolKit

SHENG Jun, WANG Yi-gang

(Institute of Graphics and Image, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: Aiming at the traditional interactive methods dependent on hardware equipments, using ARToolKit development package for secondary development of augmented reality system, a more natural approach was proposed—using finger-tip touch the virtual menu to achieve interaction. The action was decided by judging whether the finger-tip was in the virtual menu areas through finger-tip detecting in image space. Experimental results demonstrate that the system has a good result when intensity of light is moderate.

Key words: augmented reality; finger interaction; finger-tip detection

0 引 言

增强现实(Augmented Reality, AR)是虚拟现实技术(Virtual Reality, VR)的一个重要分支,是一个多学科交叉的新型研究领域。它借助计算机图形技术和可视化技术产生虚拟对象,并将虚拟对象准确置于真实环境中,借助显示设备使其融合于真实世界中^[1]。随着增强现实技术的不断进步,研究领域得到了新的拓展,目前在工程设计、现代展示、城市规划、教育娱乐、文化遗产等领域都已有了新的应用。

迄今为止,AR 应用中有许多成功的系统实例,如华盛顿大学开发的 AR 桌上系统^[2]、magicbook^[3],以及浙江大学的景观规划系统^[4],它们侧重于向用户提供一个展示浏览的平台。但是随着技术的发展,用户不仅仅满足于普通的浏览以及效果显示,而是更趋向于自然的实时交互。目前最常见的交互类型有 2 类:①传统的交互设备:鼠标、键盘以及菜单,如利用鼠标把虚拟物体放置到 AVI 视频帧的指定位置^[5];在掌上电脑中利用传统菜单来实现与虚拟物体的交互^[6];②利用一些便携式设备进行交互,如校园浏览系统中,用户背着移动 PC 和跟踪设备,头戴

收稿日期:2009-11-02

作者简介:盛 君(1985-),男,浙江杭州人,主要从事虚拟现实、图形图像方面的研究。E-mail: shengji1985@hotmail.com

通信联系人:王毅刚,男,教授,硕士生导师。E-mail: yingang.wang@hdu.edu.cn

透视式头盔显示器,手持手写笔和 PDA,在校园中漫步。AR 系统通过摄像头捕捉到校园场景,并在这些场景中添加建筑物的相关信息,从而帮助使用者了解校园情况^[7]。

以上两种交互方式,均借助于硬件交互设备来达到交互的目的,并且后者的硬件设备价格昂贵。笔者为此提出一种新的交互方式:在 AR 系统中用手指来代替原本的硬件设备,即手指交互。

1 基于 ARToolKit 手指交互系统

基于 ARToolKit 的手指交互系统是一个简单的增强现实交互系统,它利用 ARToolKit 和 OpenCV 开发包进行增强现实系统二次开发。系统使用虚拟菜单对机器小人进行控制,用户通过点击虚拟菜单中“MODEL”选项来显示机器小人,点击“PLAY”菜单项时,机器拟小人则会运动起来,而当按下“STOP”选项时,机器小人则会停止运动。

该系统原理简单,首先根据 ARToolKit 标记板进行虚拟菜单的配准,即三维注册;其次用户手指点击虚拟菜单实现交互:通过识别用户手指指尖是否在“虚拟菜单”区域内,判断是否点击了该菜单。如果指尖不在区域内,即使手指其余部分在区域内,则仍然表示没有点击该菜单。如果指尖在区域内,则表示选中,根据菜单内容,机器小人进行相应的操作。

根据实现过程,系统主要分为视频采集模块、跟踪注册模块、虚拟物体生成模块、虚实融合模块、显示模块、交互模块共 6 大模块,如图 1 所示。其中,视频采集模块主要负责真实场景的视频采集,获取场景中的信息;跟踪注册模块是用于跟踪摄像机的位置和方向;虚拟物体生成模块的作用是根据交互模块给予的指令生成三维虚拟物体;虚实融合模块是负责将虚拟信息与真实环境的信息进行配准,实现虚实融合;显示模块负责虚实融合后的显示;交互模块主要负责用户与系统的交互处理,包括指尖检测、虚拟菜单项投影以及区域检测。

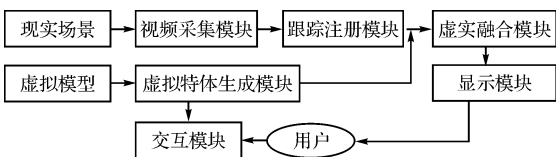


图 1 基于 ARToolKit 的手指交互系统功能结构图

2 系统实现

2.1 三维注册

如何实现三维注册是增强现实系统面临的一大挑战。系统采用 ARToolKit 开发库中原有的注册算法来进行虚实配准。首先将每帧彩色图像灰度化,在灰度图中,设定一个阈值,并将其转化成黑白二值图像,对该图像进行连通域分析,找出所有的四边形区域,并将其作为候选域,再将该候选区域与模板库中的模板进行匹配,若匹配,则认为其是标记,利用该标记区域 2D-3D 的变换来计算摄像机相对于已知标记的位置和姿态,最后根据得到的变换矩阵实现虚拟物体的注册^[8]。

而在整个过程中,如何计算标识坐标系到摄像机坐标系的矩阵转化是最为关键的,两者的转化过程为(如图 2 所示)^[9]:

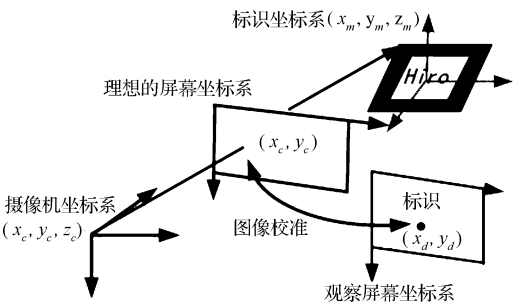


图 2 ARToolKit 坐标系

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{O} & 1 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ z_m \\ 1 \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ z_m \\ 1 \end{bmatrix} \quad (1)$$

式中 $(\mathbf{R} \ \mathbf{T})$ —摄像机外参数; $\mathbf{R}$ —旋转矩阵; $\mathbf{T}$ —平移向量。

2.2 手指交互

为了准确、快速地检测出手指指尖位置,系统采用的指尖检测方法是基于以下条件所提出的:在交互过程中,手指尽量保持笔直状态,稍许弯曲程度可接受,不考虑手指弯曲程度过大的情况;并且在交互过程中,实验环境中保持场景中不存在颜色与肤色相近的物体。

系统采用 Dhawale 等^[10]使用的肤色检测方法进行人手区域的检测。该方法基于人手肤色的先验知识,对 RGB 3 个颜色分量进行分析,满足下式的像素

被认为是人手区域(其中, R 、 G 、 B 分别表示 8 位的红色、绿色、蓝色亮度):

$$\begin{cases} (R > 40) \wedge (G > 20) \wedge (B > 10) \\ \max(R, G, B) - \min(R, G, B) > 10 \\ \text{abs}(R - G) > 10 \\ (R > G) \wedge (R > B) \end{cases} \quad (2)$$

然后,在肤色检测结果的基础上,先将图像二值化,然后进行去噪声以及区域边缘平滑处理。

由于手指尽量保持笔直,手指指尖可视为连接两平行线段的小曲线段中的最大曲率点。系统在应用 K 曲率算法的基础上,利用这一特性找到指尖位置,在所有的极大值点中排除了其余不符合该条件的极值,获得手指指尖的准确位置^[11]。

由于系统的虚拟菜单是用 3D 文字显示,每个菜单项为一个 3D 包围盒,只要判断出手指指尖位置是否落在虚拟菜单的投影区域内即可判断出用户是否点击了虚拟菜单项。

记点 P 为 3D 包围盒的顶点,在标识坐标系下坐标为 $X_m = (x_m, y_m, z_m)^T$;在摄像机坐标系中记为 $X_c = (x_c, y_c, z_c)^T$;在实际图像坐标系中记为 $m' = (x_d, y_d)^T$ 。

利用 ARToolKit 得到的相机相对于标记的位姿($R \quad T$)(见 2.1 节中公式(1)),然后将虚拟菜单(含该 3D 文字的包围盒)从 3D 空间投影到图像坐标系:

$$\begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ z_m \\ 1 \end{bmatrix} \quad (3)$$

式中 f_x, f_y —相机在 x, y 轴方向上的尺度因子(u_0, v_0)为主点; s —相机的畸变因子。

3 实现结果与分析

3.1 实验环境

该系统硬件设备简单,仅需要一个普通摄像头、一台 PC 机以及一个 ARToolKit 标记即可。笔者进行实验的配置为:

操作系统:Windows XP;

CPU: Intel Pentium 4 530;

内存:1 GB;

摄像头:Teclast WE-MS49-G35;

图像空间分辨率为 320×240 像素;

ARToolKit 标记:80 mm $\times$ 80 mm。

3.2 系统运行效果

系统初始状态时仅显示虚拟菜单,并用黑色字体显示,当用户选中一个菜单内容时,该菜单项用灰色字体标明,并实现该选项内容。系统运行结果显示,在室内光线适中、无复杂光影投影在标识上,灰度阈值设置为 120 时,系统运行良好。其效果如图 3 所示。

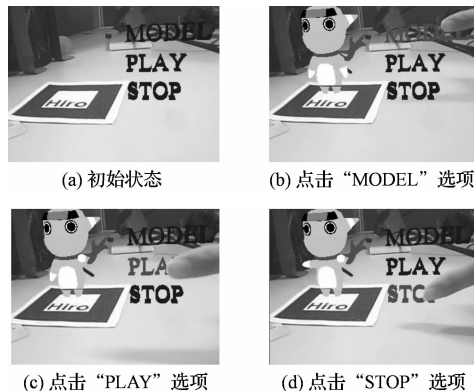


图 3 运行效果图

3.3 实验结果与分析

3.3.1 指尖检测结果

系统所提出的手指指尖交互方法是基于以下两点提出的:

(1) 在实验环境中,保持场景中不存在颜色与肤色相近的物体。

(2) 手指在交互操作时,尽量保持手指笔直状态,并且在交互过程中,手指移动的速度不宜过快。

因此只在该条件的情况下,对实验结果进行了一些考察。

指尖检测的最终显示结果如图 4 所示:去噪声和平滑处理后的二值化图像如图 4(a)所示;彩色图像如图 4(b)所示,手指用黑色轮廓显示,指尖位置则用十字区域标记。

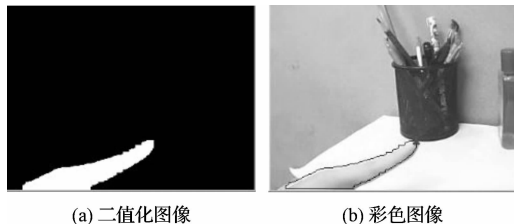


图 4 指尖检测

3.3.2 光照对实验结果影响

ARToolkit 将彩色图像进行灰度化,在灰度图中设

置一个阈值,再将其转化为黑白二值图像。笔者在不同的光照情况下,对该阈值的选取进行了实验分析(摄像头和标记板中心的连线与标志板所在平面夹角成 45°):在室内光线较暗的情况下,阈值设置为45时,系统识别标识的成功率为80%,当阈值在50~240范围内时,成功率达到95%以上;当光线较适中时,阈值取值为70,系统识别标识的成功率为80%,当阈值在75~240范围内时,成功率达到95%以上;光线明亮时,阈值设置为80,系统识别标识的成功率为80%,当阈值在85~240内时,成功率达到95%以上;当室内光线太强烈或者太昏暗,有复杂光影在标识上时,标识识别的成功率低于30%。由此可见光照情况对阈值的选取有很大的影响。

因此建议系统在光线适中的室内环境中运行,尽量将标记板放置在没有复杂光影的地方;用户可根据室内环境来自动设置阈值。

3.3.3 系统的运行时间

系统运行过程中,测定视频每帧所需时间的实验可分以下情况(如表1所示):

- (1)系统在交互过程中,用户不参与交互、只显示三维物体的情况;
- (2)用户参与交互的状态,即系统需要进行交互模块的处理:指尖检测和区域判断等。

系统分以上两组情况进行实验,每组100次实验,每次实验针对连续的500帧视频。其实验统计结果如表1所示。

表1 系统在各状态下视频每帧处理时间		
用户所处状态	不参与交互状态	交互状态
视频每帧所需时间/ms	32	68

由表1可见,系统能达到较为良好的实时效果。

4 结束语

有别于AR交互系统中采用硬件交互设备,笔者采用ARToolKit开发包进行二次开发,利用用户手指来进行“虚拟菜单”交互,它不借助任何硬件交互设备来实现与虚拟物体的交互,受到空间距离的限制较少,交互过程更加自然方便。

由于指尖提取容易受到光线以及外界其他因素的影响,在交互的过程中,系统对手指的状态以及室内环

境做了一些限制,但系统交互过程仍然较方便,并且运行效果良好。

参考文献(References):

[1] 刘越,王涌天,闫达远. 增强现实系统显示技术研究[C]//第三届全国数字成像技术及相关材料发展与应用学术研讨会. 北京:[s. n.],2004:18-32.

[2] KATO H, BILLINGHURST M, POUPYREV I, et al. Virtual Object Manipulation on a Table-Top AR Environment[C]//Proceedings of the International Symposium on Augmented Reality. Munich, Germany: [s. n.],2000:111-119.

[3] BILLINGHURST M, KATO H, POUPYREV I. The Magic-Book: A Transitional AR Interface[J]. **Computers and Graphics**,2001,25(5):745-753.

[4] 王靖滨,李明,耿卫东,等. 基于增强现实技术的景观规划系统[J]. 中国图象图形学报,2002,7(4):405-410.

[5] 黄有群,姬永成,李丹. 基于ARToolKit工具的增强现实交互操作研究[J]. 计算机与现代化,2008(9):97-100.

[6] WAGNER D, SCHMALSTIEG D. First Steps towards Hand-held Augmented Reality[C]//Proceeding of the 7th International Conference on Wearable Computers. White Plains, NY, USA: [s. n.],2003: [s. n.].

[7] HOLLERER T, FEINER S, TERAUCHI T, et al. Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system[J]. **Computers and Graphics**,1999,23(6):779-785.

[8] 任波,管涛,李利军,等. 基于ARToolKit的增强现实系统开发与应用[J]. 计算机系统应用,2006(1):81-84.

[9] KATO H, BILLINGHURST M. Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System[C]//Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality. San Francisco: IEEE CS Press,1999:85-94.

[10] DHAWALE P, MASOODIAN M, ROGERS B. Bare-hand 3D Gesture Input to Interactive Systems[C]//Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI. Christchurch, NZ: ACM Press,2006:25-32.

[11] 赵永峰,王毅刚. 采用手指的三维人机交互方法[J]. 微电子学与计算机,2009,26(6):21-24.