

文章编号: 1007-2780(2013)01-0064-07

基于光学的真三维触控定位与识别方法研究

田 晶¹, 田 丰^{1*}, 凌 晨¹, 王 鹤², 张文俊¹

(1. 上海大学 影视艺术技术学院 影视工程系, 上海 200072, E-mail: jing.tian133@gmail.com;

2. 通用电气医疗集团 应用科学实验室, 上海 201203)

摘 要: 通过介绍真三维显示技术与无标记空间定位原理, 提出了真三维两点触控实现方法, 搭建了真三维两点触控定位与识别系统。针对真三维空间定位绘制、空间两点跟踪和手势识别问题, 研究了真三维两点触控交互方法。实验证明, 在不需要佩戴任何标记和传感器的条件下, 基于光学的空间两点触控方法与系统能够满足真三维显示的人机交互要求, 且环境光的变化不干扰人机交互的操作。

关 键 词: 真三维; 光学触控; 空间定位; 手势识别

中图分类号: TN23; TP237

文献标识码: A

DOI: 10.3788/YJYXS20132801.0064

True Three-Dimension Touch Method for Orientation and Recognition Based on Optics

TIAN Jing¹, TIAN Feng^{1*}, LING Chen¹, WANG He², ZHANG Wen-jun¹

(1. Institute of Film & TV Arts and Technology, Shanghai University,

Shanghai 200072, China, E-mail: jing.tian133@gmail.com;

2. Applied Science Lab in GE Healthcare, Shanghai 201203, China)

Abstract: A two-point touch method for three-dimension space is proposed through introducing true three-dimension display and unmarked spatial orientation, while a two-point touch for true three-dimension display system is built. Interactive approach, which offers the solution of spatial orientation, tracking and gesture recognition, has been considered. Demonstrated in the experiments, optical-based two-point touch system meets the needs of human-computer interaction without any markers and sensors and changes of ambient light does not interfere with the operation of the human-computer interaction.

Key words: true 3D; optical touch; spatial orientation; gesture recognition

1 引 言

真三维显示技术(True 3D Volumetric Display Technique)是裸眼立体显示中最新的研究方向。观察者无需佩戴任何特殊的装置即可在任意角度观察到具有物理井深的立体图像, 它符合人们观察世界的真实感受, 满足所有生理和心理的

深度暗示^[1]。真三维显示技术能提供 360°裸视立体成像, 在此基础上的人机交互技术也引起了研究人员的关注。交互式真三维显示技术在计算机辅助设计、多媒体展示、教学推演和医学影像等众多领域具有广阔的应用前景和重大战略价值。

国内外对真三维显示方面的研究已经取得一定的成果。美国的 Actuality 系统公司已研制

收稿日期: 2012-09-02; 修订日期: 2012-10-17

基金项目: 上海市教育委员会科研创新项目(No. 12YZ008)

作者简介: 田晶(1988—), 女, 山东泰安人, 硕士研究生, 研究方向为人机交互。

* 通信联系人, E-mail: ouman@shu.edu.cn

Perspecta 3D 显示器^[2],可实现水平 360°垂直 180°的裸视立体影像。在国内,浙江大学已开发出基于主动发光 LED 的真三维显示系统^[3],南京航空航天大学先后制成基于振镜的真三维显示系统和基于旋转平板的真三维显示系统^[4-6]。课题组通过 5 年的研发,完成了可用于展览展示的真三维显示器,真三维图像的处理能力达到了实时人机互动的要求。

交互式手势识别已在国内外进行过广泛的研究与实现,比如早期采用数据手套来获得用户手势信号^[7-8],其后又相继有以具有特殊颜色搭配的普通手套为辅助设备的方法^[9]、在指尖佩戴特殊颜色标记的方法^[10]、以单色手套为辅助设备^[11]等方法来实现人机交互。最早关注从多方位与真三维显示系统进行人机交互问题的是 Perspecta 系统的设计师 Tovi Grossman^[12],其方法是通过在手指上佩戴特殊标记在系统半球形外壳上实现基于计算机视觉的空间多点触控人机交互。上述在三维空间内精确定位的交互方法都是通过手持特殊装置或者手指佩戴特殊标记来实现。此外,一些与用户惯用的触控手势不同的专用动作需要用户先进行学习。本文在课题组已有真三维显示器的基础上,提出一种基于光学反射的空间两点触控交互方法,实现无标记条件下的三维空间精确定位和真三维手势识别,增强了真三维立体影像的用户体验。

2 实现原理

经过数个立体显示课题的研究,课题组已获得数套可用于展示的真三维显示器。图 1 为基于投影的真三维显示系统,设计目标是应用在医学影像仿真教育中^[13]。本文研究内容即在图 1 所示真三维显示系统的基础上,基于光学反射原理,根据两组摄像模块的图像来计算三维空间内手指的空间坐标,用户手指上无需安装任何辅助装置便可实现真三维显示的交互操作。如图 2 所示,本系统采用两个基于 CMOS 器件的影像传感器,分别置于红外反射面上方左右两侧,图中红色圆圈内即左右两个光传感器,光传感器外侧安放红外光源,传感器前固定可见光滤光片。

如图 3 所示,基于光学的真三维触控系统由固定于体空间外框的红外发光二极管、CMOS 传感器以和环绕体空间的红外反射材料组成。红

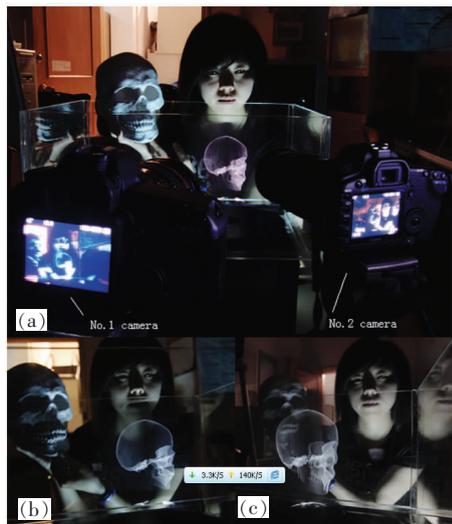


图 1 基于投影技术的真三维显示装置

Fig. 1 True 3D display device based on projector



图 2 真三维两点触控系统结构

Fig. 2 True 3D multi-touch system architecture

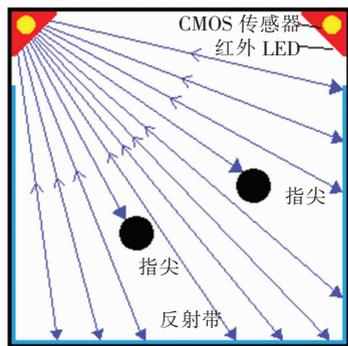


图 3 真三维触控的光学原理

Fig. 3 Optical principle of the 3D touch

外发光二极管与 CMOS 传感器经集成后组成光传感器,完成空间定位至少需要两个光传感器,即左传感器和右传感器。反射功能由位于外框左侧、底部、右侧的 3 个由反射带组成的反射面组成。利用光遮挡原理,通过红外发光二极管发射红外光,红外光线入射到反射面后沿原路径返回,CMOS 传感器接收返回的红外光线。图 3 指示

了左传感器的发光二极管发射的红外光线经回归反射后由 CMOS 接收的整个过程。当指尖进入红外光场时会产生遮挡效应,这些遮挡导致部分光线不能正常反射,从而在 CMOS 获得的图像上出现了手和指尖的阴影和形状。右传感器的探测过程与此类似,左传感器完成探测利用了底部和右侧反光获得回归光线,右传感器探测通过左侧和底部反光获得回归光线。红外光接收传感器获取指尖的遮挡信息,再利用三角定位算法对遮挡点进行定位,从而确定指尖的位置。

利用三角定位算法对遮挡点进行定位,计算得到手指指尖的空间坐标。图 4(a)所示为左摄像机拍摄的手指阴影图像,在计算空间坐标之前,先要进行图像与背景相减、图像二值化和指尖检测的预处理,根据 Dung^[14]等人提出的指尖检测算法,图 4(c)为左摄像机拍摄所得的指尖位置。计算指尖的几何中心在图像中的位置(如图 4(d)中红色十字位置即计算所得指尖中心位置),可分别得到指尖在左右两台摄像机中图像的图像坐标。

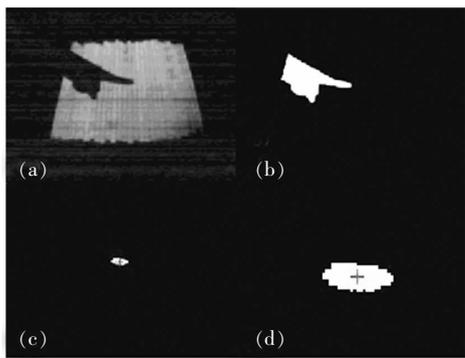


图 4 左侧手指指尖获取
Fig. 4 Left fingertip obtaining

两个摄像机空间定位原理图如图 5 所示。两摄像机成像平面坐标系 u 轴和 v 轴与摄像机坐标系的 x 轴和 y 轴平行,空间中一点 P 在左右两台摄像机获得的图像中的物理坐标分别为 $P_1(u_1, v_1)$ 和 $P_2(u_2, v_2)$,假定左侧摄像机光心处为空间坐标系的坐标原点,由三角几何关系得:

$$\begin{cases} u_1 = f \times \frac{x_w}{z_w} \\ u_2 = f \times \frac{(x_w - l)}{z_w} \\ v_1 = v_2 = f \times \frac{y_w}{z_w} \end{cases} \quad (1)$$

其中, f 为两个摄像机的焦距, l 为两摄像机之间的距离, (u_1, v_1) 和 (u_2, v_2) 分别为点在左右成像平面的坐标, (x_w, y_w, z_w) 为点 P 在空间中的坐标。可计算出 P 的空间坐标如下:

$$\begin{cases} x_w = \frac{u_1 \times l}{u_1 - u_2} \\ y_w = \frac{v_1 \times l}{u_1 - u_2} \\ z_w = \frac{f \times l}{u_1 - u_2} \end{cases} \quad (2)$$

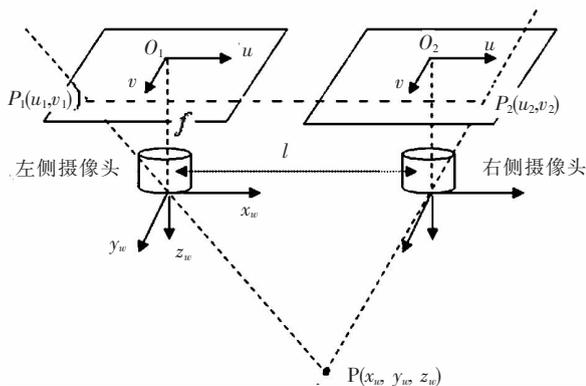


图 5 空间定位原理图

Fig. 5 Schematic of spatial orientation

3 触点跟踪

为了识别三维空间里两个手指的动作,对两个指尖的坐标位置进行跟踪,从而判别两点的移动状态。采用最小距离优先算法(MDF)^[15],分别计算每一帧中两个点的空间坐标,并将前一帧里两点坐标分别记为 $W_1(x_{i-1}^1, y_{i-1}^1, z_{i-1}^1)$ 和 $W_2(x_{i-1}^2, y_{i-1}^2, z_{i-1}^2)$, t 表示当前时刻,将当前帧里两点坐标分别记为 $P_1(x_t^1, y_t^1, z_t^1)$ 和 $P_2(x_t^2, y_t^2, z_t^2)$,计算当前帧中两点与前一帧中两点之间的距离:

$$D(i, j) = \| P_i - W_j \|_2 \quad (1 \leq i, j \leq 2) \quad (3)$$

当前帧中的点与前一帧距离最短的点是同一个手指移动获得。算法流程图如图 6 所示。

如图 6,分别得到当前两点 P_1 和 P_2 的空间坐标,并且分别计算当前 P_1 与前一帧里两个点 W_1 和 W_2 的距离,即 $D(1, 1)$ 和 $D(1, 2)$,判断二者的大小。如果 $D(1, 1)$ 小于 $D(1, 2)$,则说明当前点 P_1 与前一帧点 W_1 为同一个手指移动所得,此时 P_1 和 P_2 坐标分别赋值给 W_1 和 W_2 ,反之,如果

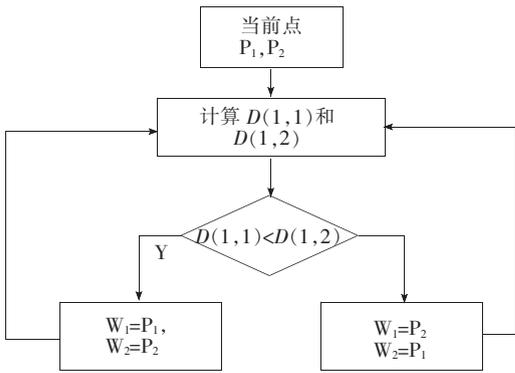


图 6 两点跟踪算法流程图

Fig. 6 Flowchart of two points tracking algorithm

$D(1,1)$ 大于 $D(1,2)$, P_1 和 P_2 坐标分别赋值给 W_2 和 W_1 。

4 激活与坐标转换

当真三维显示装置中出现多个被操作三维模型时,要特定的手势来激活需要被操作三维模型,在激活的同时,还需要对探测到的手指和影像本身进行适当的坐标转换,为激活之后的手势动作识别奠定基础。本文分两种情况来分别介绍激活方法和坐标定位方式,两种情况分别为:单点操作单个影像和两点操作单个影像。

4.1 单点操作单个影像

两组光传感器同时获取到指尖图像投影坐标,从而确定指尖的空间位置坐标。当真三维显示器中存在多个独立的三维模型时,分别计算指尖坐标 $P_i(x^i, y^i, z^i)$ 与各三维模型几何中心坐标 $M_n(x_n, y_n, z_n)$ (n 为三维模型的个数) 的距离,距离最小者即认定为要被激活的三维模型。

每个独立的三维模型都有本体几何中心为原点的模型坐标系 (X_i, Y_i, Z_i) , 此坐标系与用户自定义的世界坐标系 (X_w, Y_w, Z_w) 平行, 因此每个独立的三维模型在根据手势命令所进行的移动或

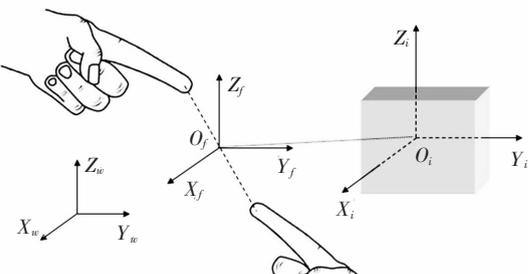


图 7 单点操作过程中各坐标系

Fig. 7 Coordinate systems in one point operating

形变的方向,都是在与其对应的空间坐标系基础上计算。图 7 所示为空间各坐标系之间的位置关系。为保证手指移动方向与三维模型位置变化方向的一致性,当单点激活真三维显示器中的一个三维模型时,被激活三维模型的模型坐标系平移到指尖的空间位置,即激活三维模型后的手指移动均为在以最初指尖位置为原点的空间坐标系 (X_f, Y_f, Z_f) 内的移动。

4.2 两点操作单个影像

实际应用中通常会用到两个手指来控制真三维显示器中三维模型的几何变换,比如缩放、旋转等,因此被激活三维模型的几何变换的基准点要与两个手指的空间位置进行准确的对应。由两组光传感器同时获取两手指初始时刻的空间位置并计算出在世界坐标系 (X_w, Y_w, Z_w) 下的空间坐标,根据上述空间各个坐标系的变换关系,我们规定两个手指指尖在空间连线的中点对应被激活三维模型的几何中心,即被激活三维模型的模型坐标系 (X_i, Y_i, Z_i) 平移到两指尖中点的空间位置,两手指的移动被认为是在以最初两指尖中点位置为原点的空间坐标系 (X_f, Y_f, Z_f) 内的移动。图 8 所示为两点操作下各坐标系之间的位置关系。

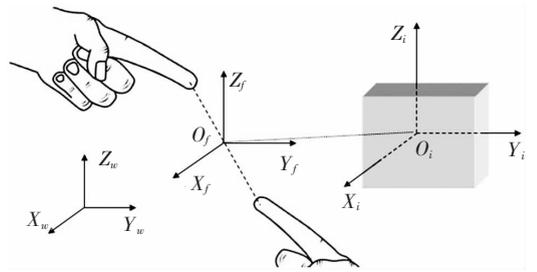


图 8 两点操作过程中各坐标系

Fig. 8 Coordinate systems in two points operating

5 手势识别

真三维显示系统具有 360° 观察视角,用户可在 360° 范围里对真三维显示器中的三维模型进行观察和控制。本文利用真三维显示系统的特点,设计符合用户习惯的触控手势来实现空间绘制和对三维模型的控制。

5.1 单点手势识别

本论文研究了用单个手指来控制被选择激活三维模型的位置平移和用单个手指在真三维显示空间中的绘图。单点操作立体影像和单个手指绘

图需要获取特定的指令信号。当光传感器获取的图像中检测无突出指尖,即一个拳头的图像,如图 9(a),表示单个手指开始绘图和绘图结束命令,图 9(b)表示中心点坐标提取,图 9(c)表示边缘检测,红色部分为提取的图像边缘。检测方法为计算手的中心点坐标与边缘各点坐标的距离,分别记录最大距离 D_{\max} 与平均距离 \bar{D} ,通过判断最大距离与平均距离的比值 D_{\max}/\bar{D} 是否大于规定阈值 τ_0 来判断是否有突出指尖。当检测到上述单个手指激活立体影像的手势(手指指向真三维显示器中的立体影像)并持续规定时间段,表示单个手指控制影像的开始和结束。图 10 所示实验过程中三维空间内指尖绘图过程。

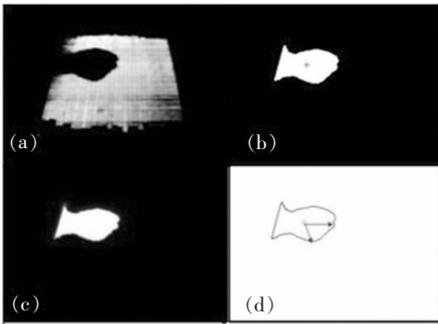


图 9 指尖绘图指令图像

Fig. 9 Image of fingertip drawing instruction

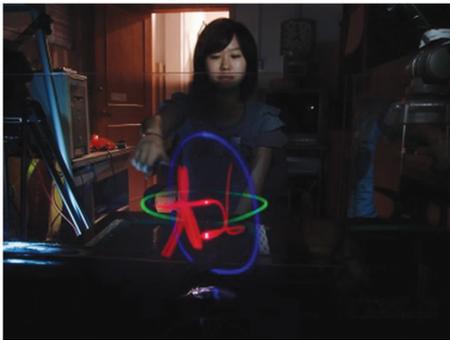


图 10 三维空间内指尖绘图

Fig. 10 Fingertip drawing in 3D space

三维模型的模型坐标系与手指指尖坐标系平行,所以在真三维显示器中三维模型的移动方向与手指的移动方向一致。计算手指指尖的空间坐标,按照一定比例映射到真三维显示器里三维模型的模型坐标系中,则三维模型的运动轨迹与手指指尖的移动轨迹存在一定的比例关系。

5.2 两点手势识别

对真三维显示器里一个选定三维模型的缩

放、旋转操作,通过在空间范围内两个手指的运动手势来实现相应命令的发送,且三维模型的几何变换与两个手指所处的空间位置有关。当光传感器获取的图像中检测出手掌的图像,如图 11(a),表示两点操作立体影像开始与结束信号,图 11(b)表示中心点坐标提取,图 11(c)表示边缘检测,红色部分为提取的图像边缘。检测方法为计算手的中心点坐标与边缘各点坐标的距离,分别记录各距离极值点(即凸点)距离 D_i 、最大距离 D_{\max} 与平均距离 \bar{D} ,当最大距离 D_{\max} 与平均距离 \bar{D} 的比值 D_{\max}/\bar{D} 大于规定阈值 τ_0 时,计算各凸点距离与平均距离的比值 D_i/\bar{D} ,判断比值是否有 5 个或 5 个以上大于规定阈值来判断图像是否有 5 个手指伸出的手掌。对于三维模型的缩放、旋转等几何变换,重点在于基准点的选择,在上述激活与坐标转换方法的描述中,已经规定两个手指指尖空间位置的变化以初始状态两指指尖中点为基准点,对应真三维显示装置中的三维模型的相应变化以其几何中心为基准点。

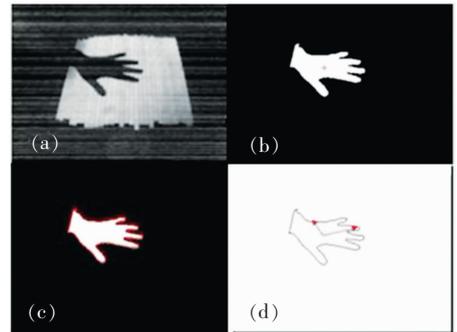


图 11 两点操作指令图像

Fig. 11 Image of two points operating instruction

三维模型的缩放通过两个手指距离的大小变化来实现。考虑到手指在空间运动过程中会有小范围的抖动,因此实验过程中人为设定一个阈值,超出阈值范围的手指运动视为有效运动。三维模型的缩放包括沿所有方向等比例缩放和仅沿一个轴的缩放,本文研究了一种情况,即通过两个手指指尖靠近和远离来实现真三维显示器中三维模型整体的等比例缩放。缩放的大小与手指指尖距离的变化有关,根据上述跟踪算法分别实时计算两个手指指尖空间移动距离 d_1 和 d_2 ,两段距离的平均值 d 的大小决定了三维模型缩放的大小比例。

$$d = \begin{cases} \frac{1}{2} \sum d_f & (1 \leq f \leq 2) \text{ if } |\Delta d| > \tau \\ 0 & \text{if } |\Delta d| \leq \tau \end{cases} \quad (4)$$

其中, Δd 为前后变化的距离差, τ 为阈值。同时两个手指指尖的间距决定了三维模型缩放变化, 当手指间距变小时, 三维模型按一定比例缩小, 当手指间距变大时, 三维模型按一定比例放大。图 12 所示 CMOS 传感器所获得的手指缩放手势。

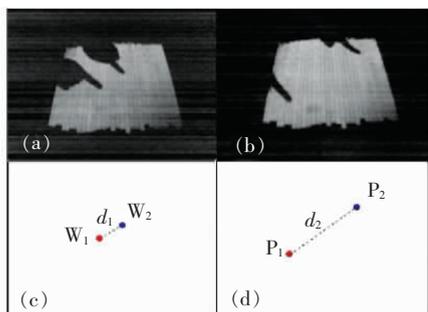


图 12 手指缩放手势

Fig. 12 Gesture of room scaling

三维模型的旋转通过两个手指的旋转来实现, 并且手指的旋转可以控制真三维显示器内的三维模型绕图像坐标系内任意一条直线旋转, 旋转轴的确定由手指移动时空间位置的变化来决定。当两个手指以其中点为旋转中心在三维空间旋转时, 由于手指不可避免地会在空间中出现任意方向小范围的抖动, 设置一定的阈值 τ_r , 两个手指旋转过程中坐标值分别与起始两点坐标距离相差在阈值范围 τ_r 时, 可近似为在一个平面内的旋转, 则旋转轴即过坐标系原点并与旋转平面垂直的直线 l_f 。由于手指运动坐标系与三维模型的图像坐标系平行, 则三维模型随手指运动而旋转的

旋转轴即通过图像坐标系的坐标原点并与直线 l_f 平行的直线 l_m 。如图 13, 图中 α 表示指尖旋转平面, 其中红色和绿色实心点表示当前时刻指尖位置, 空心点分别表示对应的上一时刻指尖位置。

三维模型的旋转方向由两个手指空间运动轨迹来确定, 其旋转方向与手指顺时针或逆时针运动方向一致。指尖的旋转方向用两个指尖旋转过程中空间向量的变化来判断, 如图 14 所示。

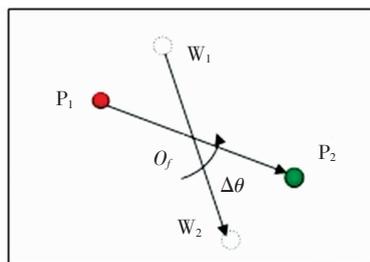


图 14 两点旋转识别

Fig. 14 Recognition of two points rotating

图中 P_1 与 P_2 分别表示当前时刻两点空间坐标, W_1 与 W_2 分别表示相应的前一时刻两点空间坐标, 计算空间向量 $\overrightarrow{P_1P_2}$ 、 $\overrightarrow{W_1W_2}$ 和二者的向量积 n :

$$\overrightarrow{P_1P_2} = (x_i^2 - x_i^1)\mathbf{i} + (y_i^2 - y_i^1)\mathbf{j} + (z_i^2 - z_i^1)\mathbf{k} \quad (5)$$

$$\overrightarrow{W_1W_2} = (x_{i-1}^2 - x_{i-1}^1)\mathbf{i} + (y_{i-1}^2 - y_{i-1}^1)\mathbf{j} + (z_{i-1}^2 - z_{i-1}^1)\mathbf{k} \quad (6)$$

$$\mathbf{n} = \overrightarrow{W_1W_2} \times \overrightarrow{P_1P_2} \quad (7)$$

计算得 n 为正数, 表示手指按顺时针方向旋转, 反之按逆时针方向旋转。三维模型的旋转角度大小通过计算手指旋转角度(即空间向量夹角 $\Delta\theta$)来确定。

6 结 论

在现有的真三维显示系统的基础上, 提出了一种基于光学的两点触控交互方法。在用户无需佩戴特殊标记条件下, 通过光学反射实现了的三维空间内手指指尖的定位, 同时实现了对空间范围内两个触点的跟踪。通过交互式手势的识别方法, 实现指尖空间绘图和对真三维立体影像的平移、缩放、旋转操作的指令解析。

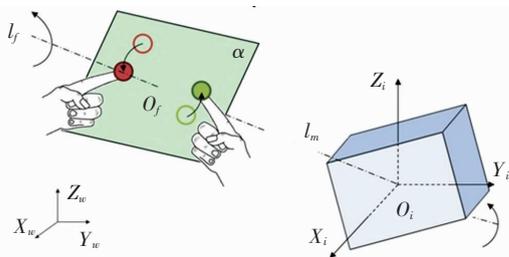


图 13 两点旋转示意图

Fig. 13 Schematic of two points rotating

参 考 文 献:

- [1] Matthew Gately, Yan Zhai, Mark Yeary. A three-dimensional swept volume display based on LED arrays [J]. *Journal of Display Technology*, 2011,7(9):503-514.
- [2] Favalora G E, Dorval R K, Hall D M, et al. Volumetric three dimensional display system with rasterization hardware [J]. *Proceedings of SPIE*, 2001,4297:227-235.
- [3] 林远芳,刘向东,刘旭,等. 基于二维旋转屏的体三维显示系统像素属性分析 [J]. *光子学报*, 2004,33(4):476-487.
- [4] 李莉,李玉峰,沈春林,等. 基于数字微镜的旋转体三维显示装置研究 [J]. *仪器仪表学报*, 2008,29(1):67-72.
- [5] 岳键,樊琼剑,沈春林. 体扫描显示系统中螺旋旋转屏及螺旋切片生成算法 [J]. *中国图象图形学报*, 2011,16(8):1466-1473.
- [6] 邢建芳,龚华军,张赵行,等. 三维数据场在扫描式体显示器中的可视化 [J]. *液晶与显示*, 2012,27(4):529-534.
- [7] Sturman D J, Zeltzer D. A survey of glove-based input [J]. *IEEE Computer Graphics and Applications (S0272-1716)*, 1994,14(1):30-39.
- [8] 李文生,邓春健,吕焱. 基于触摸显示屏的人机交互手势分析 [J]. *液晶与显示*, 2011,26(2):194-199.
- [9] Wang R Y, Popovic J. Real-time hand-tracking with a color glove [J]. *ACM Transactions on Graphics (S0730-0301)*, 2009,28(3):1-8.
- [10] Mistry P, Maes P, Chang L. WUW-wear Ur world: a wearable gestural interface [C]// *Proceedings of the 27th international Conference Extended Abstracts on Human Factors in Computing Systems*. New York, USA: ACM 2009: 4111-4116.
- [11] Just A, Marcel S. A comparative study of two state-of-the-art sequence processing techniques for hand gesture recognition [J]. *Computer Vision and Image Understanding (S1077-3142)*, 2009,113(4):532-543.
- [12] Grossman T, Wigdor D, Balakrishnan R. Multi-finger gestural interaction with 3D volumetric displays [C]// *UIST'04: Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, New York, USA: ACM 2004:61-70.
- [13] 田丰,夏雪,王鹤. 真三维显示在医学教育与仿真中的应用 [J]. *液晶与显示*, 2012,27(4):535-538.
- [14] Nguyen Dung Duc, Pham Thien Cong, Jae Wook Jeon. Fingertip detection with morphology and geometric calculation [C]// *Intelligent Robots and Systems*, St. Louis, USA: The 2009 IEEE/RSJ International Conference, 2009: 1460-1465.
- [15] Cheng H D, Lui Y M. Automatic bandwidth selection of fuzzy membership functions [J]. *Information Science*, 1997, 103:1-21.