

基于增强现实的视触觉交互算法^①

刘佳^② 钱昌宇 谢余磊 陈大鹏 郭斌

(南京信息工程大学自动化学院 南京 210044)

摘要 视触觉增强现实(VHAR)技术可在虚实结合的场景中实现力触觉交互。本文基于 Geomagic Touch 触觉设备提出了一种新的视触觉增强现实交互算法,以快速特征点提取和描述算法(ORB)与角点光流跟踪算法(KLT)算法搭建基于自然特征的增强现实环境,并以此为框架融入触觉反馈接口,实现视触觉协同反馈的交互环境。该算法首先对真实场景中的触控笔进行虚拟注册,然后基于触觉设备的前向运动模型,实现虚拟触控笔与真实触控笔的 6-DOF 协同运动;最后通过虚拟触控笔实现对虚拟物体的操控,实现了更加真实、自然的视触觉人机交互。实验结果表明,此方法具有良好的实时性、可行性和准确性。

关键词 视触觉增强现实(VHAR); 6-DOF; 空间转换; 真实交互

0 引言

增强现实(augmented reality, AR)是由虚拟现实(virtual reality, VR)技术发展延伸出来的技术领域,它突破了虚拟现实只能构建虚拟场景的局限,借助视觉技术、计算机图形学和人体交互等技术将计算机生成的虚拟信息无缝地注册到真实环境。在增强现实实现的过程中,虚拟元素必须完美地注册到目标对象上,这个过程的任何错误都会极大地降低系统的可用性,并影响用户交互的真实感^[1]。因此,与 VR 相比,AR 系统具有更高的复杂性。

目前纯视觉的技术并不能很好地满足用户的交互感和沉浸感,为了实现多通道的感知交互,研究人员提出了视触觉增强现实(visuo-haptic augmented reality, VHAR)的概念^[2-4]。VHAR 以增强现实系统为框架并添加触觉反馈接口,使用户能够看到并触摸到注册到真实环境中的虚拟物体。所以精确地校准 VHAR 系统的组件(如外部跟踪器、摄像头、触觉设备)以及它们之间的空间关系显得至关重要,尤

其在触觉设备的集成上,其核心挑战是精确确定触觉设备上触控笔的位置和方向,这是精确定位视觉增强和触觉反馈所必需的。在传统的视触觉增强现实的案例中,用户看着屏幕,触觉刺激和视觉刺激以一种分开的方式呈现,这导致了环境的割裂感。因此,本文提出一种协同共址定位的方法,即将触觉设备集成到增强现实环境中,触觉交互点与三维场景中的接触点在视觉上重合,构成了一种以同步的方法感知视觉和动态触觉刺激的视触觉增强现实体系,用户可以在相同的空间位置看到和触摸虚拟物体,增加了交互的真实感。

为了提供真实的用户体验,本文触觉设备采用 Geomagic Touch 来提供力反馈和本体运动的感觉元素,主要关键技术采用计算机图形和触控笔的精确协同共址定位。触觉设备由触控笔、手柄设备和两个相互连接的关节组成,这些关节的角度定义了万向节的位置,通常称为触觉接口点(haptic interface point, HIP)。触控笔可以绕着底座旋转,并利用 3 个关节上的编码器来计算其位置。以往的研究主要

① 国家自然科学基金(61773219)和江苏省产业前瞻与关键核心技术重点项目(BE2020006-2)资助。

② 女,1981年生,博士,教授;研究方向:计算机视觉,增强现实,机器人技术;E-mail: liujia@nuist.edu.cn (收稿日期:2020-10-25)

集中在协调系统内的触觉工作空间的校准和补偿,该系统用来减小关节角度传感器误差引起的非线性位置误差。在触觉设备集成到增强现实系统中时,往往会导致校准误差。如果虚拟工具错误地放大到触控笔手柄上,真实和虚拟物体不自然地在触觉手柄上相交,用户可以明显看到校准错误。以往交互主要集中在用户通过装置实现 3-DOF (degree of freedom) 的平移运动,但是随着交互技术的提高,3-DOF 的交互在旋转、执行抓握等多自由度操作过程中无法实现,显著影响了交互的沉浸感。而 6-DOF 不仅具备在 X 、 Y 、 Z 三轴上平移的能力,而且具备在 X 、 Y 、 Z 三轴上旋转的能力。本文的主要工作是准确地实现虚拟触控笔与真实触控笔的融合,使得触觉设备上的交互点与 3D 场景中的接触点在视觉上重合,并提出了适用于 6-DOF 应用场景的视触觉交互算法,即将场景中触觉设备的运动状态作为直接的交互方式融入到增强现实框架中,实验人员通过手握触控笔移动虚拟物体,可以明显感受到来自虚拟物体的反作用力,目的是提高用户交互体验的自然性和真实性。实验结果表明,具有协同共址定位的视觉-触觉增强现实用户界面能使用户更精确自由地进行交互^[5]。

1 相关工作

前期视触觉增强现实体系的研究主要基于传感器和视觉方法的实现。以往的研究主要是探讨触觉设备在 AR 系统中的集成,一个关键要求是触控笔与视觉增强的精确协同定位,来提供真实的力反馈并实现精确的视觉叠加^[6-8]。

在现实生活中看到和触摸虚拟物体的想法是由 Yokokohji 等人^[9]提出的。他们强调了触觉和视觉反馈的重要性,并将这个概念称为所见即所得 (what you see is what you get, WYSIWYG)。为了实现视触觉协同反馈,需要精确地计算触控笔在触觉空间的位置,以降低视触觉空间转换的共址误差。精确定位触控笔位置不仅保证了视触觉空间转换的精度,也是实现视触觉协同反馈的前提。

早期的工作是由 Vallino 等人^[3]通过仿射变换

完成触觉空间与标识空间之间的映射关系,从而将触觉空间融入到增强现实环境中,然而仿射表示法不提供任何度量信息,这阻碍了对其校准过程准确性的分析。Eck 等人^[10]和 Reinig 等人^[11]在触觉工作区放置平面校准网格用于定位触控笔的位置,该方法极大地提高了触觉笔在网格区域中的位置精度,但是拓展到整个工作区时,因触控笔远离网格,误差不断累积。后来 Eck 等人^[12]通过在触觉空间中放置 2 个垂直的平面校准网格,以改进平面网格校准方法。该方法在一定程度上降低了随着触控笔远离网格而造成的误差积累问题。Lok 等人^[13]基于触觉设备本身的机械结构对触控笔的实时定位,通过触觉设备关节传感器构建触控笔的实时运动模型,此运动模型中包含触控笔的位置和旋转信息,基于关节传感器的触控笔定位方法不仅确立了触觉空间的运动范围,而且提高了定位精度。Adcock 等人^[14]基于 ARToolkit 和 Geomagic Touch 触觉设备开发了视触觉融合的增强现实交互环境,通过投影变换将触觉工作空间映射到屏幕空间上。Ikits 等人^[15]提出了一种使用平面栅格校正触觉触针位置的触觉设备标定方法,这种方法显著提高了定标网格所在区域的位置精度。

另一种计算触觉世界转变的方法是测量外部系统测量点。Grant 等人^[16]在触觉工作区放置用于注册的标识,并使用高精度经纬仪获取标识在触觉空间中的位置,该方法成功计算出视触觉空间的相对位置关系。Harders 等人^[17]和 Bianchi 等人^[18]基于红外跟踪器搭建了视触觉增强现实系统,该系统通过红外跟踪器获取触控笔位置在世界坐标系中的位置,实现对触控笔的虚拟注册,并基于触觉设备的运动模型实现虚拟触控笔和真实触控笔的协同运动,但是在该系统中并未融入触控笔的旋转角,降低了交互的真实性。Aleotti 等人^[19]借助激光扫描器建立了一个全局世界坐标系,以此计算触觉空间和相机空间之间的转换关系,成功地对全局空间中的触觉设备进行虚拟注册,渲染出程度更高的虚实融合场景,提升了交互真实性。Wang 等人^[20]推导了具有半镜显示和定义误差度量的 VHAR 系统的共址误差的数学模型,提出了一种新的模型校准方法,并

使用精确的机械跟踪器对其进行评估。基于该共址误差模型对其视触觉增强现实系统进行校准后,虽然改善了对半反射镜显示器中视触控空间中触控笔的共址误差,但忽略了触觉设备中传感器引起的协同定位误差。Harders 等人^[21]提出了 VHAR 的共址定位方法,提出了一种补偿完整工作空间内触觉设备位置测量的非线性失真的算法,他们使用一个外部跟踪器,并将一个跟踪目标固定在触觉触控笔上,以执行开环校准。

本文采用了 VHAR 的共址定位方法,引入一种基于 6-DOF 触觉设备运动模型而搭建的视触觉增强现实交互环境^[22]。本文的工作将文献[21]的方法扩展到设备上的传感器,从而实现了触控笔和视觉覆盖的精确 6-DOF 对准,该程序会校准位置和方向,可以精确定位触控笔的位置。该系统设计消除了对硬件同步的要求,最大限度地减少了校准所需的时间和用户输入量。它不仅可以对触觉设备进行实时虚拟注册,而且融入了触觉设备的实时运动姿态,实现了更加真实的 6-DOF 人机交互场景。

2 算法结构

本文的目标是设计一个视触觉协同反馈的显示界面。允许用户操纵虚拟工具与增强现实环境中的虚拟对象进行交互,给用户一种自然真实的交互体验。在本文的工作中,所搭建的基于自然特征的增强现实系统以快速特征点提取和描述算法(oriented fast and rotated brief, ORB)和角点光流跟踪算法(Kanada Lucas Tomasi, KLT)相结合实现,并以此为框架融入触觉反馈接口,从而搭建一种视触觉协同反馈的交互环境。为了提升交互场景的虚实融合度,本文运用视触觉空间转换矩阵对场景中的真实触控笔进行虚拟注册;为了进一步提升用户对视触觉交互操作的真实感,本文提出了一种具有 6-DOF 的视触觉交互算法,即将场景中触觉设备的运动状态作为直接的交互方式融入到增强现实框架中,并通过实验证明了此方法的可行性和稳定性。

2.1 实现框架

目前,视触觉增强现实主要有两种实现模型:一

种是将触觉空间映射到屏幕坐标系中,此方式采用的是视触觉融合的非共址定位方法^[23]。非共址定位是指用一个触觉代理表示触觉接触点,用户通过操控触觉设备的虚拟探针实现触觉代理移动虚拟对象,如图 1 所示,将虚拟对象与触觉代理一并注册在真实场景中,实现标识、相机、屏幕以及虚拟设备坐标系之间的融合。另外一种是将触觉空间转换到真实的世界坐标系中,此方式采用的是视触觉融合的协同共址定位方法,即将触觉设备集成到增强现实环境中,使触觉交互点与三维场景中的接触点在视觉上重合。如图 2 所示,将其真实触控笔通过空间转换矩阵和模型视图矩阵进行虚拟注册到真实场景中,实现世界坐标系、相机、屏幕以及虚拟设备坐标系之间的融合,最后将虚实结合的场景投影到显示器上。本文采取第二种模型,并在此基础上用自然特征模板替换标识,构建一种基于自然特征的视触

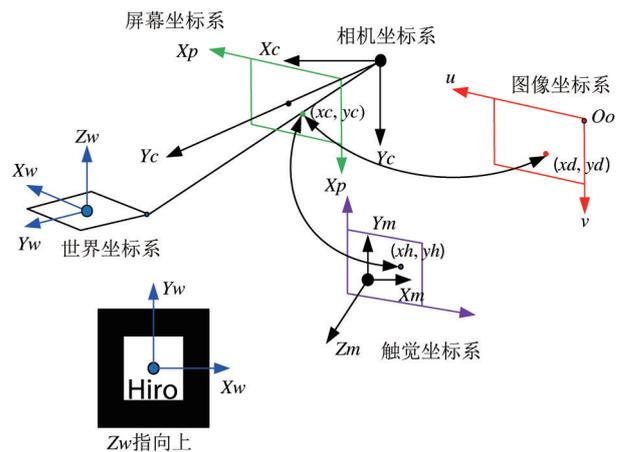


图 1 基于触觉代理的视触觉增强现实

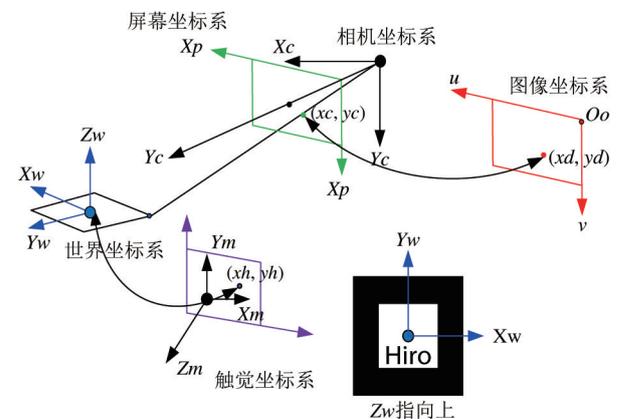


图 2 基于触控笔的视触觉增强现实

觉增强现实体系,以 ORB 与 KLT 算法搭建基于自然特征的增强现实系统,并以此为框架融入触觉反馈接口,实现视触觉协同反馈的交互环境。此外通过视触觉空间转换矩阵对场景中的真实触笔进行虚拟注册,提升交互场景的虚实融合度。为了进一步

提升用户通过触觉设备交互操作的真实感,提出了具有 6-DOF 的视触觉交互算法,成功地将触觉设备的运动姿态映射到视觉空间中。本文还将两种模型进行对比,分析两种模型的优缺点,结果如表 1 所示。

表 1 两种模型的对比表

类别	主要特征	交互场景	优缺点
基于触觉代理的视触觉增强现实	通过触觉虚拟探针完成交互动作	不依赖场景中是否有触觉设备	标识和触觉设备相对独立,交互场景不受限制。未将触觉设备引入交互场景,造成交互的距离感
基于触控笔的视触觉增强现实	通过真实触控笔完成与虚拟对象的交互操作	相机必须实时地捕获触觉设备	将触觉设备引入交互场景,增强了视触觉融合的真实感。触觉设备与标识的位置受到局限

2.2 增强现实环境搭建

目前,视触觉增强现实系统主要使用 ARToolkit 构建增强现实环境。在基于 ARToolkit 的增强现实系统中,标识存在遮挡性差且需要在场景中放置多个人工标识,从而导致在交互过程中用户无法直接、真实地与虚拟对象进行交互,造成交互不真实感。基于上述分析,本文提出了一种基于自然特征的增强现实系统框架,选取现实中常见的自然特征物体作为模板,如选取一本书作为自然特征模板,通过 ORB 和 KLT 算法对场景进行特征点检测、匹配与跟踪,实现对虚拟物体的注册,有效地克服基于标识的增强现实的弊端。基于自然特征的增强现实系统实现流程图如图 3 所示。

本文增强现实系统基于 Windows 10 系统下运行,其 CPU 主频为 3.4 GHz,内存为 8 GB,开发环境是 vs2013,使用双飞燕 1080pUSB 摄像头实时采集场景图片,视频分辨率为 640 × 480 像素,其中预设的自然特征模板是一本书的封面,如图4(a)所示。

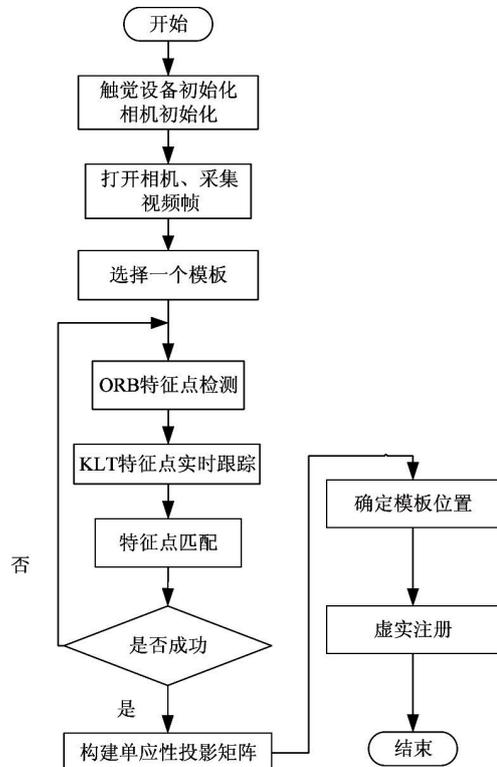


图 3 基于自然特征的增强现实算法实现流程图



图 4 基于 ORB 与 KLT 的增强现实注册结果

利用张氏棋盘法进行标定获得相机的内参,利用 ORB 和 KLT 算法对实时帧进行检测、匹配与跟踪获得相机的外参,最后将虚拟对象注册到模板图像上,实现虚实融合的效果。如图所示为基于杂志的实际注册效果,图 4(b)、(c)和(d)分别展示了算法在初始状态、旋转注册和遮挡情况下的注册结果。

2.3 视触觉空间的融合

本文主要是基于图 2 空间转换图搭建视触觉增强现实框架,将触觉反馈接口加入到自然特征的增强现实系统中。其主要过程为:(1)使用 ORB 和 KLT 算法搭建增强现实环境;(2)在触觉设备工作区的正前方放置自然特征模板 a,通过刚性转换实现触觉空间转换到基于自然特征模板 a 构建的真实世界空间,其中刚性转换矩阵为 T_1 ; (3)基于跟踪注册算法实时计算相机相对于自然特征模板 a 的相对位置,获得模型视图转换矩阵 T_2 ; (4)在场景放置另一个自然特征模板 b,该模板用于注册交互的虚拟对象,其中注册矩阵为 T_3 ; (5)通过 $T_3^{-1} T_2 T_1$ 求解用于视触觉交互的转换矩阵。其中通过 $T_2^{-1} T_1^{-1}$ 完成对触觉设备进行视觉增强。在整个交互场景中,相机、模板 a 和触觉设备位置是实时固定的。基于自然特征的视触觉增强现实模型视图如图 5 所示。

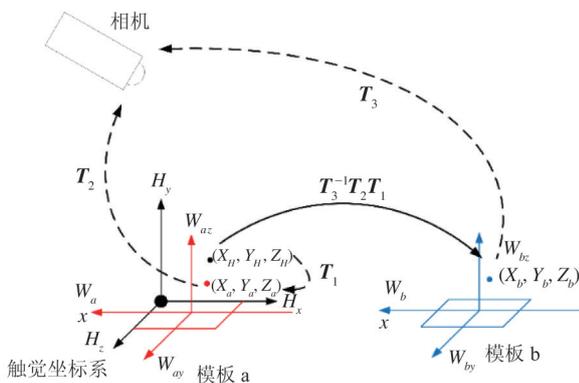


图 5 基于自然特征的视触觉增强现实模型图

根据图 2 的转换关系与图 3 的增强现实注册流程可得:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \quad \mathbf{T}] \mathbf{C}_{4 \times 4} \begin{bmatrix} X_h \\ Y_h \\ Z_h \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{f}{d_x} & 0 & u_0 \\ 0 & \frac{f}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \quad \mathbf{T}] \mathbf{C}_{4 \times 4} \begin{bmatrix} X_h \\ Y_h \\ Z_h \\ 1 \end{bmatrix} \quad (1)$$

其中, (u, v) 是成像点在像素坐标系下点, $(X_h, Y_h, Z_h, 1)$ 是触觉空间坐标系一点, \mathbf{R} 是一个 3×3 的旋转矩阵, \mathbf{T} 是一个 3×1 的平移矩阵, (d_x, d_y) 为每个像素点在 x 轴、 y 轴方向上对应的实际长度, (u_0, v_0) 为图像中心点的像素坐标, f 为成像平面到摄像机焦平面的距离, $\mathbf{C}_{4 \times 4}$ 的矩阵代表触觉空间转换到模板 a 中的刚性转换矩阵,即 $T_1[\mathbf{R} \quad \mathbf{T}]$; 实时跟踪注册的模型视图转换矩阵,即 T_3 ; 基于 $[\mathbf{R} \quad \mathbf{T}] \mathbf{C}_{4 \times 4}$ 完成视触觉空间的实时转换。

为了完成触觉坐标和标识世界坐标之间的转换,首先采集触觉探针在世界模板图像上的 4 个点位置,并分别记录触觉探针在触觉空间和世界坐标中的位置,然后通过刚性转换的方式获得视触觉空间位置之间的绝对转换关系,经过刚性转换后,基于模板可以求解出触控笔在世界模板图像中的位置。为了补充触控笔的实时姿态信息,需要实时计算触控笔绕各个轴的旋转量以用于初始化虚拟触控笔的方向角,结合刚度转换和方向角设置虚拟触控笔的模型视图矩阵,最终实现对触控笔的虚拟注册。

根据上述转换过程,构建数学转换过程如式(2)所示:

$$p_i^w = \mathbf{R}_{3 \times 3} \cdot p_i^H + \mathbf{t}_{3 \times 1} \quad (2)$$

其中 p_i^w 代表触控笔在真实世界中的三维坐标, p_i^H 表示触控笔在触觉活动空间的三维坐标。 $\mathbf{R}_{3 \times 3}$ 和 $\mathbf{t}_{3 \times 1}$ 分别表示旋转矩阵和平移向量。

$$\min_{\mathbf{R}, \mathbf{t}} \sum_{i=0}^N \| p_i^w - \mathbf{R}_{3 \times 3} \cdot p_i^H - \mathbf{t}_{3 \times 1} \|^2 \quad (3)$$

其中, $N(1 \leq N \leq 4)$ 表示分别采集真实世界坐标系中的点数量。为了获得最优的 $\mathbf{R}_{3 \times 3}$ 和 $\mathbf{t}_{3 \times 1}$ 的值,使用 Levenberg-Marquardt 算法对式(3)求解,以完成视触觉空间之间的转换。

2.4 6-DOF 姿态的实现

2.4.1 触觉设备的运动模型

为了通过触觉设备更真实地与虚拟对象交互,本文将触觉设备的前向运动模型映射到视觉空间

中,以创建更直观自然的交互场景。根据 Geomagic Touch 的运动模型对触控笔的姿势进行建模,结果显示在图 6 中,其中触觉设备的 2 个关节角度用于求解触控笔的位置,如图 6(a)所示;另外,通过触觉设备,传感器上的 3 个万向节传感器实时感知触控笔在每个坐标轴上的旋转,如图 6(b)所示。

基于触觉设备的运动学模型推导出触觉笔实时



(a) 位置参数



(b) 旋转参数

图 6 触觉设备的位姿定位

根据右手坐标系,对触觉设备定性建模分析,运动模型如图 7 所示。

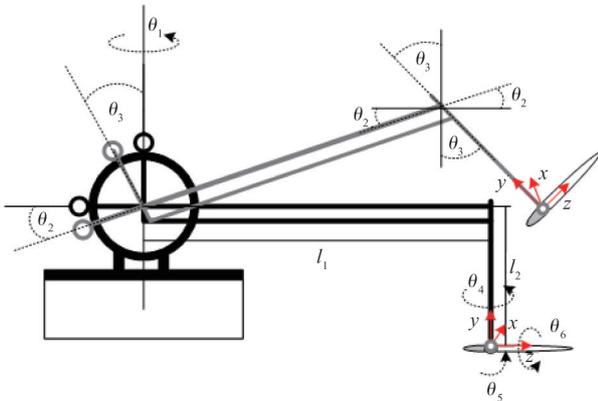


图 7 触觉设备模型

的运动状态。设 $l = (l_1, l_2)$ 代表触觉设备的臂长, $\theta_{ja}(\theta_1, \theta_2, \theta_3)$ 代表装置关节角度,其中 (l, θ_{ja}) 用于计算触控笔在触觉空间中的位置变化,从触觉工作空间原点到 **HIP** 位置 T_{HIP} 的平移由式(4)给出。

$$T_{HIP}(\theta_{ja}, l) = \begin{bmatrix} \sin\theta_1(l_1\cos\theta_2 + l_2\sin\theta_3) \\ l_2 - l_2\cos\theta_3 + l_1\sin\theta_2 \\ -l_2 + \cos\theta_1(l_1\cos\theta_2 + l_2\sin\theta_3) \end{bmatrix} \quad (4)$$

通过以下方式确定第 2 只手臂的方向:

$$\begin{aligned} R_{123}(\theta_{ja}) &= R_1 R_3 \\ &= \begin{bmatrix} \cos\theta_1 & -\sin\theta_1\sin\theta_3 & \sin\theta_1\cos\theta_3 \\ 0 & \cos\theta_3 & \sin\theta_3 \\ -\sin\theta_1 & -\cos\theta_1\sin\theta_3 & \cos\theta_1\cos\theta_3 \end{bmatrix} \end{aligned} \quad (5)$$

为了获得绕轴的角度旋转量,可以通过触觉设备上的万向节传感器获得实时旋转角度, $\theta_{ga}(\theta_4, \theta_5, \theta_6)$ 代表触控笔的绕轴方向角,即触控笔在坐标轴上的实时旋转角。根据万向节传感器可得式(6)。

$$\begin{aligned} R_4 &= \begin{pmatrix} 0 \\ 1, \theta_4 \\ 0 \end{pmatrix}, R_5 = \begin{pmatrix} 1 \\ 0 \\ 0, -\theta_5 \end{pmatrix}, \\ R_6 &= \begin{pmatrix} 0 \\ 0 \\ 1, -\theta_6 \end{pmatrix} \end{aligned} \quad (6)$$

其中 R_4, R_5 和 R_6 分别表示绕 Y, X 和 Z 轴的实时旋转角。最后可以获得触控笔在位置和方向上的实时旋转角度:

$$R_{HIP}(\theta_{ja}, \theta_{ga}) = R_{123} R_4 R_5 R_6 \quad (7)$$

结合式(6)和式(7),可得:

$$HIP_{pose}(\theta_{ja}, \theta_{ga}, l) = \begin{bmatrix} R_{HIP} & T_{HIP} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

根据式(1)将触觉设备的前向运动模型映射到增强现实空间中以用于视触觉实时交互,6-DOF 的视触觉交互如式(9)所示。

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{R}_{HRP} & \mathbf{T}_{HRP} \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{D}_{4 \times 4} \quad (9)$$

其中, $\mathbf{D}_{4 \times 4}$ 矩阵用于完成视触觉空间坐标轴的匹配, $(x_h, y_h, z_h)^T$ 是触觉空间对应的触觉坐标系一点, 从而实现 6-DOF 视触觉融合的增强现实环境。

2.4.2 实时交互的坐标转换

当通过虚拟触控笔与对象进行实时交互时, 需要将虚拟触控笔的实时状态传递交互的对象, 并通过模型视图矩阵和投影矩阵将触控笔姿态的实时变化量传递给交互对象。通过设备按钮获取: 初始位置 $\mathbf{T}_0(p_{x0}, p_{y0}, p_{z0})$, 初始旋转值 $\mathbf{R}_{3 \times 1}(R_{x0}, R_{y0}, R_{z0})$; 触笔的实时位置 $\mathbf{T}_1(p_{x1}, p_{y1}, p_{z1})$ 和旋转值 $\mathbf{r}_{3 \times 1}(r_{x1}, r_{y1}, r_{z1})$ 。基于触笔的实时移动构建交互对象的移动和旋转的模型视图矩阵, 实时的旋转量 $\mathbf{A}_{3 \times 1}$ 为

$$\mathbf{A}_{3 \times 1} = \mathbf{r}_{3 \times 1}(R_{x1}, R_{y1}, R_{z1}) \mathbf{R}_{3 \times 1}(R_{x0}, R_{y0}, R_{z0}) \quad (10)$$

然后对其进行罗德尼格斯方法构建旋转矩阵, 并对其进行齐次变换, 获得 $\mathbf{A}_{4 \times 4}$, 基于按下按钮时的旋转矩阵, 确定虚拟触笔相对于初始的旋转变换量, 如式(11)所示。

$$\mathbf{B}_{4 \times 4} = \mathbf{Rot}_{4 \times 4} \cdot \mathbf{A}_{4 \times 4} \quad (11)$$

其中 $\mathbf{Rot}_{4 \times 4} = \begin{bmatrix} \mathbf{R}_{3 \times 3} & 0 \\ 0 & 1 \end{bmatrix}$ 为当前交互对象的旋转矩阵, 基于式(11)可以获得交互对象实时旋转变换量。为了控制交互对象的平移变换, 接着将其坐标轴原点移动至交互点并以交互构建坐标轴, 将交互点作为旋转和平移的中心点, 其变换公式为

$$\mathbf{C}_{4 \times 4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -P_{x0} & -P_{y0} & -P_{z0} & 1 \end{bmatrix}$$

$$\mathbf{B}_{4 \times 4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ P_{x0} & P_{y0} & P_{z0} & 1 \end{bmatrix} \quad (12)$$

根据初始位置和实时位置计算实时位移量 \mathbf{T} , 如式(13)所示。

$$\mathbf{T} = \mathbf{T}_1 - \mathbf{T}_0 \quad (13)$$

继而根据位移量构建实时的位移转换矩阵 $\mathbf{M}_{4 \times 4}$:

$$\mathbf{M}_{4 \times 4} = \mathbf{C}_{4 \times 4} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ P_x & P_y & P_z & 1 \end{bmatrix} \quad (14)$$

最后, 将当前模型视图变换矩阵和实时变换后的模型视图矩阵相乘, 以获得包括旋转和平移的模型视图矩阵, 并将触觉设备的运动姿势传递到交互对象, 如式(14)所示。

3 实验结果分析

本文使用 Geomagic Touch 触觉设备, 个人电脑 (Intel(R) Core(TM) i7-2600 CPU @ 3.40 GHz) 和 USB 摄像头来构建具有视触觉的增强现实环境, 软件开发平台为 Visual studio。使用 USB 相机和基于 ORB-KLT 算法构建增强现实环境, 主要用于实时采集和检测场景中设置的自然特征; 当识别出场景中的自然标识时, 使用单应性变换矩阵和刚性变换矩阵对场景中虚拟触控笔进行增强。此时用户移动触觉设备, 虚拟触觉也将相应地移动, 与此同时用户按下触控笔上的按钮, 便可启用反馈力, 实现对虚拟物体的实时交互操作。

视触觉交互算法对比如表 2 所示。本文算法采用的是共定位的视触觉融合方法, 文献[24]采用基于 ARToolkit 的非共址视触觉增强现实, 如图 8 所示。由于标识遮挡性差且需要在场景中放置多个人工标识, 只能通过触觉代理点(图中圆形锥点)实现对虚拟物体的交互, 导致了在交互过程用户无法直接、真实地与虚拟对象进行交互, 造成交互的不真实感。文献[17]采用标记的方法对真实触控笔进行跟踪注册, 但需要对真实触控笔贴上标记, 且触控笔无法进行旋转, 否则将出现肉眼可见的注册偏差, 本次对比实验给触控笔笔端和笔尾贴上红色标记, 如图 9 所示。而 Ulrich 的方法中一些硬件条件无法在本实验室实现, 虽然此方法可以有效提高视触觉定向精度, 但因机械转置局限只能在特定实验室中实现, 无法在多场景中实现。而本文算法基于自然特征增强现实, 将触觉设备引入交互场景并且引入

表 2 视触觉交互算法对比

算法	显示装置	集成方式	实现过程	优缺点
Harders ^[17]	HMD	共址定位	在真实触控笔贴上标记进行跟踪注册	采用基于标识注册方法,将触觉设备集成到真实环境中;但多个标识对虚拟物体进行注册,操作复杂
Eck ^[12]	屏幕	共定位	采用两个不同的模体和光学传感器对真实触控笔进行虚拟注册	该方法改进视触觉共址的误差和提高校准的定向精度,但跟踪精度依赖于跟踪目标,且机械装置局限,硬件要求高,容易导致虚拟触控笔的注册误差增大
Huang ^[24]	屏幕	非共定位	将虚拟代理注册到现实环境中,通过触觉虚拟探针完成交互动作	使用虚拟代理代替真实触控笔,缺少真实感;通过虚拟代理完成交互,场景中的标识不能去除
本文	屏幕	共址定位	通过视觉的方法,对真实触控笔进行虚拟注册完成与虚拟对象的交互操作	将触觉设备引入交互场景,操作简单,不受场景限制,增强了视触觉融合真实感,但未能正确处理手势遮挡

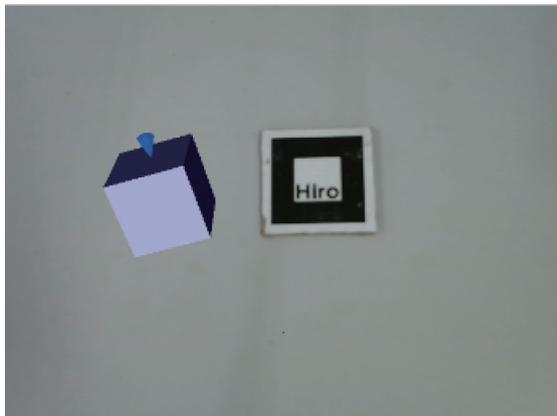


图 8 基于虚拟代理的注册效果图

6-DOF 交互算法,不仅操作简单,不需要高精度复杂的光学传感器并不受实验场景和标识限制,而且增强了视触觉人机交互的真实感和趣味性。

3.1 实验结果

为了完成视触觉空间转换和对触控笔的虚拟注册,本文使用基于自然特征的增强现实技术完成视触觉空间之间的刚性转换和单应性变换。为了验证算法注册的可行性,分别在简单的场景和复杂的场景下进行虚拟注册,结果如图 10 和图 11 所示。



(a) 对触控笔贴标记的注册结果



(b) 触控笔旋转后注册的误差效果图

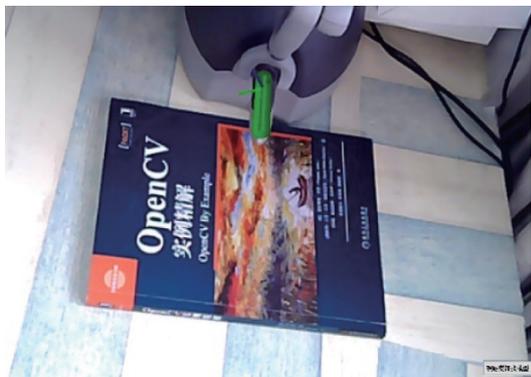
图 9 基于对触控笔标记的注册效果图

图 10(a)、10(b)、11(a)和图 11(b)的实验结果表明,当在场景中检测到目标对象时,计算机计算模型视图矩阵和视触觉空间之间的刚性转换矩阵,使用 2 个转换矩阵对相机空间中触控笔进行虚拟注

册。在初始状态时,将虚拟触控笔注册到真实触控笔,图 10(a)和 11(a)分别表示虚拟触控笔在简单和复杂的场景中注册的结果;图 10(b)和 11(b)表示用户手持虚拟触控笔在 2 种场景中移动的姿态。

当触控笔从墨水池中取出,触控笔由绿色变为淡蓝色,用户以任意方向(6-DOF)移动触觉设备时,虚拟触控笔跟随真实触控笔进行协同运动,渲染了真实的操作场景,使交互场景更加具有真实感。实验结果表明,基于自然特征的视触觉增强现实注册环境

不受场景复杂性等因素影响,这是相比于基于标识增强现实体系注册的一大优点。不过在此虚实融合的场景中由于缺少深度信息,未能处理真实对象和虚拟对象之间的遮挡关系。



(a) 简单场景注册结果

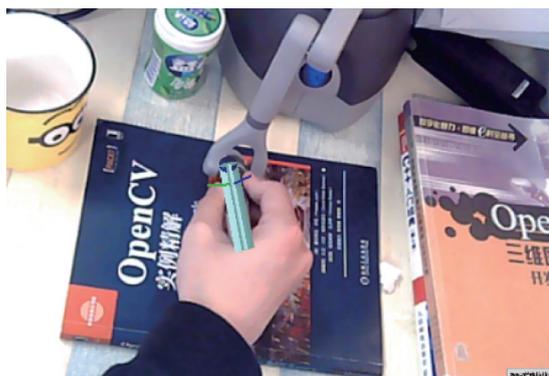


(b) 触控笔在简单场景中移动

图 10 视触觉空间转换简单场景注册效果



(a) 复杂场景注册结果



(b) 触控笔在复杂场景中移动

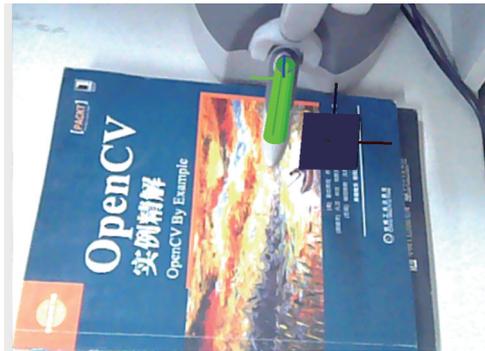
图 11 视触觉空间转换复杂场景注册效果

在图 10 实验基础上,注册一个虚拟方块并给虚拟方块引入物理引擎-重力元素。图 12 展示了真实触控笔与虚拟触控笔协同运动与场景中的虚拟方块进行交互的过程,通过真实触控笔与虚拟触控笔协同运动实现了 6-DOF 的视触觉协同交互。图 12(a) 展示了视触觉初始场景,以虚拟方块的中心绘制三维坐标轴;图 12(b) 展示了用户通过虚拟触控笔与虚拟方块进行交互操作,当触控笔离开墨水池时虚拟触控笔由绿色变为淡蓝色,按下触控笔上按钮便可牵引虚拟物体并且可以明显感受到虚拟方块反馈的重力,用户可以通过手握虚拟触控笔实现对虚拟

物体的平移和旋转,增加了人机交互的真实感,真实模拟了现实场景中手持木棒移动物体的多方位视触觉感官体验;图 12(c) 展示了交互结束的场景,当松开触控笔上按钮时,虚拟物体受重力作用,做自由落体运动,真实模拟了现实场景真实物体的运动姿态,最终展示了虚拟方块静止时最终的位姿信息。

为使设计贴近生活,本文设计了数字排序训练项目,难度较大。用户通过触觉设备在“数字键盘”中选择并移动数字方块,并将数字方块移动到固定位置。场景的注册主要使用 OpenGL 中的 GL_QUADS 绘制 10 个数字方块和 3Ds Max 绘制方格,

利用 C/C++ 代码将 3ds 文件转换成 OpenGL 可读的顶点组,并使用 OpenGL 的图形与颜色线程绘制虚拟对象,然后通过 OpenGL 的贴图技术依次对 10 个立方体进行数字贴图。本次模板注册图像采用上述书籍进行虚拟注册,用户通过移动触控笔实现对虚拟数字的空格排序,每个数字显示一个空格,操作较难,有效考验了用户手眼协调的能力,数字排序训练场景如图 13 所示。



(a) 初始场景



(b) 交互移动中

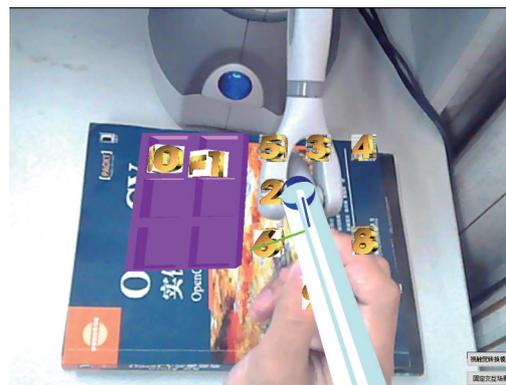


(c) 交互结束

图 12 触控笔控制虚拟对象的交互场景



(a) 初始化场景



(b) 交互场景



(c) 交互结束

图 13 数字排序训练的交互场景

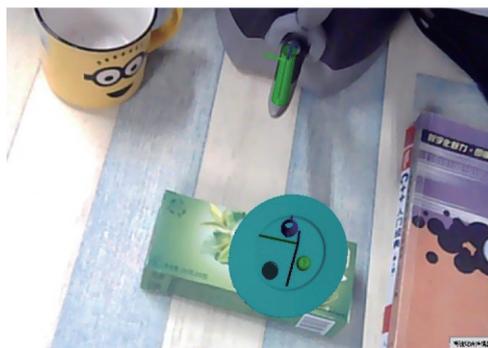
当用户从墨水池中取出触控笔控制时,虚拟触控笔触碰到数字方块时,触控笔由绿色变为淡蓝色表示交互成功,此时,用户可以按下触控笔上的按钮选择并移动该数字方块至方格中,同时对其进行排序操作;释放按钮后,则放置数字方块,在场景中放置图 13(a)用于注册交互的“数字键盘”和放置数字方块的方格,调整自然特征模板,使得整个交互场景完整地显示在窗口中,此时点击屏幕上的固定场景按钮固定交互对象。在整个交互过程中要确保摄像

头位置不变,如图 13(b)所示,固定场景后,用户可以撤走自然特征模板,此时通过移动虚拟触控笔到数字键盘附近,且通过虚拟触控笔触碰数字键盘中的某个方块,真实模拟了现实场景中手移物体的多方位视触觉感官体验。在整个练习模式中,用户可以自主选择想要移动的数字方块的个数,图 13(c)是练习结束时的场景。

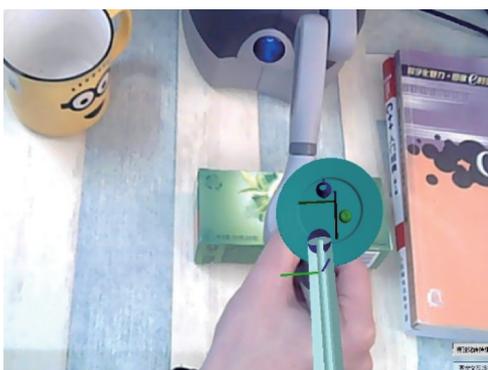
本文实验主要使用一本书作为自然特征标识物,用户也可以选取日常生活中一些彩色图像物体作为自然特征模板。这不仅为用户提供了自主选择标识物进行虚拟注册,也为用户提供参与系统开发的乐趣,进一步提升了用户的训练积极性。在此项实验中选择常见的商品包装盒进行模板替换,在此基础上增加自然场景中物品数量,增加场景的复杂性。实验结果表明,基于自然特征的增强现实不受场景约束,可以很好地满足实验的注册效果。实验注册了虚拟桌面、茶壶、黑色圆饼和绿色苹果,用户通过触觉设备控制虚拟触控笔与虚拟对象进行交互。拾取训练项目难度较大,有效锻炼了手指的捏和握的能力,实验结果如图 14 所示。

在拾取训练的交互场景中,另给触觉设备增加了无力反馈、恒力和弹性力反馈的条件来完成拾取训练,不同的反馈条件丰富了此训练的趣味性。本文实验采用弹性力反馈来进行拾取训练,图 14(a)展示了商品包装盒初始化场景,图 14(b)展示了用户通过虚拟触控笔与虚拟桌面上物体进行交互操作场景。当触控笔从墨水池取出时,虚拟触控笔接触虚拟桌面或虚拟物体时可以明显感受到来自虚拟物体的作用力,此时表明触控笔笔端已接触桌面或者虚拟物体,与此同时按下触笔上的按钮可以选择并移动虚拟对象,用户可以明显感受到实时的反馈力,而释放按钮后,则放置虚拟对象,图 14(c)展示了将虚拟黑色饼干拾取到桌面外的最终场景。

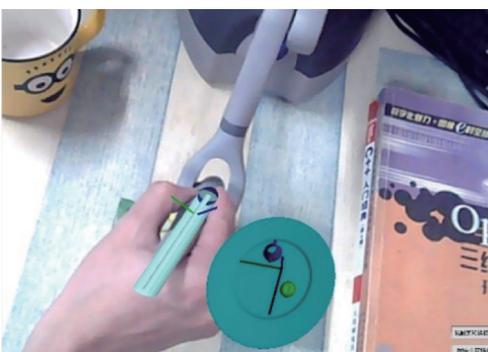
为了充分验证实验的可行性,邀请 10 名学生参与到实验中,其中 6 名男生,4 名女生,年龄在 22 ~ 25 岁。图 15 是邀请一名同学在全局视角下进行视触觉交互的场景。每位被试者依次完成拾取训练项目,所有被试者随机参与训练项目,在完成训练项目后,被试者填写一份用于评价系统可行性的问卷调



(a) 初始化场景



(b) 交互场景



(c) 交互结束

图 14 拾取训练的交互场景



图 15 全局视角下的交互场景

查,每位被试者对训练进行打分(0~10分)。评分标准如下:1.易于理解;2.场景的真实性;3.趣味性;4.易于操作;5.训练难度;6.经济成本;7.力反馈真

实性;8.交互性。拾取训练评分直方图和训练平均时间曲线图分别如图16和图17所示。

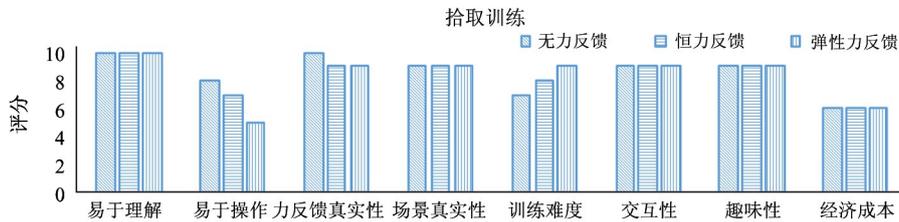


图16 拾取训练问卷评分

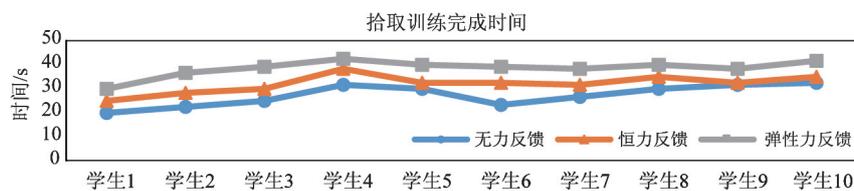


图17 拾取训练平均时间

由上图评分结果可知,视触觉增强现实系统的训练项目交互方式易于理解,且交互过程趣味性十足,所有的被试者都具有较高的积极性,并尽力去完成指定的任务。对被试者而言,因为选择的反馈力模式不同,所需要的时间也不同。通过直方图和曲线图可以清晰地看出,在此训练项目中,弹力模式下的完成时间是最长的,这主要是因为系统设置的弹性系数较大,产生的反馈力较大。另外,从图中发现,恒力模式下完成项目所需的时间与无力模式下完成项目的较为相近,这主要是因为设置的恒力系数较小。在整个训练过程中,熟悉交互场景的被试者所需要的训练时间较短。由此可见,被试者可以通过重复的训练提升完成训练项目的效率,经过多名同学实验证明此算法的可行性与趣味性。

3.2 算法的实时性和准确性分析

本文的视触觉增强现实算法在增强现实环境搭建上采用的是ORB算法,对目标进行实时检测;采用KLT算法对场景中目标进行实时跟踪注册,引入6-DOF实现姿态旋转(ORB+KLT)实现视触觉融合的交互操作。软件开发平台为Visual Studio release x64下执行编译。本文通过对ORB+BF(Brute Force)、ORB+FREAK+BF、SURF+KLT几种经典

算法和本文的算法进行特征点检测和匹配进行实时性和准确率比较。通过采集前400帧图像,分别对第50帧、100帧、250帧、350帧和400帧进行采样,对图像实时运行的帧率和正确姿态注册的图像数目进行分析。实时性的对比采用对虚拟物体注册过程耗时作为参考指标,准确率的单位采用正确姿态注册的百分比。此处所提正确姿态注册是指特征点匹配的准确对数和虚拟触控笔准确注册在真实触控笔上,实验数据结果如表3和表4所示。

从表中数据分析,使用ORB结合KLT算法三维注册方法在不同帧数条件下平均耗时相比于基于SURF和ORB结合BF算法的三维注册方法运算时间均缩短了,且本文算法注册的准确率明显高于其他算法。这是由于KLT算法能够有效、快速响应场景的快速变化,实时地比较图像序列中上一帧与当前帧在时域上的对应关系,以准确计算出两帧图像之间的运动模式及姿态。使用SURF结合算法虽然可以提高注册结果的准确性,但随着图像帧增大,算法迭代次数增大,实时性有所减低,注册失败率增大。这是因为算法本身时间复杂度高,所损耗的时间相比ORB算法远远大出一个级别。实验结果表明,本文算法具有良好的实时性、可行性和准确性。

表3 各算法的实时性分析(单位为ms)

图像帧	SURF + KLT	ORB + BF	ORB + FREAK + BF	本文算法
50	42.5	28.1	25.9	22.9
100	46.9	28.6	26.5	23.1
250	50.0	30.5	28.2	23.5
350	51.2	31.7	29.7	23.9
400	51.5	32.4	30.4	24.8

表4 各算法注册的准确率分析

图像帧	SURF + KLT	ORB + BF	ORB + FREAK + BF	本文算法
50	94%	95.2%	96%	97.5%
100	89.6%	91.5%	94.5%	96.5%
250	85.4%	88.5%	93%	95.2%
350	83.2%	84.6%	85.5%	93.4%
400	79.5%	82.5%	84.5%	92.6%

3.3 误差分析

3.3.1 系统误差

在视触觉空间转换过程中,由于视触觉空间的工作区并不是完全相同的,其中触觉和视觉的工作区间如图18所示,所以需要对手觉空间进行缩放和旋转操作,实现触觉空间与视觉活动空间的匹配。

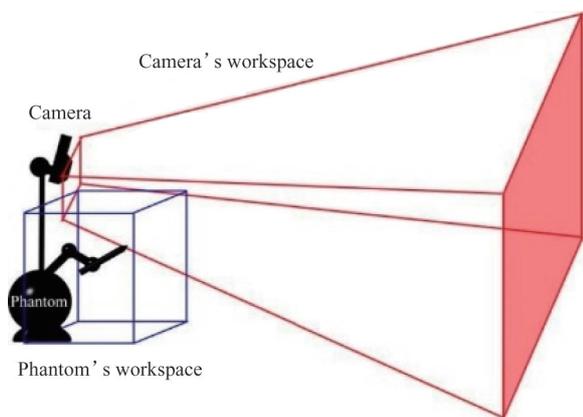


图18 视触觉空间差异图

即视觉坐标系和触觉坐标系之间存在一定的系统误差。触觉设备的真实测量值经过模型视图变换后出现偏差,在X、Y和Z轴上的系统误差如表5所示。

最后通过视口变换、投影变换和模型视图变换设置虚拟触笔的尺寸以适应视口的大小。

表5 视触觉空间的系统误差

	X轴	Y轴	Z轴
平移/mm	53.653	50.271	50.321
旋转/rad	58.221	57.120	58.145

3.3.2 视触觉空间转换误差分析

本文根据文献[20]的误差量化模型可以获得触笔真实位置的误差与虚拟视觉环境坐标系中的注册位置进行误差分析。根据误差度量模型可知,误差来自于模型视图转换过程,在Z轴上的注册误差在5.7mm;同时因为结果缺少深度信息,无法完全地对触笔的虚拟化注册,也造成了无法处理手部与虚拟触笔之间的遮挡关系。

当触笔绕Y轴旋转超过 $\pm 100^\circ$,因误差积累时,虚拟触笔与真实触笔的位置会出现偏差,导致真实触笔暴露的情况,所以本文使用视觉补偿的方法消除偏移量,即使用放大虚拟触笔完成视觉遮挡的注册效果。

4 结论

本文介绍了基于ORB和KLT搭建的视触觉增强现实环境,并且为提升交互操作的真实性,提出了一种具有6-DOF的视触觉融合算法,该算法成功地

将触觉设备的运动姿态融入到视觉空间转换中,使得用户在通过力触觉设备与对象交互时获得更加自然、真实的交互体验。实现效果显示,6-DOF 的视触觉融合算法可提升用户操作的沉浸感和真实感。未来的工作还需要对视触觉增强现实技术做进一步的优化和改进,尤其在反馈力的建模场景中,还需要对虚拟对象设置多重物理属性,以及利用双手与触觉设备搭建更加直观的视触觉增强现实训练系统,使模拟交互场景更加真实。此外本文没有解决虚实遮挡问题,所以还需要处理真实对象和虚拟对象之间的遮挡与碰撞等问题,使交互运动更加真实自然。

参考文献

- [1] Summers V A, Booth K S, Calvert T, et al. Calibration for augmented reality experimental testbeds [C] // Proceedings of the 1999 symposium on interactive 3D graphics, Atlanta, USA, 1999: 155-162
- [2] Bayart B, Didier J Y, Kheddar A. Force feedback virtual painting on real objects: a paradigm of augmented reality haptics [C] // International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, Berlin, Germany, 2008: 776-785
- [3] Vallion J, Brown C. Haptics in augmented reality [C] // Proceedings of IEEE International Conference on Multimedia Computing and Systems, Florence, Italy, 1999: 195-200
- [4] Aleotti J, Micconi G, Caselli S. Object interaction and task programming by demonstration in visuo-haptic augmented reality [J]. *Multimedia Systems*, 2016, 22 (6): 675-691
- [5] Eck U, Pankratz F, Sandor C, et al. Comprehensive workspace calibration for visuo-haptic augmented reality [C] // 2014 IEEE International Symposium on Mixed and Augmented Reality, Munich, Germany, 2014: 123-128
- [6] Olsson P, Nysjo F, Seipel S, et al. Physically co-located haptic interaction with 3D displays [C] // 2012 IEEE Haptics Symposium, Vancouver, Canada, 2012: 267-272
- [7] Saint-Aubert J, Heurley L P, Morgado N, et al. Pre-calibrated visuo-haptic co-location improves execution in virtual environments [J]. *IEEE Transactions on Haptics*, 2019: 588-599
- [8] Eck U, Pankratz F, Sandor C, et al. Precise haptic device co-location for visuo-haptic augmented reality [J]. *IEEE transactions on Visualization and Computer Graphics*, 2015, 21 (12): 1427-1441
- [9] Yokokohji Y. A visual/haptic interface to virtual environment (WYSIWYF display) and its application [C] // Proceedings of 1998 IEEE and ATR Workshop on Computer Vision for Virtual Reality Based Human Communications, Bombay, India, 1998: 99-104
- [10] Eck U, Hoang L, Sandor C, et al. Exploring the perception of co-location errors during tool interaction in visuo-haptic augmented reality [C] // Proceedings of the 2016 IEEE International Symposium on Virtual Reality, Greenville, USA, 2016: 171-172
- [11] Reining K, Tracy R, Glimore H, et al. Some calibration information for a PHAMToM 1.5 A [C] // Proceedings of the 2nd Phantom Users Group Workshop, Dedham, USA, 1997
- [12] Eck U, Sandor C. HARP: a framework for visuo-haptic augmented reality [C] // Proceedings of the 2013 IEEE International Symposium on Virtual Reality, Lake Buena Vista, USA, 2013: 145-146
- [13] Lok B, Naik S, Whitton M, et al. Incorporating dynamic real objects into immersive virtual environments [J]. *ACM Transactions on Graphics (TOG)*, 2003, 22 (3): 701
- [14] Adcock M, Hutchins M, Gunn C. Augmented reality haptics: using ARToolKit for display of haptic applications [C] // 2003 IEEE International Augmented Reality Toolkit Workshop, Tokyo, Japan, 2003: 1-2
- [15] Ikits M, Hansen C D, Johnson R. A comprehensive calibration and registration procedure for the visual haptic workbench [C] // Proceedings of the Workshop on Virtual Environments 2003, Zurich, Switzerland, 2003: 247-254
- [16] Grant B, Helser A, Taylor R M. Adding force display to a stereoscopic head-tracked projection display [C] // Proceedings of IEEE Annual International Symposium on Virtual Reality, Atlanta, USA, 1998: 81-88
- [17] Harders M, Bianchi G. Calibration, registration and synchronization for high precision augmented reality haptics [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2008, 15 (1): 138-149
- [18] Bianchi G, Knoerlein B, Szekely G, et al. High precision augmented reality haptics [C] // Proceedings of the

- EuroHaptics International Symposium, Paris, France, 2006: 169-178
- [19] Aleotti J, Denaro F, Caselli S. Object manipulation in visual-haptic augmented reality with physics-based animation[C] // 19th International Symposium in Robot and Human Interactive Communication, Viareggio, Italy, 2010: 38-43
- [20] Wang D, Zhang Y, Zhou W, et al. Collocation accuracy of visuo-haptic system: metrics and calibration[J]. *IEEE Transactions on Haptics*, 2011, 4(4): 321-326
- [21] Harders M, Bianchi G, Knoerlein B, et al. Calibration, registration, and synchronization for high precision augmented reality haptics[J]. *IEEE Transactions on Visual-ization and Computer Graphics*, 2008, 15(1): 138-149
- [22] 谢余磊. 面向手功能康复的视触觉增强现实系统研究[D]. 南京:南京信息工程大学自动化学院, 2020:20-30
- [23] 张晶晶. 面向上肢功能康复的视触觉增强现实系统研究[D]. 南京:南京信息工程大学自动化学院, 2019: 18-25
- [24] Huang C I, Tang S H. The interactive haptics technique of virtual objects on augmented reality environment[C] // Proceedings of the 11th IEEE International Conference on Control and Automation, Taizhong, China, 2014: 584-589

Visuo-haptic interaction algorithm based on augmented reality

Liu Jia, Qian Changyu, Xie Yulei, Chen Dapeng, Guo Bin

(School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044)

Abstract

Force-haptic interaction, through visuo-haptic augmented reality (VHAR) technology, could be achieved in scenes combining virtual and real. This paper, based on the Geomagic Touch haptic device, proposes a new visual and tactile augmented reality interaction algorithm. The ORB and KLT algorithms are used to build an augmented reality environment based on natural features, which is used as a framework to integrate the haptic feedback interface to realize an interactive environment of visuo-haptic collaborative feedback. The algorithm firstly carries out virtual registration for the stylus in the real scene, and then conducts the 6-DOF coordinated movement between the real and virtual styluses based on the forward motion model of the haptic device; finally, more real and natural visuo-haptic human-computer interaction is achieved in the manipulation of virtual objects by the virtual stylus. Experimental results show that this method is able to achieve good real-time performance, feasibility and accuracy.

Key words: visuo-haptic augmented reality(VHAR), 6-DOF, spatial transformation, real interaction