

面向协同装配的混合现实手眼交互技术研究

王增磊 张树生* 白晓亮

(西北工业大学现代设计与集成制造教育部重点实验室 陕西 西安 710072)

摘要 针对 MR 协同装配工作中存在的复杂人机交互问题,提出一种基于眼动追踪和手势识别的多通道交互方法。采用注视线检测和注视点检测对目标物体进行预选取。通过手势识别完成对候选目标物体的确认和相关操作。实验结果表明,该方法能实现对遮挡、微小尺寸、远距离目标对象的选取和操作。同时,协同装配实验验证了该方法在协同工作中的可行性。

关键词 混合现实 协同 手眼交互 装配

中图分类号 TP391 **文献标识码** A **DOI**:10.3969/j.issn.1000-386x.2019.01.045

MIXED REALITY HAND-EYE INTERACTION TECHNOLOGY FOR COLLABORATIVE ASSEMBLY

Wang Zenglei Zhang Shusheng* Bai Xiaoliang

(Key Laboratory of Contemporary Design and Integrated Manufacturing Technology, Ministry of Education, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China)

Abstract A multi-channel interaction method based on eye tracking and gesture recognition was proposed to solve the complex human-machine interaction problem in MR collaborative assembly work. The method pre-selected the target object by the gaze line detection and the gaze point detection. The gesture recognition was used to accomplish the confirmation and related operation of the target object. The experimental results show that the method can achieve the selection and operation of occlusion, micro-size and long distance target objects. Collaborative assembly experiment verifies the feasibility of the proposed method in collaborative work.

Keywords Mixed reality Collaboration Hand-eye interaction Assembly

0 引言

随着“工业 4.0”和《中国制造 2025》的提出,现代制造业迫切需要向智能化发展。实现智能制造的前提之一在于信息连通性,基于混合现实技术 MR (Mixed Reality) 的协同制造可令工作人员透过混合实境装置取得背景数据无缝、直观的视觉线索,从物联网和人工智能系统获取数据,协助员工实时在工作中将关键数据可视化。MR 技术作为一种高级人机交互技术,将在智能制造系统中负责提供人与智能设备之间传递、交换信息的媒介和对话接口。计算机科学技术的发展促使人机交互方式不断改变,早期的 MR 人机交互技

术主要采用手柄或键盘装置^[1]、触控面板^[2]、手势识别技术^[3-4]、语音识别技术^[5-6]、3D 交互技术^[7-9]、触觉反馈技术、眼动跟踪技术^[10-11]、笔交互技术^[12-13]等单一方式来实现,该类方式在日趋成熟的过程中形成了一定的优势,但在处理某些复杂的交互问题(例如物体的旋转、平移、缩放)时仍然存在一定困难。近年来,多通道协同交互技术取得了长足的进展,其中以微软 HoloLens 的手势、语音和“头部追踪”相结合的协同交互方式为代表,可完成简单的目标点击、选择等交互动作。但该方法需要用户不断移动头部进行目标选取,长时间操作会引起颈部的疲劳与不适,且其语音交互方式响应速度慢、容易受到环境的干扰等缺陷,使得其应用场景受到限制。谷歌公司的 Google glass 是另

一个协同交互产品的代表,它采用语音和触控面板完成交互动作,这种交互方式只注重交互功能的实现而忽略了交互效率与用户体验问题,所完成的也是非常简单、轻量的交互操作,对复杂交互无能为力。此外,现有的协同交互方式对有前后遮挡关系、较小的目标物体等的选取仍然存在困难。

为了提高复杂场景的交互效率,缩短作业周期,本文提出了一种基于眼动追踪和手势识别的MR场景手眼协同交互技术方案,设计了一种目标物体的眼动拾取和手势确认算法,解决了MR环境中三维物理对象拾取操作困难的问题。利用二者在人机交互中的固有特点产生优势互补,解决了远距离场景下,物体相互遮挡以及小尺寸目标的选择拾取与操作问题。

1 手眼交互方法流程

本文手眼交互(Hand-Eye Interaction),以下简称HEI方法分为四个主要阶段:预选取阶段、选取阶段、物理操作阶段,以及追踪阶段。图1展示了在MR环境中手眼交互方法的实现流程。

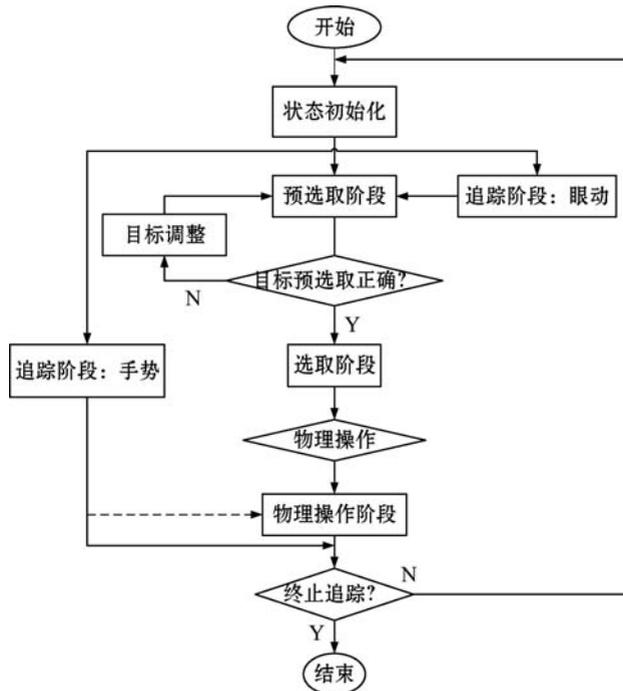


图1 手眼交互方法流程

在预选取阶段,主要是通过眼动信息(注视点坐标)来识别目标物体,相当于传统交互方式中的“搜索”过程。在选取阶段,主要是通过手势识别完成对候选目标物体最终的确认操作。在物理操作阶段,需要对选择的对象进行平移、旋转、缩放等操作。通过将目标物体的位置、姿态与手模型绑定,建立两者之间的映射关系,以达到同步变换的目的。追踪阶段贯穿于

整个交互过程,实际上在交互过程中需要实时监测眼动状况以及手势变化,只是在不同的阶段手眼信息响应值各不相同。在预选取阶段,系统主要针对眼动信息做出响应,在此过程中手势识别传感器处于工作状态并实时检测、更新手势数据库,一旦系统监测到选取阶段中的任意一个手势,即进入选取阶段;在选取阶段为了防止眼动和手势指令的“重叠触发”,对眼动信息采取“只追踪、不处理”,直到触发返回指令或者进入物理操作阶段。在进入物理操作阶段后要用的只有手势信息,同时整体选择拾取已经完成,这意味着眼动追踪可以暂时停止,直到再次回到“初始化状态”。

1.1 眼动追踪与目标物体预选取

眼动信息在MR环境中主要负责虚拟物理对象的预选取。根据目标对象和场景状态的不同,主要有注视线和注视点两种检测方式。

1.1.1 注视线检测

注视线检测是利用人眼注视线来检测目标物体的。在Unity 3D引擎中,相机代表人的双眼,相机所获取的帧画面即为人眼能够看到的MR场景。在本文所搭建的硬件设备环境中,HTC vive头显即是真实环境中相机的映射,换句话说,头显的位置与姿态即代表了相机的位置与姿态。我们选择将相机的坐标位置设置为注视线的起点,由硬件设备所计算得到的眼动信息中为注视点,相机坐标点起点到注视点的连线即为注视线,如图2所示。

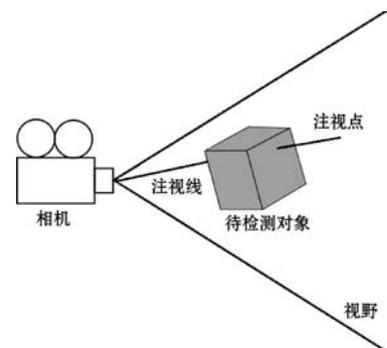


图2 注视线检测示意图

在Unity 3D交互引擎中,通过API接口获取注视点坐标 $G(x, y, z)$ (这里指完成可视化以后的注视点); HTC vive定位系统能够实时定位头盔显示器在真实工作空间中的位置,该位置信息也就是相机坐标 $P(x, y, z)$, Unity 3D提供获取该坐标的API接口,坐标点 G, P 连接成的直线即为注视线 L 。将注视线作为对象的一种进行处理,为该对象添加物理属性;待检测对象作为MR场景中的虚拟模型,系统引擎在初始化该模型时默认赋予后者真实的物理特性。从相机位置处发出的

物理射线,在碰撞到带有碰撞属性的物理模型时停止向前发射,并反馈该模型的数据信息。该数据存储于 Unity 3D 本地数据库中,利用 API 接口用户就可以即可获取目标物体并对其进行下一步操作,见图 3。

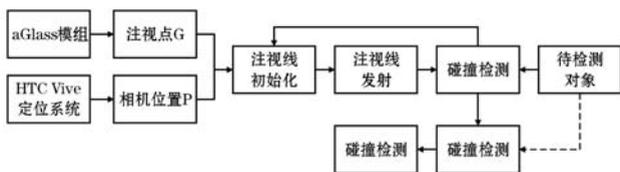


图 3 注视线检测示意图

1.1.2 注视点检测

注视点检测与注视线检测类似,均是利用眼动信息来完成目标对象的选择拾取,区别在于:注视线检测要求射线对象必须与目标对象发生物理碰撞,而注视点检测则摆脱了这种物理属性的约束。从理论上来说注视线检测方式适用于任何场景,但是由于眼动追踪的精度问题以及眼动所固有的跳动与“漂移”,在目标物体较小(射线难以与之发生碰撞)或者需要长时间、不间断监测时,该方式存在一定操作困难。对于这种问题,本文又设计了注视点检测的方法:以注视点所在位置为圆心,设置检测半径 R ,生成球形检测区域,获取处于该球形区域内部所有对象的信息。

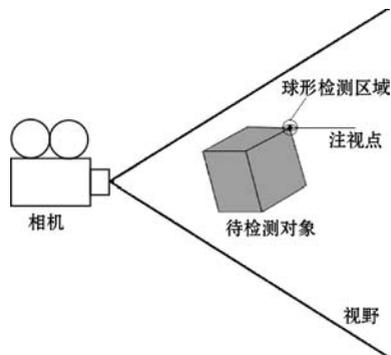


图 4 注视点检测示意图

注视点检测涉及两个物理参数:一是注视点位置 $G(x, y, z)$,经过标定校准之后的眼动模组能够提供;二是球形区域半径 R ,该值从理论上来说并不受限,但是需要根据待检测对象的位置、尺寸来确定。本文注视点检测方法主要针对设计过程中对于物理模型点、线等小尺寸对象,均为小范围检测,因此该值设置为 3,同时对点、线等对象添加球形物理碰撞属性。当注视点落在目标物体周围时,该球形检测器与之产生交集随之碰撞发生,Unity 3D 交互引擎能敏感地捕获该类碰撞,并获取关于碰撞的全部信息。

注视线检测与注视点检测两种方法能够解决交互

场景中的选择拾取问题。在具体的交互设计中我们发现,注视线检测的方法多适用于聚集程度不高、单个物体尺寸较大的选取,而注视点检测则适用于尺寸较小、小区域内聚集多个对象的场景。

1.2 手势识别与目标物体选取及操作

本文手势识别的全部实现均建立在硬件设备 Leap Motion(LM)的基础之上。如图 5 所示,LM 通过对获取到的视频图像进行处理,能够定位到图像中手指指尖、关节、手掌、手腕在三维物理空间中的位置。然后在建立的 MR 环境中叠加人手模型预制体到该位置,通过关节和指尖点、手掌、手腕等共 23 个关键点来确定手与手掌的位置、姿态。

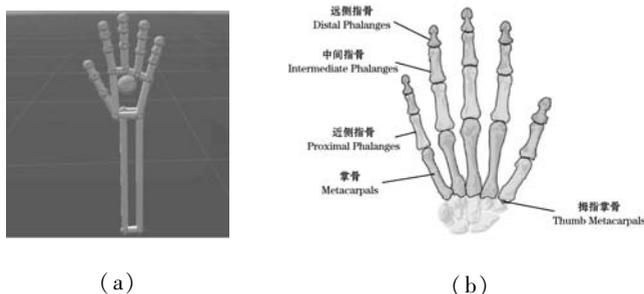


图 5 Leap Motion 手模型

通过 LM 提供的 API 接口,可以比较方便地获取到手指节点的位置、手指弯曲程度、手掌移动速度等信息。在手眼交互中除了需要对人手进行追踪外,还需要能够识别足够数量的手势以帮助完成相关指令操作的触发。

针对手眼交互场景需要,本文在手势设计中遵循以下原则:

- (1) 同一场景中手势数量不超过 5 个,手势数量过多会造成指令在触发上的混乱。
- (2) 手势的简单与自然性,过于复杂的手势对于使用者来说不仅不易于学习,也会降低交互的自然流畅感。
- (3) 不同手势的之间的高区分度,避免手势的误触发。

LM 自有的刷新频率是 0.01,也就是说场景将保持 0.01 s 刷新一帧的速度。在研究中我们发现,如果在此速率对每一帧图像中的手势变化进行监测,整个交互过程将变得“僵硬”。为了提高交互的柔性,本文在手势算法设计中加入监测帧阈值参数 α_{LM} ,当某一手势指令连续触发帧数不小于 α_{LM} 时视作该手势触发,调用对应的事件函数完成对特定事件或条件的响应,单个手势定义流程如图 6 所示。

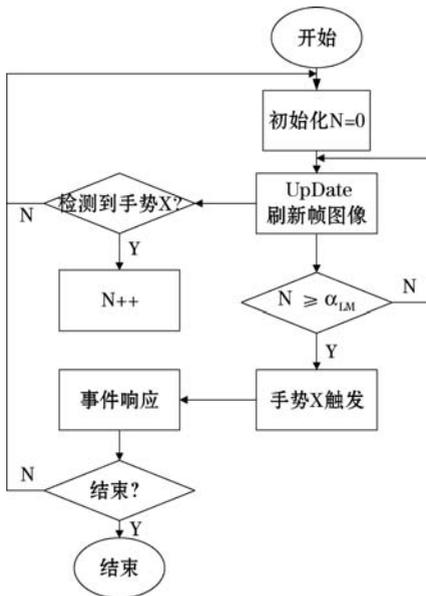


图6 基本手势流程

如表 1 所示,本文所用到的手势在功能上主要有两部分:一是完成指令触发,在眼动对目标物体进行预选取之后通过手来完成最终的选中指令;二是完成眼动所无法实现的,对目标对象的物理操作,包括移动、旋转、缩放等。

表 1 手势设计

手势	动作示意图	指令	触发动作
Y		确认	紧握手掌呈“拳头”状
F		向前	五指并拢,手掌竖直展开,整体向右挥动
B		向后	五指并拢,手掌竖直展开,整体向左挥动
TM		移动模式	伸出食指,其他手指弯曲
RM		旋转模式	伸出食指与中指并呈 30°,其他手指弯曲
ZM		缩放模式	五指张开,手掌静止不动
Z		缩放比例	五指指尖由聚拢模式同时向外扩张,或由扩张向内收缩
OM		偏移模式	拇指与小指弯曲,其他手指伸直

Y 手势表示“确认”指令,该手势可以完成最终目标物体的选取。Leap Motion 并不直接支持对 Y 手势的检测,它只提供关于手模型的相关参数,例如手掌抓取力度、手指的弯曲程度等。手掌抓取力度代表了手掌紧握的程度,当手掌自然伸开时该参数值为 0,当手掌紧握呈“拳头”状时该参数值为 1。本文根据交互需求,设置阈值 $Grab_Threshold = 0.8$ 。

$$hand. grabStrength > Grab_Threshold \quad (1)$$

当手掌紧握成表 1 中的“Y 手势”时,满足式 (1),表示触发手势命令。手指弯曲程度参数值的变化范围为 0 ~ 3.14,对应手指自然伸直到完全弯曲的状态变化,通过监测该参数值判断手指是否伸直,通过计算处于伸直状态的手指数量可设计不同手势指令。

$$hand. extendFinger. num = k \quad (2)$$

当 k 值为 1、2、3、5 时分别对应表 1 中的 TM (Translation-Model)、RM (Rotation-Model)、OM (Offset-Model)、ZM (Zoom-Model) 四种手势。获取手掌移动速度矢量参数,将该参数分解得到在水平方向的分量,判断该分量的大小以及方向可得到手势 F (Forward)、手势 B (Back)。

$$\begin{cases} hand. palmVelocity. x > Speed_Threshold (F) \\ hand. palmVelocity. x < -Speed_Threshold (B) \end{cases} \quad (3)$$

式中:参数 $Speed_Threshold$ 表示速度阈值,本文取 $Speed_Threshold = 0.7 \text{ m/s}$ 。对于缩放比例,需要计算五根手指指尖的距离,本文将相邻指尖距离值之和作为手掌指尖距离参数 $hand. tipPosition$:

$$hand. tipPosition = Z_H \quad (4)$$

式中: Z_H 为手指捏合度参数。

1.3 混合现实中手眼交互实现

1.3.1 HEI 对遮挡对象的选择

对于遮挡情况,利用手眼交互技术可以方便地解决。对比常规对象选择情形,遮挡情形的拾取难度在于不能准确、快速识别目标物体,其根本原因是周围物体对目标物体的干扰。从物理空间角度上来说,干扰物体与目标物体之间的距离小于眼动拾取的最小可区分距离,如图 7 所示。这种物理空间上的限制一部分表现为距离数值上的大小约束,另一部分则是直接呈现局部或是全部遮挡。因此,手眼交互在处理这种遮挡关系时,直接将干扰物体与目标物体间的距离放大,使其处于交互可区分距离范围以内。

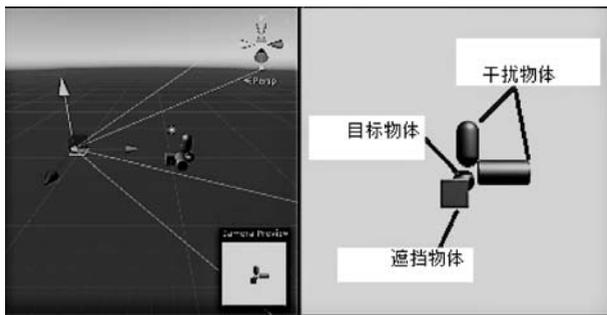


图 7 遮挡与干扰

对于完全遮挡情形,按照注视线检测算法理论以及手眼交互算法流程,在预选取阶段首先采用注视线检测方法得到存在遮挡关系的所有目标物体(即位于注视线方向上的所有物体),这些物体将按照与注视线碰撞的先后顺序、以对象的形式存储于一个对象数组中。对于对象数组中的第 i 个成员(从 0 开始),以注视线与物体碰撞的交点为原点、在垂直于注视线平面上对该成员做偏移操作,如图 8 所示。

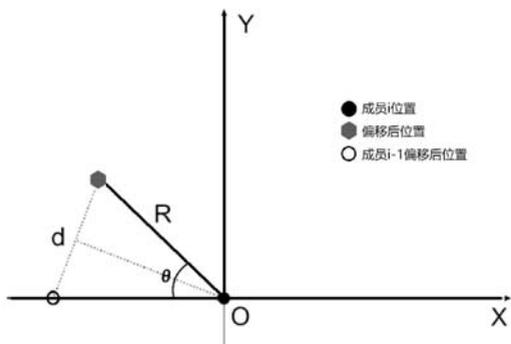


图 8 遮挡偏移示意图

偏移后对象成员的位置坐标为:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -\cos\theta \\ 0 & 1 & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ R \end{bmatrix} \quad (5)$$

式中: (x, y) 为注视点坐标, z 为成员对象深度坐标, 为偏移角度, R 为偏移半径。设 n 为对象数组大小, d 为交互最小可区分距离, 则第 i 个对象的偏移角度和偏移距离计算公式为:

$$\begin{cases} \theta = \frac{360^\circ}{n} \cdot i \\ R = \frac{d}{2 \cdot \sin \frac{\theta}{2}} \end{cases} \quad (6)$$

对于未完全遮挡,但由于聚集程度过高而难以选取的情形,注视线检测方法不能穿过所有的干扰物体,因此无法获取到干扰对象集合。对于这种情况,考虑采用注视点检测的方法,利用注视点定位到聚集程度过高的区域中心,通过球形检测器即可得到局部范围

内所有对目标物体产生干扰的对象。之后依照上述同样的偏移方案即可完成降低聚集度。

在遮挡关系消除后,预选取阶段的另一个任务是普通单一对象的预选取。对于普通对象的选取操作,一种方案是利用注视线检测的方法(射线穿过物体)实现对单个物体的预选取,通过颜色的变化来提示预选取到的目标;另一种方案是,通过手势 F 和手势 B 来对对象集合成员循环遍历,直至找到目标物体,同样通过物体颜色的变化来提示预选取操作的进度。具体方案的选择视具体场景而定,一般对于候选目标较多的情形选择第一种方案,因为眼动拾取更加直接、高效;如果目标物体尺寸较小则选择第二种方案。

在选取阶段,通过触发手势 Y 和手势 N 来做最终的选择操作,如图 9 所示。(a) 表示初始遮挡状态,存在部分遮挡与聚集干扰;(b) 表示偏移后的位置状态。

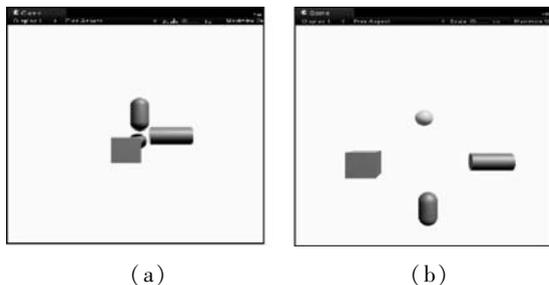


图 9 遮挡对象的选择拾取

1.3.2 HEI 对微小尺寸对象的选择

在交互中除了对常规尺寸对象进行选择拾取以外,往往还需要处理尺寸较小(在场景中所占视野角小于眼动可选最小阈值 0.5°) 的对象,例如角点、边等。一般眼动装置的追踪精度在 0.5° 左右,但实际操作证明由于人眼分辨率的限制,对于 MR 场景中距离较远的物体,即使设备可以捕捉到 0.5° 的视野差,人眼也比较难实现该差值范围内的视线移动。因此,注视线很难准确碰撞到小尺寸物体,注视线检测拾取的方法不再适用,考虑采用注视点检测拾取方法。

在预选取阶段,通过追踪定位眼动注视点位置,确定球形检测区域的球心;球形检测区域的半径设置为交互最小可区分距离,即 $R = d$ 。目前眼动设备所提供的眼动信息大多只包含二维注视点坐标,即注视点在视平面上的投影,而球形检测区域球心需要三维坐标信息。对于二维注视点坐标与三维球心坐标之间的矛盾,本文采用线性搜索的办法:球心坐标由相机位置开始沿注视线方向以一定速率运动,实时搜索当前区域内是否存在待检测对象,直到待检测对象数组不为空。如图 10 所示,注视点在靠近小尺寸目标物体的过程中,检测球从相机所在位置出发朝注视点可视化对象所在位置运动,运动速度代表了检测灵敏度,

依据注视线长度调整该值至合适大小能够提升交互体验。

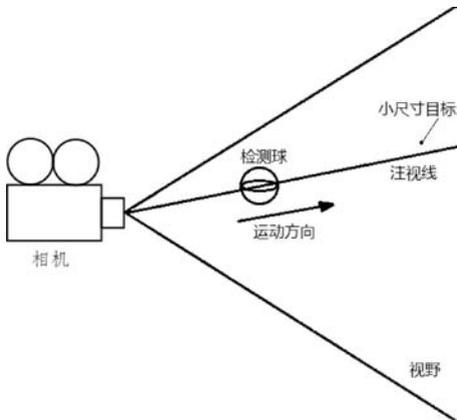


图 10 小尺寸对象检测示意图

当点对象位于球形检测器内部时,检测器返回点对象数据信息,并将该点高亮显示。Unity 3D 默认长边碰撞器为长矩形,这使得边很难整体进入球形检测区域;如果扩大球形检测区域的半径,又会导致一次返回多个小尺寸目标,需要进行二次循环选取操作,增加交互工作量。本文考虑将小尺寸对象碰撞器统一设置为球形,放置于对象的物理中心处,其半径与一般点碰撞器一致,这就表示在预选取阶段,我们需要尽量移动注视点到达目标对象物理中心附近来完成预选取。

在选取阶段,通过手势 Y 来完成最终对预选取目标的确认操作。如果产生误确认,利用手势 N 可取消确认,返回预选取阶段重新选择。无论是在预选取阶段还是选取阶段,本文均采用不同的颜色指示当前目标所处状态,如图 11 所示。(a)表示初始阶段;(b)表示预选取阶段,线条被选中;(c)表示选取阶段,线条被加粗显示。

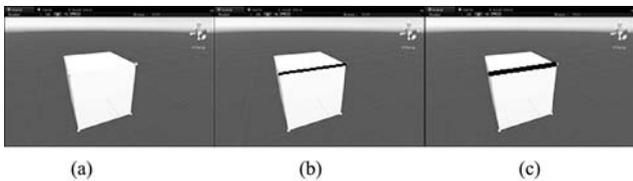


图 11 小尺寸对象的选择拾取

1.3.3 HEI 对远距离对象的操作

在物理操作阶段,主要是通过对手和手势的追踪与识别同步映射到目标物体,从而实现对目标物体的三维物理状态(包括位置与姿态)的改变。移动模式中,掌心位置的移动映射到目标物体位置的改变,由于 LM 的“视野”成一个倒立金字塔型,且掌心位置与目标物体存在一定距离,因此不宜采用 1:1 映射比例。图 12 给出了二者真实的映射关系示意。

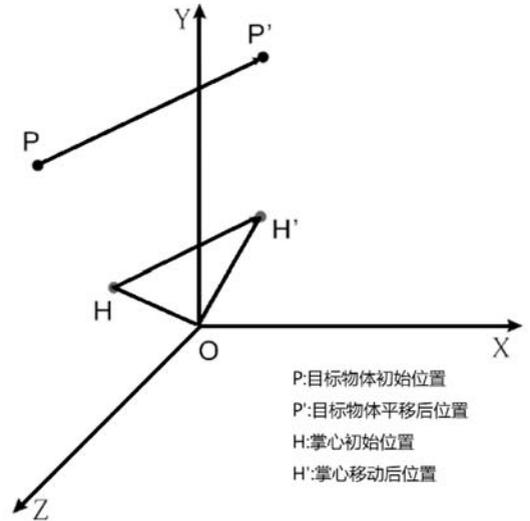


图 12 掌心与目标物体映射关系示意
掌心位置移动的单位向量为:

$$\Delta = \frac{H \cdot H'}{\|H \cdot H'\|} \quad (7)$$

目标物体位移向量为:

$$P - P' = \Delta \cdot k \quad (8)$$

结合式(7)和式(8)可以得到目标物体移动后的位置为:

$$P' = P - k \cdot \frac{H \cdot H'}{\|H \cdot H'\|} \quad (9)$$

式中: k 表示移动变换因子,理论取值范围为 $k \geq 1$ 。由于 k 代表着移动变换映射的精度,其实际取值需根据具体场景要求确定,本文实验 k 值为2。

物理操作阶段的旋转与缩放模式相较于移动来说更加直观。LM 提供获取手掌旋转向量参数的 API 接口,将该旋转向量与目标物体的旋转向量对齐,即可得到目标物体跟随手掌旋转变换的效果,如公式所示:

$$R_p = s \cdot R_H \quad (10)$$

式中: s 表示旋转变换比例因子,其值越大旋转变换效果越明显、灵明度越高,本文实验为该因子取值为3。至于缩放模式,本文将缩放原点设置为物体中心, Z_H 手势的捏合程度 Z_H 代表物体的缩放比例 S_p ,在手掌完全捏合时($Z_H = 0$)物体处于最小比例状态(其缩放比例为 S_{min}),当手掌完全张开($Z_H = 1$)时物体处于最大比例状态(其缩放比例为 S_{max}),因此物体缩放比例与手掌捏合度之间的关系为:

$$S_p = (S_{max} - S_{min}) \cdot Z_H + S_{min} \quad (11)$$

其中 S_{min} 、 S_{max} 代表着缩放灵敏度,可依据具体场景需求取值。

2 结果与讨论

2.1 硬件配置

本文 MR 硬件环境采用 HTC 公司生产的 HTC vive, 选用 aGlass-DK2 来进行眼动追踪。如图 13 所示, 将 aGlass-DK2 安装在 HTC vive 头盔显示器内部, 利用 USB3.0 数据线与 HTC vive 连接完成数据传输。



图 13 眼动追踪模块组件

选取美国 Leap 公司生产的 Leap Motion 作为手势追踪设备。如图 14 所示, 将 LM 安装在 HTC vive 头显外部的正前方, 同样使用 USB 数据线与头显连接完成数据传输, 用来追踪手势的变化。



图 14 手势追踪模块组件

2.2 基于手眼交互的协同装配

本节基于前文所介绍的 MR 手眼交互技术, 搭建了一套协同装配原型系统。如图 15 所示, 该系统分为两部分, 主要包括工艺设计人员工作场景和工装现场。在该系统中, 工艺设计人员头戴加装眼动追踪模组和手势识别传感器的 HTC vive 头盔显示器, 利用手眼交互方法能够完成对场景中虚拟物体的选择拾取操作; 装配工人在装备 3D 投影仪的作业车间, 按照投影仪

的投影提示信息完成作业。3D 投影仪架设于工作台上(本文实验中架设于天花板, 投影仪距离工作桌面 1.7 m), 通过数据线连接到 PC 客户端, 能够将虚拟模型场景投影到工作桌面, 在视觉上呈现出三维模型效果, 图 15 中工作台上的光斑部分即为 3D 投影仪的投影内容。



图 15 协同装配工作场景

工作台上的 3D 投影仪将每个零部件的虚拟模型投影到桌面上, 工艺设计人员根据图 16 所示流程操作虚拟模型, 装配工人则根据投影亮斑的变化来选取对应的零部件进行装配。由于本文研究的重点在于协同装配工作中的手眼交互技术, 为了简化装配过程、避免不必要的人为因素影响, 我们给出如下假定条件:

(1) 假定工艺设计人员具备工艺工装的先前经验知识, 能够对虎钳装配进行正确的流程引导。

(2) 假定工人无任何虎钳装配的先前经验知识, 且无法从外观或其他除引导以外的任何渠道获取虎钳装配流程信息。

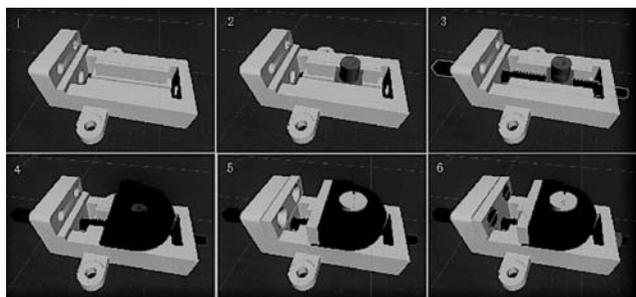


图 16 装配流程示意图

装配工人在进行作业时能够获取的唯一指导信息来源于工艺设计人员的手眼交互操作, 工艺设计人员按照预先设定的装配流程来操作虚拟模型进行装配引导。如图 17 所示, 工艺设计人员在进行操作前首先需要切换到相应的操作模式(平移、旋转、缩放、遮挡处理), 该指令的触发主要由手势来完成, a1 ~ d1 给出了不同模式对应的手势指令。在平移操作中, 工艺设计

人员通过眼动注视点来预选取目标物体,预选取中的物体颜色由白色变化为蓝色,如图 17 中的 a2 所示。若确认预选取无误则伸出右手并将手掌紧握成“拳头”状,即代表确认选中预选取物体为该阶段的操作目标物体,该物体颜色由蓝色变化为红色,如图 17 中的 a3 所示;物体选取完成后,工艺设计人员移动手部,目标物体将同步进行移动,如图 17 中的 a4 所示。在旋转、缩放模式中的目标物体预选取以及最终确认的方式同平移类似,图 17 中的 b2、b3、c2、c3 分别给出了示意。在旋转操作阶段,工艺设计人员将手部绕手腕进行旋转,目标物体将围绕自身坐标系原点进行同步旋转,如图 17 中的 b4 所示。在缩放操作阶段,需要用到表 1 中的缩放手势(即 Z 手势),通过调整手指指尖距离来完成缩放,如图 17 中的 c4 所示。虚拟场景中的螺钉相互之间存在聚集程度较高的问题,我们将该模式定义为偏移模式,在该模式中通过眼动选取存在遮挡关系的对象集合,如图 17 中的 d2 所示。然后工艺设计人员右手握拳完成位置偏移,如图 17 中的 d3 所示。

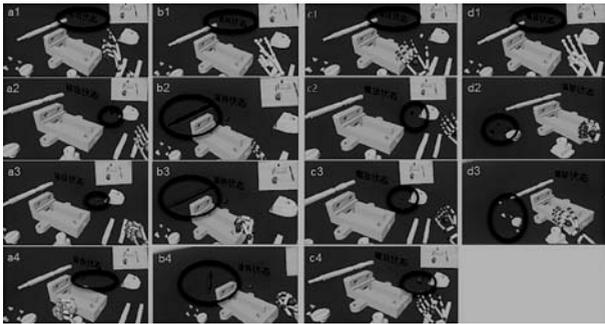


图 17 工艺设计人员引导装配过程

在选取与操作过程中,工艺设计人员有可能产生误操作,针对此问题本文设计返回手势,工艺设计人员通过向左挥动右手来实现“返回上一步”的操作(表 1 中的 B 手势)。该手势可令工艺设计人员退出对当前目标物体的操作,目标物体的颜色由红色变为白色,工艺设计人员可通过手眼交互重新选择目标物体。

3 结 语

本文针对 MR 协同工作环境中的交互问题,在分析具体的交互场景特点之后,提出结合眼动追踪和手势识别的手眼交互技术。研究了眼动生理特点,提出了两种有利于优化眼动数据的信息预处理方法,设计了两种眼动拾取算法(注视线检测和注视点检测)和 8

种易操作的手势识别模型,实现了手眼交互对远距离对象的三维物理操作。以手眼配合的生理机制为基础,提出了手眼交互的方法流程,针对具体交互问题设计了三种手眼配合的交互模式,最后搭建了基于手眼交互的协同工作系统,验证了本文提出的手眼交互方法在协同工作中的可行性。

参 考 文 献

- [1] Fiorentino M, Uva A E, Gattullo M, et al. Augmented reality on large screen for interactive maintenance instructions [J]. *Computers in Industry*, 2014, 65(2):270-278.
- [2] Morrison A, Oulasvirta A, Peltonen P, et al. Like bees around the hive: a comparative study of a mobile augmented reality map [C]//*Sigchi Conference on Human Factors in Computing Systems*. ACM, 2009:1889-1898.
- [3] Gurevich P, Lanir J, Cohen B, et al. TeleAdvisor: a versatile augmented reality tool for remote assistance [C]//*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012:619-622.
- [4] 黄俊,景红. 基于 Leap Motion 的手势识别在虚拟交互中的研究 [J]. *计算机应用研究*, 2017, 34(4):1231-1234.
- [5] 李云. 快速语音识别算法的研究 [J]. *信息技术*, 2017(7):118-120.
- [6] 柯登峰,徐波. 互联网时代语音识别基本问题 [J]. *中国科学:信息科学*, 2013, 43(12):1578-1597.
- [7] Takala T M, Rauhamaa P, Takala T. Survey of 3DUI applications and development challenges [C]//*2012 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 2012:89-96.
- [8] 纪连恩,张凤军,付永刚,等. 虚拟环境下基于语义的三维交互技术 [J]. *软件学报*, 2006, 17(7):1535-1543.
- [9] 纪连恩,张凤军,王亮,等. 基于场景语义的 3D 交互体系结构 [J]. *计算机辅助设计与图形学学报*, 2006, 18(8):1236-1242.
- [10] 韩俊霞,康健楠,欧阳高翔,等. 孤独症儿童脑电与眼动追踪研究 [J]. *科学通报*, 2018(15):1464-1473.
- [11] 李晶,郁舒兰,刘玮. 基于眼动追踪的数控界面布局认知特性评价 [J]. *计算机辅助设计与图形学学报*, 2017, 29(7):1334-1342.
- [12] 戴国忠,田丰. 笔式用户界面 [M]. 合肥:中国科学技术大学出版社, 2009.
- [13] Jiang Y. Structuring and manipulating hand-drawn concept maps [C]//*Proceedings of the 14th international conference on Intelligent user interfaces*. IEEE, 2009:457-462.