



单位代码 10006

学 号 19373345

分 类 号 TP391.9

# 北京航空航天大学

BEIHANG UNIVERSITY

## 毕业设计 (论文)

虚拟现实头眼协同对象操纵  
方法设计与实现

学 院 名 称 计算机学院

专 业 名 称 计算机科学与技术专业

学 生 姓 名 刘兆薰

指 导 教 师 王莉莉

2023 年 05 月

# 北京航空航天大学

## 本科生毕业设计（论文）任务书

### I、毕业设计（论文）题目：

虚拟现实头眼协同对象操纵方法设计与实现

---

---

### II、毕业设计（论文）使用的原始资料（数据）及设计技术要求：

针对虚拟现实设计并实现一套基于头眼协同的高效易用的对象操纵方法

---

使用 C# 在 Unity3D 引擎中进行交互系统逻辑与代码开发

---

使用 SteamVR 进行相关虚拟现实场景开发

---

适用 VIVE Focus 3 眼动仪获取、分析眼动数据

---

### III、毕业设计（论文）工作内容：

头眼协同的场景与目标浏览

---

头眼协同的目标选择

---

头眼协同的对象操纵

---

两个先导实验

---

两个用户实验

---

#### IV、主要参考资料：

M. Daniel, C. Ariel, G. Andrea, et al. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments: Survey on 3d virtual object manipulation[J]. Computer Graphics Forum, 2018.

C. Liu, A. Plopski, J. Orlosky. Orthogaze: Gaze-based three-dimensional object manipulation using orthogonal planes[J]. Computers & Graphics, 2020.

K. Kyungyoon, L. R. L., K. Nikki, et al. Anatomical 2d/3d shape-matching in virtual reality: A user interface for quantifying joint kinematics with radiographic imaging[C]// 2017 IEEE Symposium on 3D User Interfaces. [S.l.: s.n.], 2017.

计算机 学院 计算机科学与技术 专业 190614 班 学生 刘兆薰

毕业设计(论文)时间： 2023 年 03 月 01 日至 2023 年 06 月 01 日

答辩时间： 2023 年 06 月 01 日

成 绩： \_\_\_\_\_

指导教师： \_\_\_\_\_

兼职教师或答疑教师（并指出所负责部分）：

\_\_\_\_\_  
\_\_\_\_\_

\_\_\_\_\_系（教研室）主任（签字）： \_\_\_\_\_

注：任务书应该附在已完成的毕业设计（论文）的首页。

## 本人声明

我声明，本论文及其研究工作是由本人在导师指导下独立完成的，在完成论文时所利用的一切资料均已在参考文献中列出。

作者： 刘兆薰

签字：

时间： 2023 年 05 月



## 虚拟现实头眼协同对象操纵方法设计与实现

学 生： 刘兆薰

指导教师： 王莉莉

### 摘 要

本篇文档主要介绍北航毕业设计论文  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  模板使用和相关软件环境的安装配置，以及本模板所遵循的开源协议等。

**关键词：** 人机交互，虚拟现实，头眼协同，多模态操纵，无手操纵



# **Design and Implementation of a Head-eye Collaborative Object Manipulation Method in Virtual Reality**

Author: Zhaoxun Liu

Tutor: Lili Wang

## **Abstract**

Here is the Abstract in English. And this is a test sentence, just for a test to see how the buaathesis works. You can just ignore this.

This is another paragraph.

**Key words:** Human-computer Interaction, Virtual Reality, Head-eye Collaboration, Multi-modal Manipulation, Hands-free Manipulation



## 目 录

1 绪论 .....	1
1.1 课题来源 .....	1
1.2 研究背景 .....	1
1.3 研究意义 .....	1
1.4 研究目标 .....	2
1.5 论文结构 .....	2
2 研究现状 .....	3
2.1 基于手部(含手柄)的操纵 .....	3
2.2 基于语音的操纵 .....	7
2.3 基于眼动的操纵 .....	8
3 方法设计与实现 .....	12
3.1 交互系统有限状态机 .....	12
3.2 场景与目标浏览 .....	12
3.3 目标选择 .....	12
3.4 对象操纵 .....	12
4 实验设计 .....	13
4.1 先导实验 .....	13
4.1.1 实验 1: 眼动确认信号筛选 .....	13
4.1.2 实验 2: 眼动操纵视线驻留检测优化 .....	13
4.2 用户实验 .....	13
4.2.1 实验 1: 单物体位移对接实验 .....	13
4.2.2 实验 2: 单物体操纵对接实验 .....	13
致谢 .....	14
参考文献 .....	15



## 1 绪论

### 1.1 课题来源

课题来源于北京航空航天大学虚拟现实技术与系统国家重点实验室。

### 1.2 研究背景

虚拟现实(Virtual Reality)技术,简称VR,是一种利用计算机技术来模拟生活环境或创造虚拟现实的新型多媒体技术,是扩展现实(Extended Reality)技术的一个分支。目前主流的VR设备可利用头戴式显示器建立起一个完全虚拟的三维空间。使用者在这个虚拟的环境里进行交互操作时,计算机可以立即进行高度实时的、复杂的运算,将精确的三维影像传回,让使用者身处完全的沉浸式视觉环境中。该技术整合了计算机图形学、仿真模拟、人工智能以及并行计算等技术的最新发展成果,是一种融合多种先进技术的模拟系统。

### 1.3 研究意义

VR已经在影视娱乐、教研教学、设计辅助等领域颇有建树,然而学界和工业界依旧留存着许多非常关键的问题亟待解决,拥有非常大的研究价值。

就目前而言,限制VR普及和发展的较为直接的阻碍,除开较高的市场售价,则来自于其仍旧较低的易用性,即虚拟环境中对物体的操控和交互的准确度依旧不容乐观,或是操作指令和交互动作过于复杂繁琐。这个缺陷直接降低了用户对VR技术的接受度和使用期望。因此,VR中的对象操纵方法的优越性是提高其使用体验和普及度的基本问题之一。许多研究者已经进行了大量的研究,但仍有较大的提升空间。常规的对象操纵动作包括点击、按压、抓取、移动、释放等,其对应对象的直接具体表现主要为创建、销毁、位移、形变、旋转和缩放。对象操纵的速度、准确性、学习成本、使用压力和多样性将直接影响应用程序的效果,而在虚拟环境中实现高效且易用的对象操纵具有一定的挑战性。因此,本研究拟针对虚拟现实对象操纵的关键问题进行研究,旨在提出相较于目前国际一流水准方法更加高效易用的基于头眼协同的虚拟现实对象操纵方法。





## 1.4 研究目标

本研究的主要探究问题是如何确定一套快速、准确、易用的虚拟现实头眼协同对象操纵方法。这个问题的主要难点包含：(1) 通过眼动可获取的信号有限；(2) 眼动信号不稳定，并且由于本能动作（如眨眼）干扰，眼动信号解析难度大；(3) 眼动操纵使用负担非常大，现有方法操作流程复杂并且需要视线高度集中，容易产生眼球生理性疲劳。

基于此考虑，我们需要以尽可能少的头眼动作组合和尽可能小的专注度要求完成尽可能多样的对象操纵任务。因此，本研究的主要内容是：(1) 头眼协同的场景与目标浏览；(2) 头眼协同的目标选择；(3) 头眼协同的对象操纵。大致研究步骤是：(1) 确定针对操纵速度、准确度和使用负担的评估指标；(2) 探索比较多种头眼协同交互方法；(3) 确定一套高效易用的基于头眼协同的虚拟现实对象操纵方法，支持操纵对象位移、旋转和缩放；(4) 根据评估指标对比目前国际一流方法（baseline）并作结果分析。

## 1.5 论文结构



## 2 研究现状

对象操纵是 VR 技术中最为基本且必要的交互行为之一。针对对象操纵这一特定的研究主题,扩展现实的所有门类,甚至传统计算机图形界面的研究成果皆可启发虚拟现实的更新和发展。在过去的二十年里,国内外许多相关领域学者致力于对象操纵的相关研究。根据已有的研究内容,扩展现实中的对象操纵方法主要由以下三大思想构成:(1) 基于手部(含手柄)动作;(2) 基于语音;(3) 基于眼动<sup>[1][2]</sup>。

### 2.1 基于手部(含手柄)的操纵

在基于手部(含手柄)追踪的方法研究早期,学界主要的三大思路为在虚拟环境中单手直接操纵、延长用户手臂和射线广播(ray-casting),见图 1。虚拟延长手臂思路的代表研究来自于 Ivan Poupyrev 团队在 1996 年发表的 Go-Go 沉浸式交互方法<sup>[3]</sup>。Go-Go 使用交互式增长用户手臂的元函数和非线性映射来指定和操纵远处的物体。与同时期的其他技术不同的是,Go-Go 允许对附近的和远处的物体进行无缝直接操纵。然而,Go-Go 技术提出的物体选择和操控模式并不能完全作为一个完整的交互方法来供人们使用;Go-Go 应该被视为以同时期技术为基础的一个补充,而不能完全取而代之。射线广播的思路和虚拟延长手臂类似;射线广播的思路是将射线束从使用者的手中延伸出来,从而指定操纵物体。然而,射线广播思路存在比较明显的弊端。由于在射线广播的对象操纵中物体是被连接到射线末端的,所以除了以射线本身为轴可以完成的动作,许多操纵都是难以简单实现的,因为只有一个自由度(围绕射线轴的旋转)可以用射线广播的方式独立控制。比如,若用户希望以与射线方向垂直的轴向旋转一个物体,单纯以射线广播是无法完成的。此外,射线广播还缺乏一种控制物体与用户之间距离的方法,导致用户无法准确地将物体拉近或推远,而这也是对象操纵的基础功能之一。

当把 Go-Go 方法和射线广播以及其他方法(例如通过将虚拟手臂延伸到无限远来改进 Go-Go 的 Stretch Go-Go)进行比较时,并没有明显的赢家<sup>[4]</sup>。用户评估结果显示,所有技术都有明显的缺点。在这次评估中,HOMER 方法被提出来了;这是一种以手为中心的基于射线广播的对象操作技术<sup>[4]</sup>。HOMER 使用射线来选择物体,在选择物体后,它将虚拟的手移动到物体上;用户的身体和手之间的当前距离被映射为到虚拟物体的距离。因此,HOMER 方法操纵对象的方式与 Go-Go 技术类似,但缩放系数是针对每个选

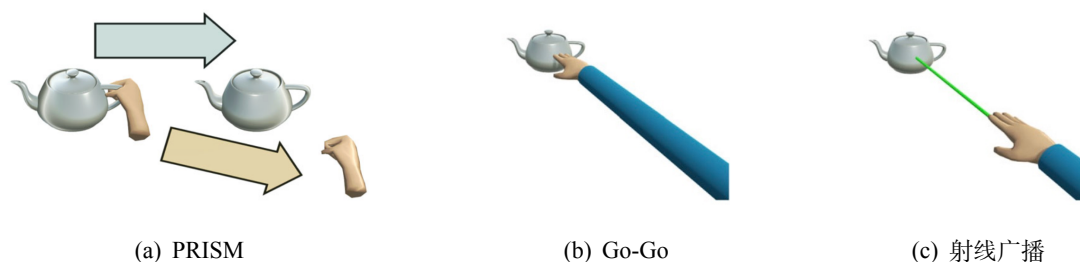


图 2.1 研究早期三种主要的基于手部动作追踪的对象操纵方法

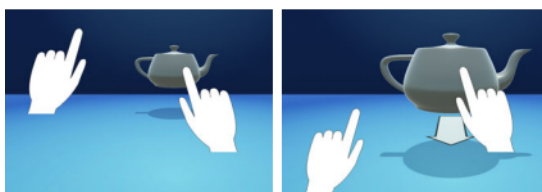
定的物体独立计算的。

到了 21 世纪, 学界将研究重心放在虚拟环境对象操纵的精确性上。一个具有代表性的方法是 Scott Frees 团队在 2005 年提出的 PRISM<sup>[5]</sup>, (见图 1(a)) PRISM 是一种通过缩放操作进行精确和快速交互的方法, 它在同时期是一种非常新颖且具有开创性的交互技术。PRISM 主动根据用户在虚拟环境中的行为特征来确定他们所想的操纵目标是明确还是不明确的。当操纵目标明确时, PRISM 动态地调整“控制/显示”比例来提高对象操纵的精确度。该比例决定了物理手部运动和受控虚拟物体运动之间的关系, 降低了传感器对手部运动监测不必要的灵敏度, 从而减少操作误差。使用 PRISM, 用户始终完全控制着被操纵物体的位置。与像 Go-Go 这样的技术相比, PRISM 在能力范围上也有很大提升, 最明显的进步在于 PRISM 扩大了手部运动的规模以允许“远距离”操纵, 同时在特定情形中可以主动缩小手部运动的幅度以提高精确度。

在这之后, Curtis Wilkes 等人将 PRISM 与 HOMER 相结合, 在 2008 年提出了融合了两种方法精华的 Scaled HOMER<sup>[6]</sup>。Scaled HOMER 使用基于速度的缩放, 允许用户在近距离和远距离进行更为精确的操作。它比原始的 HOMER 在各种任务条件下, 尤其是有关需要高度精确、远距离放置物体或大运动距离的任务中的性能都有所提高。2015 年, 在 Go-Go 和 PRISM 研究之后, Chris Auteri 等人将这两种技术结合起来, 以提高延伸的三维操作的精确性<sup>[7]</sup>。该方法首先将 PRISM 直接应用于用户的手(基础光标)的运动, 从而基于运动速度计算出一个新的光标位置(PRISM 光标)。然后, PRISM 光标移动的距离被基于 Go-Go 距离的启发式方法所放大。与 PRISM 和 HOMER 的结合一样, Go-Go 和 PRISM 的结合提供了一些改进, 尤其是在任务完成的成功率和精细度上。

基于双手的操纵在 2008 年被 Noritaka Osawa 团队提出<sup>[8]</sup>。该团队提出了一种用于在沉浸式虚拟环境中精确定位 3D 虚拟物体的单手和双手控制技术。这个方法提出了一种位置调整策略, 包括一个类似于 PRISM 的用于减缓手部运动的比例系数以及一个被

动的视角调整。该交互系统会自动将视角接近抓取点,使被操纵的物体看起来更大,从而更易于操控。为了有效控制这些调整,该团队提出了两种技术。第一种是基于单手操纵的;因为当用户想精确地操纵一个物体时,他们的手会慢慢移动,所以通过对单手的速度监测,系统可以判断当前对象是否需要精确操纵。另一种是基于两手间距离的;当用户两手之间的距离很小时,调整就会被激活。通过用户评估,位置和视点的调整比禁用这种调整有更好的操纵效率和用户体验。此外,该团队的测试结果还显示,双手控制比单手表现更好。承接双手直接操纵的方法, Martinet 团队提出了两种移动 3D 对象的技术<sup>[9]</sup>。第一种扩展了许多 CAD (Computer-aided Design, 计算机辅助设计) 应用程序中的视窗概念;它引入了四个视窗,每个视窗显示 3D 对象的不同视图。在其中一个视窗中触摸并拖动物体,可以在与该视窗平行的平面上平移物体。第二种方法被称为 Z 技术;Z 技术只使用场景的一个视图(见图)。在这种技术中,第一次触摸触发在平行于视图的平面上移动物体,第二次触摸触发垂直于视图平面的前后运动。Martinet 的初步评估表明,用户更喜欢 Z 技术。Martinet 等人在 Z 技术的基础上进行了改进,推出了 DS3,一种基于 DOF 分离的三维对象操纵技术<sup>[10]</sup>。与 Z 技术类似,一次直接触摸可以在屏幕平面上移动物体,间接触摸可以操纵物体深度,两次直接触摸可以实现旋转。Martinet 将 DS3 与之前的类似方法,比如 Hancock 团队提出的 Sticky Tools 方法(见图)和 Reisman 团队提出的 Screen-Space 方法进行了比较,结果显示 DOF 分离导致了更好的结果<sup>[11][12]</sup>。



(a) 基于双手直接操纵的 Z 技术



(b) 基于双手直接操纵的 Sticky Tools 方法

图 2.2 两种具代表性的基于双手动作的对象操纵方法

另外一种值得一提的基于手部动作追踪的方法是外加虚拟控制柄(见图);虽然这类方法暂未被应用到虚拟环境中,但是它们对于对象操纵的研究是颇具启发的。Houde

团队在 1992 年开发了一种基于“操纵盒”的方法；这种方法由一个围绕着物体的边界长方体框组成，拖动长方体即拖动物体，另外还有三个旋转柄用于围绕其中心轴旋转物体<sup>[13]</sup>。Conner 团队也采用了设置虚拟控制柄来进行对象操纵；他们的方法允许完整的 9-DOF (Degree of Freedom, 自由度) 控制（平移、旋转和缩放）甚至诸如扭曲的其他变形<sup>[14]</sup>。该方法的虚拟控制柄的两端有一个小球体，它们将几何变换约束在一个平面或轴上；用户拖动其中一个球体可以平移、旋转或缩放物体。秉承着这两个方法的思想，Mendes 团队在 2016 年提出了相似的一个基于外加虚拟控制柄的方法。他们从实验结果中提出了一套发展准则：（1）直接操作很适合粗略的变换；（2）位移和旋转操作应尽可能分离，以防止不需要的变换；（3）单一的 DOF 分离对于精确的变换是非常理想的，通常用于细粒度的调整。

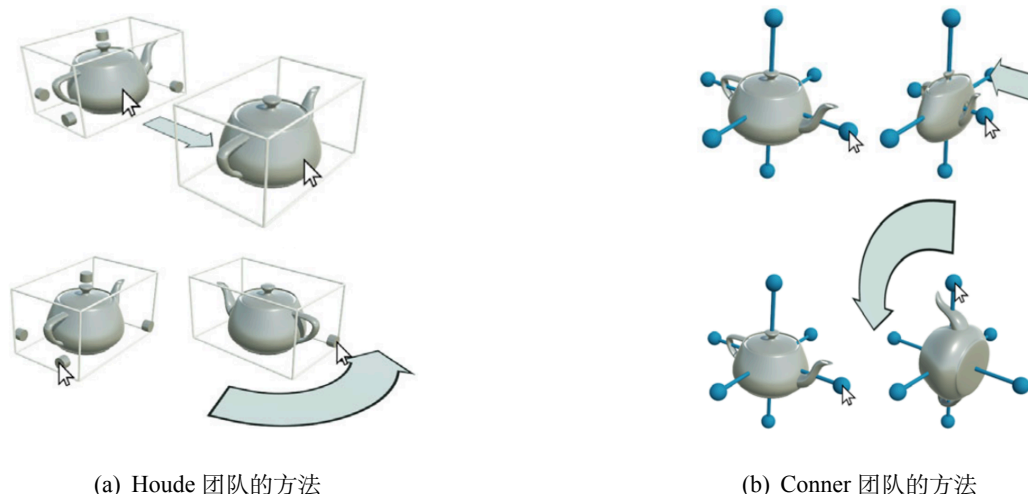


图 2.3 两种具代表性的基于外加虚拟控制柄的对象操纵方法

2010 年以后，基于手部动作（含手柄）追踪的非直接对象操纵方法开始出现。其中较有代表性的两个方法为 Mendes 团队在 2011 年提出的 LTouchIt 和 Kin-Chung Au 团队在 2012 年提出的多点触摸方法（见图）。LTouchIt 虽然使用了直接操纵平移的方法，但在 DOF 分离之后，它可以控制物体在不超过两个维度上的位置，并使用旋转手柄一次围绕一个轴进行旋转；用户可以选择一个手柄来定义一个旋转轴，并通过另一只手的操作来指定旋转角度<sup>[15]</sup>。Au 团队利用多点触摸表面的高输入带宽，将标准变换部件的操作能力委托给多点触摸手势。这使得使用单一的多点触控动作就能对约束和变换操作进行无缝控制。用户可以用两个触摸点选择一个候选轴，通过按住并移动两个手指来进行物体的变换<sup>[16]</sup>。



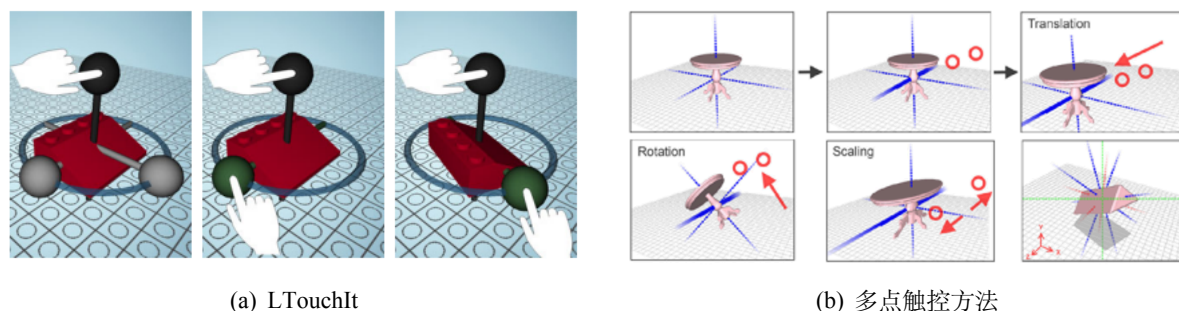


图 2.4 两种具代表性的基于外加虚拟控制柄的对象操纵方法

目前, 基于手部动作(含手柄)的追踪的对象操纵方法的 SOTA (state-of-the-art, 最先进方法) 为 Gloumeau 团队在 2020 年提出的 PinNPivot 方法<sup>[17]</sup>。这个方法使用“销钉(pin)”来约束 1DOF/2DOF/3DOF 旋转; PinNPivot 还支持 6DOF 操纵和 3DOF 平移, 具体的一个操作流程可参考图 6。在该团队与以往技术的比较表明, PinNPivot 拥有更准确和更快的操纵效率。

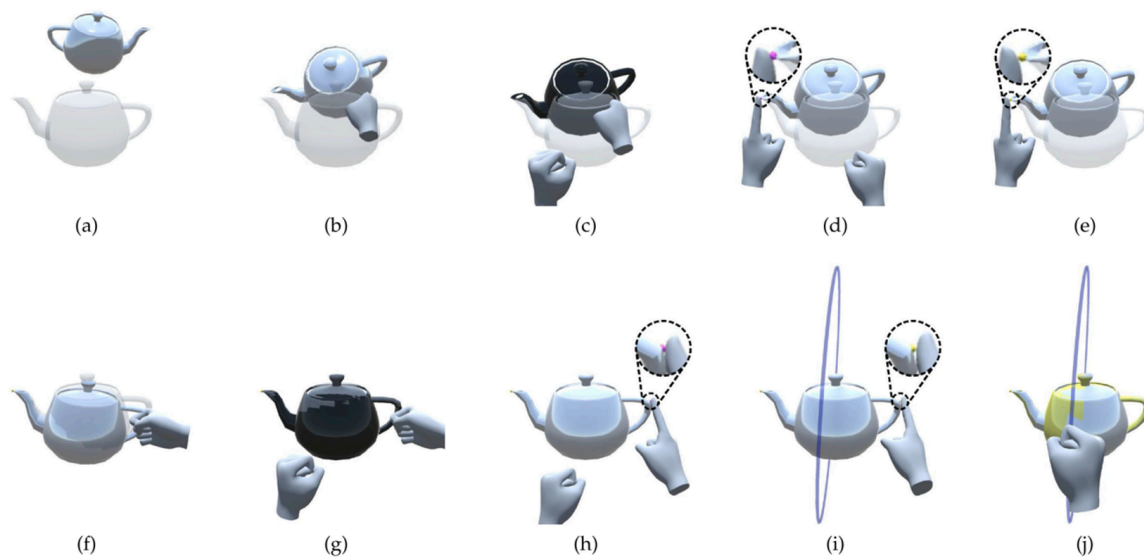


图 2.5 目前基于手部动作(含手柄)的对象操纵最优方法: PinNPivot

## 2.2 基于语音的操纵

同样在 2010 年以后, 基于语音交互的对象操纵方法出现<sup>[18]</sup>。该类方法代表是 Carruth 团队在 2019 年提出的一种使用 HMD 对小型无人驾驶地面车辆的语音操作进行沉浸式培训方法<sup>[19]</sup>、Mirzaei 团队在 2014 年提出的一种使用自然语言和用户手势在真实规模的

数字模拟中进行导航和互动的方法<sup>[20]</sup>和 Ferracani 团队在 2017 年提出的一种通过语音指令游览虚拟博物馆的方法<sup>[21]</sup>。然而,基于语音的交互模式存在许多不理想的因素。

首先,语音交互对用户的要求相对较高,用户必须以标准的发音、强逻辑性和强简洁性陈述自己的语音指令来与系统交互,这无可避免地提高了用户使用语音交互系统时的负担;其次,语音交互的干扰(如噪音)很难消除,错误率较高;再者,语音交互的局限性较大,因为在某些需要安静和隐私的特殊环境(如办公室,学生宿舍等)语音交互系统是不会被考虑的;最后,语音交互难以实现撤回指令,用户无法以简易的方式更正操作失误。

### 2.3 基于眼动的操纵

在 2015 年以后,眼球追踪在头戴式虚拟现实显示器中的应用越来越多,各种集成了眼球追踪器的头盔已经在市场上销售。根据 Adhanom 团队在 2023 年发表的研究,眼动跟踪在虚拟现实中的应用是高度多样化的,并且跨越了多个学科<sup>[22]</sup>。因此,近年来基于眼动追踪的对象操纵方法应运兴起。

在 VR 中实现基于眼睛注视的指向的最常见方法是使用眼动仪提供的 3D 注视方向向量,并观察场景中的哪些对象与方向向量相交<sup>[23]</sup>。通常,射线是基于方向向量投射的,并且射线相交的第一个物体被认为是被指向的项目(见图 7)。这与射线广播的基本思想是一致的。各种研究表明,基于凝视的指向比基于手的指向更快,因为我们能够比我们的手更快地将目光移向目标<sup>[24]</sup>。然而,由于眼球运动的固有生理特性和眼动追踪的技术限制,与其他常见的指点界面相比,基于眼睛注视的准确度还是稍显逊色的<sup>[25][26]</sup>。基于眼睛注视的指向界面中的不准确性主要有两种形式,一是由眼睛跟踪数据中的自然噪声引起的,二是由眼睛跟踪数据质量不稳定引起的。

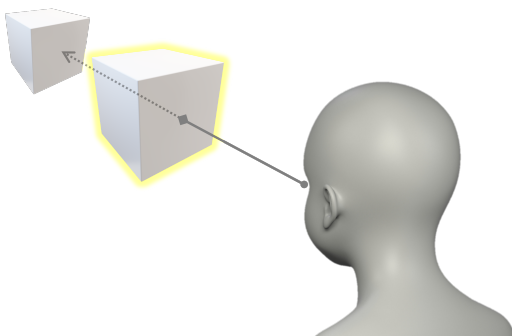


图 2.6 基于眼睛注视的指向选择



基于眼动追踪的操纵方法的一大难点是目标对象选择。仅通过眼睛注视进行选择是一项相对具有挑战性的任务,需要实施更为精密的机制以在虚拟环境中使用基于眼睛的交互。根据以往的方法,我们可以实施其他选择确认技术来辅助眼动交互。这样做的一个额外好处是可以解决人机交互领域经典的“点石成金”问题,即“无论你看哪里,都有东西被激活;你不能在没有发出命令的情况下看任何地方”<sup>[27][28]</sup>。

目前已有许多学者使用各种技术来实现虚拟环境中基于注视的交互选择确认。Hansen 团队在 2018 年提出了一种基于眼睛注视的停留进行选择确认的技术<sup>[25]</sup>。Sidenmark 和 Gellersen 在 2019 年实施了两种头部辅助的眼动交互技术,第一种是 Eye & Head Dwell,第二种是 Eye & Head Convergence<sup>[23]</sup>。Eye & Head Dwell 是一种停留以确认的技术,其中停留计时器仅由头部支持的凝视转移触发,但可以通过仅眼睛凝视暂停和恢复;Eye & Head Convergence 是一种用于快速目标确认的替代技术,它允许用户通过将眼睛指示器和头部指示器对准目标来确认选择。Kumar 和 Sharma 团队在 2016 年提出了一种使用眨眼进行选择确认的技术<sup>[29]</sup>。Pfeuffer 团队在 2017 年提出了一种手眼协统的选择确认方法;这个方法允许用户用眼睛注视物体并同时加以一种“捏合”手势来辅助确认选择<sup>[30]</sup>。Pai 团队在 2019 年提出了另外一种协同辅助操纵技术;用户可以用目光指向目标,并使用肌电图检测到的手臂肌肉收缩来触发操纵动作<sup>[31]</sup>。Qian 和 Teather 团队在 2017 年提出了一种通过键盘按钮按下进行选择确认辅助,并使用眼睛注视进行指向选择的方法<sup>[32]</sup>。最近, Sidenmark 团队在 2020 年提出了 Outline Pursuits 方法,它利用平滑追踪来允许用户在虚拟环境中选择被遮挡的对象<sup>[33]</sup>。

伴随选择技术而来的另外一个值得关注的问题是反馈技术。一个完整的虚拟环境中的对象操纵技术应该向用户提供反馈,让用户能够清楚地了解系统的状态<sup>[34]</sup>。由于眼睛对视野中的视觉变化很敏感,它们会本能地尝试将注意力转移到这些视觉变化上。因此,在向用户提供反馈时应该格外小心,因为视觉上突出的反馈机制可能会产生意想不到的后果,即转移用户的视线以产生不期望的交互动作。Boyer 团队在 2017 年提出的一种非视觉反馈方法,他们使用听觉反馈来避免不必要的视线转移<sup>[35]</sup>。然而,学界依然有很多基于视觉的反馈方法:Blattgerste 团队在 2018 年提出了一种突出显示所选对象的反馈方法;Mohan 团队也在 2018 年提出了一种在所选对象周围显示确认标志的方法;Sidenmark 团队在 2020 年提出了一种在所选对象周围显示轮廓的反馈方法<sup>[36][37][33]</sup>。

根据目前基于眼动追踪的对象操纵方法的研究数量,眼动追踪将很快成为 HMD 系统不可或缺的一部分。因此,我们预计围绕 HMD 眼动追踪的研究和开发将在未来几年



加速和扩展。然而,目前大部分眼动追踪方法依旧存在许多值得优化的问题。除了硬件限制,交互动作所带来的生理性不适也需要得到改善。大多数基于眼动追踪的操纵方法都包含眨眼命令(包括眨眼、双眼眨眼和眨眼眼球运动)<sup>[29]</sup>。要求用户改变他们的自然眨眼频率可能会导致用户眼睛疲劳、眼睛干涩和眼睛疲劳<sup>[38]</sup>。Kumar 和 Sharma 团队在 2016 年的研究结果也表明,频繁眨眼和眨眼会导致用户眼睛疲劳<sup>[29]</sup>。而且,基于眨眼的界面往往不准确,因为下意识的眨眼很难与自然眨眼区分开来,所以系统往往需要用户做出完全下意识的眨眼动作,如快速多次眨眼。然而,长时间眨眼有明显的缺点,例如减慢交互流程并在长时间眨眼期间阻挡用户的视线。因此,基于眼睛注视的系统控制需要在交互动作上做出合理的优化。

一个值得学习的基于凝视的对象移动方法是 Liu 团队在 2020 年提出的 OrthoGaze,它允许用户只需用眼睛凝视就能直观地操纵虚拟物体的三维位置(见图 8)<sup>[39]</sup>。该方法利用了三个可选择的正交平面,其中每个平面不仅有助于引导用户在任意虚拟空间中的注视,而且还允许对物体位置进行 2DOF 操作,是一种完全不依赖手部的对象操纵方法。

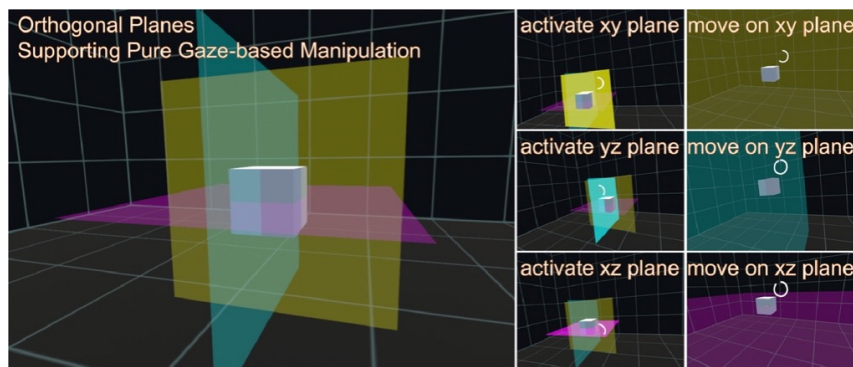


图 2.7 一个基于眼动和正交平面的对象操纵方法: OrthoGaze

目前涵盖眼动追踪的对象操纵方法的 SOTA 为 Yu 团队在 2021 年提出的一种基于凝视和手部动作的三维物体操纵方法<sup>[40]</sup>。该方法的整个操纵任务可以分解为四个阶段:指示、确认、操作和释放(见图 9)。该研究表明,当所有目标都位于用户前方且在手臂可触及的距离内时,基于凝视的交互对于物体操作没有明显的性能优势;但对于有远处物体的较大空间,凝视输入可以减轻手臂的疲劳问题。眼动和其他模态的不同整合、协调和过渡策略可以为构建更高效的物体操纵技术提供优势。然而,这个方法依旧依靠手部动作,并不能被视为一个完全无手(hands-free)的操纵方法。

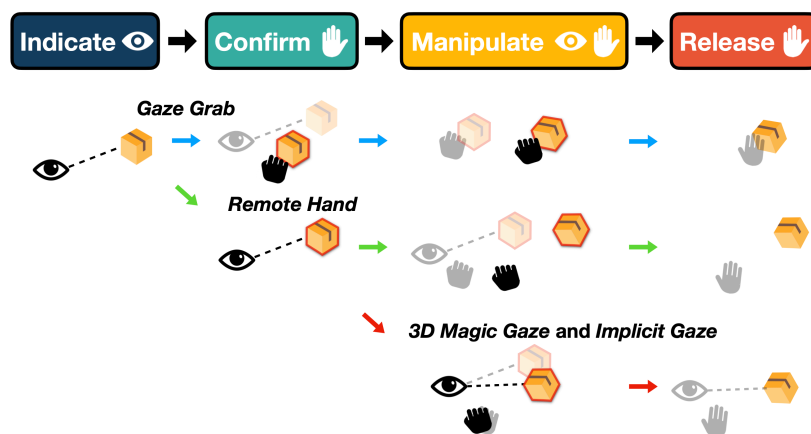


图 2.8 目前涵盖凝视的物体操纵最优方法



### 3 方法设计与实现

3.1 交互系统有限状态机

3.2 场景与目标浏览

3.3 目标选择

3.4 对象操纵



## 4 实验设计

### 4.1 先导实验

#### 4.1.1 实验 1: 眼动确认信号筛选

#### 4.1.2 实验 2: 眼动操纵视线驻留检测优化

### 4.2 用户实验

#### 4.2.1 实验 1: 单物体位移对接实验

#### 4.2.2 实验 2: 单物体操纵对接实验



## 致谢

我测你们码的这个格式咋这么难调。



## 参考文献

- [1] M. Daniel, C. Ariel, G. Andrea, et al. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments: Survey on 3d virtual object manipulation [J]. Computer Graphics Forum, 2018.
- [2] R. Jaziar, M. T. A., F. Jennifer, et al. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda [J]. Computers & Education, 2020.
- [3] P. Ivan, B. Mark, W. Suzanne, et al. The go-go interaction technique: Non-linear mapping for direct manipulation in vr[C]//Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology. [S.l.: s.n.], 1996.
- [4] D. A. Bowman, L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments[C]//Proceedings of the 1997 symposium on Interactive 3D graphics (I3D '97). [S.l.: s.n.], 1997.
- [5] S. Frees, G. D. Kessler. Precise and rapid interaction through scaled manipulation in immersive virtual environments[C]//IEEE Proceedings. VR 2005. Virtual Reality. [S.l.: s.n.], 2005.
- [6] C. Wilkes, D. A. Bowman. Advantages of velocity-based scaling for distant 3d manipulation[C]//Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology (VRST '08). [S.l.: s.n.], 2008.
- [7] A. Chris, G. Mark, F. Scott. Increasing precision for extended reach 3d manipulation[J]. International Journal of Virtual Reality, 2015.
- [8] N. Osawa. Two-handed and one-handed techniques for precise and efficient manipulation in immersive virtual environments[J]. roceedings of the 4th International Symposium on Advances in Visual Computing (ISVC '08), 2008.
- [9] M. Anthony, C. Géry, G. Laurent. The design and evaluation of 3d positioning techniques for multi-touch displays[C]//2010 IEEE Symposium on 3D User Interfaces (3DUI). [S.l.: s.n.], 2010.
- [10] M. Anthony, C. Géry, G. Laurent. The effect of dof separation in 3d manipulation tasks



- with multi-touch displays[C]//Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology. [S.l.: s.n.], 2010.
- [11] H. Mark, ten Cate. Thomas, C. Sheelagh. Sticky tools: Full 6dof force-based interaction for multi-touch tables[C]//Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces. [S.l.: s.n.], 2009.
- [12] R. J. L., D. P. L., H. J. Y. A screen-space formulation for 2d and 3d direct manipulation [C]//Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology. [S.l.: s.n.], 2009.
- [13] S. Houde. Iterative design of an interface for easy 3-d direct manipulation[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92). [S.l.: s.n.], 1992.
- [14] C. B. D., S. S. S., H. K. P., et al. Three-dimensional widgets[C]//Proceedings of the 1992 Symposium on Interactive 3D Graphics. [S.l.: s.n.], 1992.
- [15] M. Daniel, L. Pedro, F. Alfredo. Hands-on interactive tabletop lego application[C]//Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology. [S.l.: s.n.], 2011.
- [16] A. O. Kin-Chung, T. Chiew-Lan, F. Hongbo. Multitouch gestures for constrained transformation of 3d objects[J]. Computer Graphics Forum, 2012.
- [17] G. P. Christopher, S. Wolfgang, H. JungHyun. Pinnpivot: Object manipulation using pins in immersive virtual environments[J]. IEEE Transactions on Visualization and Computer Graphics, 2021.
- [18] M. Pedro, G. Guilherme, C. Hugo, et al. Hands-free interaction in immersive virtual reality: A systematic review[J]. IEEE Transactions on Visualization and Computer Graphics, 2021.
- [19] C. Daniel, H. Christopher, B. Cindy, et al. Using hmd for immersive training of voice-based operation of small unmanned ground vehicles[M/OL]. Springer International Publishing, 2019. [https://link.springer.com/chapter/10.1007/978-3-030-21565-1\\_3](https://link.springer.com/chapter/10.1007/978-3-030-21565-1_3).
- [20] M. M. A., C. J.-R., M. F., et al. Navigation and interaction in a real-scale digital mock-up using natural language and user gesture[C]//Proceedings of the 2014 Virtual Reality International Conference. [S.l.: s.n.], 2014.



- [21] A. Ferracani, M. Faustino, G. X. Giannini, et al. Natural experiences in museums through virtual reality and voice commands[C]//Proceedings of the 25th ACM International Conference on Multimedia. [S.l.: s.n.], 2017.
- [22] A. I.B., M. P., F. E. Eye tracking in virtual reality: a broad review of applications and challenges[J]. Virtual Reality, 2023.
- [23] S. Ludwig, G. Hans. Eye & head: Synergetic eye and head movement for gaze pointing and selection[C]//Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology. [S.l.: s.n.], 2019.
- [24] T. Vildan, J. R. J. K. Interacting with eye movements in virtual environments[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. [S.l.: s.n.], 2000.
- [25] H. J. Paulin, R. Vijay, M. I. Scott, et al. A fitts' law study of click and dwell interaction by gaze, head and mouse with a head-mounted display[C]//Proceedings of the Workshop on Communication by Gaze Interaction. [S.l.: s.n.], 2018.
- [26] L. F. Lopez, S. Veronica. A comparative study of eye tracking and hand controller for aiming tasks in virtual reality[C]//Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications. [S.l.: s.n.], 2019.
- [27] J. R. J. K. What you look at is what you get: Eye movement-based interaction techniques [C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. [S.l.: s.n.], 1990.
- [28] J. Rob, S. Sophie. What you look at is what you get: Gaze-based user interfaces[J]. Interactions, 2016.
- [29] K. Devender, S. Amit. Electrooculogram-based virtual reality game control using blink detection and gaze calibration[C]//2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI). [S.l.: s.n.], 2016.
- [30] P. Ken, M. Benedikt, M. Diako, et al. Gaze + pinch interaction in virtual reality[C]//Proceedings of the 5th Symposium on Spatial User Interaction. [S.l.: s.n.], 2017.
- [31] P. Y. Suen, D. Tilman, K. Kai. Assessing hands-free interactions for vr using eye gaze and electromyography[J]. Virtual Reality, 2019.
- [32] Q. Y. Yuan, T. R. J. The eyes don't have it: An empirical comparison of head-based and





- eye-based selection in virtual reality[C]//Proceedings of the 5th Symposium on Spatial User Interaction. [S.l.: s.n.], 2017.
- [33] S. Ludwig, C. Christopher, Z. Xuesong, et al. Outline pursuits: Gaze-assisted selection of occluded objects in virtual reality[C]//Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. [S.l.: s.n.], 2020.
- [34] M. Päivi, B. Andreas. Eye tracking and eye-based human-computer interaction[M/OL]. Springer London, 2014. [https://doi.org/10.1007/978-1-4471-6392-3\\_3](https://doi.org/10.1007/978-1-4471-6392-3_3).
- [35] B. Eric, L. Jean, P. Arthur, et al. Continuous auditory feedback of eye movements: An exploratory study toward improving oculomotor control[J]. *Frontiers in Neuroscience*, 2017.
- [36] B. Jonas, R. Patrick, P. Thies. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views[C]//Proceedings of the Workshop on Communication by Gaze Interaction. [S.l.: s.n.], 2018.
- [37] M. Pallavi, G. W. Boon, F. Chi-Wing, et al. Dualgaze: Addressing the midas touch problem in gaze mediated vr interaction[C]//2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). [S.l.: s.n.], 2018.
- [38] H. Teresa, C. Maurice, R. Enrico, et al. A survey of digital eye strain in gaze-based interactive systems[C]//ETRA '20: 2020 Symposium on Eye Tracking Research and Applications. [S.l.: s.n.], 2020.
- [39] C. Liu, A. Plopski, J. Orlosky. Orthogaze: Gaze-based three-dimensional object manipulation using orthogonal planes[J]. *Computers & Graphics*, 2020.
- [40] Y. Difeng, L. Xueshi, S. Rongkai, et al. Gaze-supported 3d object manipulation in virtual reality[C]//Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. [S.l.: s.n.], 2021.
- [41] M. Daniel, R. Filipe, F. Alfredo, et al. The benefits of dof separation in mid-air 3d object manipulation[C]//Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology. [S.l.: s.n.], 2016.
- [42] P. D. A. G. F. D. et al. Types of eye movements and their functions[M/OL]. Sinauer Associates, 2001. <https://www.ncbi.nlm.nih.gov/books/NBK10991/>.
- [43] K. Kyungyoon, L. R. L., K. Nikki, et al. Anatomical 2d/3d shape-matching in virtual



---

reality: A user interface for quantifying joint kinematics with radiographic imaging[C]//  
2017 IEEE Symposium on 3D User Interfaces (3DUI). [S.l.: s.n.], 2017.