



单位代码 10006

学 号 19373345

分 类 号 TP391.7

密 级 公开

北京航空航天大学  
BEIHANG UNIVERSITY

# 毕业设计(中期报告)

## 虚拟现实头眼协同 对象操纵方法设计与实现

院（系）名称	计算机学院
专 业 名 称	计算机科学与技术
学 生 姓 名	刘兆薰
学院指导教师	王莉莉
工业部门导师	无

2023 年 4 月



## 1. 论文选题依据

### 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 国内外研究现状

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）动作的追踪；（2）基于语音的交互；（3）基于眼动的追踪<sup>[1][2]</sup>。

#### 1.4.1 基于手部（含手柄）动作的追踪

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>[5]</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>[5]</sup>。

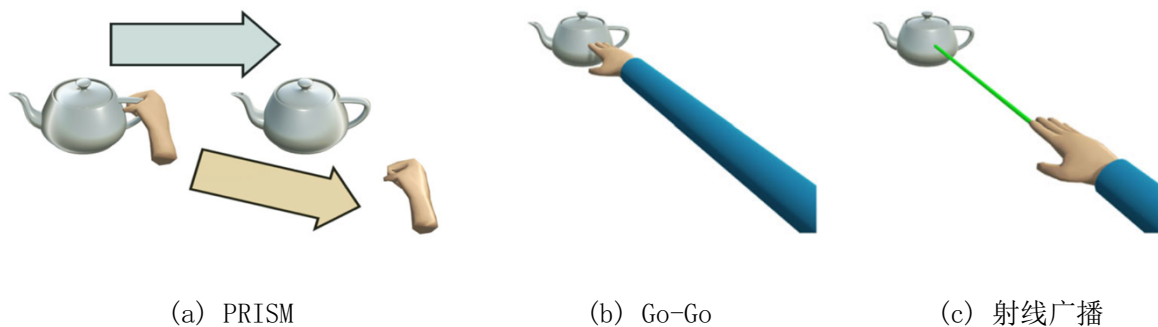


图 1 研究早期三种主要的基于手部动作追踪的对象操纵方法



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>[4]</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

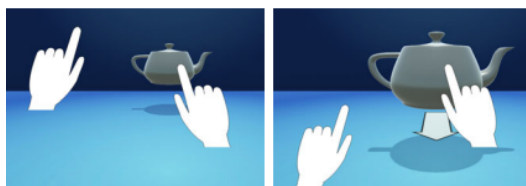
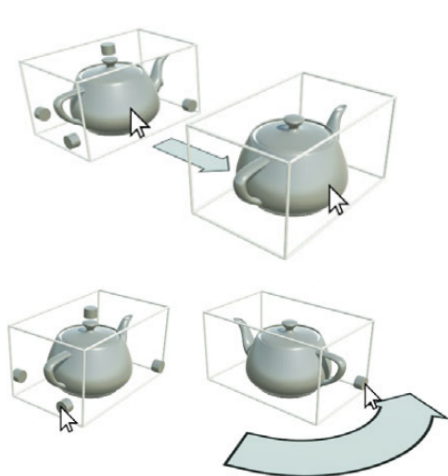


图 2 基于双手直接操纵的 Z 技术

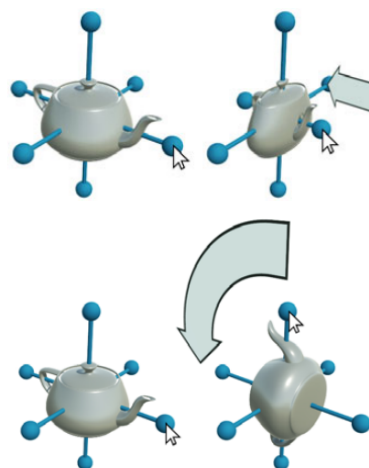


图 3 基于双手直接操纵的 Sticky Tools 方法

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>[12]</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>[13]</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



(a) Houde 团队的方法



(b) Conner 团队的方法

图 4 外加虚拟控制柄的两种代表方法

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>[14][15]</sup>。

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>[9]</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>[10]</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。

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>[16]</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>[17]</sup>。

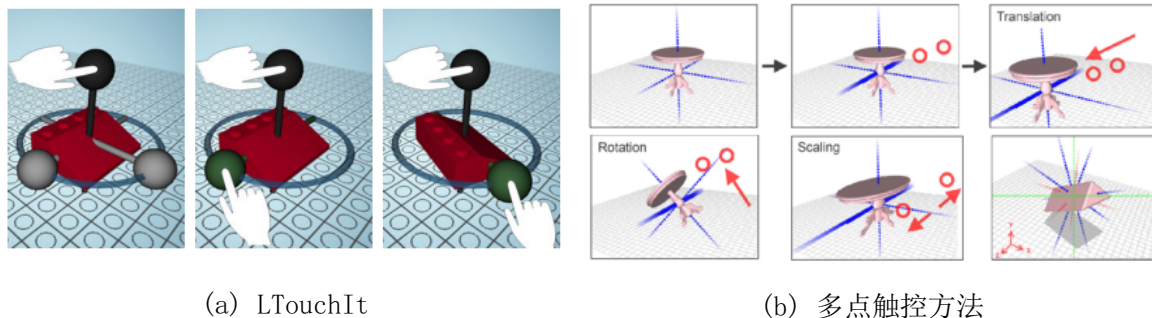


图 5 两种基于手部动作(含手柄)追踪的非直接对象操纵方法



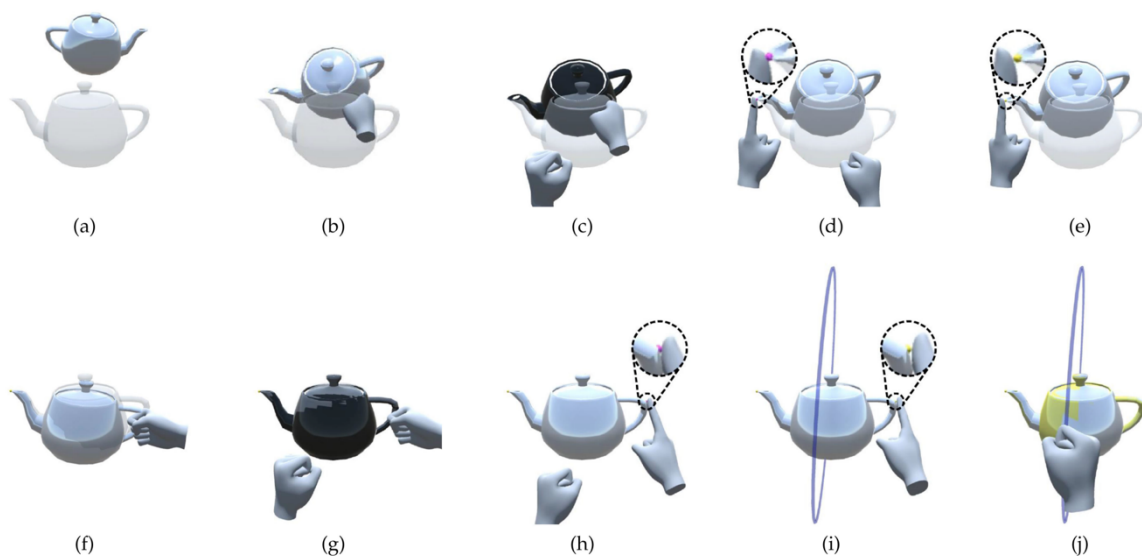


图 6 PinNPivot 完成物体转移的过程：(a)初始位置和目标位置；(b)6DOF 操纵迅速将物体翻转过来；(c)按比例的 3DOF 平移准确地放置壶口的尖端；(d)一个销钉被创建；(e)销钉被锁定；(f)物体在 3DOF 中快速旋转；(g)物体在按比例的 2DOF 中被精确旋转；(h)创建第二个销钉；(i)第二个销钉被锁定，出现一个环；(j)该物体在按比例的 1DOF 中旋转。当目标变成黄色时，表示与目标位置吻合良好。

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>[39]</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。

#### 1.4.2 基于语音的交互

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。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。

### 1.4.3 基于眼动的追踪

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>[21]</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>[22]</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>[23]</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>[24][25]</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。

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>[26][27]</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

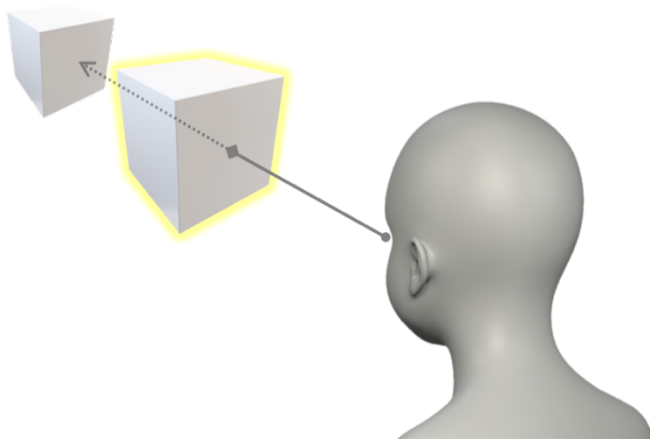


图 7 基于眼睛注视的指向选择



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>[24]</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>[22]</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>[28]</sup>。Pfeuffer 团队在 2017 年提出了一种手眼协同的选择确认方法；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>[29]</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>[30]</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>[31]</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>[32]</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>[33]</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>[34]</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>[35][36][32]</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

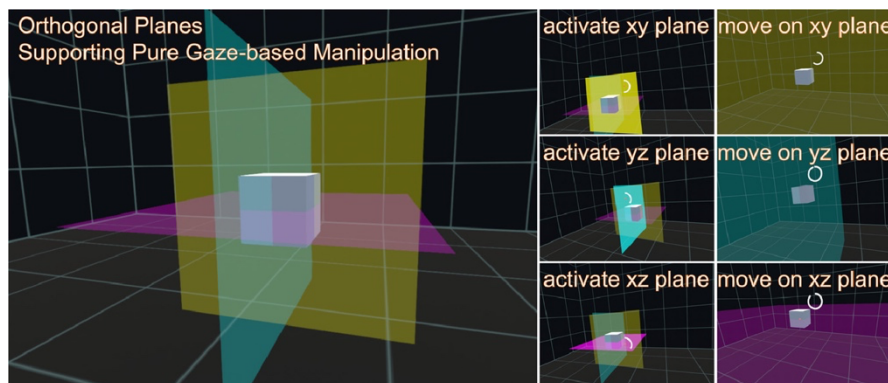


图 8 OrthoGaze 的正交平面

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>[28]</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>[37]</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>[28]</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<sup>[41]</sup>(见图 8)。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。

目前基于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。

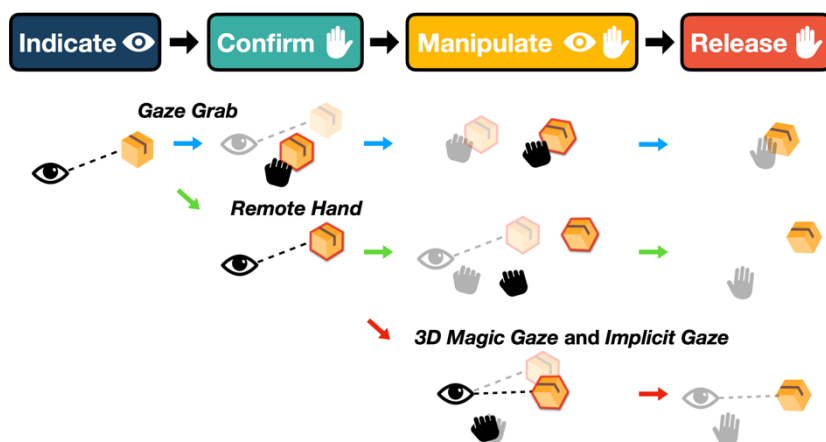


图 9 Yu 团队提出的基于凝视的三维物体操纵方法操作流程

## 2. 研究方案

### 2.1 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。

### 2.2 研究内容（已完成）

#### 2.2.1 有限状态机

交互系统的运作遵循有限状态机（见图 10），该状态机确保交互系统的稳定运行。

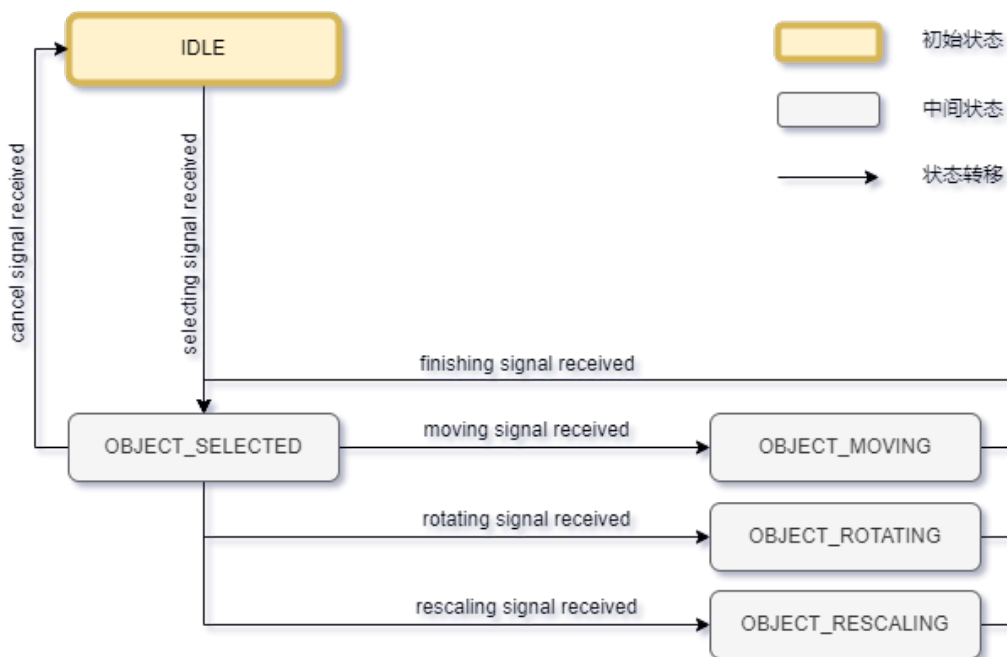


图 11 交互系统的有限状态机



### 2.2.2 场景与目标浏览

在场景与目标浏览时，我们首先需要计算实时视点。视点的计算涉及到两个方面：视角和视位。视角是指用户在虚拟环境中的观察方向，通常由头部的旋转角度决定。视位是指用户在虚拟环境中的位置，通常由身体的移动或手柄的操作决定。为了实时地获取用户的视角和视位，需要使用一些传感器或追踪器，例如红外摄像头等。这些设备可以测量用户头部或身体的运动数据，并将其传输给计算机。根据用户的视角和视位，我们可以获取用户当前所期望观察、选中以及操作的物体。

在判断驻留点时，我们使用射线广播（ray-casting）的方法。射线广播是一种在虚拟现实中选择对象的常用技术；在我们的交互系统中，它利用用户的头部方向来发射一条射线，与场景中的对象进行碰撞检测，从而获取驻留点。射线广播的优点是简单、直观、高效、学习成本低并且不容易产生眩晕感和迷惑感。当驻留点在某个物体上时，该物体将被高亮，以消除对象选择时的歧义。

### 2.2.3 目标选择

在状态机的 IDLE 状态下，我们使用眼动信号来选择目标。该方法的主要挑战包括如何准确地捕捉用户的眼动数据，如何有效地识别用户的选择意图，以及如何提供合适的反馈和提示。

最为常见且自然的主动眼动信号有单眼眨眼和快速两次眨眼。我们首先设计了一个前导实验（Pilot Study）以从这两种眼动信号中，确定一个最为高效并且带来最小使用压力的眼动选择以及确认信号，这将在之后的实验设计章节中作更细致的介绍。我们最终确定使用快速两次眨眼来作为我们的目标选择以及确认信号。

在处理眼动数据时，我们使用了中位值平均滤波算法，以处理由于硬件限制而较为不稳定的眼动脉冲信号。中位值平均滤波算法是一种常用的数字信号处理方法，它结合了中位值滤波和算术平均滤波的优点，能有效地抑制脉冲噪声和周期性干扰，提高信号的平滑度和稳定性。中位值平均滤波算法的基本思想是：对于给定的一组采样数据，先去掉其中的最大值和最小值，然后对剩余的数据求算术平均值，作为该组数据的滤波输出。中位值平均滤波算法的优点有以下几个方面：1. 它能有效地消除由偶然出现的脉冲性干扰所引起的采样值偏差，保证了信号的真实性和完整性；2. 它对周期性干扰有良好的抑制作用，能够保留信号的基本特征；3. 它具有较高的平滑度，适用于高频振荡的系统；4. 该算法不需要排序，计算成本较低且反馈速度较快。针对我们的交互系统的某一时刻，我们对其前 SAMPLE 帧采用中位值平均滤波算法，见算法 1。

---

**Algorithm 1** MedianAverageFilter

---

```
1: SAMPLE  $\leftarrow 5$ 
2: dataBuffer  $\leftarrow$  EMPTY LIST of NUMBER
3: currentBufferSize  $\leftarrow$  dataBuffer.Count()
4: if currentBufferSize < SAMPLE then
5:   for i in Range(SAMPLE - currentBufferSize) do
6:     dataBuffer.Enqueue(GetData())
7:     Delay()
8:   end for
9: else
10:  dataBuffer.Dequeue()
11:  dataBuffer.Enqueue(GetData())
12:  Delay()
13: end if
14:
15: max  $\leftarrow -INF$ 
16: min  $\leftarrow INF$ 
17: for value in dataBuffer do
18:   max  $\leftarrow$  value if value > max else max
19:   min  $\leftarrow$  value if value < min else min
20: end for
21: dataBuffer.RemoveFirst(max)
22: dataBuffer.RemoveFirst(min)
23:
24: sum  $\leftarrow 0$ 
25: for value in dataBuffer do
26:   sum  $\leftarrow$  sum + value
27: end for
28: return sum / (SAMPLE - 2)
```

---

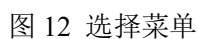
算法 1 针对交互系统的中位值平均滤波算法

在目标选择时，单纯依靠眼动或者头动的射线广播来选择物体是不理想的；这样的动作是不够自然的，会引入不必要的使用负担。因此，我们引入了一种头眼协同的目标选择方式，为选择过程引入更小的负担，使交互流程更加自然。同时，为了对抗干扰而触发更加精确的动作，我们通过以下方式优化用户的眼动操纵注视停留。我们规定在  $t_0$  时刻的眼动操纵注视停留  $OE$  被计算为在时间段  $n$  内的眼动向前射线和头动向前射线的配合角度偏移：

$$OE_{t_0} = \frac{1}{n} \sum_{t=t_0-n}^{t_0} |\widehat{eye}_t \cdot \widehat{head}_t - \widehat{eye}_{t-1} \cdot \widehat{head}_{t-1}|$$

其中， $\widehat{eye}$  是视线的单位向量， $\widehat{head}$  是头部注视的单位向量， $\cdot$  是向量内积运算符。如果  $OE$  小于某一阈值，则表示用户正在尝试注视。我们的一个先导实验表明，这个优化是必要的；它可以使交互流程更加自然，引入更小的使用负担。

在选择确认后，系统的状态转移到 OBJECT\_SELECTED，并且会提供给用户一个基于听觉的声音反馈信号。这个反馈信号相对于视觉是异模态的，目的是提高交互的效率、流畅性和准确性以及增强用户的信任感和舒适度。



在状态机的 `OBJECT_SELECTED` 状态下，我们可以通过一个选择菜单（见图 12）来选择进入到具体交互模式：空间位移、空间旋转，分别对应系统有限状态机中的 `OBJECT_MOVING`、`OBJECT_ROTATING` 和 `OBJECT_RESCALING` 三个状态。在具体的交互模式下完成操纵任务后，用户通过再次快速两次眨眼以确认操纵，并且返回到 `OBJECT_SELECTED` 状态；这时，用户即可以继续选择进入下一个具体的交互模式，也可以选择返回到 `IDLE` 状态以浏览、操纵其他对象。这个过程可以在我们规定的系统有限状态机中体现。

在选择菜单中的选择确认是通过注视停留实现的。如果要选择进入某一种模式，用户只需将视线向其对应方位移动。在尝试注视选择时，用户界面上会出现一个确认进度提示符。在提示符进度加载期间，用户可以通过回归视线来取消选择；当提示符进度加载完成时，系统则进入对应的状态。

图 13 信号增强函数（蓝色为增强前信号，红色为增强后信号）



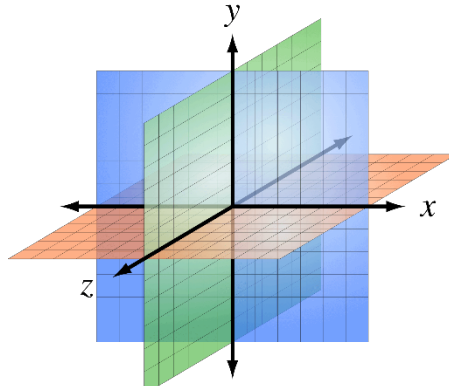


图 14 操纵空间中的空间直角坐标系

动信号以对抗不稳定脉冲信号的干扰。同时我们引入一种特殊的信号增强函数  $A(v) = 5 \cdot v^5$ ,  $-1 \leq v \leq 1$ , 让用户在小幅度动作时可以对对象进行微调, 在大幅度动作时可以让对象作快速动作, 见图 13。为了减小操作负担, 我们都尽可能采用眼动进行全权的操纵, 并且尽量不引入更多的操纵模态。我们在对象操纵空间中设置了一个三维直角坐标系, 见图 14。

在空间位移时, 对于 X-Y 平面的移动, 对象跟随眼动向前射线在 X-Y 平面的投影距离作出相应的动作; 假设眼动向前射线在 X-Y 平面的投影坐标为  $(x, y)$ , 对象在 X 轴向和 Y 轴向的移动为  $\delta_x = \begin{cases} 0, & \text{if } |x| < T \\ x \cdot C, & \text{default} \end{cases}$ ,  $\delta_y = \begin{cases} 0, & \text{if } |y| < T \\ y \cdot C, & \text{default} \end{cases}$ , 其中  $T$  和  $C$  分别是预先规定的阈值和比例系数。对于 Z 轴方向的移动, 对象跟随头部以 Z 轴为旋转轴的角度-距离映射作出相应的动作; 假设头动绕 Z 轴的旋转角度为  $\omega$ , 对象在 Z 轴方向的移动为  $\delta_z = \begin{cases} 0, & \text{if } |\omega| < T \\ \sin \omega \cdot C, & \text{default} \end{cases}$ , 其中  $T$  和  $C$  分别是预先规定的阈值和比例系数。

在空间旋转时, 对于以 X 轴的旋转, 对象跟随眼动向前射线在 Y 轴的投影距离-角度映射作出相应的动作; 对于以 Y 轴的旋转, 对象跟随眼动向前射线在 X 轴的投影距离-角度映射作出相应的动作; 假设眼动向前射线在 X-Y 平面的投影坐标为  $(x, y)$ , 对象在绕 X 轴向和 Y 轴向的旋转为  $\omega_x = \begin{cases} 0, & \text{if } |x| < T \\ \sin^{-1} x \cdot C, & \text{default} \end{cases}$ ,  $\omega_y = \begin{cases} 0, & \text{if } |y| < T \\ \sin^{-1} y \cdot C, & \text{default} \end{cases}$ , 其中  $T$  和  $C$  分别是预先规定的阈值和比例系数。。

在空间缩放时, 对象跟随眼动向前射线在 X 轴的投影距离-缩放系数映射作出相应的动作; 假设眼动向前射线在 X 轴的投影坐标为  $(x, 0)$ , 对象的缩放系数  $K = \begin{cases} 0, & \text{if } |x| < T \text{ or } x \leq -1 \\ 2, & \text{if } x \geq 1 \\ 1 + x, & \text{default} \end{cases}$ , 最终具

体表现为  $Scale' = Scale \cdot K$ 。



## 2.3 先导实验（已完成）

### 2.3.1 先导实验 1：眼动确认信号筛选

此先导实验的目的是从最为常见且自然的两种主动眼动信号（单眼眨眼和快速两次眨眼）中确定一个最为高效并且带来最小使用压力的眼动确认信号。

我们招募了 6 名参试者；为了保证公平性和均一性，我们的此次实验组由 3 名男性和 3 名女性构成。参试者在开始实验前填写第一个个人信息调查问卷，以询问性别、工作机构、年龄、VR 使用经验和能否自然地完成单眼眨眼的动作。

每名参试者将使用不同的两个眼动确认信号分两次完成同样的一个任务。该任务内容是，在虚拟环境中会依次在随机时刻和随机位置出现 20 个小球，参试者需要使用当次实验规定的眼动信号（单眼眨眼或快速两次眨眼）选择以消灭它们。为了保证每次实验结果的可比较性，我们规定：每个小球只有在参与者瞄准并且发出确认信号后才能消灭；下一个小球只有在当前小球被消灭后才生成；在每次发出确认信号前，参与者必须按下一次键盘上的空格键，以记录实际的信号尝试次数。

在实验完成后，每名参试者将分别针对两种确认信号完成一个 NASA-TLX 任务负担表。我们将针对以下三个客观度量评估实验结果：1. 任务完成时间（毫秒）；2. 任务负担（NASA-TLX 分数）；3. 信号反馈精确系数（Feedback Accuracy Index, FAI）。其中，FAI 的计算规则是：

$$FAI = \frac{\text{消灭的小球数量}}{\text{实际的确认信号尝试次数}}$$

本次先导实验的实验结果汇报为：

参试者	来源	单眼眨眼			快速两次眨眼		
		任务完成时间 (ms)	NASA-TLX	FAI	任务完成时间 (ms)	NASA-TLX	FAI
男, 22岁, 用过 VR, 无法单眼眨眼	北航计算机学院	187621	92	0.43	105832	71	0.79
男, 21岁, 熟练 VR, 可以单眼眨眼	北航计算机学院	80122	71	0.87	113475	68	0.88
男, 22岁, 用过 VR, 无法单眼眨眼	滑铁卢大学计算机学院	161298	83	0.39	92103	79	0.91
女, 20岁, 没有VR 经验, 可以单眼眨眼	北航经济管理学院	89172	62	0.91	125764	73	0.76
女, 21岁, 用过 VR, 可以单眼眨眼	滑铁卢大学计算机学院	90817	77	0.81	109675	91	0.86
女, 21岁, 用过 VR, 无法单眼眨眼	多伦多大学计算机学院	80012	89	0.49	98122	80	0.83

表 1 眼动确认信号筛选先导实验结果



我们采用曼-惠特尼 U 检验 (Mann-Whitney U Test) 来分析数据的显著差异性; 该检验假设两个样本分别来自除了总体均值意外完全相同的两个总体, 目的是检验这两个总体的均值是否有显著的差异。对于任务完成时间, 检验结果是  $U = 18, p < 0.05$ ; 对于任务负担, 检验结果是  $U = 42, p < 0.05$ ; 对于 FAI, 检验结果是  $U = 15.5, p < 0.05$ 。可以发现, 三组结果皆存在显著差异; 而快速两次眨眼均值表现更为理想, 故此可以确定快速两次眨眼显著优于单眼眨眼。

### 2.3.2 先导实验 2: 眼动操纵注视停留检测优化

此先导实验的目的是是否有必要引入眼动操纵注视停留检测的优化。

此次实验的实验组和先导实验 1 一致。每名参试者将分两次完成同样的一个任务; 在两次的实验中, 我们会随机在某一次引入优化以消除主观的心理作用干扰。该任务内容和先导实验 1 一致; 我们规定在此次先导实验中, 两次任务都使用快速两次眨眼作为确认信号。

在实验完成后, 每名参试者将分别针对两次任务完成一个 NASA-TLX 任务负担表; 由于用户在实验过程中并不知情某次任务是否引入优化, 所以我们可以认为结果是客观的。我们将针对任务负担评估实验结果。

本次先导实验的实验结果汇报为:

参试者	来源	引入优化组别 (用户不可见)	含优化	不含优化
			NASA-TLX	NASA-TLX
男, 22岁, 用过 VR, 无法单眼眨眼	北航计算机学院	1	73	80
男, 21岁, 熟练 VR, 可以单眼眨眼	北航计算机学院	2	47	53
男, 22岁, 用过 VR, 无法单眼眨眼	滑铁卢大学计算机学院	1	68	62
女, 20岁, 没有VR 经验, 可以单眼眨眼	北航经济管理学院	2	75	83
女, 21岁, 用过 VR, 可以单眼眨眼	滑铁卢大学计算机学院	2	70	79
女, 21岁, 用过 VR, 无法单眼眨眼	多伦多大学计算机学院	1	82	86

表 2 眼动操纵注视停留检测优化先导实验结果

我们同样采用曼-惠特尼 U 检验来分析数据的显著差异性。对于任务负担, 检验结果是  $U = 32, p < 0.05$ 。可以发现, 两组结果存在显著差异; 而引入优化的任务负担均值表现更为理想, 故此可以确定引入优化是必要的。

## 2.4 用户实验（未完成）

为了消除偶然性、保证多样性，我们预计招募 20 名参试者，并且尽可能确保男女比例相当。

### 2.4.1 用户实验 1：单物体位移对接实验

我们将 OrthoGaze 中作为一个基准线（baseline），并且复用其中的一个物体位移用户实验。该用户实验设计为：参与者位于起始位置  $(0, 1[m], 0)$  处，并尝试将大小为  $0.5[m] \times 0.5[m] \times 0.5[m]$  的白色立方体从固定的起始位置  $(-1[m], 0.5[m], 5.5[m])$  移动到多个目标位置。在每次实验中，在目标位置上将出现一个半透明的、与白色立方体大小相同的绿色立方体；参与者必须将白色立方体与绿色立方体对齐。目标位置总是位于一个具有  $2N[m]$  边长的立方体空间的角落，其中心与白色立方体的初始位置中心重合，如图 15 所示。我们使用了 8 个不同大小的立方体空间，其中  $N \in \{0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$ 。为了保持立方体对接的数量合理，该立方体空间的每个偏移方向都将被选择两次，并与每次出现的不同距离配对，从而确保每个偏移距离和方向在实验期间出现两次；这样，实验期间总共产生 16 个不同的目标位置。由于完全对齐白色和绿色立方体很困难，所以如果在用户确认放置时两个立方体之间距离小于阈值距离  $0.2m$ ，则认为对接成功。作为额外提示，当两个立方体在阈值距离内时，我们将目标立方体从半透明绿色变为红色。请注意，为了在一个试验中成功，参试者不仅需要将立方体移动到目标位置，还需要成功放置立方体。如果试验成功，参试者将收到一个声音反馈来确认。

在开始实验之前，参试者将进行 10 次对接以练习该交互方法。在正式实验开始后，每一组任务从白色立方体和目标位置出现在参与者面前开始，并在对接成功后结束；如果参试者在 30 秒内没有对齐立方体，则本次任务将被视为失败。

我们将从三个主观度量和三个客观度量来评估每一个参试者的实验结果。主观度量：眩晕感（SSQ）、可用性（Usability）、任务负担（NASA-TLX）。客观度量：

- 成功率：成功率是针对每个参与者计算的，是成功任务的次数与所有任务次数的比率。这



图 15 用户实验：单物体位移对接实验



评估了操作物体的总体效率，因为要成功完成一个任务需要全面考虑准确性和速度。

- 完成时间：在每个成功完成的任务中，完成时间将被记录；对于此项度量，我们不会将失败的任务纳入考虑。我们会记录两组实验数据：包括和不包括凝视停留的时间。
- 最终距离：此项度量仅针对失败的任务。最终距离是指在任务失败时刻白色立方体到目标位置的欧几里得距离（Euclidean Metric）；我们将其归一化为最终距离除以初始距离。这种归一化反映了参试者在将物体移动到目标位置时相对于其初始位置的靠近/远离程度。

#### 2.4.2 用户实验 2：单物体操纵对接实验

以单物体位移的对接实验为基础，加入旋转和缩放的操纵以实验所有对象操纵动作，并且以目前最优的基于手的 PRISM 方法[42]为 baseline，与之作最终的实验结果对比。我们的方法和基于手的 PRISM 方法同样可以完成单个对象的位移、旋转和缩放，所以有条件作为对比对象。用户将分别以我们的基于头眼协同的方法和 PRISM 方法完成一次与用户实验 1 中相同的任务。我们同样将从成功率、完成时间和最终距离三个客观度量来评估实验结果。

### 3. 后续研究计划

- ✧ 2023.4 – 2023.5：完成所有用户实验任务并且分析数据
- ✧ 2023.4 – 2023.6：论文撰写



#### 4. 参考文献

- [1] Mendes, Daniel & Caputo, Ariel & Giachetti, Andrea & Ferreira, Alfredo & Jorge, Joaquim. (2018). A Survey on 3D Virtual Object Manipulation: From the Desktop to Immersive Virtual Environments: Survey on 3D Virtual Object Manipulation. *Computer Graphics Forum*. 38. 10.1111/cgf.13390.
- [2] Radianti, Jaziar & Majchrzak, Tim A. & Fromm, Jennifer & Wohlgenannt, Isabell. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*. 147. 10.1016/j.compedu.2019.103778.
- [3] Poupyrev, Ivan & Billinghurst, Mark & Weghorst, Suzanne & Ichikawa, Tadao. (1998). The Go-Go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. *Proc. of UIST'96*.
- [4] Frees, Scott & Kessler, G.D.. (2005). Precise and rapid interaction through scaled manipulation in immersive virtual environments. *Proceedings - IEEE Virtual Reality*. 99-106. 10.1109/VR.2005.1492759.
- [5] Bowman, Doug & Hodges, Larry. (1999). An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. *Symposium on Interactive 3D Graphics*. 182. 10.1145/253284.253301.
- [6] Wilkes, Curtis & Bowman, Doug. (2008). Advantages of velocity-based scaling for distant 3D manipulation. 23-29. 10.1145/1450579.1450585.
- [7] Auteri, Chris & Guerra, Mark & Frees, Scott. (2015). Increasing Precision for Extended Reach 3D Manipulation. *International Journal of Virtual Reality*. 12. 10.20870/IJVR.2013.12.1.2859.
- [8] Osawa, Noritaka. (2008). Two-Handed and One-Handed Techniques for Precise and Efficient Manipulation in Immersive Virtual Environments. *International Symposium on Visual Computing*.
- [9] Stephanie Houde. (1992). Iterative design of an interface for easy 3-D direct manipulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. Association for Computing Machinery, New York, NY, USA, 135–142. <https://doi.org/10.1145/142750.142772>
- [10] Conner, D. & Snibbe, Scott & Herndon, Kenneth & Robbins, Daniel & Zeleznik, Robert & van Dam, Andries. (1992). Three-Dimensional Widgets. *Computer Graphics (1992 Symposium on Interactive 3D Graphics)*. 25. 10.1145/147156.147199.
- [11] Mendes, Daniel & Relvas, Filipe & Ferreira, Alfredo & Jorge, Joaquim. (2016). The Benefits of DOF Separation in Mid-air 3D Object Manipulation. 10.1145/2993369.2993396.





- 
- [12] Martinet, Anthony & Casiez, G ry & Grisoni, Laurent. (2010). The Design and Evaluation of 3D Positioning Techniques for Multi-touch Displays. 3DUI 2010 - IEEE Symposium on 3D User Interfaces 2010, Proceedings. 115-118. 10.1109/3DUI.2010.5444709.
- [13] Martinet, Anthony & Casiez, G ry & Grisoni, Laurent. (2010). The Effect of DOF Separation in 3D Manipulation Tasks with Multi-touch Displays. Proceedings of VRST'10, the 17th ACM Symposium on Virtual Reality Software and Technology. 10.1145/1889863.1889888.
- [14] Hancock, Mark & Cate, Thomas & Carpendale, Sheelagh. (2009). Sticky tools: Full 6DOF force-based interaction for multi-touch tables. 133-140. 10.1145/1731903.1731930.
- [15] Reisman, Jason & Davidson, Philip & Han, Jefferson. (2009). A screen-space formulation for 2D and 3D direct manipulation. UIST 2009 - Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology. 69-78. 10.1145/1622176.1622190.
- [16] Mendes, Daniel & Lopes, Pedro & Ferreira, Alfredo. (2011). Hands-on interactive tabletop LEGO application. 19. 10.1145/2071423.2071447.
- [17] Au, Oscar & Tai, Chiew-Lan & Fu, Hongbo. (2012). Multitouch Gestures for Constrained Transformation of 3D Objects. Computer Graphics Forum. 31. 651-660. 10.1111/j.1467-8659.2012.03044.x.
- [18] P. Monteiro, G. Gon alves, H. Coelho, M. Melo and M. Bessa. (2021). "Hands-free interaction in immersive virtual reality: A systematic review," in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 5, pp. 2702-2713, doi: 10.1109/TVCG.2021.3067687.
- [19] Carruth, Daniel & Hudson, Christopher & Bethel, Cindy & Pleva, Matus & Ond s, Stanislav & Juh r, Jozef. (2019). Using HMD for Immersive Training of Voice-Based Operation of Small Unmanned Ground Vehicles. 10.1007/978-3-030-21565-1\_3.
- [20] MIRZAEI, Mohammad & Chardonnet, Jean-Remy & Merienne, Fr d ric & Genty, Ariane. (2014). Navigation and interaction in a real-scale digital mock-up using natural language and user gesture. ACM International Conference Proceeding Series. 2014. 10.1145/2617841.2620716.
- [21] Adhanom, Isayas & MacNeilage, Paul & folmer, eelke. (2023). Eye Tracking in Virtual Reality: a Broad Review of Applications and Challenges. Virtual Reality. 1-24. 10.1007/s10055-022-00738-z.
- [22] Sidenmark, Ludwig & Gellersen, Hans. (2019). Eye&Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection. 1161-1174. 10.1145/3332165.3347921.
- [23] Tanriverdi, Vildan & Jacob, Robert. (2001). Interacting With Eye Movements In Virtual Environments.



Conference on Human Factors in Computing Systems - Proceedings. 10.1145/332040.332443.

[24] Hansen, John & Rajanna, Vijay & MacKenzie, I. & Baekgaard, Per. (2018). A Fitts' law study of click and dwell interaction by gaze, head and mouse with a head-mounted display. 1-5. 10.1145/3206343.3206344.

[25] Luro, Francisco & Sundstedt, Veronica. (2019). A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. 1-9. 10.1145/3314111.3318153.

[26] Robert J. K. Jacob. (1990). What you look at is what you get: eye movement-based interaction techniques. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90). Association for Computing Machinery, New York, NY, USA, 11–18. <https://doi.org/10.1145/97243.97246>

[27] Jacob, Rob & Stellmach, Sophie. (2016). What you look at is what you get: gaze-based user interfaces. interactions. 23. 62-65. 10.1145/2978577.

[28] Kumar, Devender & Sharma, Amit. (2016). Electrooculogram-based virtual reality game control using blink detection and gaze calibration. 2358-2362. 10.1109/ICACCI.2016.7732407.

[29] Pfeuffer, Ken & Mayer, Benedikt & Mardanbegi, Diako & Gellersen, Hans. (2017). Gaze + pinch interaction in virtual reality. 99-108. 10.1145/3131277.3132180.

[30] Pai, Yun Suen & Dingler, Tilman & Kunze, Kai. (2019). Assessing hands-free interactions for VR using eye gaze and electromyography. Virtual Reality. 23. 10.1007/s10055-018-0371-2.

[31] Qian, Yuan Yuan & Teather, Robert. (2017). The eyes don't have it: an empirical comparison of head-based and eye-based selection in virtual reality. 91-98. 10.1145/3131277.3132182.

[32] Sidenmark, Ludwig & Clarke, Christopher & Zhang, Xuesong & Phu, Jenny & Gellersen, Hans. (2020). Outline Pursuits: Gaze-assisted Selection of Occluded Objects in Virtual Reality. 10.1145/3313831.3376438.

[33] Majaranta, Päivi & Bulling, Andreas. (2014). Eye tracking eye-based human-computer interaction. Advances in Physiological Computing. 39-65. 10.1007/978-1-4471-6392-3\_3.

[34] Boyer, Eric & Lorenceau, Jean & Portron, Arthur & Bevilacqua, Frédéric. (2017). Continuous Auditory Feedback of Eye Movements: An Exploratory Study toward Improving Oculomotor Control. Frontiers in Neuroscience. 11. 10.3389/fnins.2017.00197.

[35] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. (2018). Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. In Proceedings of the Workshop on Communication by Gaze Interaction (COGAIN '18). Association for Computing Machinery, New York, NY, USA, Article 1, 1–9. <https://doi.org/10.1145/3206343.3206349>



- 
- [36] Mohan, Pallavi & Goh, Boon & Fu, Chi-Wing & Yeung, Sai-Kit. (2018). DualGaze: Addressing the Midas Touch Problem in Gaze Mediated VR Interaction. 79-84. 10.1109/ISMAR-Adjunct.2018.00039.
- [37] Hirzle, Teresa & Cordts, Maurice & Rukzio, Enrico & Bulling, Andreas. (2020). A Survey of Digital Eye Strain in Gaze-Based Interactive Systems. 1-12. 10.1145/3379155.3391313.
- [38] Purves D, Augustine GJ, Fitzpatrick D, et al., editors. (2001). Types of Eye Movements and Their Functions. Neuroscience. 2nd edition. Sunderland (MA): Sinauer Associates
- [39] P. C. Gloumeau, W. Stuerzlinger and J. Han. (2021). "PinNPivot: Object Manipulation Using Pins in Immersive Virtual Environments," in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 4, pp. 2488-2494, doi: 10.1109/TVCG.2020.2987834.
- [40] Yu, Difeng & Lu, Xueshi & Shi, Rongkai & Liang, Hai-Ning & Dingler, Tilman & Velloso, Eduardo & Goncalves, Jorge. (2021). Gaze-Supported 3D Object Manipulation in Virtual Reality. 1-13. 10.1145/3411764.3445343.
- [41] Liu, Chang & Plopski, Alexander & Orlosky, Jason. (2020). OrthoGaze: Gaze-based Three-dimensional Object Manipulation using Orthogonal Planes. Computers & Graphics. 89. 10.1016/j.cag.2020.04.005.
- [42] K. Kim, R. L. Lawrence, N. Kyllonen, P. M. Ludewig, A. M. Ellingson and D. F. Keefe, "Anatomical 2D/3D shape-matching in virtual reality: A user interface for quantifying joint kinematics with radiographic imaging," 2017 IEEE Symposium on 3D User Interfaces (3DUI), Los Angeles, CA, USA, 2017, pp. 243-244, doi: 10.1109/3DUI.2017.7893362.