老师同学们好！很高兴能有此次机会与大家分享此次作业我收获的内容。

这是我分享的两大块内容。第一是我的读书报告。我选的文章是Going, Going, Gone: Exploring Intention Communication for Multi-User Locomotion in Virtual Reality，这部分分为了引言、相关工作、研究方法、实验结果和我的感悟这五个部分调研报告的主题是从视听探究AI技术在HCI领域中的运用，该部分分为引言、相关工作和背景概述、具体介绍和专家观点和总结与个人体会四个部分。

首先来看一下我的读书报告。随着科技的迅速发展，虚拟现实技术(VR)已经成为了现实生活中不可或缺的一部分。其中多人共同探索虚拟世界、多人共同完成一项活动这一主题更是增加了社交成分。探索虚拟世界中多用户运动的意图交流是本文章的中心点。现有的VR技术包括以下三个方面：单用户定位工作，指的是研究和开发用于在虚拟现实环境中让单个用户进行移动和导航的技术和方法。但是它有社交交互限制；VR中的群体定位，指的是在多个用户之间协调和管理他们在虚拟环境中的位置和移动，但面临着协调复杂性、资源消耗和可能导致混乱等挑战；运动的社交方面，指的是虚拟现实中的运动具有丰富的社交层面，涉及到与他人的社交互动、群体协作、社交氛围等方面。

在本文的实验中设计了一个虚拟现实迷宫任务，供两名参与者完成。任务务分为个体和集体两个阶段。在个人阶段，参与者需要自由移动，不能有相应的交流，穿过迷宫找到目标点之一。在集体阶段，他们需要协作返回起点。参与者可以口头交流。同时，该实验也相应地设置了自变量和因变量。自变量为不同级别的可视化水平，这包括了目标、方向和群体；因变量有IOS、(用来评估对另一名玩家的心理亲近感)GEQ(用来衡量玩家对对方的关注程度)。适应性在于在集体阶段的实验过程中以分别以领导者和跟随者的角度探究了三种意图沟通方法和两种基准方法的效果，这样使得实验结果更具有全面性。

文章得出了以下结论：方向技术的优势：对于跟随玩家来说，方向技术比目标技术更有助于感知其他玩家的位置并更容易保持在一起。

完整可视化的优势：完整可视化技术能够提供关于其他玩家意图的全部信息，包括目标位置和传送轨迹。

角色对技术偏好的影响：玩家的角色在他们对不同技术的偏好上有很大影响。

接下来分享我的感悟：首先我对虚拟现实中多人虚拟现实交互的理解有了进一步的认识，多人虚拟现实交互是指多个用户能够在虚拟现实环境中实时互动、共享体验的技术和应用。其次是一些问题和分析，在我看来，多用户之间交互的体验是很重要的，所以要协调单用户移动和团体运动对于用户自由度、空间感知以及沟通和协作的影响；其次，用户之间的协调和沟通是非常重要的，所以在未来利用更加高效的方向技术可以更容易地指引玩家找到位置或者是导航。

接着来看看我的调研报告。人机交互技术是人类与计算机系统之间进行信息交换与互动的过程中的技术手段，近几年来，随着人工智能异军突起，市面上出现了许多令人惊奇的人工智能模型，这些模型有着强大的功能。这些功能也同样被人机交互技术所吸收，出现了许多新奇的产品。

在传统的观念中，人机交互需要的是一个实体介质，需要人们对计算机的IO部分进行触碰。计算机视觉和计算机听觉是AI技术在HCI领域中的两个重要方面，它们分别涉及计算机对视觉和听觉信息的处理和理解，以实现更智能、更自然的人机交互体验。计算机视觉涉及计算机对图像、视频等视觉信息的理解和处理比如虚拟拟现实(VR)和增强现实(AR)；计算机视觉涉及计算机对声音、语音等听觉信息的处理和理解。语音识别和自然语言理解(如Siri、Alexa、Google Assistant等)

如果我们将语音识别与图像识别，手势识别结合起来，那有会怎么样？在最近，苹果公司推出了首个空间计算设备——Apple Vision Pro

这是一台集合计算机视觉、计算机听觉为一体的人机交互设备机器，为每个用户带来身临其境的即时体验。

ARKit是苹果公司推出的增强现实(AR)开发框架，首先ARKit使用设备的摄像头、运动传感器和深度传感器来获取环境的实时数据。这些传感器提供了设备位置、方向、移动速度以及周围环境的深度信息等数据；再次通过分析传感器数据，识别环境中的平面、物体和人体等场景元素，并构建出环境的3D空间模型；接着根据环境的理解和用户的交互，将虚拟对象叠加在真实世界中的适当位置。

如下图，ARkit强大的功能为Apple Vision Pro的视觉方面打下了坚实的基础。

在语音识别任务中，循环神经网络(RNN)和长短期记忆网络(LSTM)是两种常见的深度学习模型，用于处理序列数据。LSTM是RNN的改进，通过引入记忆单元和门控机制，LSTM 可以有效地处理长序列数据。如下图，我们利用所训练的两个模型来对一段相同的语言进行识别，结果我们可以看见，LSTM所带来的识别效果随着时间的推移准确率会比RNN所带来的识别效果好。同样的，虽然苹果公司SIRI背后所用的训练神经网络没有被完全地公开，但是毋庸置疑的是强大的神经网络训练系统使得语音识别在用户交互的准确性上有极大的提升。

当然，Apple Vision Pro在快速火热的同时，也引来了不少专家对其进行测评。数码测评师钟文泽表示Apple Vision Pro在硬件层面可以在4K的显示效果下完成图片渲染，在软件层面搭载了众多APP，可以给用户以完整的交互体验。

本篇调研的主要方向还是AI技术中视觉、听觉方面在HCI中的应用，上文的举例分析中不难看出，计算机视觉、听觉可能还存在着一些问题。比如计算机视觉系统往往只能对图像进行浅层的特征提取和物体检测(引入复杂神经网络、数据集)人机交互技术也需要关注用户数据的隐私保护和安全性。

最后，不管是Apple Vision Pro还是别的用户交互技术产品，我们都可以发现，人工智能是一项极为高端的技术，对它的直接操控和调整优化是由技术员执行的，只有通过人机交互技术的人工智能才能真正地被大众所使用。正如Richard H. R. Harper在他的文章《The Role of HCI in the Age of AI》指出，未来不仅仅是人工智能，而是必须由人机交互技术所搭载的人工智能。

谢谢大家，我的pre到此结束。