

基于全景漫游技术的可视化交互设计理论与实践研究

马申彦

2017 年 2 月 26 日

目录

第一章 全景漫游技术与设计	3
1.1 全景漫游的目的	3
1.2 全景漫游的现状	3
1.3 全景漫游与设计结合的必要性	3
1.4 全景漫游中可视化交互理论研究的意义	4
第二章 全景漫游行业的市场现状	5
2.1 全景漫游技术总览	5
2.1.1 场景呈现硬件技术	5
2.1.2 场景捕获硬件技术	5
2.1.3 场景处理技术	5
2.1.4 场景还原与增强技术	6
2.2 全景漫游市场前景	6
2.3 全景漫游产业类别	6
2.4 全景漫游现有 APP 分析	7
2.4.1 Ascape	7
2.4.2 Fulldive	7
2.4.3 VR X-Racer	8
2.4.4 暴风魔镜 VR	10
2.5 全景漫游行业现状总结	10
第三章 全景漫游与人的关系	12
3.1 全景漫游的生理特性	12
3.1.1 全景漫游与视觉特征	12
3.1.2 全景漫游与听觉	14
3.1.3 全景漫游与其他感觉	14
3.2 全景漫游与肢体动作	16
3.2.1 关节运动	16
3.2.2 手部活动	17
3.3 全景漫游与人的心理特征	19
3.3.1 知觉	19
3.3.2 意识	21

3.3.3 认知	22
第四章 全景漫游的交互形式	24
4.1 全景漫游的信息架构	25
4.1.1 信息架构形式	25
4.1.2 信息架构组件	27
4.1.3 信息架构可视化	27
4.2 全景漫游的功能架构	27
4.2.1 漫游功能	27
4.2.2 导航功能	28
4.2.3 搜索功能	28
4.2.4 记忆功能	28
4.3 全景漫游的功能模块	29
4.3.1 导航功能模块	29
4.3.2 场景特殊模块	29
4.3.3 多输入交互模块	29
4.3.4 热点模块	29
第五章 全景漫游的交互实现	30
第六章 待写部分	31
第七章 参考文献	32

第一章 全景漫游技术与设计

1.1 全景漫游的目的

古语云：眼观六路，耳听八方

这句话本意是形容人机制灵活，遇事能多方观察分析。而如今科技进步迅猛、社会发展气象万千，古时人们所向往的“千里眼”和“顺风耳”早已成为了现实（卫星雷达和手机电话），信息获取效率相比于古代有了巨大的提升。人们真正地做到了“眼观六路，耳听八方”。但这些技术往往只是片面地延展了人的感官能力，只是将人与物间某些距离拉近了，少有将人作为感官系统的中心来进行考量。在现代信息通讯过程中，人被排除在计算机的屏幕之外，人和真正的信息之间有一层薄薄的但难以直接触碰的玻璃阻隔。

古希腊的哲学家普罗泰戈拉说过：“人是万物的尺度，是存在的事物存在的尺度，也是不存在的事物不存在的尺度。”。借助现有科学技术，可以轻易地观摩到深入海底数公里的珊瑚礁分布、险象环生的热带雨林内生物的栖息环境、甚至是浩瀚无垠的太空中星象奇观。面对如此美景，只是观赏单一的图像帧序列来企图感受自然的奥妙总是觉得少了一份身临其境的体验。

人并不是一定要盯着一块显示屏。显然可得的想法就是利用显示屏将人的视线包裹起来，这样观察起来具有良好的主动性和体验感，同时规避了将人置于一些特殊环境所带来的附加成本或是风险因素。虽是虚拟，胜似现实，这就是“虚拟现实”的意义。而“全景漫游”则是“虚拟现实”中最重要的一部分。

1.2 全景漫游的现状

全景漫游是当今互联网行业发展最为迅速的技术之一，它是虚拟现实（即 VR 技术）的基础。沉浸性和交互性是其得到众多用户青睐的最主要原因，可供用户身临其境般体验的图形化界面，以及不同于传统平面媒介的交互手段共同促进了全景漫游技术的推广。全景漫游解决了传统工业界对于空间概念呈现形式不足的痛点，通过定义场景的 360 度无缝全景图像，仿佛使人置身于制作者脑海中设计出的环境内，配合移动设备内置陀螺仪支持的重力感应功能，足不出户即可感受任何可以用图像表达的场景。80 天环游地球不再只是一个梦想，而且现今全景播放设备层出不穷，最便宜的 Google Cardboard 制作成本不足一元钱，人人 VR 的时代即将到来。

1.3 全景漫游与设计结合的必要性

设计行业向来紧跟时代发展潮流，全景漫游技术已经发展地比较成熟，因其本身具有互联网基因，故将其移植至设计产业的成本相较传统的工业技术会显著降低。传统产品原型技术将产品制作模型交付演示的时代可能即将被虚拟化技术所取代，相伴随的是设计成本的减轻以及设计迭代的加速。将全景漫游技术引入设计

行业的难点在于将全景展示的程式化模型转化为用户可探索、可理解、可学习的交互模型，现阶段的问题在于用户进入场景之后因场景与以往二维界面的割裂而导致心理产生的不适应性使其手足无措。值得研究的重点环节是降低用户的初次体验成本，加深用户记忆中对操作流程感悟的交互模型，最终达到用户无缝体验设计师所要传达的设计意图的目的。

1.4 全景漫游中可视化交互理论研究的意义

本课题旨在探索全景漫游场景中应用程序接口与用户心智模型的转化过程，通过分析场景程序执行的逻辑，用户操作的普适性规律以及人类学习行为的通常过程，将三者有机结合，以场景插件程序的形式进行封装，并提取出相应的 API 接口，以供本课题其后实例开发及其他对本课题感兴趣的开发者使用。可视化技术对于工业发展的提升是显然的，从传统的制图学到近代的计算机辅助制图，再到现今的三维模拟技术，工业界向来拥抱可视化技术的前沿。可视化交互的意义就在于将图形化表述的产品按照其功能有机地组合起来，以可体验的形式呈现给用户/开发者。不再需要对产品的每部分组成有基础的了解也可认知其完成状态，也就意味着机械产品设计的成本进一步降低，普及到每个想要创造的人身上进而使其创造出更多产品的可能性变得现实。1.4

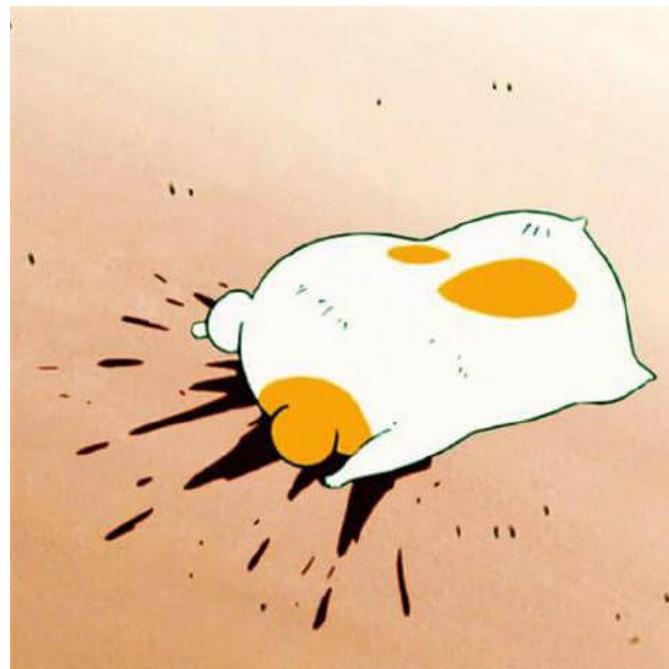


图 1.1: test

第二章 全景漫游行业的市场现状

2.1 全景漫游技术总览

全景漫游技术按载体分类可分为：

1. 场景捕获硬件技术：视频捕获设备及其保障设备，用以采集全景漫游的视频、图片、音频等素材。
2. 场景呈现硬件技术：主要包括 VR 眼镜、VR 互动手柄、VR 主机等。
3. 场景处理技术：场景规划设计、场景素材的保存与传输以及相应程序开发等。
4. 场景还原与增强技术：将场景素材与功能模块有机地整合，提供给用户身临其境般的全景漫游体验。

2.1.1 场景呈现硬件技术

相对而言，硬件技术相比于后两者进入门槛更高。目前市场上三大 VR 硬件厂商 OculusRift、HTC Vive 和 PlayStationVR 几乎垄断了高端 VR 播放设备市场。之后加入 VR 市场的企业（如暴风、大朋、3Glasses、蚁视等）几乎都放低了姿态，推出了价格低廉但基本满足播放功能的入门级 VR 眼镜作为卖点。但随着 Google Cardboard 的推出，这款几乎没有成本可言的开源“硬件”迅速占领了低端 VR 硬件行业相当客观的市场份额。

2.1.2 场景捕获硬件技术

场景捕获技术的难点在于实时捕获，合成全景照片是一般手机都拥有的功能，大致原理就是捕获数张连续且相近的照片并通过算法进行合成球形场景。但实时场景捕获硬件的成本因其同步的特性更为高昂，目前市面上已有的厂商例如 Google JUMP、NOKIA OZO 等均是将现有平面镜头进行堆叠排布并通过后期处理合成虚拟场景，且这种方式因需要较多的镜头来堆砌同时刻的画面所以硬件成本非常高。目前市面上尚只有堆叠摄像头这一种方式进行全景场景捕获，而光场摄影等技术仍处在技术攻关阶段，相信不远的将来有可能出现消费级的场景捕获硬件。

2.1.3 场景处理技术

场景处理是目前全景漫游软件领域需要重点攻克的最大难关，但市面上各企业已基本做到自给自足的技术支持。全景漫游所依赖的基础是图像信号的捕获、加工与储存，而这些方面已有较多成功经验。目前 Web 端视频播放协议有以下两种：Real Time Messaging Protocol（实时消息传输协议，简称 RTMP）和 HTTP Live Streaming（HTTP 渐进下载，简称 HLS）。这两者是孑然不同的两种协议，而且 iPhone 等手机由于不自带支持 Flash 播放，一般考虑全端支持的流媒体播放会选用 HLS 作为传输协议，即对视频做切片，边播放边加载下一时段的切片。

2.1.4 场景还原与增强技术

场景还原与增强是与用户直接相连的部分，可以说前面的技术都是起为其保驾护航的功能。其中场景还原部分为计算机图形学的范畴，例如图像在空间上的曲率计算等，主要涉及到的技术有 OpenGL／WebGL 等，用于将二维图像还原成球状的场景。而场景增强技术则是本文所讨论的重点，其又分前处理和后处理两种，但因场景搭建的过程比较复杂，所以两种处理模式界限并不是特别清晰。

场景前处理代表：Unity3D Unity3D 是由 Unity Technologies 开发的一个让玩家轻松创建诸如三维视频游戏、建筑可视化、实时三维动画等类型互动内容的多平台的综合型游戏开发工具，是一个全面整合的专业游戏引擎。¹

场景后处理代表：Krpano Krpano 是一个小巧灵活的用来呈现各种全景图像和交互式的虚拟之旅的高性能全景查看器。可作为 Flash 和 HTML5 应用程序在 Web 上使用。并附有利用拖拽全景生成场景的 Krpano Tools 以供开发者快速生成用以展示的全景场景。本文将应用其进行部分设计案例的制作与演示。

场景后处理代表：Aframe Aframe 是一个用 Web 技术构建 VR 体验的框架。使用 HTML 语言及实体组件来构建场景，可应用于 web／mobile 和其他多种设备端。本文将应用其进行部分设计案例的制作与演示。

2.2 全景漫游市场前景

2016 年，随着三大 VR 设备开始出售，未来 2-3 年 VR 设备普及率将快速提升。到达 2020 年，虚拟现实生态圈将初步形成，内容、服务等盈利模式逐步成熟，全球 VR 市场规模将达到 404 亿美元，VR 游戏市场规模将达到 149.5 亿美元。²

虚拟现实（VR/AR）产业市场具有良好前景，2015 年中国虚拟现实行业市场规模为 15.4 亿元人民币。融资方面，国内 VR/AR 领域投资活跃度从 2015 年开始显著提升，2015 年第四季度和 2016 年第一季度的融资额均接近 10 亿元。其中显示设备融资占据首要地位，2015 年融资案例数量占比 30%，融资额占比 69%；内容制作的融资案例数量虽然占比 22%，但融资额仅占比 6%³。可见 VR 及全景漫游方向的内容制作生产领域仍处于起步阶段，与硬件厂商等差距较大，但同时其中也孕育了巨大的商机。

VR 内容开发受市场认可，线下体验馆增长迅速。由于中国 VR 市场主流设备仍以移动端 VR 眼镜为主，VR 视频内容的开发数量要远多于 VR 游戏内容。VR 平台上已有约 2700 款视频和 800 款游戏。预计 2020 年中国 VR 设备出货量 820 万台，用户量超过 2500 万人，见图2.1。

2.3 全景漫游产业类别

全景漫游产业按出发点可大体分为两类：

- 以视频、游戏为主的虚拟内容提供商／平台商
- 涉及电商、教育、医疗、建筑等传统行业的行业融合应用服务商

短期而言，市场上比较活跃的是虚拟内容提供商，但长远而言，ToB 模式更利于产生更为健壮的全景漫游服务体系，对于高质内容的生产、分发和变现的模式也偏向于有传统大规模企业的行业服务商。

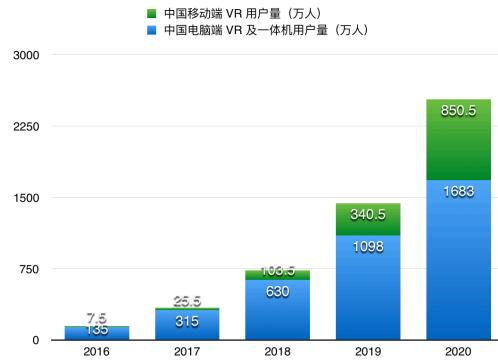


图 2.1: 2016-2020 年中国 VR 用户规模

2.4 全景漫游现有 APP 分析

2.4.1 Ascape

Ascape 这款应用通过 360° 无缝全景视频让你通过景色的立体画面切身体验视频中呈现的著名景点，探索只有当地人才知道的隐藏宝藏的经历。并且镜头会伴随着你的移动，画面的位置和显示的角度都会随之改变，见图2.2。



图 2.2: 简洁的全景视频播放

同时，这款应用在 APP 传统界面上也继承了欧美一向以来的简约风格：“探索”页面采用卡片设计模式直观展示了新奇的场景；“发现”页面通过地图和名单两种模式展示了全球各地的场景选项；“我的旅行”也通过卡片模式展示了已下载的全景场景，见图2.3。

2.4.2 Fulldive

Fulldive 是一个智能手机与虚拟场景连接的平台，进入 APP 后即进入了一个全景漫游的世界。场景内包含常用的网络视频、本地视频／图片、VR 相机、VR 浏览器等应用，同时支持从其内部市场下载更多基于其



图 2.3: 简洁的全景视频播放

开发的 VR 应用，见图2.4。

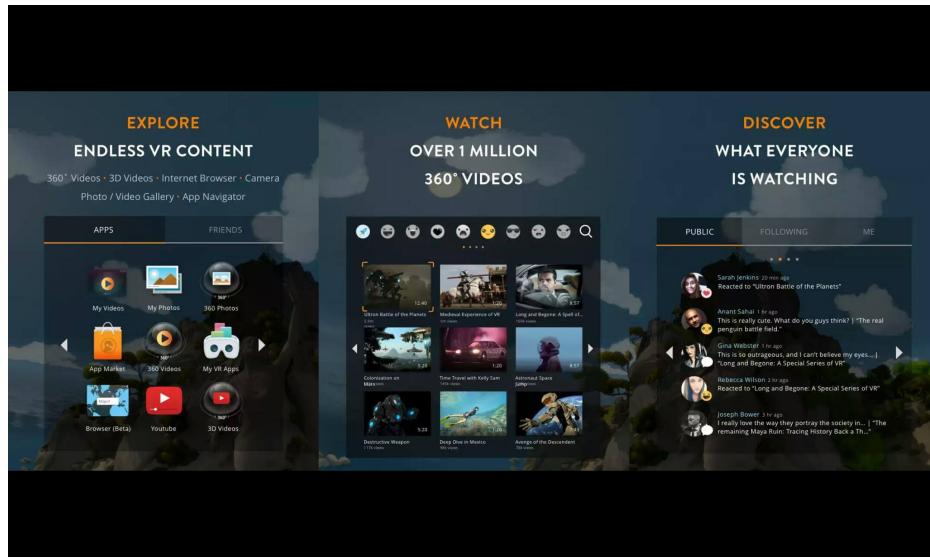


图 2.4: Fulldive 内置的应用列表

与 Ascape 不同之处在于 Fulldive 的操控完全是在虚拟全景中，所以在播放视频及操作选项时 Fulldive 有一些针对配戴 VR 眼镜者所特别提供的操作方式，见图2.5。例如，将视线聚焦在某个按钮上超过 3 秒后即视为按下了该按钮。这是一种与桌面鼠标和触屏点按都完全不同的交互形式，借鉴了鼠标的双击或是触屏的长按这两种操作，见图2.6。

当然，这种模拟点击的操作只适用于“确认／取消”这种布尔判断的操作，用户需要输入大段的文字时则需有更高效的方式。Fulldive 结合了已日益成熟的语音识别技术，图2.7为实际操作 Fulldive 搜索功能并口述“中国地质大学“后的 APP 截屏。

2.4.3 VR X-Racer

VR X-Racer 是一款简单的 VR 操控类飞行游戏，其操作方式为晃动设备（如 VR 眼镜或手机）来控制屏幕上的飞机躲避障碍物，见图2.8。这是全景漫游技术在游戏中最直接的体现：几乎没有多余的操作，在游

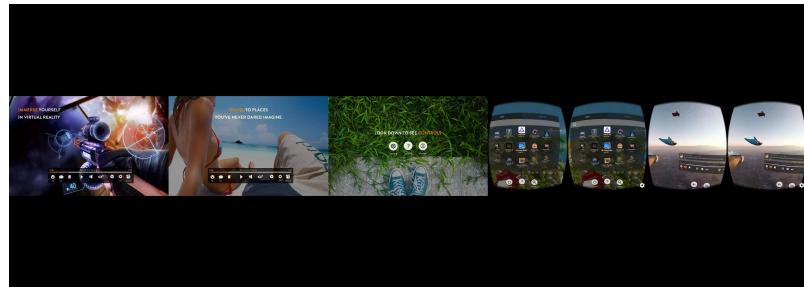


图 2.5: Fulldive 特有操作形式



图 2.6: Fulldive 视线停留模拟点击



图 2.7: Fulldive 语音识别

戏中飞机撞到障碍物坠毁后若干秒即自动重新开始游戏。但其最大的缺陷是无法长时间使用，不断晃动的全景屏幕容易使人产生眩晕、恶心等不良反应。全景漫游与 3D 影片的体验是完全不同的，全景漫游将用户完全包裹在视频所构建的封闭环境中，而 3D 眼镜只是对屏幕这种有限区域进行了折射。两者的区别在于 3D 影片由于有周围环境作为铺垫不易使人完全沉浸其中，全景漫游则是在很短的时间内（20-30 秒）就使人沉浸在虚拟世界中，直至摘掉全景设备重新适应周围环境。



图 2.8: X-racer 游戏截图

2.4.4 暴风魔镜 VR

暴风影音是国内在全景漫游生态领域探索前沿的产品之一，其高端产品售价可达数千元，甚至与国外高端 VR 设备如 Oculus 和 GearVR 等价格相齐。与其硬件设备配套的则是一款叫做“暴风魔镜 VR”的手机 APP。

国内 APP 的发展方向一直是“大而全”，这款 APP 也不例外。暴风魔镜 VR 力图包括网络／本地视频、影音播放与 VR APP 等多种应用，形成自己的 VR 平台体系。同时支持普通的 APP 界面操作模式和全景漫游的 VR 模式。

在交互形式上与上文所列举的 Fulldive 类似，均为视线聚焦停留数秒视为确认，但缺少了语音识别输入大段文字的功能（在页面模式下支持手机输入法输入），如图2.9。总体使用可满足基本的全景漫游体验，但识别速度较慢或误识别是比较严重的问题。

值得注意的一点是，暴风魔镜 VR APP 内场景下方有一个“归位”图标，触发后将会调整屏幕主视域至正对当前屏幕的位置，可类比于传统 APP 的“回到顶部”功能。这个功能很方便地起到了定位自身的作用，让用户不必盲目转动来寻找起始时正对着的界面。

2.5 全景漫游行业现状总结

全景漫游正处在行业发展的成长期，相关技术发展迅速，但同质化较为严重。众多厂商基于经济利益考量注重硬件设备的研发和更新换代，但软件质量与硬件质量仍有一定的差距，主要体现软件交互设计的质量上。

在全景漫游的交互设计上，国内外 APP 均有较多针对与不同与传统人机界面交互的创新点，例如前文列举的“视线停留数秒确定”、“语音识别”和“底部 Docker 栏”等。但交互可操作性相比于传统人机界面还有一定的差距，例如无法很快定位到所需内容、运动过程中不易操控等问题。同时，在全景漫游的现有设计中，仍旧无法脱离面向传统人机界面的设计语境，在考虑保留用户习惯的同时难以创新出更为独特的交互模式，可以说是全景漫游交互设计亟待研究的难点。

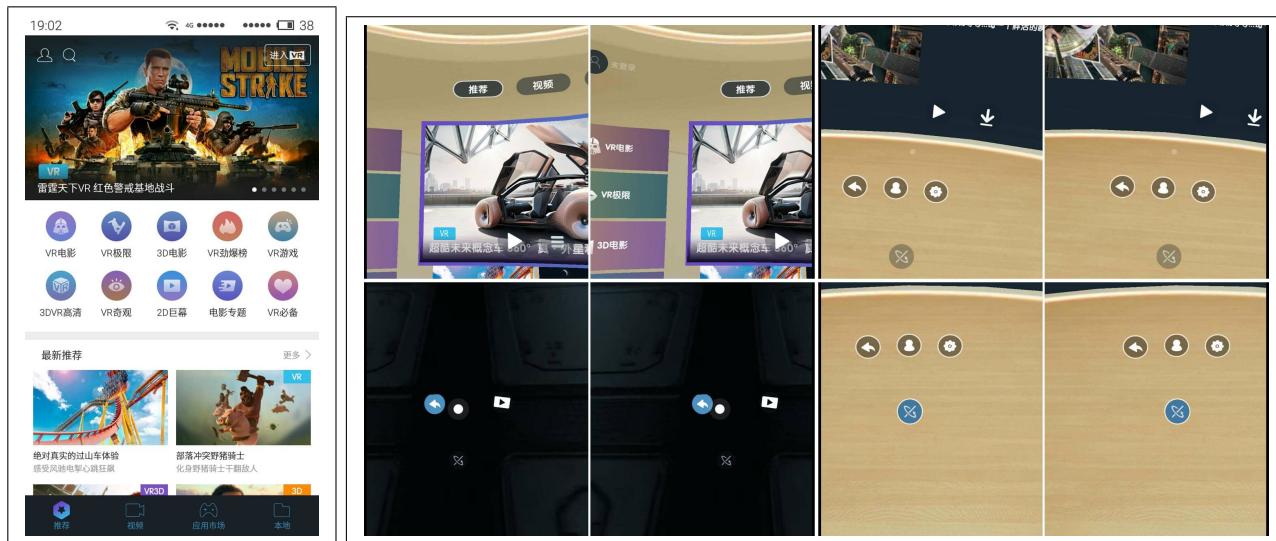


图 2.9: 暴风魔镜 VR 操作方式

全景漫游的行业发展离不开全景漫游的软硬件及生态环境的良好结合，只有如此才能留存有效用户并产生价值，故行业目前的重点就是在软硬件上均体现出应有的价值提升。其中，软件方面可在全景漫游领域的交互方向结合相关理论作出新的设计以改善用户体验。

第三章 全景漫游与人的关系

3.1 全景漫游的生理特性

实现全景漫游与传统人机界面最大的区别在于全景漫游时的设备（包括 VR 眼镜或专用眼罩）距离人眼只有 2 ~ 3 公分距离，设备整体对人眼呈包裹状态。使用者仅可通过听觉和其他一些不够灵敏的感觉来感受外界环境，使用环境的舒适性就变得非常重要。

全景漫游与人体关联的感觉和生理运动大致有：视觉、听觉、平衡感等。如图3.1。

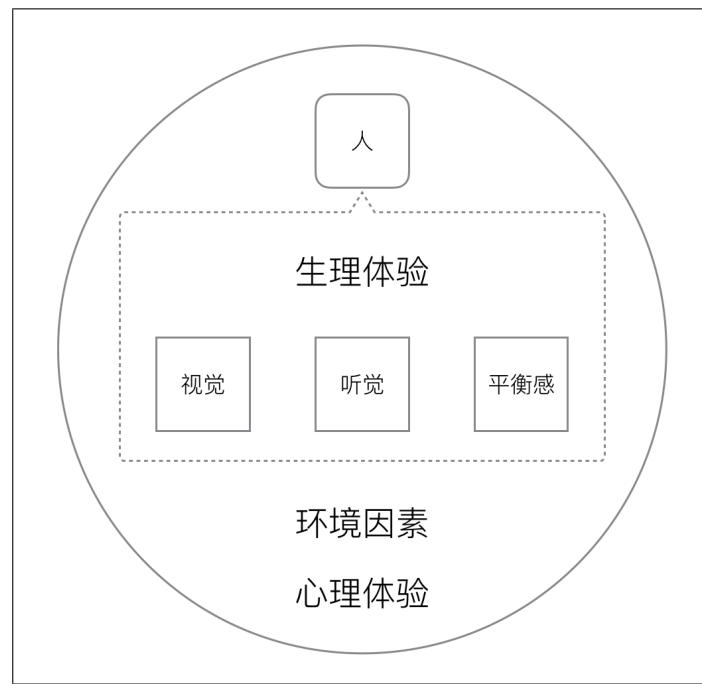


图 3.1: 全景漫游与人体知觉

3.1.1 全景漫游与视觉特征

全景漫游对人体最大的刺激来源即是视觉刺激，其特征为距离眼睛距离近，色彩、明暗、动作变化等刺激较一般显示屏幕会更为剧烈。

视角与视野

人的视角是确定被看物尺寸范围的两端点光线射入眼球的相交角度。视角的大小与观察距离及被看物体上两端点的直线距离有关，其计算公式3.1如下：

$$\alpha = 2 \arctan \frac{D}{2L} \quad (3.1)$$

其中， α 为视角，用 ('') 表示，即 $(1/60)^\circ$ 单位；D 为被看物体上两端点的直线距离；L 为眼睛到被看物体的距离。如果设眼球距镜片距离为 1.5cm，而镜片上一物体显示高度为 2cm，则带入公式可得3.2

$$\alpha = 2 \arctan \frac{2\text{cm}}{2 * 1.5\text{cm}} \approx 67.38^\circ = 2 \times 38.69^\circ \quad (3.2)$$

即上下视角各约为 38.69° ，查阅数据可知人在垂直平面的视野最大视区为视平线以上 50° 和视平线以下 70° ，颜色辨别界限为视平线以上 30° ，视平线以下 40° 。

由上述计算可知，在使用全景漫游设备观看场景时，假设高为 2cm 的显示物体在人视野中上部已超出颜色辨别界限，下部也已接近颜色辨别界限。物体上部比下部更接近垂直平面内的视野极限，如图3.2。由此可知，全景漫游场景设计中应将主要元素置于屏幕中央，次要元素尽可能置于屏幕下方以便识别操作。同时，色彩丰富且饱和度高的元素也应置于视野中部，而一些操作类的元素则不应用过于丰富饱和的色彩缀饰，以免让使用者应难以分辨色彩而操作出错。

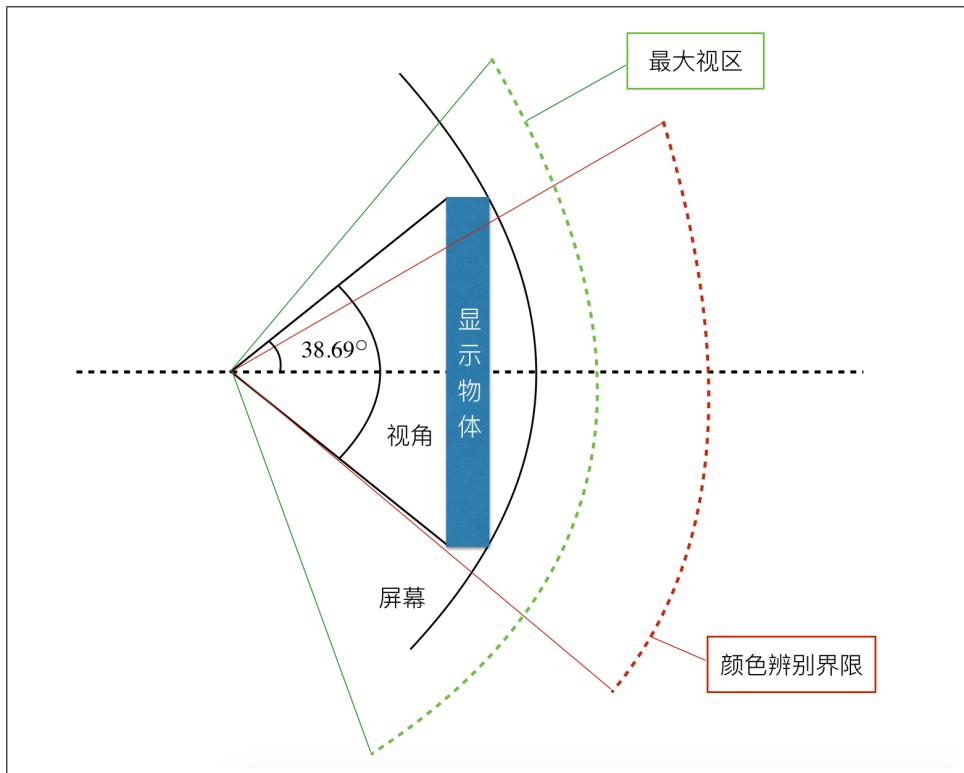


图 3.2: 全景漫游中视野的角度关系

视觉适应

配戴全景漫游设备进行观察时，屏幕距眼球距离不足 2cm，属于视距过近状态，应当避免长时间操作，否则易引起眼睛疲劳。同时由于屏幕与眼球距离过近，屏幕亮暗对眼球刺激比通常工作环境更为强烈。当屏幕上场景切换间明暗变化过快时，人眼需要一定的适应时间才能正常分辨，但由于全景漫游设备等封闭性，无法较轻易地脱离这种明暗急剧变化的环境，则眼睛无法得到充足有效的适应时间，则很容易产生视觉疲劳，影响视觉能力。

3.1.2 全景漫游与听觉

全景漫游的视觉是主要人体感觉，但听觉是仅次于视觉的重要感觉，尤其是在视觉被全景场景完全包裹住的情况下。

现有全景漫游设备有的自带耳机设备，也有的需要自行配备耳机。但长远来看，配备耳机可能是必然的选择。因为人耳具有“方向敏感度”（或称“双耳效应”），即：声波在进入鼓膜时，受人体的反射、折射和衍射而扭曲变形，大脑能根据经验判断出声源的方位和距离。而前文所说的“耳听八方”即说的是这种能够重建场景音效的能力。

双耳效应

为了“制造”出能够“蒙骗”大脑使其误以为使用者身处虚拟场景中的体验，大致方向是重建出空间任意一点至双耳处的滤波器函数。我们不必去关心空间中某一点的声波是如何传播到人耳中的，只需要将该点声波信号经过滤波器便可以得到它在人耳处的声波信号，如图3.3。而这个信号是可逆的，即被人耳捕捉到之后便由大脑还原成相应位置的“声信号”，从而达到“欺骗”大脑的目的。这个领域的研究对象叫做 HRTF（Head Related Transfer Function）：头相关变换函数，是一种音效定位算法。但每个人的听觉能力都是不同的，所以 HRTF 对人类整体并没有通解，而只有对于人类个体经过反复实验测试后的特解。

国内外已有厂商根据相关理论设计并制作出了 VR 耳机，可配合 VR 眼镜使用，也可独立使用，它会追踪使用者的姿态、位置变化等，进而计算并还原出模拟三维场景内的声音分布，让使用者仿佛置于真实的场景之中。

视觉的「补充」

虽然声音不是可视化领域的研究重点，但良好的听觉体验的确有利于信息渠道的建立。视觉与听觉互为补充、相互作用、最终是人的知觉更为敏锐而准确。可以设想以下场景：「用户配戴全景漫游头盔进行一场探险类的游戏，用户扮演的主人公正在探索未知的世界，正当主人公小心翼翼地穿过一片沼泽时，后方突然传来了一声尖叫……」如图3.4看不见的事物通过音效来表现，可以提升场景所渲染出的气氛效果。听觉与视觉不同，不受方向影响，能够用来表现一些视觉上无法做到的特殊效果。

3.1.3 全景漫游与其他感觉

众所周知的是，人之所以可以站立行走、骑行脚踏车或进行高空作业依赖的是人与生俱来并在后天不断磨合的平衡感。平衡感由内耳的淋巴液控制，同视觉相互配合，使人能安全地四处走动。而在全景漫游过程中，人的视觉几乎全部集中在眼前的屏幕上，即很难看到外界真实的场景，故视觉此时便难以匹配内耳中内淋

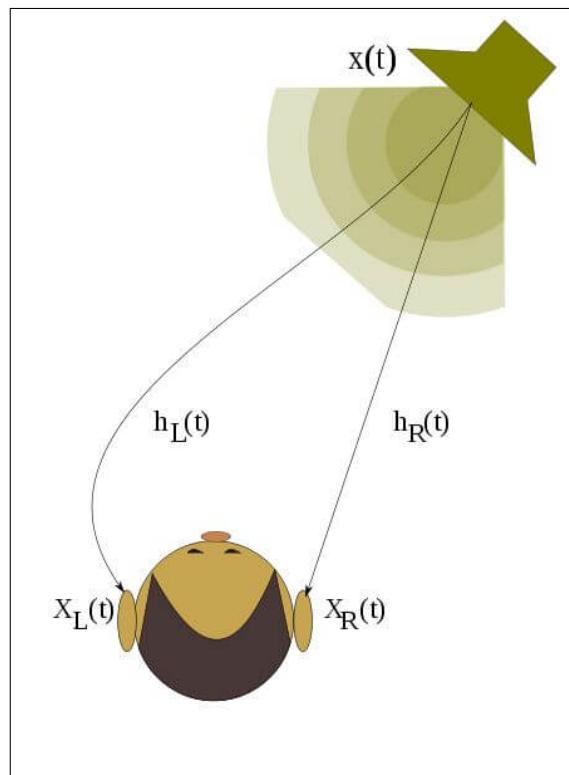


图 3.3: 头相关变换函数 HRTF

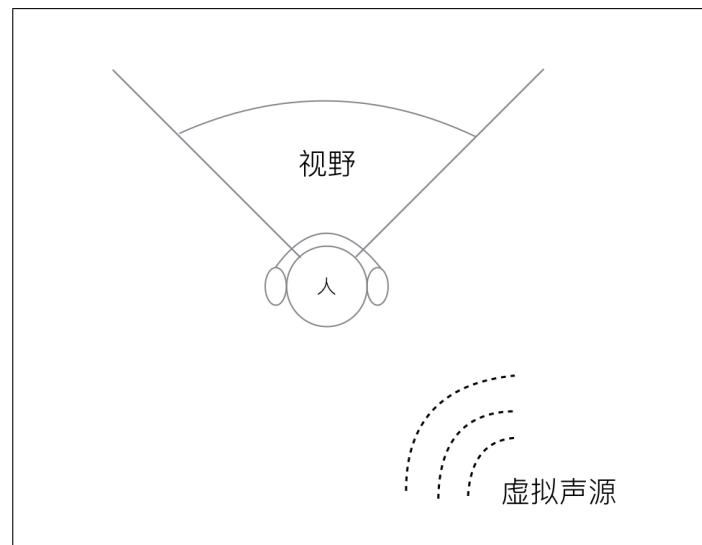


图 3.4: 用音效来表现「看不见」的事物

巴液流动引起的细胞兴奋，对人来说就是平衡感的缺失。失去平衡感会使人晕眩，严重时会使人倾倒，有造成意外伤害的可能，故在进行全景漫游时应尽可能减少乃至避免对平衡感造成的影响。

维持平衡感有两种方式，一种是降低运动的速率（即减小在全景漫游中需要运动的强度）视为主动维持平衡感，另一种是通过某些手段来减少视觉或其他感觉中平衡感的减少，即运动补偿，可视为被动维持平衡感。

现有全景漫游设备部分带有陀螺仪、部分利用手机内置陀螺仪用以感应方位信息，可做运动补偿以模拟人在环境中的运动，但人的眼球可以根据平衡来进行前庭—眼动反射（vestibulo-ocular reflex）控制眼球作补偿性运动，使之匹配头部的转向，以求维持人视野的稳定。这套机制在不戴全景漫游设备时可以正常运作，但在配戴了全景漫游设备后因眼镜视野与外界隔离，故设备与外界的运动匹配度就变得格外关键。

现有普通手机所内置的是微机电陀螺仪（MEMS），是利用物理学的科里奥利力，在内部产生微小的电容变化，通过测量电容，计算出角速度，从而判断物体转向和速度，这种芯片价格低廉，体积小巧，但缺点是因其电磁特性易受环境干扰，而且精度即使能够满足一般手机应用（如转屏、翻转等）但在更为精细的全景漫游的方位感应上还是显得力不从心。屏幕显示画面本身就有约 $1/60s$ 即 $16ms$ 左右的刷新时差，再加上陀螺仪反馈信息与实时状态也有偏差，两者累积后在使用者运动时画面很难及时更新至准确的位置上，容易造成失去平衡的感觉。甚至在人停止运动后，全景漫游设备错误地输出带有运动补偿的画面以致人眼误以为人体还在做运动，造成认知上的偏差，导致非运动状态下的感官刺激，这种刺激往往对使用者造成较长时间的影响，延续到使用完全景设备后的数分钟乃至若干小时后。

全景漫游技术的发展在维持人体平衡感方面尚有很大的发展空间，现有高端全景漫游设备即使是支持全方向漫游的平台也只允许使用者慢步使用，而运动补偿机制仍沿用着手机等终端的设计（大多数全景漫游设备并没有自己的陀螺仪，而是直接调取了与其连接的手机的重力感应接口）在需要高精度分辨的场合表现就达不到使用要求了。

3.2 全景漫游与肢体动作

全景漫游，其一是“全景”，其二则是“漫游”。漫游，意为随意游玩、漫无目的地行走。全景漫游的体验重在参与感的营造，用户不仅能够“身临其境”般地体验全景场景，更可以亲身参与其中，与场景中的虚拟事物进行互动。

3.2.1 关节运动

全景漫游人体主要运动可分解为躯干的平面运动和头部的转动，这两方面配合眼球的转动就构成了全景漫游中人个体在场景中的位移和转向，如图3.5。

但躯干的平面运动一般受场地的制约，并不是绝对的自由行进，而且由于人在看不到外界环境时处于一种对外界警惕的状态，实现行走不是全景漫游最优的运动方式。头部的转动则是一种比较好的选择，首先它的运动幅度较小，易于控制，其次头部转动也是人本能的反应，符合人对场景活动的预期。

人头部转动的形式有以下几种：低头与仰头、左歪与右歪、左转与右转。低头仰头的最大范围为 75° ，另两种形式都为 110° ，但低头舒适范围为 25° ，仰头舒适范围为 12° 。由此可得，全景漫游所用观看形式仍应以正面为主，辅以左右延展开的场景以供转头观察，这样可使人在舒适范围内使用全景漫游设备。同时，除主要场景外的功能部件也因设置在上下方向，以方便使用者通过低头仰头就可捕捉到所需要的部件。

同时，因为环境等不确定因素，易使用户丢失通过视觉等多种方式建立起的方位感。为避免方位感丢失对用户造成的迷茫，随时随地可以回到初始状态的功能也需处处设计在全景漫游中。

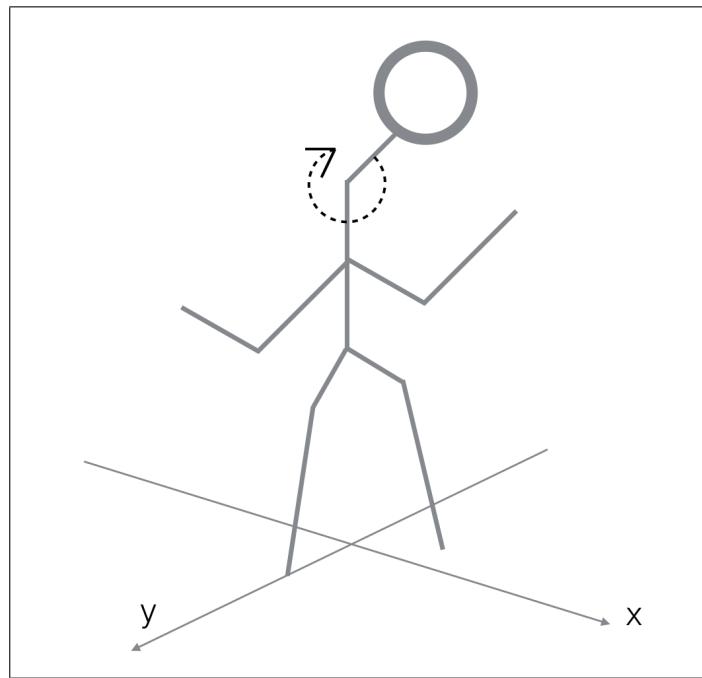


图 3.5: 躯干的平面运动与头部的转动

3.2.2 手部活动

漫游只是满足人观察的需求，需要与场景互动仍需要更多样化的操作形式。手势操作自然而然成了人控制的最佳选择，现有手柄、传感手套、指环甚至肌肉传感器可以作为全景漫游的输入装置。

手：手柄

通过手柄来体验场景可以说是最符合全景漫游现状的一种操作方式了，如图3.6。无需过多说明，使用者会自然而然将手柄与游戏中呈现的自己的“双手”在脑海中联结起来。例如，简单的射箭游戏中，使用者在若干次尝试后就能“学会”通过左手握住手柄并保持，右手握住手柄并后拉来张开场景中的“弓”，如图3.7。手柄的优点是操作简便，其与全景漫游屏幕可类比成鼠标与计算机显示屏，但缺点是交互形式略微单一（摆动手臂，按下按钮等）。



图 3.6: 用以全景漫游中操作的手柄



图 3.7: 使用手柄控制全景漫游中的射箭动作

手指：传感手套

手柄无法做到的操作形式之一是捕获手指的作用，而显然可知的是，人手进行各种动作时绝大部分操作都是通过手指来完成的，将人手指的操作追踪到系统中将极大地提高全景漫游中操作的便利性，如图。但目前手套类套件的局限性也是非常明显，比如手套相比手柄而言与手接触面积更大，在配戴时难免造成一定的不适应。而从控制计算而言，手套所需要的传感器数量将远远超过手柄，其开发相关的程序逻辑复杂性也会大大增加，故当前全景漫游配套手套尚未得到普及。



图 3.8: 使用手套可在全景漫游中进行更为精细的操作

单个手指：指环

指环作为手柄和手套的中间替代品，相比后两者而言，指环更为小巧。相比于操作手柄，通过指环与场景进行交互会更为简便，不必过于考虑把持手柄所带来的累赘感，指环如配合更为敏感的传感器或可实现相比手柄更为灵敏的操作，如图3.9。相比手套而言，指环显得不易沾染污垢、细菌等，而且由于其又可作为配饰携带，拥有者会更倾向拥有一枚可以长期配戴的指环，而不是仅在全景漫游时可用的手套。

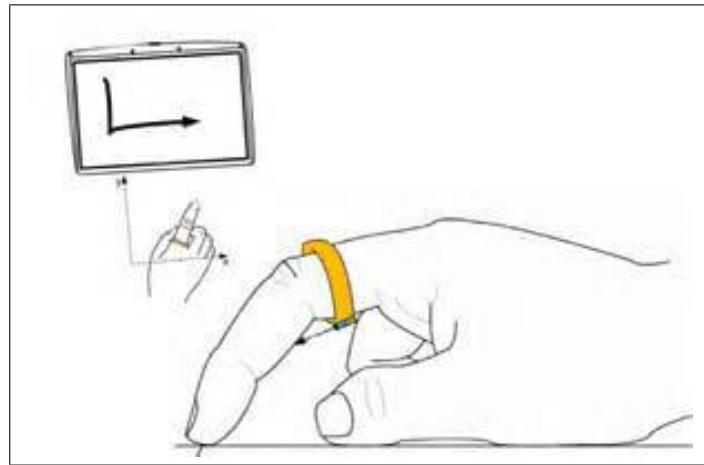


图 3.9: 配戴指环通过手指进行操作

手部操控方式比较

全景漫游目前较为成熟的手部操控方式即手柄、手套和指环这三种，此外还有肌肉传感器、甚至根据脑电位进行的脑机接口接口等，但相关技术尚未成熟。手柄、手套和指环在全景漫游中应用特点各有所长，但也有一定的使用局限性，表3.1为它们之间就一些关键性能的定性对比，以供参考。

表 3.1: 全景漫游手部操控设备比较

	手柄	手套	指环
尺寸	较大	适中	小巧
舒适度	差	较好	好
灵敏度	一般	好	较好
便携性	差	差	好
可清洁性	好	差	较差
成本	低	高	偏高

3.3 全景漫游与人的心理特征

3.3.1 知觉

人们通过感官得到了外部世界的信息，这些信息经过头脑的加工（综合与解释），产生了对事物整体的认识，就是知觉（perception）。⁴知觉按照是由人体感官或是人脑后天习得的事物特性可分为两类：前文所列举的视觉、听觉和平衡感觉等都属于第一类，而下文所涉及到的空间知觉和时间知觉则属于后天根据既往经验认知得到的第二类感觉，其特征是受环境影响大，易被外在因素所影响。在进行全景漫游活动中，后一类知觉更易受场景画面、环境声响甚至温度湿度等变化而发生改变。

空间知觉

空间知觉是指人对物体、环境等单体或多个在空间关系上的认知。其包含形状、大小、距离、方位等，在人辨别与周围环境的相互关系中起重要作用。

人可以分辨距离的能力主要来源于双眼视差和眼球变焦。双眼视差是指物体在人的左眼及右眼中分别投射的影像是不同的，根据成像距离的不同将使人产生与物体间的距离感，这也是包括 3D 眼镜在内的所有模拟立体成像技术的基本原理，如图3.10。

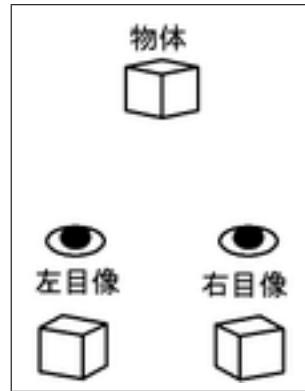


图 3.10: 双眼成象

眼球变焦是指人在观察一定距离外的物体时，需要调整睫状肌带动眼中的晶状体来进行对焦。而当晶状体发生改变时，眼球中接收到的光线就是不同距离和角度的光线了，称为眼球的“适应性调节”。人眼同时需要转动一定角度以适应这种对焦引起的改变，这种调节过程称为“视觉辐辏”。上述两者一般是成双成对地出现，这是一种正常的生理反应，如图3.11。但当配戴上全景眼镜时，人眼在注视不同物体时仍会产生上述两种反应，但因为全景图像没有景深，即所有场景内物体的焦距都是相差无几的，这时便造成了视觉辐辏和适应性调节后无法准确对焦物体，造成了视觉冲突。这种冲突称为“视觉辐辏调节冲突”，将会在一定时间后带来晕眩不适感。

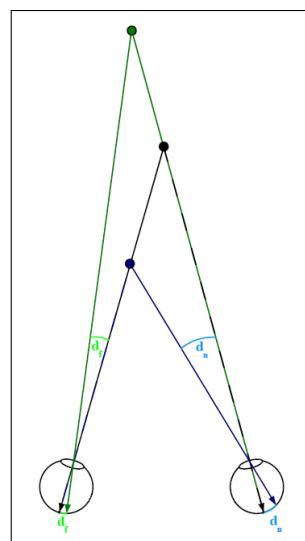


图 3.11: 视觉辐辏和适应性调节

目前能够完美解决这种冲突的方式是还原出物体的真实景深信息，使得双眼在屏幕上不同焦距时可以看到成不同像的物体。但现阶段该技术仍处在研究状态，当前可以缓解这种全景图像没有景深的方式是尽量在不影响画面完整度的情况下缩小场景内物体间景深的差距，以使眼球的视觉辐辏与适应性调节产生的偏差尽可能小。

时间知觉

时间流逝的速率在宇宙可观察的尺度上是恒久不变的，但人对时间的概念却因很多因素而发生改变。人对时间的判断依赖于人脑对事物事件的发生次序和连贯性的综合评价，也与人当前的生理状态有关。

在全景漫游中，场景中发生的所有事件都一一呈现在眼前，人脑长时间暴露在大量信息的处理过程中，会觉得时间过的很快（类似于沉迷于网络游戏的青少年永远也不觉得自己玩游戏时间过长）。而且全景场景中事物的一般运动在距人眼非常近的情况下变得非常明显，更是进一步加剧了人对时间流逝的估计。

单纯从体验角度来观察使用者在使用全景漫游设备时的时间知觉意义并不明显，但结合现有关于使用全景漫游设备体验后造成不适的例子可以看出，大部分体验者在摘下眼镜后的数分钟甚至数十分钟内均有一定的不适应感，表现为对外界环境的陌生感和行为上的迟钝感。

若要在全景漫游中保持人正常的时间知觉能力，应当适当降低场景中事物的数量和并发量，以减少人脑所需处理的信息量，使人能处于一种比较平和的体验过程。这样既可以缓解人的紧张情绪以使人可以更长时间且不疲劳地进行全景漫游体验，又能让人在场景中注意到更多细节之处，提升体验质量。

3.3.2 意识

意识可以被认为是个体对客体的一系列相关的知觉，是一种心理状态，对信息的觉察和感知，甚至还具有能动性和调节作用。

选择性注意

选择性注意是指在多种刺激中选择一种进行注意，同时忽略或减少注意其他的刺激。现有研究大多是经实验分析推测选择性注意的过程，比较著名的理论有“过滤器理论”、“衰减理论”、“后期选择理论”和“多阶段选择理论”。而上述理论都是假设注意是一个容量有限的通道，则注意的过程像是一把筛子，筛选出那些需要感觉器官反应的刺激。认知资源理论认为人的认知系统对刺激的识别需要占用认知资源，越复杂的刺激需要越多的认知资源；为了处理更多“被认为是重要的刺激”，人脑需要将有限的认知资源向更重要的刺激分配，因而忽视掉那些不是很重要的刺激。双加工理论认为人类的认知加工有两种：自动化加工和受意识控制的加工，两者并行运作，各有各自的认知过程，且可以互相转换、影响。

综合上述理论可知，人对刺激的注意程度是不一样的，人会对那些“自己觉得重要”或是“超过一般强度”的刺激进行注意。在全景漫游中，交互设计应注意刺激的强度和持续时间，不宜出现过于突然的刺激，以免令使用者产生过大的注意，而忽视正常的安全防范或作出过当的反应（如全景漫游体验过山车或造成没有准备的体验者过于激烈的心跳反应甚至危害健康）。同时也不宜使人长期处在过大的刺激中，有可能提高了人注意的反应阈限以致在体验过程后产生认知失调的现象。

持续性注意与意识动摇

持续性注意是指在一段时间内保持注意某事物的现象。人为了满足生产工作学习的需要，必须有或长或短的注意持续时间。而与之对应的是注意动摇，即短时间内注意力发生变化。

需要明确的是，注意动摇并不总是一种不好的现象，因为人持续聚精会神的时间越长越容易发生注意动摇，所以日常活动或是游戏中常见有特意设置分散注意力的动作以调整人的注意力。比如图3.12是某款手机跑酷游戏，通过设置特殊道具来吸引玩家注意力，以避免无意识的注意动摇。



图 3.12: 某手机游戏截图

3.3.3 认知

人类认知是非常复杂的过程，以下只列举了具有代表性的几种心理过程。认知在日常活动中起到了重要的作用，可以说没有认知过程人就无法进行任何需要动脑才能完成的工作和任务。

记忆

记忆是指头脑中积累和保存个体经验的心理过程，运用信息加工的术语讲，就是人脑对外界输入的信息进行编码、存储和提取的过程。记忆按信息保持时间的长短可分为感觉记忆（即瞬时记忆）、短时记忆和长时记忆。在全景漫游中，这三种记忆都非常重要。

感觉记忆只有 $0.25 \sim 2$ 秒，是屏幕图像组合成人脑中运动画面的关键因素。短时记忆则是漫游过程的直接材料，比如上文提到的射箭游戏，即是利用人短时建立起来的关于“手柄 = 弓和箭”的记忆，使得使用者在一次成功尝试后便可利用该记忆完成后续数十次操作以完成游戏。而短时记忆在脱离记忆场景后会进入长时记忆中，但不会一直被唤起。因此使用者在结束全景漫游后不会继续将其他手柄也认作为是游戏中的“弓和箭”进行操作，但在下一次游戏时使用者便可不经提示就能直接上手操作手柄完成游戏中“拉开弓并射出箭”的操作，这时短时记忆就变成了长时记忆。

思维

思维是人对外通过多种方式表现出的对客观事物的认识，是认识的高级形式。如果以相同输入而言，人对信息的加工过程就体现了思维的不同。思维的形式和理论有很多，在此只列举一个与本文贴近的假设以做参考。

人脑中基本的认知单位是概念，是人对客观事物的本质特征的认识。概念又可分为自然概念和人工概念两类，如何形成人工概念的途径也有很多假说。有布鲁纳等人提出的假设检验（hypothesis test model）认为，概念形成的过程是不断提出假设、验证假设的过程。人不会无端形成一种概念，其在不被告知正确概念的时候可以通过试错的方式来检验自身的假设是否正确，如果得到正确的反馈则验证了假设，反之则促使进行更多的假设。而成功的假设可能形成记忆，将认知保留在长时记忆中以供后续使用。

较为典型的案例是，现在各种电子产品的说明书越来越薄，甚至于全部功能都留给用户自己来发掘。当用户在第一次使用触屏手机浏览图片时，发现没有熟悉的滚轮缩放图片大小的功能，会尝试性地通过各种方

式来放大图片，而在某一次尝试中会发现“双指缩放”可以达到这种效果从而成功验证了假设，完成了概念的形成，如图。此后，即使是到了其他设备上，用户也能够利用这一概念顺利地完成类似操作。

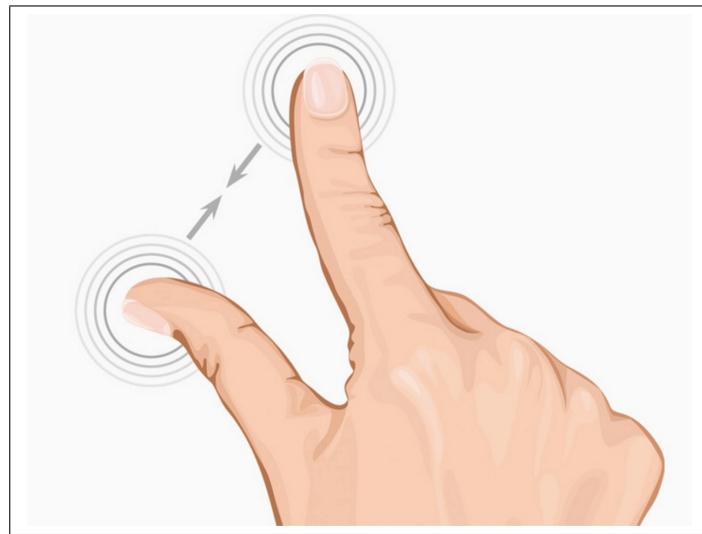


图 3.13: 双指缩放

动机

动机是由目标或对象引导、激发和维持个体活动的一种内在心理过程或内部动力。动机能够激发行为，且能引导行为的发展。马斯洛的需求层次理论认为，人的生理需求处于金字塔的末端，而自我实现和尊重的需要处于金字塔的顶端。人天生是一类勇于探索的生物，从海洋到陆地，从森林到草原，从石器到电气，人类进步最大的动机就是探索未知的世界。一个人若想自我实现所能达到的最高成就，便是在人类未知的领域去迈出自己的一步。古往今来，哥伦布、麦哲伦、阿姆斯特朗等都是代表人类迈出了探索世界的脚步，探索的精神已经深深植入人类的心底，戴上眼镜，进入全景漫游的世界，对于个体而言即是相当于自我探索的一大步，这也就是全景漫游技术所带给体验者最好的礼物。

第四章 全景漫游的交互形式

全景漫游的交互实现有可视化界面、可配戴设备的操作以及语音指令等多种形式，本文所涉及内容主要为可视化交互这种形式，辅以其他交互形式作为参照。可视化图形化人机界面即 GUI 界面由来已久，而最早的计算机则是需要科学家通过输入带孔纸带，计算机通过读取纸带上的小孔获取输入信息来进行运算，现今计算机用户甚至连 70 ~ 80 年代常见的命令行界面都未曾见过，如图4.1。

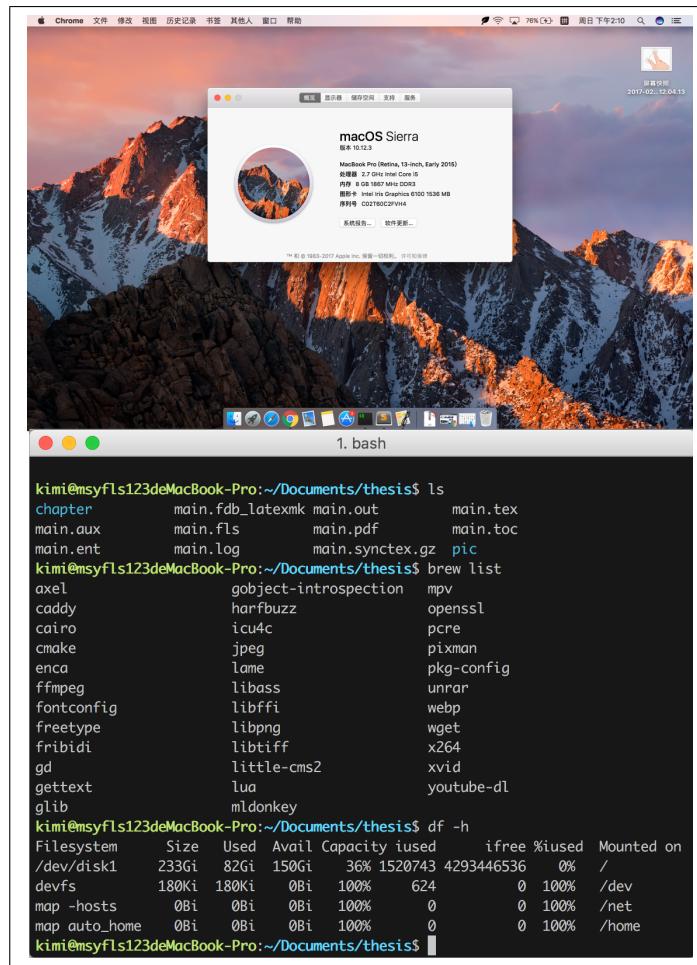


图 4.1: 图形化人机界面与命令行界面

从命令行程序到图形化人机界面，用户获得的信息变得更多，但输入信息却减少了很多。无数复杂的程序逻辑隐藏在一次又一次的鼠标点击或是手指点按后面，用户已经不是当年那群使用纸带操作计算机的科学家了。换而言之，用户无法准确记住那么多复杂的计算机命令，更不用提理解计算机内部的程序逻辑，他们只用

理解计算机可以提供给他们的功能，例如通过点击某个按钮而无须记住类似“:w”这样的程序命令即可保存文件，并善加利用即可。

而图形化人机界面毕竟信息量相比与命令行界面信息量多出好几个数量级，例如一个全屏 shell 窗口的信息量不会超过 500 字节，但是一个屏幕却有高达 $1920 \times 1080 = 2073600$ 个像素，而且图形化人机界面的信息变化效率高，可以用来表现视频动画等信息，其信息复杂度又提高了很多。

综上所述，在图形化人机界面时代随着信息爆炸而造成的信息过载（即信息的处理反馈速度低于信息生产的速度，而造成信息沉积）促使信息界面的设计制作者需从信息架构的角度出发，减少冗余信息，整理信息的分布情况以使用户能更直观高效地获取使用信息。

本章将从信息架构、功能架构两条主线出发，探究全景漫游中可视化交互的可行技术路线，并结合相关理论尝试总结出部分经过实践检验或因技术选择而必然选用的功能模块，以供设计参考。

4.1 全景漫游的信息架构

全景漫游按形式可类比成游乐园，使用者就像游客一样在场景中漫游，由此可以大致描绘出全景漫游的整体信息结构，如图4.2。这种假象的设计可以将使用者置于一种需求和环境相匹配的使用情景中，使用者可以借鉴平时游玩游乐场所积累的记忆，从场景中找寻到熟悉的“场所”以完成体验的过程。这种模式能够尽可能减少乃至消除用户对踏足陌生环境的恐惧感，人的行动预期也可以被规划在场景的设计中。

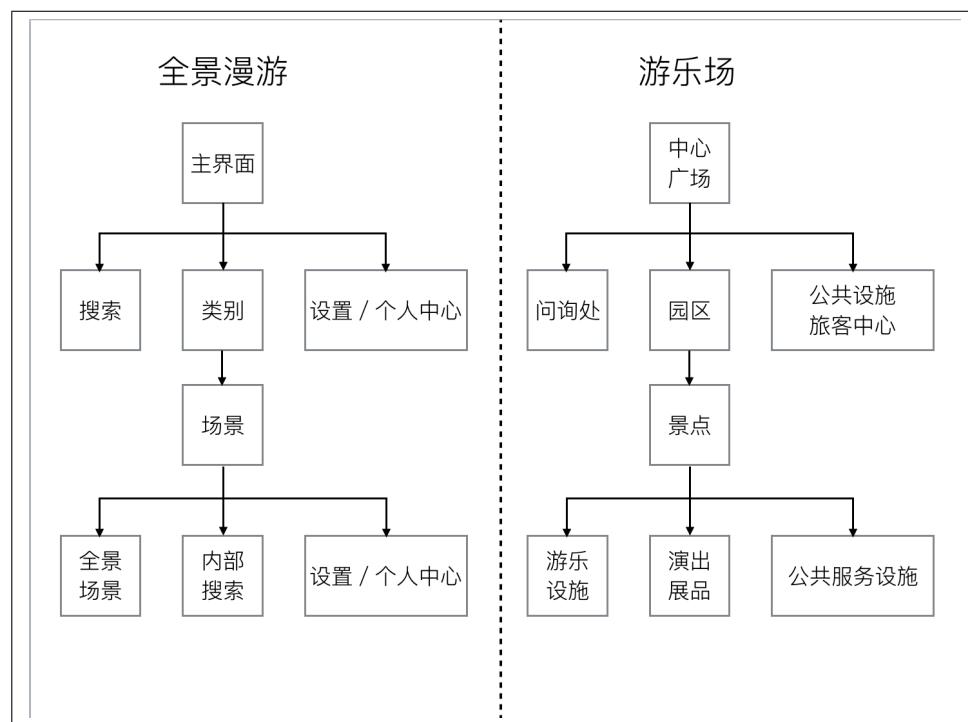


图 4.2: 全景漫游与游乐场的信息（组织）结构示意

4.1.1 信息架构形式

信息架构的形式多样，一般而言可分为：自顶向下、自底向上和不可见的这三种形式。当前数以亿计的网站、应用和系统的信息架构都可以用它们来概括⁵。

自顶向下的信息架构是通常网页所采用的形式，其核心思想是用户已经有了一个需求概念和一个立足点（比如网站的首页），用户从首页出发，通过浏览展现的各种信息选择与自己的需求最为匹配的那个进行浏览（通常是进入下一个页面），然后在下一页继续重复上述步骤直至找到自己所需的信息。自顶向下的信息架构非常清晰，所有类别目录都被呈现在用户面前，用户只需明确自己的需求就一定能找到“最接近”的那个页面。但缺点是层级嵌套过深时容易使用户丧失进一步浏览的冲动，而往往此时离最后的页面只差一步之遥。自顶而下的信息层级以不超过三级为宜，在全景漫游中，“主界面-类别-场景”的层级刚好是三级，所以大部分导航类的场景（页面）采用这种信息架构是可行的。

自底向上的信息架构主要应用于服务于单个目的的应用中，主要通过建立事物间的联系来帮助使用者在条目间切换。维基百科就是一个很好的例子，每一个词条都有众多相关链接指向其他的百科页面，通过不同的页面联结就构成了完整的知识体系。这种联系的方式也是现代搜索引擎排序的方式，即通过互相引用的计数来判断一个信息的权重，引用计数越高说明这条信息的真实性和准确性越高。全景漫游中单个场景的设计中可应用这种自底向上信息架构，例如在某个著名景点的全景漫游中，在浏览至某壁画前时通过点击其上的热点按钮，可以弹出关于该壁画的相关信息（如图4.3），这种信息代入方式相比传统的图文模式更容易使人沉浸于情景中，寓教于乐的效果更为理想。

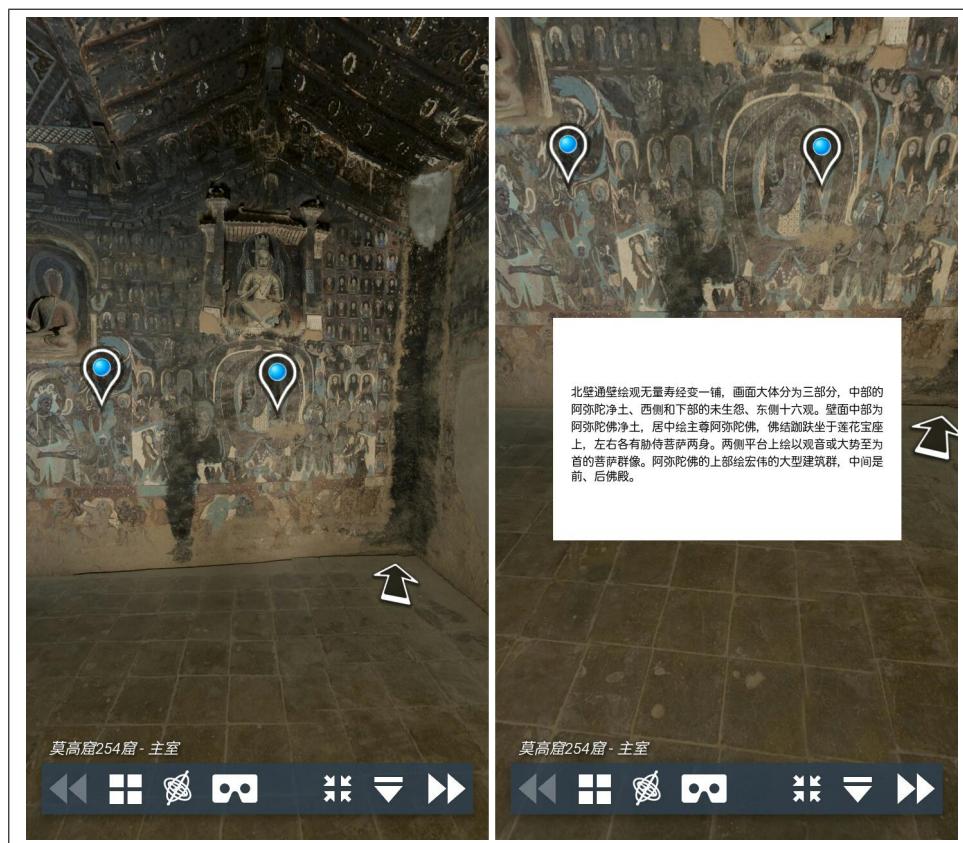


图 4.3: 全景漫游与景点介绍结合

不可见的信息架构则包含用户输入的部分转换成特定的结构，例如用户的搜索功能等。这种信息架构有时也会影响前两者，比如通过自定义类别或提高相对信息权重来引导用户去关注某一类内容，这一点赋予了信息架构的设计管理者在设计制作完成后不断完善丰富交互内容的能力。

4.1.2 信息架构组件

信息架构的层次间可以互动，用户一般不常接触直接的层次，而是通过与信息架构组件间的交互完成信息互动。常见的组件有浏览帮手、搜索帮手、内容与任务，以及“不可见的”组件。

浏览帮手如网站的目录、导航栏、网站向导等，相当于文章的摘要和目录等，用以引导用户理解网站的大体结构。搜索帮手则是帮助带有特定目的浏览网站的用户更直接更快速地通过检索相关信息，直达自己所需信息的页面。内容和任务则是帮助带有探索目的的用户，以完成任务的形式去发现网站的功能，带有这种组件的网站一般为功能工具型网站而非展示类网站。“不可见的”组件即网站背后的算法和公式等，它们指导前台组件将用户导向更为合理高效的浏览途径上去。

在场景漫游中，内容与任务在场景内起主导作用，而在场景外则是导航与搜索起到引导用户踏足更多场景等作用。

4.1.3 信息架构可视化

信息架构不是藏在界面后面的东西，相反它出现在界面的各个角落。信息是一切操作的出发点和目的地，显然可见的信息架构有助于对此感兴趣的用户去进行深入体验。常见的信息架构的可视化体现在导航栏、搜索功能等上，但其他功能也会需要信息架构的支持，所以信息架构也被用来支持其他一切有功能的架构，以一种无形的形态在其他功能模块里实现自身的可视化。

4.2 全景漫游的功能架构

以功能目的为分类标准可讲全景漫游的功能分为：漫游功能、导航功能、搜索功能和记忆功能等。在实际应用中，功能的分块可因业务逻辑的复杂而增加或减少（如增添支付功能等）。以纯粹的体验角度而言，以上四项功能构成了使用者在全景漫游中绝大部分的行为模式，故在此将它们列举出来讨论是较为合适的。

4.2.1 漫游功能

全景漫游，重在漫游。虽是全景漫游的绝对重点，但其实是设计中最不需要考虑的一点。因为对于该功能的使用，人完全是通过日常经验作出的本能反应，设计只能迎合这种需求。例如，想要让人可以观察到一个景点的全部景致就需要获取到这个地方的全景图（如图4.4），而不是简单地提供一张平面图片去让用户想象。而要做到良好的全景漫游体验，只是提供一张全景图片是远远不够的，需要加入众多的场景互动，例如上文的点击热点展开景点的相关信息等。这部分功能可以用另一种全景体验来形容，称为“增强现实”。所谓“虚拟现实”，就是在模拟出全景场景后，再添加“增强现实”的元素，以达到以假乱真甚至超过现实一般的体验。



图 4.4: 全景图像示例

4.2.2 导航功能

在场景间切换需要导航功能，这种导航与常见网站与手机应用内的导航基本类似，但因全景漫游的场景是三维空间形式的，所以在场景内也通常会设置类似于一般游戏中的小地图，如图4.5。这种地图主要起到定位使用者在虚拟场景内位置的作用，原因是真实世界内使用者的位移和转向与游戏中并不完全对应，且因全景漫游体验的封闭性，人很难找到一个参考物，而地图这种形式可以方便地指明场景中角色于其他物体间的参照关系。当使用者转向时，小地图上的代表角色的图标也会随之转向，通过仔细观察并反向演算可以恢复自身的朝向感。

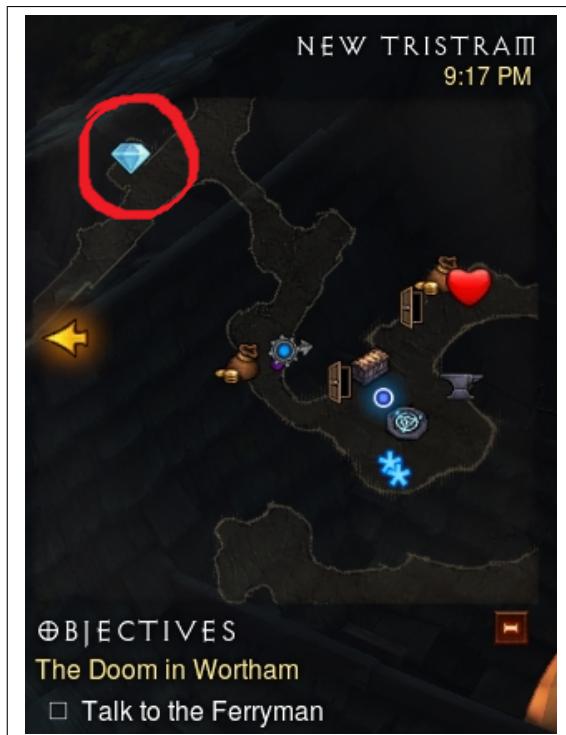


图 4.5: 一般游戏中的小地图

4.2.3 搜索功能

搜索功能是非常常见的功能，在全景漫游中搜索功能与其他的基本类似，但因手部操作形式较为简单，故其输入文字信息的功能主要为语音识别来完成。同时，场景内的搜索功能可通过对物件的注视以获取悬浮提示的方式来选择搜索目标，即使用该物件作为搜索对象进行相关信息等检索。

4.2.4 记忆功能

记忆即储存相关信息，在探索行为中是非常重要的。记忆用户行为等目的不只是给予用户可查看的历史记录，更重要的是全景漫游中用户的操作目的性有时并不是那么明显，易造成误操作，通过还原场景功能可以撤销用户最近的若干次操作，同时可以保证场景漫游的流畅性。

系统记忆个人用户的身份信息也是有所必要的，因为每个人使用习惯与生理特性不同，系统应根据账号记忆不同账户使用者的使用习惯及设置，从使用过程中自适应用户的操作，以减少用户进行复杂设置的需求。

4.3 全景漫游的功能模块

4.3.1 导航功能模块

4.3.2 场景特殊模块

4.3.3 多输入交互模块

4.3.4 热点模块

第五章 全景漫游的交互实现

第六章 待写部分

1. 以人为中心，角度，远近等
2. 上下摆动头部超过一定角度浮现 Docker 栏，例如购物车等

第七章 参考文献

¹<http://t.cn/RJDzA1F>

²<http://vreyes.baijia.baidu.com/article/595977>

³<http://www.elecfans.com/vr/444138.html>

⁴彭, 聂龄. 普通心理学 [M]. 北京师范大学出版社, 2004.

⁵罗森菲尔德, 莫尔维莱, 阿朗戈, 等. 信息架构: 超越 Web 设计 [M]. 电子工业出版社, 2016.