

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 3 月 5 日

目录

第一章 全景漫游技术与设计	4
1.1 全景漫游的目的	4
1.2 全景漫游的现状	4
1.3 全景漫游与设计结合的必要性	4
1.4 全景漫游中可视化交互理论研究的意义	5
第二章 全景漫游行业的市场现状	6
2.1 全景漫游技术总览	6
2.1.1 场景呈现硬件技术	6
2.1.2 场景捕获硬件技术	6
2.1.3 场景处理技术	6
2.1.4 场景还原与增强技术	7
2.2 全景漫游市场前景	7
2.3 全景漫游产业类别	7
2.4 全景漫游现有 APP 分析	8
2.4.1 Ascape	8
2.4.2 Fulldive	8
2.4.3 VR X-Racer	9
2.4.4 暴风魔镜 VR	11
2.5 全景漫游行业现状总结	11
第三章 全景漫游与人的关系	13
3.1 全景漫游与人的生理特性	13
3.1.1 全景漫游与视觉特征	13
3.1.2 全景漫游与听觉	15
3.1.3 全景漫游与其他感觉	15
3.2 全景漫游与人的肢体动作	17
3.2.1 关节运动	17
3.2.2 手部活动	18
3.3 全景漫游与人的心理特征	20
3.3.1 知觉	20
3.3.2 意识	22

3.3.3 认知	23
第四章 全景漫游的交互形式	25
4.1 全景漫游的信息架构	26
4.1.1 信息架构形式	26
4.1.2 信息架构组件	28
4.1.3 信息架构可视化	28
4.2 全景漫游的功能架构	28
4.2.1 漫游功能	28
4.2.2 导航功能	29
4.2.3 搜索功能	29
4.2.4 记忆功能	29
4.3 全景漫游的功能模块	30
4.3.1 导航功能模块	30
4.3.2 场景特殊模块	30
4.3.3 多输入交互模块	31
4.3.4 热点模块	31
第五章 全景漫游的交互模型	32
5.1 设备交互	32
5.1.1 信息流与反馈等待	32
5.1.2 信息修正	33
5.2 菜单交互	34
5.2.1 上下文感知交互	34
5.2.2 背景无刷新交互	34
5.2.3 三屏交互	34
5.3 场景交互	34
5.3.1 直观的交互	34
5.3.2 熟悉的定位点	34
第六章 全景漫游的交互实现	36
6.1 需求定义	36
6.1.1 价值需求分析	36
6.1.2 功能需求分析	36
6.1.3 数据需求分析	39
6.2 功能架构	39
6.2.1 发现功能	40
6.2.2 场景漫游功能	40
6.3 交互模型	40
6.4 原型设计	41
6.4.1 类别界面设计	41

6.4.2 漫游界面设计	41
6.4.3 设置界面设计	41
6.4.4 未完待续	41
6.5 程序示例	41
6.5.1 示例代码	43
6.5.2 未完待续	44
第七章 全景漫游设计评价	45
第八章 总结	46
第九章 参考文献	47

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。

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。

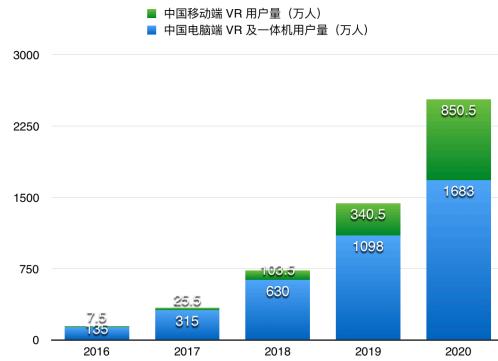


图 2.1: 2016-2020 年中国 VR 用户规模

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

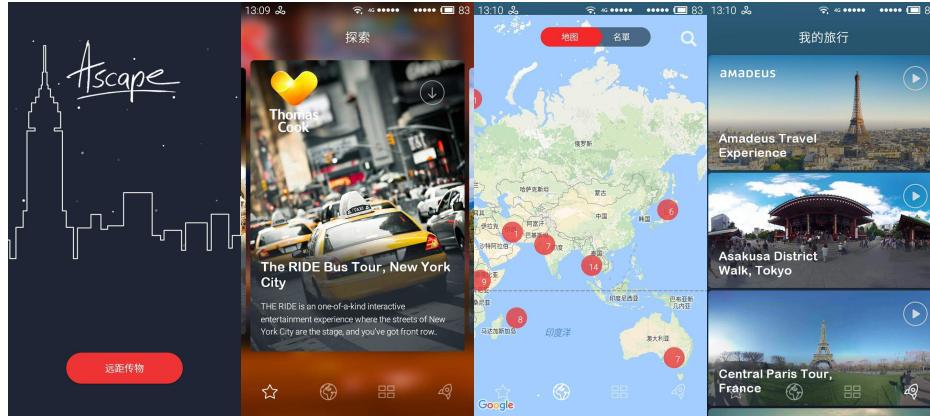


图 2.3: 简洁的全景视频播放

开发的 VR 应用，见图2.4。

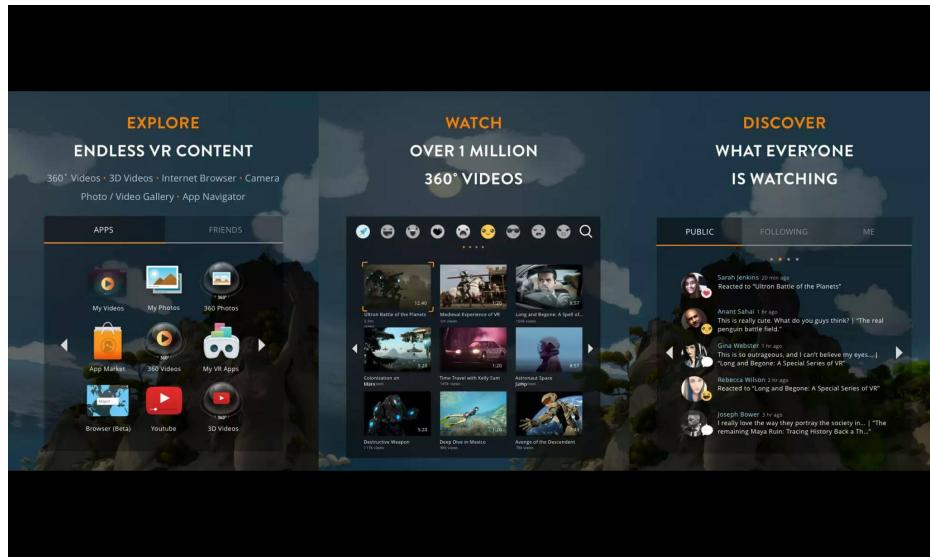


图 2.4: Fulldive 内置的应用列表

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

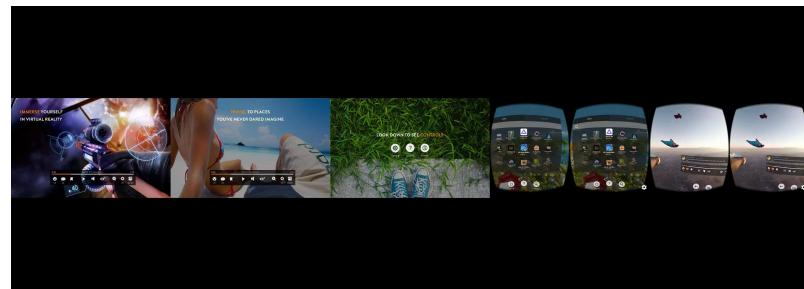


图 2.5: Fulldive 特有操作形式



图 2.6: Fulldive 视线停留模拟点击

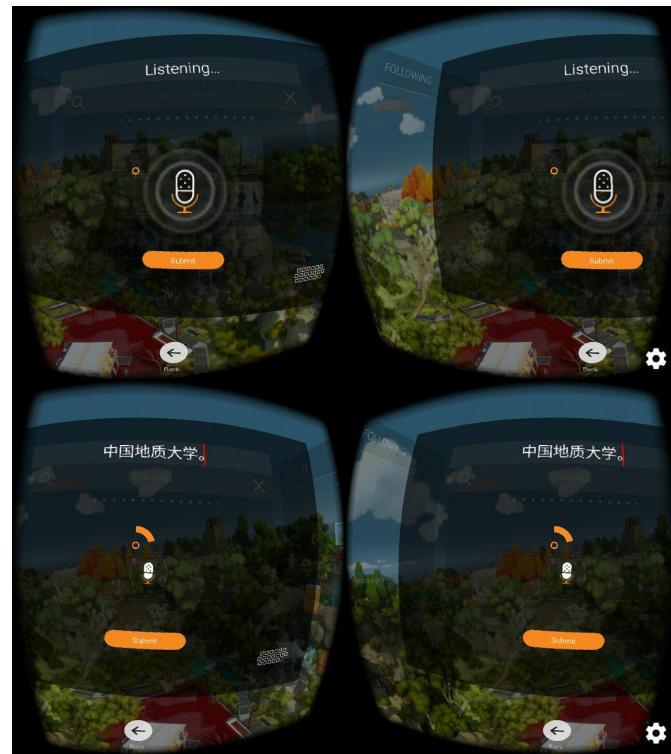


图 2.7: Fulldive 语音识别

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。

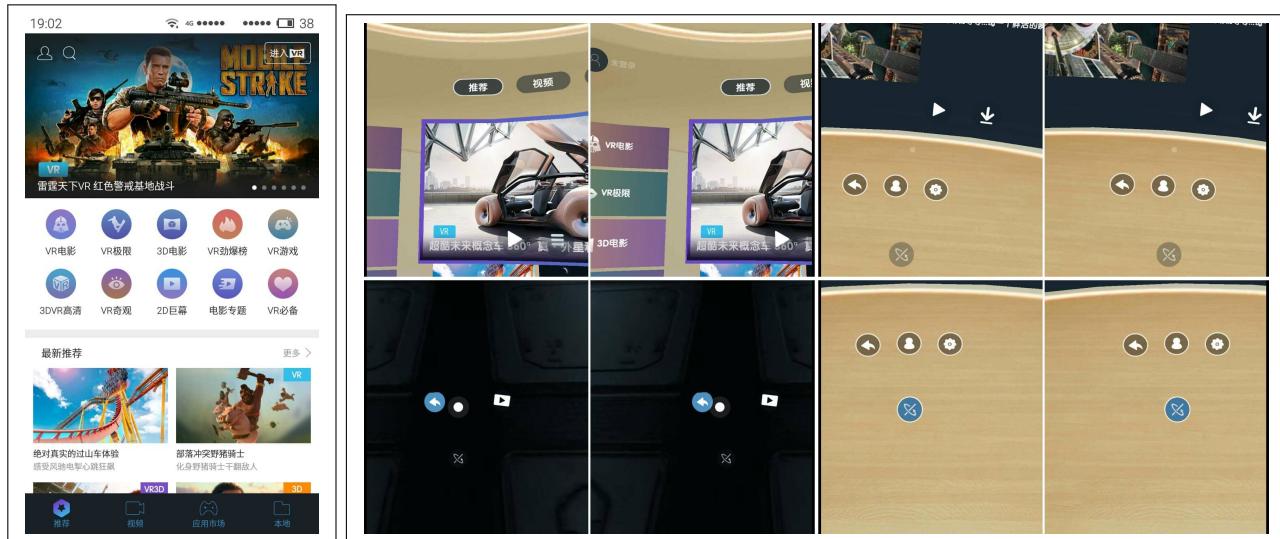


图 2.9: 暴风魔镜 VR 操作方式

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 全景漫游与人的生理特性

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。

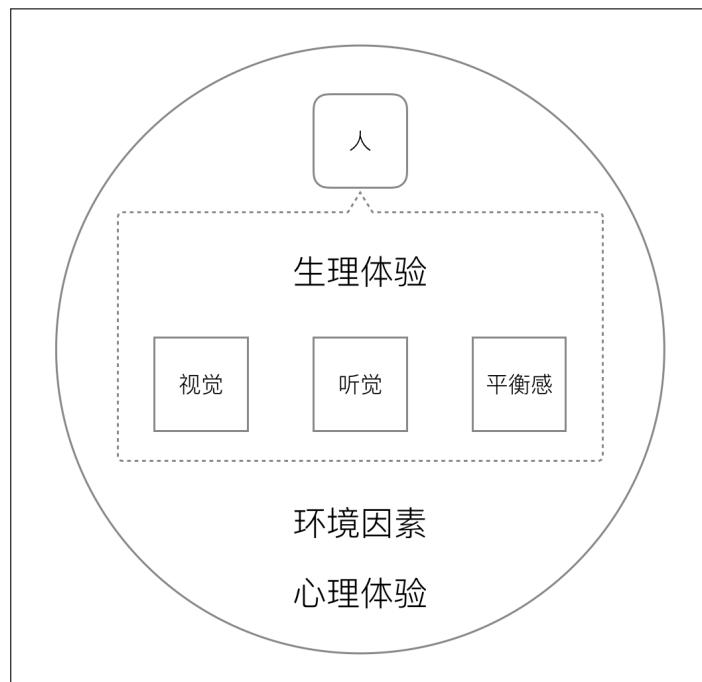


图 3.1: 全景漫游与人体知觉

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。

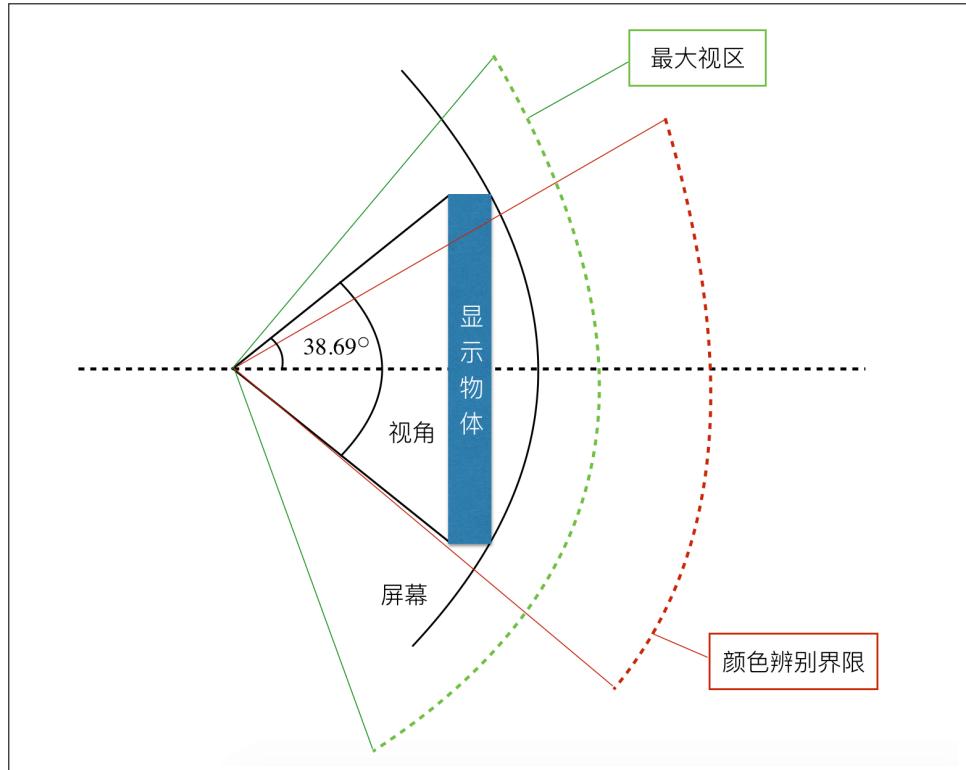


图 3.2: 全景漫游中视野的角度关系

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

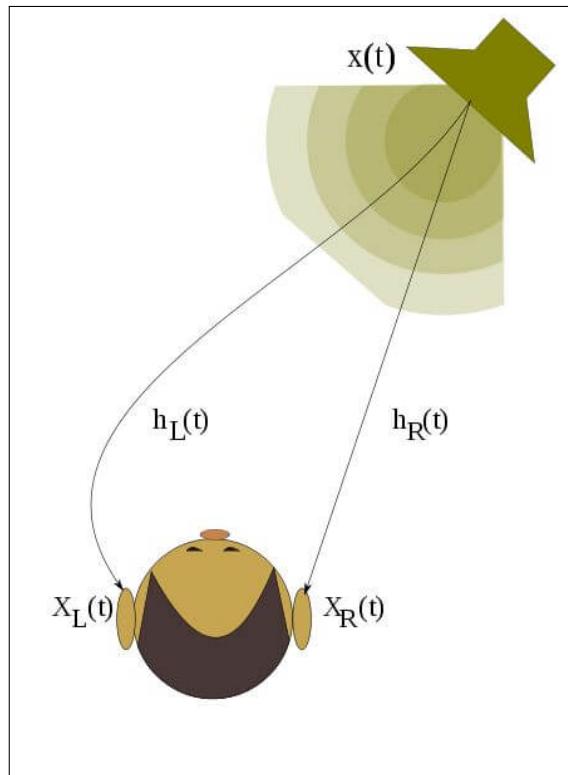


图 3.3: 头相关变换函数 HRTF

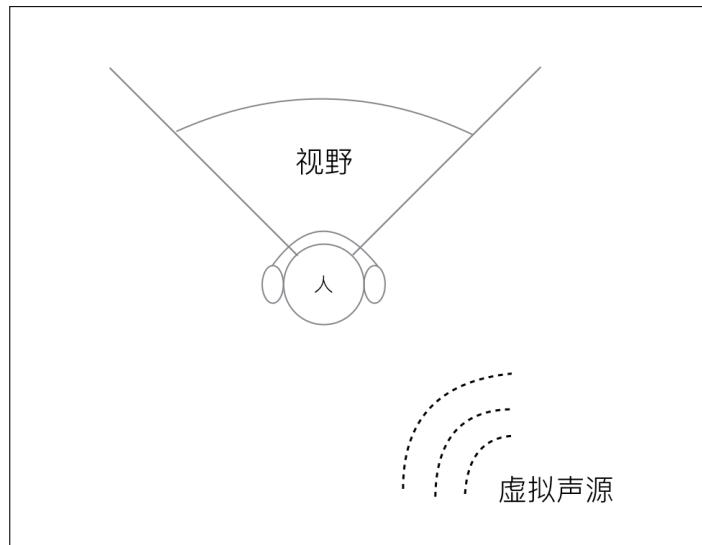


图 3.4: 用音效来表现「看不见」的事物

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 全景漫游与人的肢体动作

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。

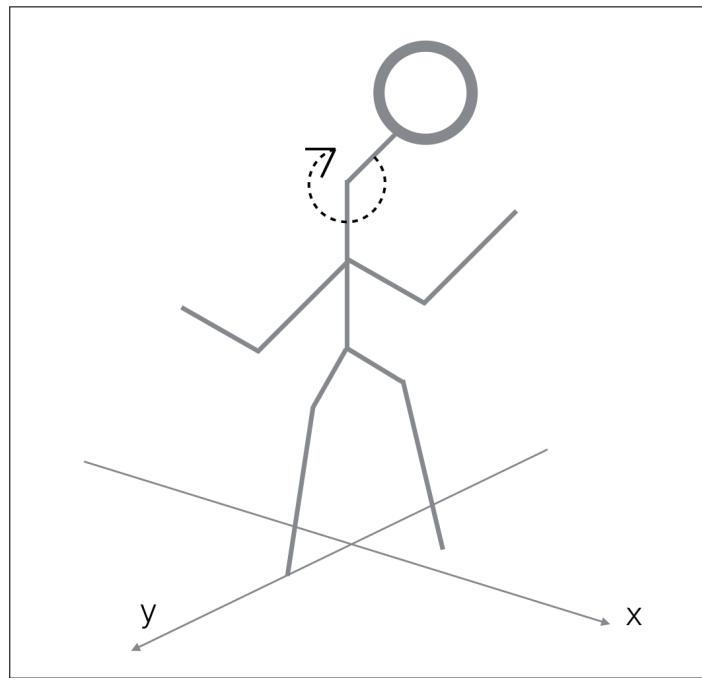


图 3.5: 躯干的平面运动与头部的转动

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。

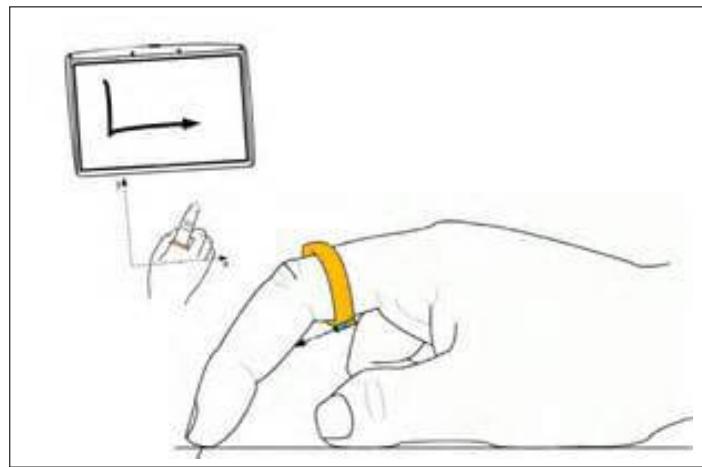


图 3.9: 配戴指环通过手指进行操作

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。

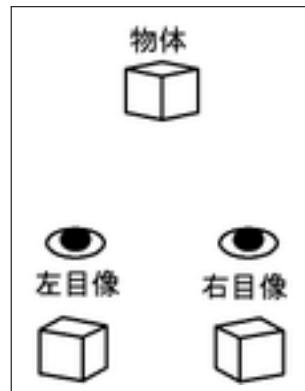


图 3.10: 双眼成象

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。

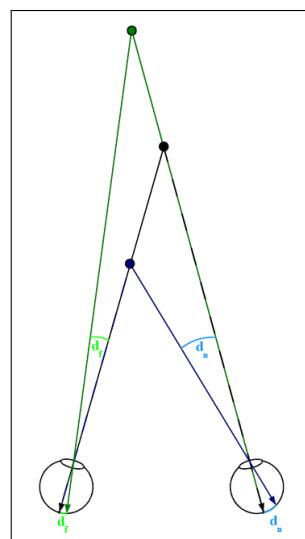


图 3.11: 视觉辐辏和适应性调节

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。

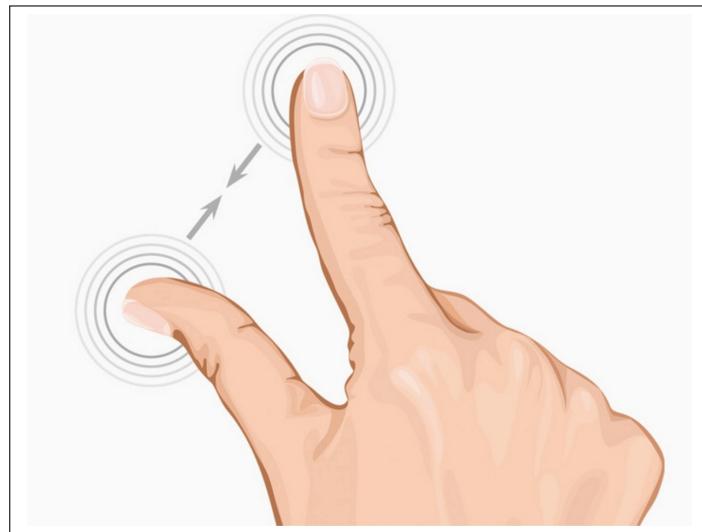


图 3.13: 双指缩放

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。

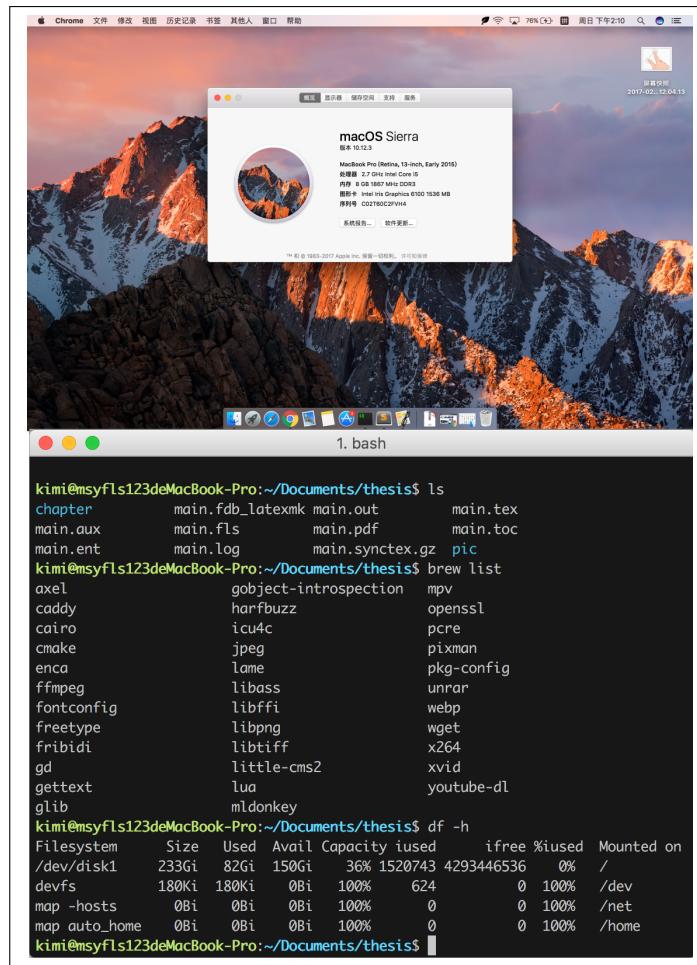


图 4.1: 图形化人机界面与命令行界面

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。

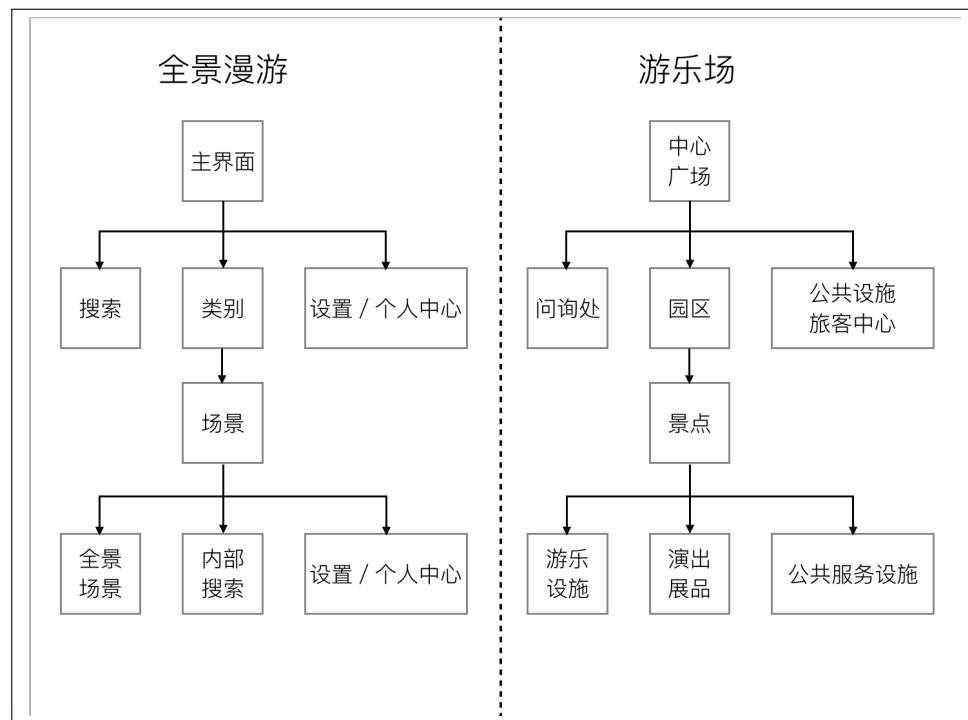


图 4.2: 全景漫游与游乐场的信息（组织）结构示意

4.1.1 信息架构组织形式

信息架构的组织形式多样，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。

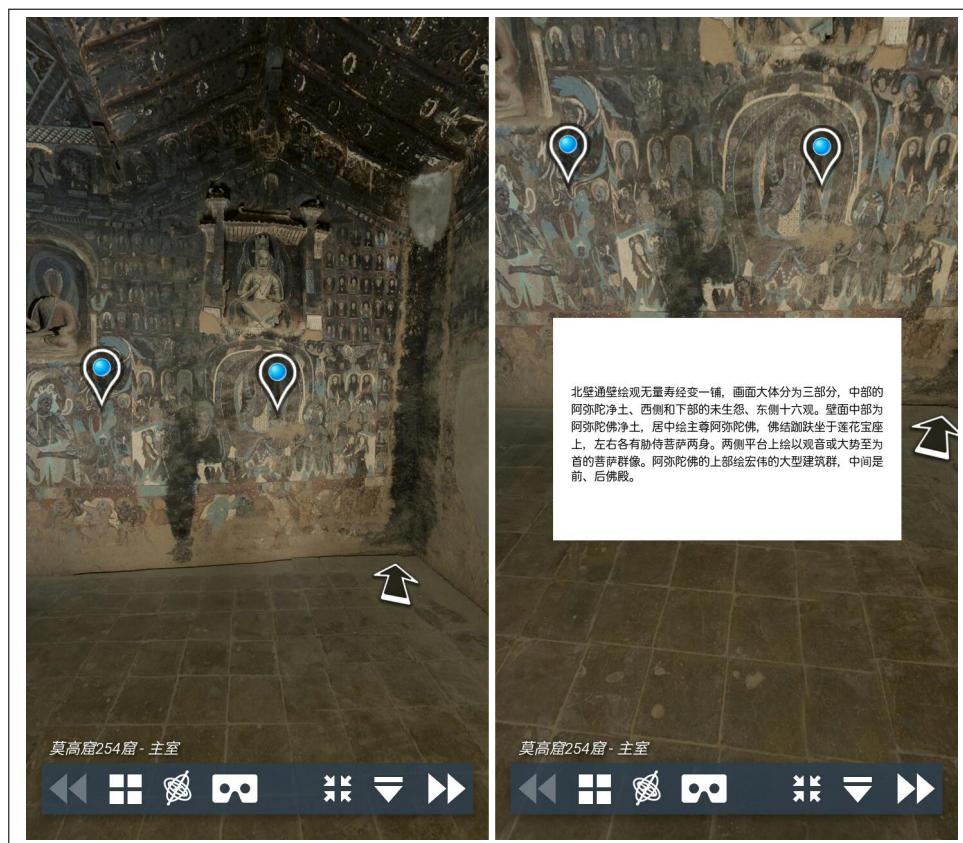


图 4.3: 全景漫游与景点介绍结合

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空间形式的，而使用者在其中并不能通过观察周围环境（虚拟场景中的地理信息通常是无效的，且脱离使用者真实所处的场景）分辨场景中的朝向，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。

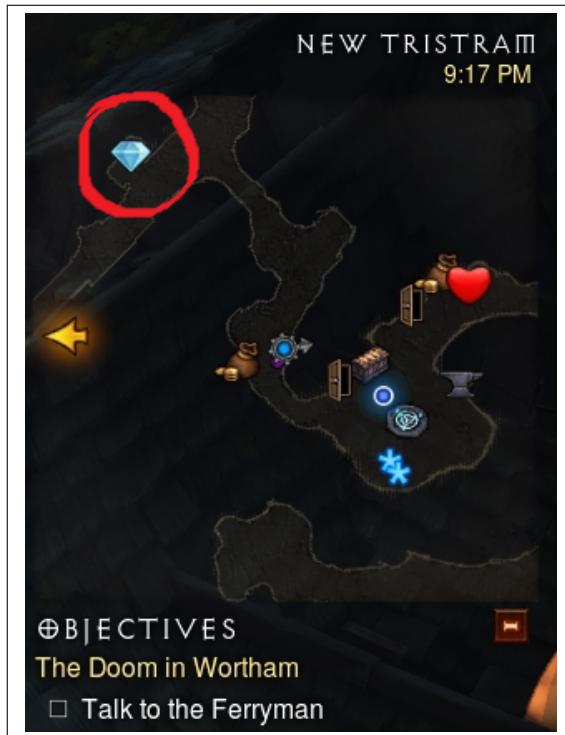


图 4.5: 一般游戏中的小地图

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

为了实现上述功能，需要在场景中插入很多组件。但组件的排序因遵循一定的规律，比如按照格式塔法则，人在注意一些相邻事物时会主动将相似的事物看作一个整体（即一个模块）进行考虑。在设计时即应按照模块化的思维对设计进行分类，模块的设计对于整个交互模式是极为重要的，好的模块设计有利于使用者准确分辨并使用不同模块的功能。根据上文所述全景漫游的功能，可分为以下四个功能模块：

- 导航功能模块
- 场景特殊模块
- 多输入交互模块
- 热点模块

4.3.1 导航功能模块

导航功能在场景漫游中使用频率并不高，主要是因为现今全景漫游大部分都只是小范围的活动，导航在整个场景漫游中起到的作用并不明显。但导航的不足会给使用者带来迷茫和困惑，甚至令使用者不再愿意使用全景漫游。全景漫游力图给使用者每次带来的都是全新的历程，所以在探索新环境时，导航类的功能会给使用者提供一种情景和舒适感。

导航的规则

导航即是将东西整理并告知使用者其所在位置的过程。导航的基本原则是在不影响清晰度和阅读效率的情形下展现尽可能多的信息层级结构，同时也需要指出用户当前所处的位置信息。导航除了基本的设计法则外，其规则基本遵循经验规则，如网站导航中经常采用的面包屑导航等。导航功能模块不是一个整体，通常情况下是由一个主要的导航界面附带多个从属的局部导航功能模块所构成。优质的导航服务不会处处显露给用户其全部内容，而只是恰好在用户需要时能够在比较近的区域内找到所需的导航元素。

导航的内容

导航模块在不同场景下应该提供不同的内容。例如，主界面下就应当提供类别切换功能，而场景界面里应提供回到主屏幕的返回功能等。具体导航内容可见表4.1。

4.3.2 场景特殊模块

相比于普通页面的功能模块而言，场景的特殊模块需要额外考虑自身在屏幕内的适应性。例如一个悬浮在场景内的模块，需要在与场景其他部分重叠或接触时可被分辨并识别出来。模块的设计需要考量的因素相当之多，甚至到具体实施的阶段仍会不断调整模块间的适配性。关于模块定义的具体内容会在下一章内以实例的形式进行说明，本段只对一些基本的因素作出说明。

表 4.1: 界面导航内容

界面	导航内容	全局导航
主界面	类别导航、搜索功能、本地内容、历史记录	否
类别	类别切换、场景列表	有
搜索	无	有
本地内容	类别切换、内容列表	有
历史记录	类别切换、记录列表	有
单个场景	场景内导航（小地图等）	有

模块的单义性

单个模块不会出现歧义，但模块一旦超过一定规模，歧义现象就不可避免地出现了，表现为重复定义模块的部分或全部。当模块间所表现出的功能相近时，使用者便会产生疑惑感继而不敢做出操作或是进行重复操作。例如，最常见的是签到功能中日历当日选项和“一键签到”的按钮，其功能意义相近，但点日历也许会有查看当日信息的操作可能，所以应该无论用户点击其中任意一项都可完成签到。在设计中应当尽可能避免这种容易迷惑用户的选项，而根据不同功能设置多个单义的功能模块，这样即方便使用者识别，也降低了开发过程中的代码耦合度。

模块的可组合性

模块并不是独立存在的，模块间也会进行组合，表现为模块互相调用依赖等。比如，搜索模块也需要唤起语音输入模块来完成文本信息的录入，否则就失去了搜索的能力。同时，模块有时也会如队列般排列，比如重要操作都会需要二次验证，最简单的形式就是弹出一个确认框模块，只有用户选择了“确认”后才会继续完成后面的流程。

4.3.3 多输入交互模块

主要表现为语音输入模块、重力感应模块等。使用者通过多种方式进行输入，设计时需要考量模块总输入的过滤筛选，并加以整合进行输出。

值得注意的是，全景漫游功能是建立在重力感应的技术基础上的，故最贴近其使用形式的交互模式可以采用摆动头部这类动作的形式。可以通过实验方式采集每种基础头部动作的三维运动加速度，经过预处理后作为头部动作的特征量，然后通过建立隐马尔可夫模型后对实际运用过程中采集的头部动作数据进行识别，转化为基础头部动作的序列，达到头部动作实现交互操作的功能⁶。

4.3.4 热点模块

场景中有很多非通用但也具有交互能力的模块，统一将其设计称为一个触发器，当使用者选择后便给它发出一个激活信号，令其执行预定的操作。热点模块类似于生活中电器的开关，随处可见非常普通但也很重要。

第五章 全景漫游的交互模型

全景漫游中的用户交互属于典型的多通道交互，根据前文所列举的关于人的意识的相关理论可知，人在面对多通道输入的刺激时，会或主动或被动地选择一种进行注意。如果这些刺激可以达到被人记忆并在需要时主动唤起的效果，那有可能在多通道输入中自动加工处理，减少对人的脑力资源的占用。同时，经过良好规划的交互模型也能避免用户出错的概率，提高网络资源的利用率，降低后台服务器程序处理的负担。

基于上述假设，尝试通过认知理论对全景漫游中设备端交互、菜单交互和场景交互，分别构建交互模型。以期通过多种形式展现信息，达到更高效表达场景的意图，结合人与设备的关系切实提升用户操作体验。

5.1 设备交互

交互从人发出的信息至被设备捕获转换为数字信号是一个明显的“编码／解码”的过程。人的动作相当于人通过分析相关信息，并试图理解机器的运作模式，认识到自己所需表达的信息，最后发出了一个按照“所认为机器能够理解的方式”进行编码的“信号”。而设备则是处在各种信息的海洋中，一刻不歇地从外界收集信息，当成功捕获到一个匹配其预设规律的“信号”时，通过确定的解码方式，还原出人最初想表达的信息。

5.1.1 信息流与反馈等待

设备无时无刻不在等待着用户的信息输入，而用户的输入是多通道的，在不同的通道上都会捕获到用户的输入流。多输入流经过处理后在主信息流上合并成一条输入流，由设备程序进行集中处理。如图5.1。

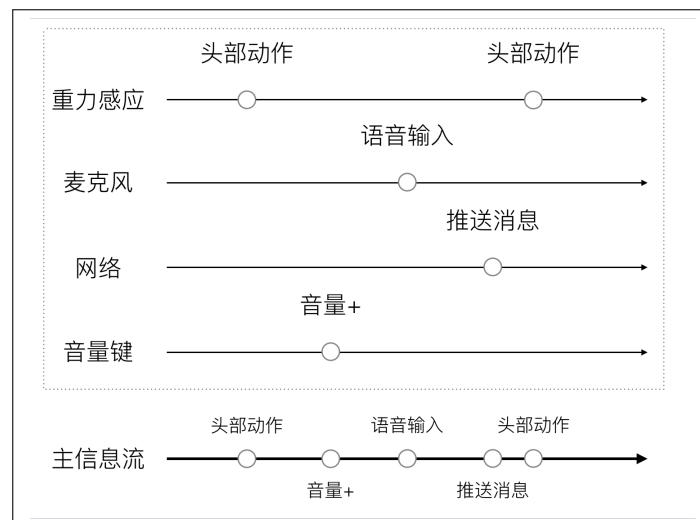


图 5.1: 信息流

但常见的场景是，交互行为的发生与处理完成得到反馈并不是即时的，而要经过一定的时间，其等待时间因网络延迟或设备响应而不一。这时用户等待的过程就会形成一阵交互效果上的“盲区”，也就是常见的“设备无响应”状态，如图5.2。“设备无响应”情形在一般设备上出现并不会造成过大的后果，但是在全景漫游中用户的视野处在封闭的空间中，长时间得不到合理反馈易造成用户心理恐慌，是需要通过设计极力避免的情形。

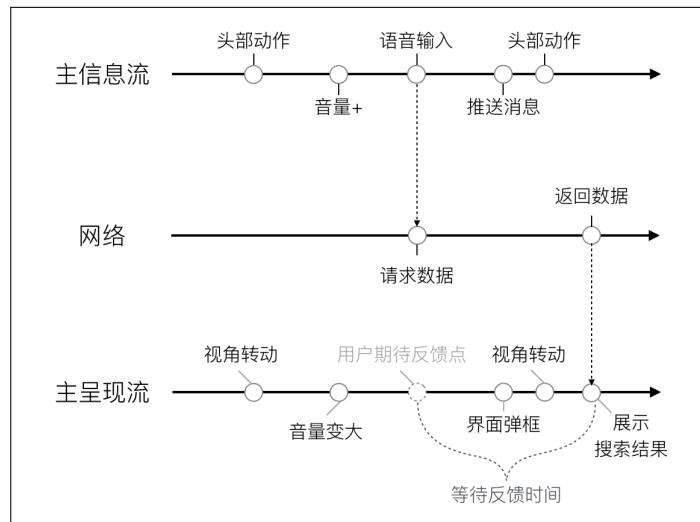


图 5.2: 交互中因其他元素造成的等待

即时反馈信息是最理想的方案，但现实情况必须对此作出解释以避免用户等待无反馈而产生迷茫情绪。常见的方案是对用户的输入做出即时反馈，当反馈不能即时给到时给出等待反馈的提示。前文所提到的通过注视(gaze)一定时间而做出确定选择的过程就是一种非即时的操作，在等待确认完成过程中通过一些标示告知用户：此处正在等待完成的行为是什么，预计完成的时间是多少等。

5.1.2 信息修正

信息流的出现帮助解决了信息加工过程中信息传达时间不一致的问题，但用户有可能得到的信息不准确或是很模糊，需要用户去判断信息是否符合预期，在无法判断信息的有效性时，如果能够提供给用户一个可以自我纠正的语境，将有利于用户自己辨别信息⁷。

例如，在导航过程中提供给用户当前朝向与正确方向的箭头图示，用户会尝试向左或向右转向，当向左转时会发现偏离变大，而向右转时会发现偏离变小，用户自然会选择向右转以回到正确方向上去，如图5.3。

在全景漫游中与设备进行多通道交互是可行且必要的，人的感官可以从不同通道间获取信息，使人获得的信息更为全面。更为关键的是多通道信息往往能够起到互补的作用，使得信息交互的鲁棒性大大提高，也使用户对于系统和环境的适应性得到了提升。而全景漫游因与普通的设备交互有所不同，故应加强信息流的建立，避免用户无谓的等待，提高交互的流畅性。而在交互的结果上注重信息的可修正性，帮助用户通过自我反馈建立准确有效的信息流。

5.2 导航交互

全景漫游中导航功能的重要性不言而喻，但场景中元素的可识别要求又限制了导航功能的复杂度。导航应符合用户对信息整合的需求，而又对场景内容起烘托作用。由前一章信息架构的组织形式可知，导航应包括

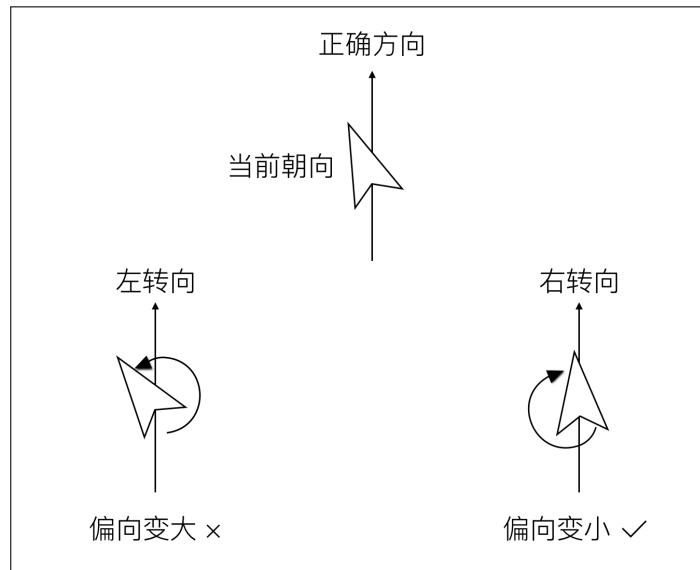


图 5.3: 自我修正的交互行为

自顶向下、自底向上、不可见联结这三种基本信息组织形式，它们分别体现为导航层级的上下文感知、单个内容增强信息展示和多任务切换。

5.2.1 上下文感知

上下文是指当前环境中区别于当前本体的因素。上下文按与用户关系可简单分为主动上下文和被动上下文两种。主动上下文指环境中可以直接改变系统行为的状态或变量。被动上下文指那些虽不能直接产生作用但可引起用户兴趣继而通过用户改变系统行为的状态或变量⁸。上下文感知是指计算实体能够根据上下文环境的变化及时调整自身行为，使用户从信息的管理和输入中解放出来，专注于要执行任务的本身⁹。

上下文感知的作用非常重要，没有上下文用户很难预计进行操作时即将发生的变化，容易在使用过程产生困惑。在通过列表批量展示内容时都会同时展示该内容的多项信息，包括名称、录入时间、信息量等（可能因内容类型不同而信息名称有所差异）。例如商城类网页会展示商品的名称、产地、出厂时间等，这些信息就是上下文。通过阅读理解相关信息，用户可以在不打开具体页面的情况下对相应项目得到一个基本的认识。

5.2.2 增强信息

在单个页面里尽可能多地铺陈相关信息，包括其他相关内容，减少用户反复查找的负担。

5.2.3 多任务切换

类似 MacOS 的最大化最小化。

5.3 场景交互

5.3.1 直观的交互

WYSIWYG，即“所见即所得”是一种常见的编辑模式，它便是一种很理想的设备交互模型，用户不必关注程序背后的逻辑，甚至不必关心自己编辑的文章真实存储的数据是何种形式，只需要看到屏幕上呈现的内容就是别人所能看到的内容。

全景漫游展示的正是这样一种过程，由不直观的概念转换为直观的概念。如图5.4所示，人对房屋整体概念的认识观念是从具象至抽象，只有抽象简洁的概念才能方便地存储于人脑的记忆中，但涉及到具体房屋的理解时，是从单纯的“房屋”名词，到草图的展示，再到带有色彩的照片。而全景漫游要实现的是一个身临其境的房屋场景，是比照片更贴近真实房屋的信息载体。全景漫游的内容注重交互形式和体验感受的留存，不注重内容的记忆，故直观的交互符合其交互需要。

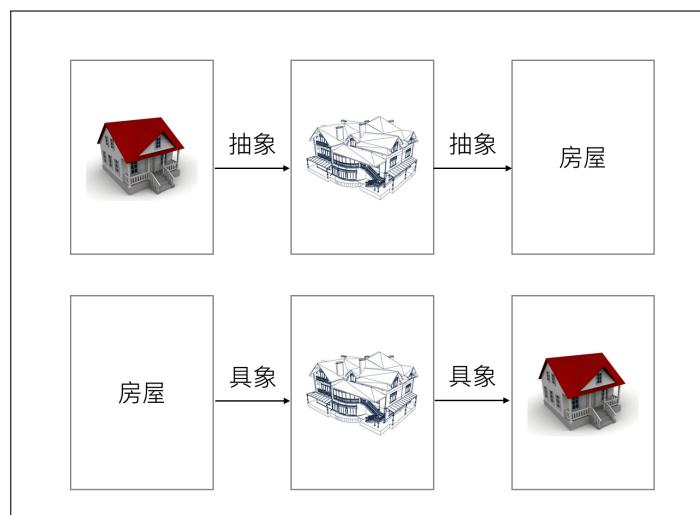


图 5.4: 具象和抽象

5.3.2 熟悉的定位点

游戏中我来过的地方可以作为参照物。

第六章 全景漫游的交互实现

上文中通过对全景漫游与人的生理心理特性相互关系进行分析，梳理了以信息架构和功能模型为主导的可视化设计思路，为实际开发全景漫游项目奠定了基础。结合上述理论支持与作者实习实践所得经验，本章就某公司全景漫游系统开发为可视化交互示范案例进行分析。本案例将从需求定义、功能架构、交互模型定义出发，以实例开发为主线，通过界面设计和程序逻辑部分详细说明全景漫游的可视化实现过程，将理论与实践结合，从而提升全景漫游整体的使用交互体验。

6.1 需求定义

本例中的研究对象是一个包含商城内容、视频内容和基本交互场景的基于 web 技术的全景漫游系统。原始需求列表如下：

- 满足基本全景漫游功能。因 web 环境受限，以直接在浏览器内打开所支持的操作为限，暂不考虑开发多种操作形式功能。
- 带有商城功能，实现互动场景的付费／免费使用。
- 支持全景视频的播放，基本功能类似普通播放器。
- 带有基本的用户设置功能。

根据以上原始需求，从价值需求、功能需求、数据需求等三个角度出发进行需求分析。

6.1.1 价值需求分析

该全景漫游系统的主要用户群为使用计算机或手持移动设备进行网页浏览的青少年群体。其用户特征为对新事物充满好奇心，愿意为良好的用户体验付出金钱成本，但缺点是对事物的专注度低，易于分心。

以该群体作为主要用户对象进行价值评估，按功能和用户的对应需求结合匹配，其价值需求如表6.1。

根据分析，目标群体所期望的是一种不带有过多使用负担、随到随用、用完可离开也可进行分享的交互体验¹⁰。

6.1.2 功能需求分析

据上述价值需求分析可得该全景漫游系统的功能需求，系统内的各种分类或是查找功能等，目的都是通过数据整理来让用户更好地发现场景并与其进行交互，如图6.1。经过整理分析以及结合软件开发难度等因素，可得功能需求表6.2。

表 6.1: 目标群体价值分析

功能点	用户需求	价值
新鲜感	体验到与其他应用不同的内容	高
操作性	可以方便地了解并获取到所需体验的内容	中
可信度	能够理解场景传达的信息并作出肯定答复	中
丰富性	内容类别丰富、有不同难易层次可供选择	中
可搜索	可以通过文本信息检索需要的项目	低
可回溯	查看历史回放、收藏夹功能	低
社交性	体验后愿意与别人分享使用体验	高

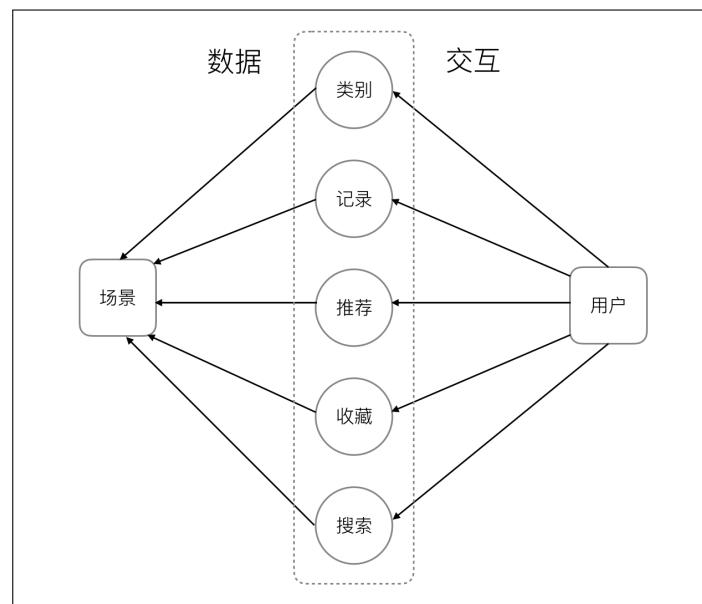


图 6.1: 用户通过多种形式发现场景

表 6.2: 功能需求表

类别	功能点	描述
主界面	推荐	进入推荐场景
	类别	按类别查看场景视频, 可翻页, 可进入类别界面
	应用	被固定的第一三方开发的全景应用
类别界面	列表	查看场景视频
	标签	切换类别
应用界面	列表	展示第三方应用
	标签	切换类别
设置界面	通用设置	提醒、关于
	个人偏好	账户设置、收藏夹、使用记录
全局导航	返回	返回上一层
	主页	返回主界面
	搜索	唤起全局或当前类别搜索
场景	定位	回到初始角度
	标签	标记为喜爱或讨厌、帮助改进推荐质量
	分享	分享页面至社交媒体
	解锁	解锁付费内容

6.1.3 数据需求分析

根据上述功能需求可知，整个系统中最小的个体单位是场景，数据集合的建立是以场景为中心的。则该全景漫游系统的数据需求以单个场景的数据集出发，进行数据字段的定义，见表6.3。

表 6.3: 场景数据集

类别	字段	描述
基础信息	名称	场景名称
	性质	视频、游戏、应用等
	标签	场景分类（如科幻、太空等）
	介绍	场景简短介绍
自然信息	容量	场景占用空间大小，以 kb 计
	创建日期	场景创建日期
	时长	仅对视频场景有效
	地理信息	场景录入地点
人为信息	价格	免费、收费、部分收费
	分级	年龄限制、内容限制
	可用	可用或不可用

而作为交互界面，场景是用来服务于使用其的用户的，故数据需求也包含用户信息数据的模型，见表6.4。

表 6.4: 用户数据集

类别	字段	描述
基本信息	名称	账户名称
	密码	账户密码
	邮箱	账户邮箱
被动收集信息	观看记录	场景、时间、时长
	付费记录	场景、应用、视频即其价格
	操作记录	使用场景的有效信息
主动输入信息	标签	标记喜恶、收藏
	偏好设置	使用设备、情景模式等

6.2 功能架构

据上述需求分析，该全景漫游系统等功能架构表现两种核心功能：“发现功能”和“场景漫游”。

6.2.1 发现功能

- 浏览功能：使用者在包括主界面内的各种类别界面内浏览场景相关信息，发现自己所想要体验的场景。
- 搜索功能：使用者需要即时搜索得到自己想要的信息。
- 记录功能：使用者可以查看自己的浏览历史和收藏内容。
- 付费功能：使用者需要支付一定成本来查看某些场景或场景的局部。
- 自定义功能：使用者可以通过设置来自定义系统偏好。

6.2.2 场景漫游功能

- 漫游功能：使用者可以通过转向、仰头低头、注视等方式漫游整个场景。
- 互动功能：使用者可以通过转向、仰头低头、注视等方式漫与场景内特定物件进行互动。
- 视频功能：使用者可以以全景的形式观看特殊全景视频。
- 标记功能：使用者将场景标记自己的评价、收藏或是分享给其他人。

6.3 交互模型

交互模型即是一系列交互方式、模式、行为的总称，它定义了交互过程中交互的主体对象和行为方式。交互模型并没有准确的定义，因为根据不同的产品而言，交互行为的目的并不相同。例如，对于网站管理员而言，编写并输入网站内容是使用网站的目的，而对于一般浏览器而言，浏览网页上提供的信息则是其在该网站上行为的目的，故从不同角色和视角看待交互模型都会是不同的。而构建交互模型的主要目的则是指导交互设计的进行，则构建一个能同时被用户接受并符合一般软件开发规律的交互模型是十分重要的。

交互模型需要符合用户心智模型并在后续可以方便进行扩充。在该全景漫游系统中，其交互模型如图6.2。

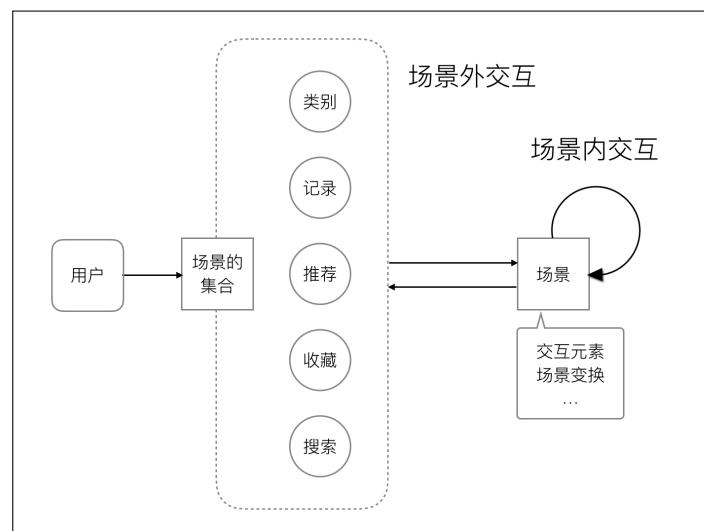


图 6.2: 全景漫游系统的交互模型

6.4 原型设计

6.4.1 类别界面设计

上方为类别菜单，下边横长条为标签栏，可切换分类。底部为通用导航栏。如图6.3。

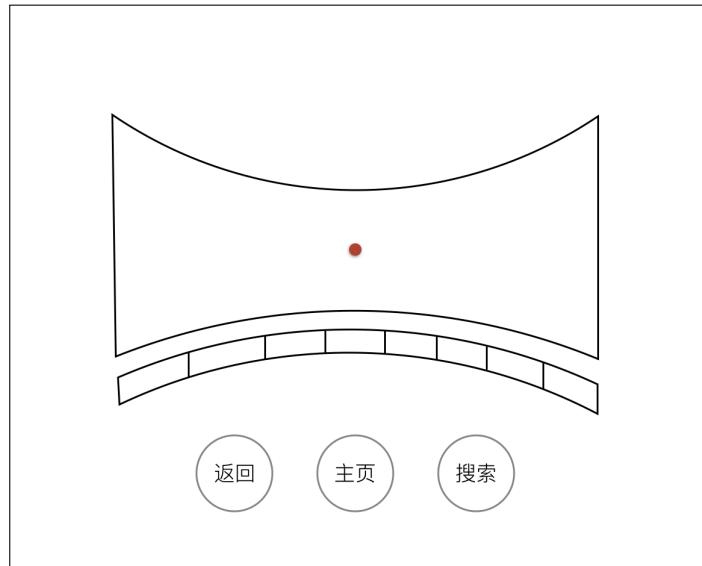


图 6.3: 类别界面设计

6.4.2 漫游界面设计

- 漫游界面带有紧急退出区域，只需注视 1 ~ 2 秒以上就可以紧急退出场景，避免眩晕加重。
- 可通过地面上的指示图标切换场景。
- 可通过场景中的热点区域查看详细信息。

如图6.4。

6.4.3 设置界面设计

设置界面可通过注视相关范围控件以调节参数，如图6.5。

6.4.4 未完待续

6.5 程序示例

本全景漫游系统采用 Mozilla 开源的 Aframe 框架进行开发，可满足普通计算机及移动设备通过浏览器进行使用。

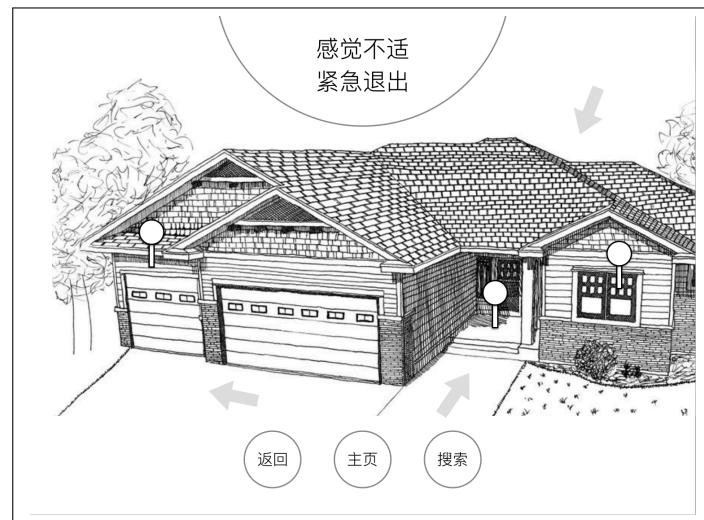


图 6.4: 漫游界面设计

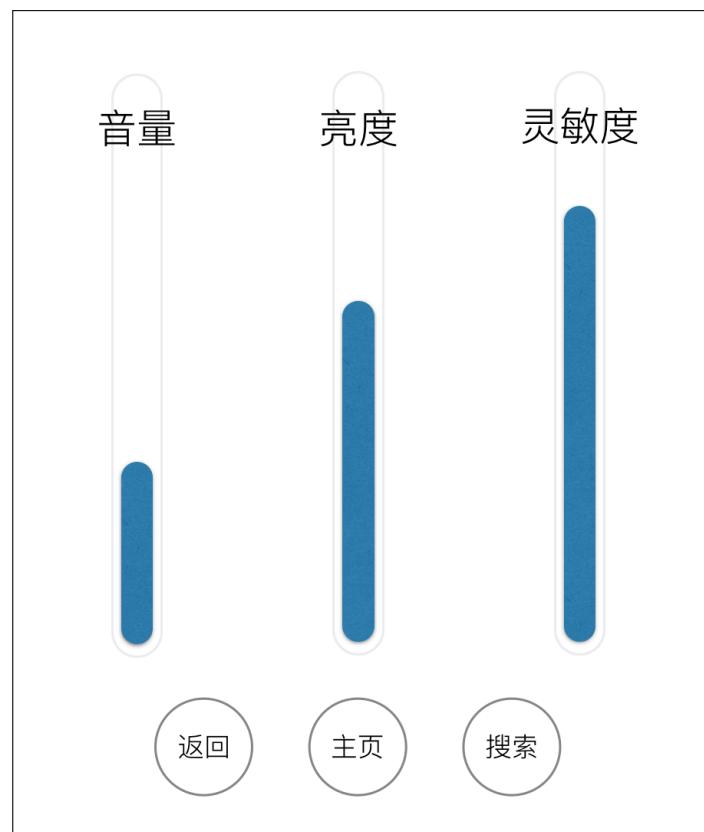


图 6.5: 设置界面设计

6.5.1 示例代码

实现一个简单的可用鼠标或注视（Fuse）操作的小场景。

HTML

```
1 <a-scene>
2   <a-entity
3     id="box" cursor="fuse: true; fuseTimeout: 1000"
4       geometry="primitive: box"
5       material="color: blue">
6     <a-box
7       src="img/logo.jpg"
8       position="0 3 -5"
9       rotation="45 0 45"
10      scale="2 2 2"
11      id="box1">
12      <a-animation
13        attribute="position"
14        begin="focus"
15        to="0 2.2 -5"
16        direction="normal"
17        dur="2000">
18      </a-animation>
19    </a-box>
20  </a-entity>
21
22  <a-camera>
23    <a-cursor></a-cursor>
24  </a-camera>
25
26  <a-plane
27    position="0 0 -4"
28    rotation="-90 0 0"
29    width="4"
30    height="4"
31    color="#7BC8A4">
32  </a-plane>
33  <a-sky color="#ECECEC"></a-sky>
34</a-scene>
```

JavaScript

```
1 document.querySelector('#box').addEventListener('click', function(e){
```

```
2 |     document.querySelector('#box1').emit('focus')
3 | })
```

6.5.2 未完待续

第七章 全景漫游设计评价

待补充，主要方式为在程序中埋点自动记录用户使用习惯：

1. 用户初次打开页面至进入第一个场景的时间
2. 用户在第一个场景中停留的时间
3. 用户第一次学会使用视线停留触发物件效果的时间
4. 用户触发全局导航栏内任一功能时间
5. 用户是否分享、分享到何处
6. 用户是否第二次打开应用、相隔时间
7. 用户有无打开多个应用或场景、个数

第八章 总结

本文从全景漫游发展的现状出发，列举并阐述了现有全景漫游技术的特点及需求，通过比较全景漫游多种应用的操作方式及交互流程发现其有待加强的地方。从人体生理与心理的相关特征特性出发，结合人机工程学和心理学观点深入剖析了全景漫游与人结合的难点和问题，力图以人为设计中心，加强人在全景漫游体验中的参与作用。从图形化人机界面到信息架构再到功能模块，以交互研究的方式结合相关实例说明交互模式在人机界面的重要作用。以实例分析的方式解释并总结了某一例全景漫游系统的交互设计及程序实现的过程，从需求分析到功能架构，以交互模型为指导原则进行可视化设计，完成了漫游系统的开发。在设计完成后，以收集用户体验数据的形式得到了宝贵的第一手资料，分析并论证了该交互设计中应用相关设计理论的可行性与成效。

第九章 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, 聪龄. 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.
- ⁶孔俊其. 基于三维加速度传感器的手势识别及交互模型研究 [D]. 苏州大学, 2009.
- ⁷路璐, 田丰, 戴国忠, 王宏安. 融合触、听、视觉的多通道认知和交互模型 [J]. 计算机辅助设计与图形学学报, 2014, (04):654-661.
- ⁸张磊. 上下文感知在导游系统中的应用 [A]. 中国计算机学会、中国图象图形学学会、ACM SIGCHI 中国分会、清华大学计算机科学与技术系. 第一届建立和谐人机环境联合学术会议 (HHME2005) 论文集 [C]. 中国计算机学会、中国图象图形学学会、ACM SIGCHI 中国分会、清华大学计算机科学与技术系;, 2005:4.
- ⁹王守芳, 金浩, 魏鲲, 等. 上下文感知综述 [C]// 建立和谐人机环境联合学术会议. 2005.
- ¹⁰陈圆. 手机应用软件界面体验设计研究 [D]. 哈尔滨工程大学, 2013.