

VR 眩晕问题的产生机理与解决方案探究调研

——18373528 杨凌华

一、现状阐述

虚拟现实技术在过去二十年间，有了较为快速的发展，各方面的性能有了逐步的完善，许多行业大鳄在过去的十年间相继推出了虚拟现实的相关产品和服务，虽然得到了一定程度的商用，但距离大众化普及却还有着很大的距离，因为在某些关键技术上仍然存在着较大的瓶颈未能得到有效突破。

当前虚拟现实所面临的瓶颈主要分为三大方向：

1) 大范围多目标的精确实时定位和数据实时共享

当前大多数产品还只能针对一个独立有限的狭小空间内的单个个体进行位置姿态的精确实时测定，而未来的应用场景需要针对大范围多目标的精确实时定位以及数据实时共享，其首要目标就是提高计算速度和准确度并降低数据传输的延时。

2) 感知的延伸

视觉系统是人体最复杂、最重要、精确度最高的感知系统，因此目前虚拟现实的大部分产品都针对于单一视觉来提高用户的沉浸感，但这必然是有局限的，还需要在听觉、嗅觉、触觉等感知系统上做进一步的延伸。

3) 眩晕感和人眼疲劳

目前几乎所有的在售产品，都未能有效摆脱引起使用者产生眩晕感的问题，用户往往在持续使用 5~20min 后就会开始产生不适，因此有必要从机理上分析产生眩晕的原因，并采取有效的解决措施尽可能降低用户的眩晕程度，延长产品的使用时间。

在上述技术瓶颈中，大范围多目标的精确实时定位和数据实时共享，其突破只是时间的问题，当前 GPU、CPU 的计算性能都在随着时间稳步增长，5G 技术也在日趋完善，这都将为大范围多目标的虚拟现实场景提供有力的技术支撑。感知的延伸还存在着较大的技术难度，尤其是在触觉的精度把握上，还远远无法和真实的触觉相媲美，但这并不是阻碍虚拟现实往前发展的决定性因素，因为目前人们对于触觉的精度要求还并不是那么迫切。相比之下，眩晕感和人眼疲劳的问题

目前为止还没有得到有效解决但人们对其又有着迫切的需求,因此它才是阻碍当前虚拟现实技术大众化普及化的重要因素。

本文想通过调研报告的形式,从感知机理层面分析眩晕产生原因,并汇总当前人们所提出的有效可行的解决方案,并对其进行分析与评价,最后对虚拟现实的未来进行展望。

二、问题分析

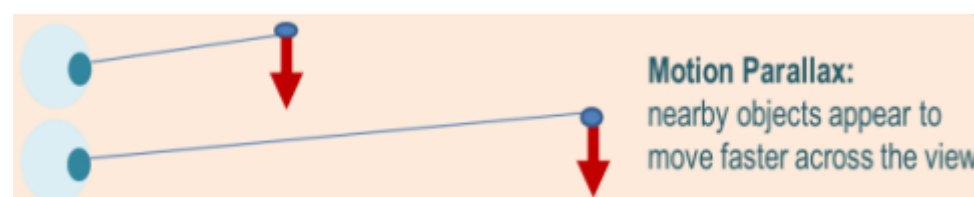
2.1 人眼感知三维的视觉机理

人眼区分二维画面和三维空间场景的主要方式,是通过利用双眼特有的视觉深度感知功能,人能通过双眼来获取物体的深度信息(depth cues),所谓的深度信息包含以下三个方面:

1) 双目视差(binocular parallax),也称为立体视差(stereoscopic vision),是指由于正常的瞳孔距离和注视角度不同,造成左右眼视网膜上物象存在一定程度的水平差异。在观察立体物体时,因为两只眼相距约 60mm,物体离眼睛越近,双眼观察到的同一物体的图像相差就越大,而物体离眼睛足够远时,这种差异则几乎可以忽略不计,因此大脑视觉中枢会根据双眼看同一物体的差异大小,来判断物体的远近程度,即估测物体的深度信息。因此这种方式需要双眼的协同参与。

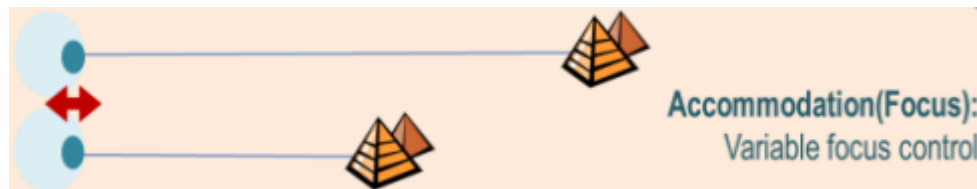


2) 移动视差(motion parallax),当观察视点改变,即观察人产生位移之后,所观察的物体也会相应的在人眼中产生一定的位移,远近不同的物体在人眼中产生的位移会不同,远处的物体产生的位移更小,近处物体产生的位移会更大。大脑视觉中枢可以通过位移视差的相对大小,来估测物体的相对远近程度。这种视差估测方式无论是双眼还是单眼,都可以完成。



3) 聚焦模糊(focus-blur),将人眼比作照相机,那么人眼的睫状肌就起到

收放镜头调整镜头焦距的作用，当睫状肌紧绷收缩时，人眼聚焦在近处平面，此时近处物体清晰，远处物体模糊；当睫状肌放松舒张时，人眼聚焦在远处平面，此时远处物体清晰，近处物体模糊。大脑视觉中枢可以结合不同物体的模糊程度以及睫状肌的收缩程度来判断物体的相对大小，这种深度信息的获取方式通过单眼即可完成。



2.2 产生眩晕的原因

1) 人体与硬件设备的交互匹配不同步

可以分为两类：

i) 身已动而画面未动

当硬件设备获取用户姿态和位移信息之后，若存在延迟或反应不灵敏，即画面的变换和过渡跟不上人体姿态和位移的变化，那么大脑在处理视觉信息和肢体运动信息时就会产生冲突，从而在一定程度上导致眩晕不适。

ii) 画面已动而身未动

当硬件设备对于用户姿态和位移信息的反应过于灵敏时，一个微小的抖动就会导致画面产生过量的变换，而肢体实际上并没有产生对应幅度的运动，从而会在大脑中产生肢体运动信息和视觉信息的冲突，画面变化越快，这种冲突越明显。

以上两种眩晕都是由视觉信息与肢体运动信息不匹配造成的，统称为晕动症。影响晕动症产生的技术因素主要有：

i) 空间位置定位和姿态角度定位的精度和速度。

ii) 显示器件的刷新频率。

iii) 图像渲染的时延。

VR 设备图像变换过渡的灵敏度取决于以上三个指标的最低值，而当前以上三个指标均能达到 90Hz 以上了，这远高于人眼视觉暂留的刷新阈值(24Hz)，但当前的设备还是会使用户产生明显的眩晕感，其根本原因当属如下：

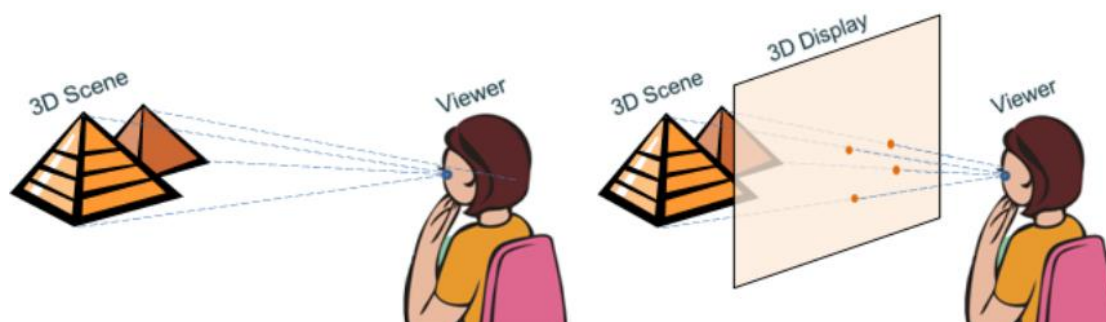
2) 聚焦与视差冲突

对于上文提到的双目视差、移动视差、聚焦模糊这三种典型的深度信息，目前的设备对于前两种都有了较为成熟的足以以假乱真的模拟方法，而对于第三种，即聚焦模糊，其难度和复杂度远在前两者之上，尚处于实验室阶段。现阶段大部分的虚拟现实头显设备，都只提供了单一景深的图片，并且图片的景深固定，即图像不会因为人眼聚焦的远近变化而改变图像的模糊程度，而设备播放的画面是存在远近差别的，这能很轻易的通过双目视差和移动视差来判断得出，也就是说，用户根据双目视差、移动视差得出来的深度信息，与聚焦模糊得出来的深度信息是相冲突的，这种冲突学术上称为“聚焦与视差冲突”（accommodation-convergence conflict, ACC）。当大脑检测到 ACC 时，会迫使睫状肌调节到新的屈张程度使之与双目视差所提供的深度信息相匹配，但这样一来，因为聚焦错乱，图像会变得模糊，而人眼会本能地趋向于使所观察的物体尽量清晰，因此大脑又会迫使睫状肌改变屈张程度恢复到之前的水平，如此循环往复周而复始，大脑就会不堪重负，如同 CPU 超负荷被“烧掉”，进而就会使人体产生眩晕等紊乱不适的反应。除了眩晕之外，长期使用头显设备，还会使睫状肌弹性下降，失去自主调节的能力，诱发近视，因此它对眼睛的损耗程度甚至比一般的电子产品更严重，因此很难在低龄未成年人群体中得到普及，而所谓的广泛应用到中小学教育领域，也会因为这个问题不能得到有效解决而成为泡影。

三、解决方案

3.1 理想模型

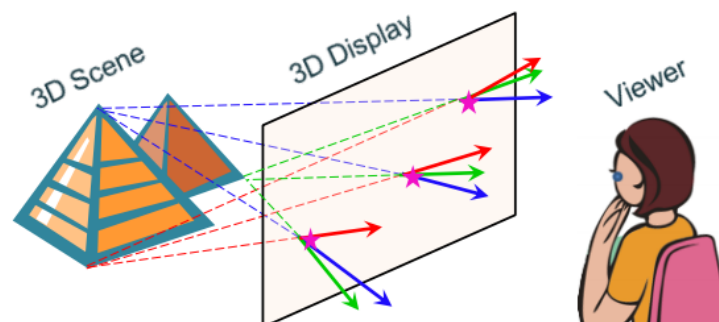
如果存在完美的解决方案，那么其效果应该如同下面这张图片所展示的：



即观察者通过三维显示器所看到的场景应该和透过一层窗户所看到的真实场景完全一致。

3.2 动态光场

我们不妨将头显设备的屏幕看作是这样一个“窗户”，我们看到的都是设备通过计算产生的高度模仿真实场景的投在这个屏幕上的画面，我们不妨将这个屏幕看做一个 $H \times W$ 的矩形，矩形上每个点都有一个坐标 (x, y) ，每个点对应一个等效点光源，每个等效点光源会朝屏幕所正对方向的半球范围内的任何一个方向发出一束光线，这个方向可以用水平和竖直方向上的两个夹角 (θ, ρ) 来表示。



如图，因为每一个等效点光源存是场景中多个部分发射出的光线叠加得到的，因此等效点光源在每一个方向上发出的光的颜色是不一样的，也就是光的波长 λ ，是关于 (x, y, θ, ρ) 的函数，而如果场景是随时间动态变化的，则该函数就有五个参数，即 $\lambda = F(x, y, \theta, \rho, t)$ ，这些光线的集合，称为光场。

而值得注意的是，头显设备的屏幕实际相对佩戴者来说是静止的，并且屏幕上一个 (x, y) 点像素只能发出一种颜色的光，因此这里的 θ 和 ρ 所代表的角度，是“虚”角度，它实际并不存在，只是在计算机中用来模拟仿真人眼相对于虚拟物体的姿态和位置，因此其值会根据人的位移和姿态而动态改变，故这种光场称为动态光场。

3.3 眼动追踪

要解决眩晕问题，关键还是要能够捕捉到眼球的运动和聚焦，从而让画面同步产生聚焦和模糊，从而让用户感受到景深，使其睫状肌收缩程度处于与双目视差和位移视差相匹配的稳定状态，因此当前可以认为眼动追踪技术是解决虚拟现实头显设备眩晕的关键突破之处。Oculus 创始人帕尔默·拉奇曾称眼动追踪技术为“VR 的心脏”，因为它对于人眼位置的检测，能够为当前所处视角提供最佳的 3D 效果，使 VR 头显呈现出的图像更自然，延迟更小。同时，由于眼动追踪技术可以获知人眼的真实注视点，从而得到虚拟物体上视点位置的景深。

在调研过程中了解到，随着机器学习的广泛应用，其实当前眼动追踪已经因

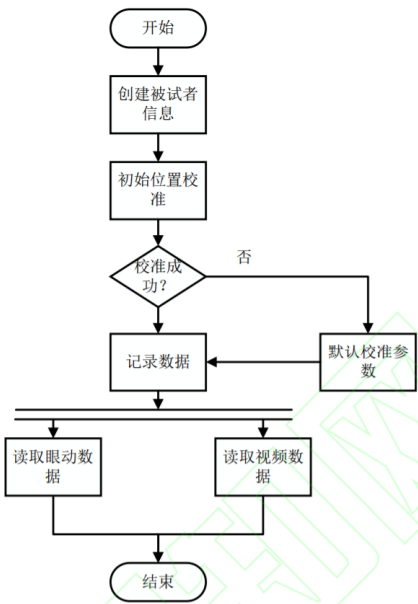
为利用了机器学习技术而有了一定程度的发展，并且已经有一些公司研发并投入商用了一些产品，比较典型的是 Tobii 公司，该公司目前专门针对眼动追踪技术已经生产了大量的硬件与软件解决方案，并且目前已经广泛应用到了心理学与神经科学研究、婴幼儿发展研究、市场与用户调研、工业领域与人的效能等诸多领域。

该公司生产的 Tobii Pro Glasses 2，它是一款具备无线实时观察能力的可穿戴式眼动仪，是市面上重量最轻(仅 45 克)、最自然的可穿戴式眼动追踪系统，可获得最自然的视觉行为数据，可实时呈现被试者在任何真实场景中自由移动时的视线位置，眼动仪采样率 50Hz 或 100Hz。



该眼动仪包括穿戴数据采集单元和数据记录单元。采集单元外形类似眼镜，通过其镜框一周的红外摄像头获得瞳孔不同角度的图片，并通过缆线传至记录单元；记录单元通过对图片的分析解算得到眼睛的注视点，并通过无线传输到计算机。此外，该采集单元还可获得人眼视角的场景视频流。

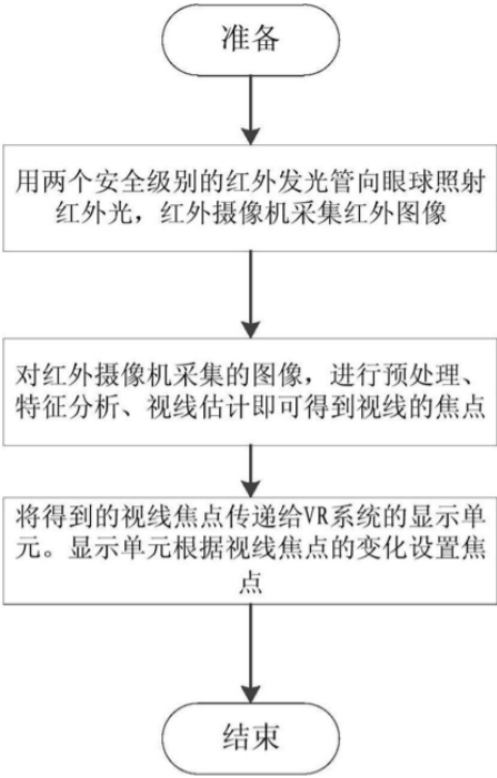
该眼动仪数据采集软件的流程如下：



可以根据由眼动仪获取到的人眼的注视方向，选择性的模糊掉人眼并不关注的像素块，从而造成一种人眼可以主动选择聚焦的效果，从而有效避免眩晕的产生。

将上述眼动仪采集眼动数据与头戴显示设备显示场景相结合，就得到如下能有效避免眩晕的方法：

其主要的实现思路是，向虚拟现实显示输出设备的用户佩戴区域发射红外光，采集红外图像，红外图像包括用户佩戴区域反射红外光所形成的光斑，对所采集的红外图像中的光斑进行视线估计确定用户视线焦点，对用户视线焦点进行跟踪。红外图像由红外摄像机采集，红外摄像机设置在显示屏前端且正对用户佩戴区域中用户眼球的角膜曲率中心的预估位置，红外光由红外发光管发射，红外发光管包括设置在虚拟现实显示输出设备的显示屏两端的两个显示屏红外发光管以及设置在所述红外摄像机上的摄像机红外发光管。



3.4 动态光场与眼球焦点追踪结合

以用户启动设备的位置为原点 0，则虚拟场景中每一个部位都有一个三维空间的坐标，从而可以计算出场景中每一个部位相对于用户的距离，同时眼球追踪器会计算出眼球当前注释的方向以及焦点的位置，通过焦点的位置，将其对应景

深的物体清晰度取高，选择性的模糊掉人眼并不关注的像素块，理论上这就能形成人眼透过一层窗户看真实景象的效果，不过这种方式对设备的实时性要求极高，因为人眼球的运动速度都是毫秒级，也就是需要设备在毫秒之内同时获取并校准好人的位置、方向角度以及眼睛的注视方向和角度。当前这样高实时性的设备还没有普及，成本也相对较高，可能还是需要将降低成本的重任交给计算性能的提升，这或许是我们计算机专业学生将来需要考虑的事情了。

四、结论

本文通过对当前虚拟现实技术发展现状的调研，阐述了当前虚拟现实技术所面临的主要的三个瓶颈，即大范围多目标的精确实时定位和数据实时共享、感知的延伸、眩晕感和人眼疲劳。

通过对三者迫切程度以及用户对其的需求程度的对比，得出当前阻碍虚拟现实大众化普及化所面临的最大的阻碍以及最亟待解决的问题，是用户佩戴虚拟现实头显设备所产生的眩晕问题，接着通过对人眼成像的机理进行分析，并对比人眼看自然景物和看虚拟成像的景物的不同，得出人眼感知景物深度的方式有双目视差、移动视差、聚焦模糊三种，而双目视差和移动视差都已经得到了很好的解决，但眩晕感仍然存在，因此遂将调研的重点放在如何模拟聚焦模糊上来。

关于聚焦模糊问题，当前在理论层面也已经有了一些突破，本文就阐述了其中比较可行的方案，即将动态光场与眼球焦点追踪相结合，这样在理论上能够完全模仿真实场景的呈现，但这种实现方式当前大致也还只在实验室阶段，其成本依旧很高，但比较乐观的是，其中的眼动追踪技术已经有一些科技公司开发出了一些可以给用户使用的产品，精确度也得到了一定的保证，但将其运用到虚拟现实的比较成熟的产品还不是很多，成本是其一，实时性的要求也相对比较苛刻，因此还需要计算速度的提升，但相比这也只是时间的问题，相信不久之后，高逼真度并且无眩晕感的虚拟现实设备会得到大众化普及化。

参考文献

- [1] 曹焯. 虚拟现实的技术瓶颈[J]. 科技导报, 2016, 34(15):94-103.
- [2] Zhang Z , Geng Z , Zhang M , et al. An interactive multiview 3D display

system[J]. Proceedings of Spie the International Society for Optical Engineering, 2013.

- [3] 张卿, 王兴坚, 苗忆南, 王少萍, Alexander I. Gavrilov. 基于眼动、位姿及场景的人体运动方向预测方法 [J/OL]. 北京航空航天大学学报 :1-11 [2021-01-18]. <https://doi-org-443.e1.buaa.edu.cn/10.13700/j.bh.1001-5965.2020.0316>.
- [4] 张超. 防眩晕方法及虚拟现实显示输出设备 [P]. 天津: CN105892634A, 2016-08-24.
- [5] 耿雨晴. 自适应无眩晕虚拟现实眼镜的光学设计与仿真 [D]. 华南理工大学, 2019.
- [6] Hwang A D , Peli E . Instability of the perceived world while watching 3D stereoscopic imagery: A likely source of motion sickness symptoms[J]. i-Perception, 2014, 5(6).
- [7] Bando T , Iijima A , Yano S . Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review[J]. Displays, 2012, 33(2).