

清华大学

# 综合论文训练

题目：基于增强现实的术前规划协作交互系统

系 别：生物医学工程系

专 业：生物医学工程

姓 名：杨煜行

指导教师：王广志 教授

2023 年 6 月 20 日



## 关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内 容，可以采用影印、缩印或其他复制手段保存该论文。

(涉密的学位论文在解密后应遵守此规定)

签 名： \_\_\_\_\_ 导师签名： \_\_\_\_\_ 日 期： \_\_\_\_\_



## 中文摘要

术前规划可以使医师在外科手术前了解病灶及其周围组织的信息，制定合适的手术方案，提高手术效率和安全性。近年来 3D Slicer 等计算机辅助医学图像分析软件可以让用户更清晰直观地观察重建的三维影像，但交互受到二维界面限制。

增强现实技术的出现实现了在三维空间中操纵三维数据，进一步提高观察和操作的效率。HoloLens 2 等头戴式显示器作为增强现实技术的实现载体，相比于其他载体极大提升了使用者的感知能力和交互主动性。近二十年来增强现实术前规划领域得到了诸多有益的进展。但在交互工具精度和交互系统效率等方面仍然存在可以提升的空间。

本文以提升术前规划的效率和精度为研究目标，以开放式脑胶质瘤切除手术为应用场景，提出并实现了：

1. 基于红外被动标记的精准多面体交互工具。利用 HoloLens 2 的深度传感器识别红外反光标记物的特性制作了十二面体追踪标记物，实现在使用场景下精准定位工具尖端位置。自主设计了电路结构，丰富了工具和三维医学图像数据的交互方式。该工具为系统提供了精准的交互作用点。
2. 基于多模态术前影像的规划协作交互方法。创新性地结合了二维医学图像软件和三维增强现实系统，提出了结合两者优势进行术前规划协作的完整流程，有望提升术前规划效率。
3. 增强现实术前规划协作系统构建。以 Unity 3D 作为图形渲染器，3D Slicer 作为医学影像数据处理器，HoloLens 2 作为虚拟图像显示器，构建了一套高效、实时的术前规划协作系统。

**关键词：**增强现实；术前规划；人机交互

## ABSTRACT

Preoperative planning enables surgeons to obtain the information of the targeted area and its surrounding tissue before surgery, make a precise surgical plan, and improve the efficiency of surgery. Computer-aided medical image analysis softwares such as 3D Slicer allow users to observe reconstructed 3D images more clearly and intuitively. However, 2D interface limits the interaction, which leads to low interaction efficiency.

The emergence of augmented reality technology realizes manipulating 3D data in 3D space, so as to further improve the efficiency of observation and operation. As one of display devices of augmented reality, head-mounted displays such as HoloLens 2 can greatly improve users' perception ability and interactive initiative.

Aiming at improving the efficiency and accuracy of preoperative planning and taking the glioma resection as one of application cases, we proposed and achieved high precision multi-face interaction tool based on infrared passive fiducial markers, collaborative and interactive methods of preoperative planning based on multi-modal preoperative images and a collaborative preoperative planning system in augmented reality.

**Keywords:** Augmented Reality; Preoperative Planning; Human Computer Interaction



# 目 录

第 1 章 绪 论 .....	1
1.1 脑胶质瘤手术 .....	1
1.1.1 脑胶质瘤疾病背景 .....	1
1.1.2 脑胶质瘤手术术前规划背景 .....	1
1.1.3 术前影像学评估 .....	2
1.1.4 切口与入路设计 .....	3
1.1.5 小结 .....	3
1.2 增强现实在术前规划中的应用 .....	3
1.2.1 数字重建三维图像 .....	3
1.2.2 增强现实简介 .....	4
1.2.3 增强现实的术前规划应用 .....	5
1.2.4 小结 .....	7
1.3 研究目的 .....	7
1.4 论文结构 .....	8
第 2 章 基于红外被动标记的精准多面体交互工具 .....	10
2.1 引言 .....	10
2.2 工具设计 .....	12
2.2.1 概述 .....	12
2.2.2 尾部标记物设计 .....	12
2.2.3 控制电路设计 .....	16
2.2.4 外壳设计 .....	18
2.3 算法设计与验证 .....	18
2.3.1 概述 .....	18
2.3.2 数据准备 .....	19
2.3.3 标记点检测与特征提取 .....	20
2.3.4 定义算法设计 .....	23

2.3.5 定义算法评估 .....	24
2.3.6 工具尖端中心标定 .....	25
2.3.7 定位算法设计 .....	26
2.3.8 定位算法评估 .....	31
2.3.9 尖端误差评估 .....	31
2.3.10 尖端位姿优化和轨迹误差评估 .....	34
2.4 总结与讨论 .....	35
<b>第 3 章 基于多模态术前影像的规划协作交互方法 .....</b>	<b>37</b>
3.1 引言 .....	37
3.2 交互方式概述 .....	38
3.2.1 交互分类 .....	38
3.2.2 交互平台与载体 .....	39
3.3 协作方式概述 .....	39
3.4 获取和处理术前多模态影像 .....	40
3.4.2 MRI 数据处理 .....	40
3.4.3 DTI 数据处理 .....	41
3.4.4 CT 数据处理 .....	41
3.4.5 MRA 数据处理 .....	41
3.4.6 fMRI 数据处理 .....	41
3.5 观察影像建立感知 .....	41
3.5.2 多模态影像渲染技术与渲染结构 .....	43
3.5.3 多模态影像渲染模式 .....	48
3.5.4 多对象渲染与半透明渲染 .....	48
3.5.5 小结 .....	51
3.5.6 截断盒和截断面 .....	52
3.5.7 感兴趣区域盒 .....	53
3.5.8 切片平面 .....	54
3.5.9 绘制与擦除 .....	55
3.5.10 自由标记 .....	56
3.5.11 距离与角度测量 .....	57

3.6 入路和切口规划 .....	58
3.7 风险评估 .....	59
3.8 影像处理软件与增强现实端的协作 .....	62
3.9 总结与讨论 .....	63
<b>第 4 章 增强现实术前规划系统构建 .....</b>	<b>65</b>
4.1 引言 .....	65
4.2 系统框架 .....	65
4.3 交互工具 .....	66
4.3.1 按键状态传输 .....	66
4.3.2 按键功能分配 .....	66
4.3.3 虚像工具的必要性 .....	67
4.4 增强现实显示器 .....	67
4.4.1 位姿数据的获取与传输 .....	68
4.4.2 远程渲染 .....	68
4.5 影像数据处理器 .....	68
4.5.1 增强现实显示器指令通道 .....	68
4.5.2 数据与指令通道 .....	68
4.5.3 用户界面 .....	69
4.6 渲染服务器 .....	70
4.6.1 按键状态接收 .....	70
4.6.2 用户界面 .....	70
4.7 系统验证 .....	71
4.7.1 交互工具精度验证 .....	71
4.8 总结与讨论 .....	74
<b>第 5 章 总结与展望 .....</b>	<b>75</b>
5.1 论文成果与创新点 .....	75
5.2 论文未来工作展望 .....	76
<b>插图索引 .....</b>	<b>77</b>
<b>表格索引 .....</b>	<b>80</b>
<b>参考文献 .....</b>	<b>81</b>

致 谢 .....	85
声 明 .....	86
附录 A 外文资料的书面翻译 .....	83
附录 B 其他内容 .....	96
在学期期间参加课题的研究成果 .....	99

## 主要符号表

ANOVA	方差分析 (Analysis of Variance)
AR	增强现实 (Augmented Reality)
CNC	精准数控机床 (Computer Numerical Control)
COV	协方差 (Covariance)
CPU	中央处理器 (Central Processing Unit)
CT	电子计算机断层扫描 (Computed Tomography)
DTI	弥散张量成像 (Diffusion Tensor Imaging)
DVR	直接体渲染 (Direct Volume Rendering)
fMRI	功能磁共振成像 (functional Magnetic Resonance Imaging)
GPIO	通用型输入输出 (General-purpose Input/Output)
GPU	图形处理器 (Graphics Processing Unit)
HMD	头戴式设备 (Head-Mounted Device)
ICP	迭代最近点算法 (Iterative Closest Point)
IP	网际互连协议 (Internet Protocol)
IVR	间接体渲染 (Indirect Volume Rendering)
MIP	最大密度投影 (Maximum Intensity Projection)
MRA	磁共振血管造影 (Magnetic Resonance Angiography)
MRI	磁共振成像 (Magnetic Resonance Imaging)
MSE	均方差 (Mean Square Error)
PCB	印刷电路板 (Printed Circuit Board)
RMSE	均方根误差 (Root Mean Square Error)
SNR	信噪比 (Signal-to-Noise Ratio)
TCP	工具中心点 (Tool Center Point)
TCP	传输控制协议 (Transmission Control Protocol)
VR	虚拟现实 (Virtual Reality)
2D	二维 (Two Dimension)
3D	三维 (Three Dimension)





# 第 1 章 绪 论

## 1.1 脑胶质瘤手术

### 1.1.1 脑胶质瘤疾病背景

脑胶质瘤是一种常见的原发性中枢神经系统肿瘤<sup>[1]</sup>。根据《中国恶性肿瘤学科发展报告（2021）》描述，在中国神经胶质瘤的发病率为 3-6.4/10 万，在中枢神经系统肿瘤中约占 23.3%，占脑部恶性肿瘤的比例高达 78.3%<sup>[2][3][4]</sup>。

世界卫生组织在《中枢神经系统肿瘤分类》的蓝皮书中将其分为四个级别<sup>[5]</sup>，其中 I 级和 II 级，属于低级别的胶质瘤，术后三年存活率在 55-87% 间；而 III 级和 IV 级，属于偏恶性的高级别肿瘤，术后生存率仅为 11-22%<sup>[6]</sup>。目前临床实践结果表明，脑神经胶质瘤五年病死率仅次于胰腺癌和肺癌<sup>[7]</sup>。

多个国际公认的治疗规范指南均指出，脑胶质瘤应以手术切除为主，再按需结合放疗、化疗等方法<sup>[8]</sup>。手术切除是目前应用最广泛的脑胶质瘤治疗解决方案，也是最直接、有效的解除由颅内肿瘤引起的物理占位效应的方式<sup>[9]</sup>。故本文探讨的术前规划流程基于脑胶质瘤切除手术。

由于脑胶质瘤生长区域和形态的特殊性，在术前常常需要获得精确的关于肿瘤位置及其与毗邻的组织的信息，从而确定最佳手术策略，预测完全切除的可行性，并确定术中所需的额外工具以最大限度地切除肿瘤<sup>[10]</sup>。

### 1.1.2 脑胶质瘤手术术前规划背景

由于脑胶质瘤常常生长于脑功能区，在形态上缺乏清晰的边缘，脑胶质瘤手术的一大难点就在于需要在完全切除肿瘤组织和尽量减少脑功能丧失之间进行权衡<sup>[6]</sup>。最大限度切除肿瘤组织能有效降低复发率和病死率，同时如果手术对脑功能区造成了损害，可能会导致失语、偏瘫，甚至情感和认知功能的损失<sup>[11]</sup>。因此在术前规划时如何辅助外科医生对解剖结构和附近脑区功能进行充分了解并制定手术方案成为一个重要议题。

脑胶质瘤切除手术术前规划步骤主要包括术前评估和手术准备两部分，术前评估包含影像学评估、神经功能评估和术前癫痫评估三个步骤，而手术准备需要考量切口与路径设计、麻醉方式等细节<sup>[7]</sup>。本文构建的系统主要旨在改善和提高影像学评估和切口与入路设计这两个步骤的效率与精度。

### 1.1.3 术前影像学评估

在术前影像学评估中，外科医师常常借助各种模态的医学图像对病灶位置进行观察。在脑胶质瘤手术前常用的反映结构的影像模态是磁共振成像 (Magnetic Resonance Imaging, MRI)；反映病灶附近功能特性的成像模态有功能磁共振成像 (functional Magnetic Resonance Imaging, fMRI)、和弥散张量成像 (Diffusion Tensor Imaging, DTI) [11]。最主要使用的是能提供肿瘤位置和大小信息的 MRI，其他成像模态在不同方面对脑胶质瘤评估和手术规划也有着重要意义。

磁共振扫描获取的 T1 加权成像 (T1-weighted imaging, T1WI)、T2 加权成像 (T2-weighted imaging, T2WI)、和液体衰减反转恢复序列 (Fluid Attenuated Inversion Recovery, FLAIR) 成像能够提供肿瘤病变范围、水肿和恶性程度的信息。另一种基于 MRI 用于血管成像的技术叫磁共振血管造影 (Magnetic Resonance Angiography, MRA)，用于对动脉和部分静脉进行现象，以评估它们是否存在异常。在脑胶质瘤切除手术中，大血管所在位置也是术前规划的重要考量之一。MRA 能提供有效的血管位置和走向的信息，所以也可以被用于脑胶质瘤手术术前规划中。MRI 是初步了解脑肿瘤解剖学表征的金标准，软组织细节和在高级别肿瘤病症中常见的被破坏的血脑屏障区域在 MRI 中得到突出 [11]。

虽然 MRI 可以提供肿瘤位置和其解剖学特征，但它们无法提供皮质区域位置的功能信息，而以下几种模态有助于确定功能特性，从而明确和希望手术后保留的如语言和运动等关键脑功能对应的脑区。

fMRI 是一种非侵入性评估人类大脑活动的影像工具，通过检测神经活动时不同区域的血流量识别皮质激活区域，血流的相对差异可以映射涉及运动和语言功能的皮质区域。fMRI 识别语言区域的灵敏度和特异性分别在 37–91% 和 64–83% 之间，而运动区域的识别灵敏度和特异性分别为 95–100% [12][13]。fMRI 可以帮助手术前更好地规划手术路径，避免损害功能区。

DTI 是一种基于 T2 的 MRI 技术，通过检测水分子中氢原子沿轴突运动的方向还原白质纤维束结构。通过在手术中避免破坏 DTI 所显示的白质纤维束有可能降低对神经功能的损害 [14]。

综上所述，MRI 用于明确肿瘤的结构信息，MRA 模态用于显示血管，DTI 用于检测纤维束信息，fMRI 用于关联病灶附近的功能信息。本文构建的系统旨在从 T1、MRA、fMRI 和 DTI 几种模态的影像中提取重要的结构或功能信息，有机结合显示在三维空间中，在脑胶质瘤切除手术术前规划任务中向观察者传达更多的信息。

#### 1.1.4 切口与入路设计

在术前规划中外科医师需要综合考量病灶和功能区的位罝设计切口和路径。需要考虑的因素有 (1) 暴露病灶和附近的功能区, 方便术中检测 (2) 考虑病情复发二次手术的可能 (3) 考虑功能区分布的个体差异性 (4) 考虑血管、硬脑膜静脉窦、颅神经等重要的组织和结构<sup>[7]</sup>。

在切除策略和入路设计方面, 应遵循保留重要功能结构的前提下, 先切除非功能区肿瘤, 再逐步切除功能区附近肿瘤的原则<sup>[7]</sup>。

#### 1.1.5 小结

神经胶质瘤切除手术本身就对其术前规划系统提出了两个需求。一个是精度要求, 由于神经系统的复杂性, 尽可能高的精度可以减少对患者不必要的损害。因此本系统设计的工具定位算法, 影像的三维显示到交互都力求精准。

其次是为了弥补单模态存在的缺陷, 常常需要结合多模态的影像, 从而同时获取功能和结构信息。故在本文提出的交互方式中也综合考量了如何融合多模态影像, 以传达更多的有效信息。

### 1.2 增强现实在术前规划中的应用

#### 1.2.1 数字重建三维图像

近年来出现了很多如 3D Slicer<sup>①</sup>等计算机辅助医学图像分析软件, 通过一定的算法将断层图像重建为三维图像。传统手术中医师通过临床经验结合从二维断层图像中获取的信息对患者病灶进行定位和分析, 这种方法可能会引起定位偏差, 存在着颅骨开窗过大、无意破坏其他重要组织的风险<sup>[15]</sup>。而三维的可视化相比起传统断层图像更符合视觉直觉, 使得使用者能更清晰直观地观察病灶。另外一些商业化软件还引入如标记、分割、路径规划等术前规划功能, 辅助医师进行更安全的手术规划。

但在二维界面进行三维操作并不符合直觉。如鼠标和键盘等普通输入设备的操作能力与复杂三维数据集的部分体积操作和可视化任务不匹配, 可能会带来较大的误差<sup>[16]</sup>。因此探索在三维空间中对重建后的医学影像的交互方式成为一个新的研究方向。增强现实 (Augmented Reality, AR) 就是能够实现三维观察、操作的技术之一。增强现实技术能提供在三维空间中对三维对象进行三维交互的体验。

---

<sup>①</sup> <https://www.slicer.org/>

但并不是所有的三维交互方式都优于二维交互方式，二维交互的一大特点在于把复杂的信息简化使得人更容易操作，同时这样的简化也带来了信息量的损失。而三维交互的优势在于提供观察者三维空间信息，观察者得以综合三维因素进行交互。所以如何平衡两者也成为了构建交互系统的重要考量之一，这也是本文在下文会着重讨论的话题。

### 1.2.2 增强现实简介

增强现实是一种结合现实世界和计算机生成内容，使得人们可以在现实中与虚拟物体互动的交互技术<sup>[17]</sup>。内容可以跨越多种感官方式，包括视觉、听觉、触觉、体感和嗅觉，其中视觉为最受关注的领域。

在 AR 技术出现以前，人们通过计算机显示器构建并观察虚拟的三维模型。然而这是一种局限于二维空间的展示技术，人们只能观察显示器上的画面并通过鼠标、键盘与之交互。这种交互方式的缺陷是常常让人觉得没有代入感，对于感官习惯于三维世界的人类而言，这种方法缺乏足够完善感知的具身体验。具身认知 (Embodied Cognition) 理论认为身体的感知觉和运动强烈地影响着人的认知<sup>[18]</sup>。

具体来说，在现实世界中人类观察一个物体可以通过移动自身位置改变观察角度，此时人自身的肌肉运动和视觉观察都能对物体与自身的相对位置的计算产生贡献；或者在对物体本身进行旋转平移的过程中，通过手部的触觉和肌肉运动形成相对位置的认知能帮助人在思维中进一步构建物体的形状。而二维空间中键鼠的操作方式极大削弱了人的运动主动性，从而降低了人对物体的认知能力。

而增强现实技术，尤其是以头戴式设备 (Head-Mounted Device, HMD) 为技术载体的增强现实提供了与实际物体的交互几乎无异的和虚拟物体的交互方式，这是符合人类手眼协调直觉的、准确的靠谱及方式。头戴式设备 Microsoft HoloLens 2<sup>②</sup>作为增强现实市场占有率第一的设备，提供了空间映射、六自由度追踪、即时定位与地图构建 (Simultaneous Localization and Mapping, SLAM)、眼球追踪和全息投影等功能，为开发者搭建了一个相对完善的增强现实开发平台与架构，已被广泛应用于广泛应用于教育、培训、医疗、设计、娱乐等领域，并仍处于发展期，具有巨大的市场潜力。本系统将采用 HoloLens 2 作为增强现实的显示载体，从而为系统使用者提供更强的空间体验。

---

<sup>②</sup> <https://www.microsoft.com/zh-cn/hololens/hardware>

### 1.2.3 增强现实的术前规划应用

在传统的术前规划中，外科医生需要在脑海中将二维断层成像图像转化为三维图像，形成一个心理表象。他们还需要在工作记忆中保留之前的情节，并将重构的数据与患者对齐，结合临床经验确定手术入口及手术路径。然而这种空间重构过程对脑力资源的要求很高，工作记忆负担过重可能导致识别过载，降低效率。增强现实系统最显著的作用是减轻了医生的思维负担，另外某种程度上也提升了规划的准确性和有效性。此外 HMD 为载体的增强现实可以将术前影像原位显示，能有效减少医生在导航屏幕和病灶区之间的视线切换，从而解决术中手眼不协调问题<sup>[19]</sup>。此外，还可以将术前规划好的路径以及关键点在虚拟图像上标记，引导医生进行手术过程。

术前规划增强现实系统能让外科医生用三维工具以立体形式操纵解剖特征在虚拟显示器中使用三维工具，以立体形式操作解剖特征。首先在显示方面需要做到能区分和突出颅神经、动脉和不同的皮质结构。另外外科医生可以修改模型的渲染阈值、着色和透明度等参数使其在不同环境下方便观察。实现这些功能的技术路线很多，近二十年来不同的研究者和工程师都在这个领域进行了探索，得到了很多有益的进展。

最早进行商业化的虚拟现实术前规划系统是最先于 2000 年在欧洲推出的 Dextroscope 系统<sup>[16]</sup>，外科医生可以在其中规划神经外科手术和其他外科手术。这是第一个使用虚拟现实原理并且摒弃了鼠标和键盘的商业手术计划系统，Dextroscope 使用一块将双手和相关工具建模在虚拟世界的立体显示器，提供自然舒适的界面以帮助用户长时间处理多模态的医学图像。外科医生通过手势对感兴趣的对象进行旋转等操作，并可以对器官和结构进行虚拟分割、精确测量等（图 1.1 (a)）。此外该系统还配备一个带有开关的手柄和一只手写笔，按下开关时就可以自由移动 3D 图像，并且手柄还可以用作钻具、量具、刀具等，而手写笔可以用于选择体素和其他例如去除骨骼和软组织等精细操作<sup>[21]</sup>。值得注意的是该系统被多国批准为医疗器械，并用于实际的临床应用中。多项研究表明，使用该系统进行术前规划案例中的患者的术后各项正向指标都有所提升<sup>[22][23][24]</sup>。但由于工作空间受限、视野小、没有头部追踪等系统缺陷，用户的互动受到限制。

Hinckley 等人<sup>[25]</sup>提出了一种更直观、空间限制更小的交互方法，即用一只手拿着一个小比例的头部模型，另一只手拿着一块塑料板（见图 1.1 (b)）。塑料板代表的平面可以将显示器上显示的病人体素虚拟地切开。但每当操作者的视角和虚拟切片发生错位时，这种互动会显得不自然。

Abhari 等人<sup>[26]</sup>则通过一副 AR 眼镜，一个外部光学跟踪器和一个自制的带红外反光标记的笔搭建并评估了一套脑肿瘤手术术前规划系统（见图 1.1 (c)）。这种系统相比于上述两种系统实现了虚拟图像和假体（患者）的注册，也就是原位显示。这是常见的由外自内（Outside-in）的 AR 追踪系统，通过计算不同组件的相对位置关系，实现虚拟图像的原位显示。相对的，如果用于定位的光学传感器位于眼镜上，则为自内而外（Inside-out）的定位方式。

另一种允许原位可视化的显示器是 AR 窗口<sup>[27][28]</sup>，它利用放置在病人和外科医生眼睛之间的半透明镜子，将实际场景与自动立体屏幕创建的虚拟图像合并。然而这可能会导致较差的图像质量，限制外科医生的有效观察范围，并且在进行手术时也会阻碍外科医生。

除了头戴式设备和 AR 窗口外，还有使用手机、平板等移动设备<sup>[29][30][31]</sup>作为显示器的案例。同样使用的也是自外而内的定位原理。移动设备可以有效避免观察范围被限制的问题，但同时也存在无法手眼协调的缺陷。

从上述的例子可以看出，AR 术前规划系统的实现设备种类繁多，所用技术也有所差异。显示设备一般有头戴式设备、移动设备、AR 窗口、一般显示屏等，根据其设备差异，所需的定位技术和显示技术也会有所不同。而交互工具也是各式各样，常见的有被动标记工具、手势以及一些商用定位工具。

AR 系统所需要的定位技术在近些年来日益成熟，于是人们逐渐将目光投向了和人更为相关的感知和交互部分。通过什么样的技术和细节能够提升人对虚拟图像的感知能力，什么样的交互方式能让用户感到更符合直觉，使得使用效率提高，这些问题都成为了这个领域最新的研究热点。

Allgaier 等人<sup>[32]</sup>从交互工具的角度出发，进行严谨的用户实验对比使用不同交互工具完成相同任务的精度和效率（见图 1.1 (d)）。该工作以肝脏手术计划和开颅手术培训为实验场景，对 VR 控制器、笔状 VR Ink<sup>®</sup>、手套和带有标记物的开颅钳的性能进行了定量比较，评估了有关任务负荷、晕动症以及设备可用性和适用性的定性数据。

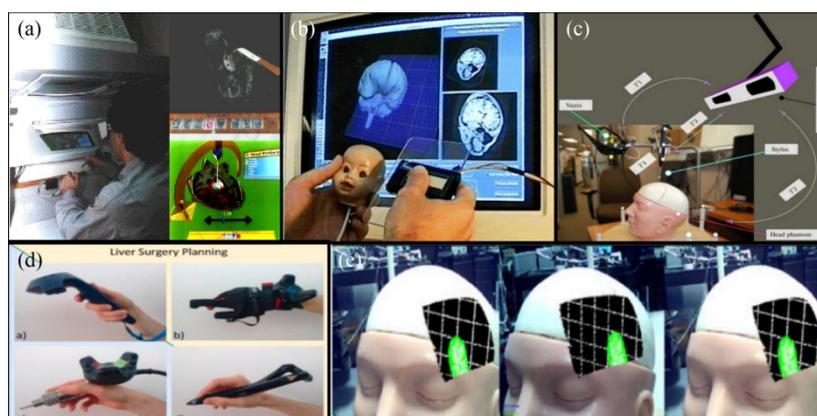
而 Hellum 等人<sup>[33]</sup>则对比了几种差异较大的输入方式，这些输入方式的基本元素是：注视、头部运动、控制器、虚拟键盘和语音。通过用不同方式完成脑肿瘤手术中常用的点标记和注释两个功能，得到了基于控制器的交互比基于眼睛追踪

---

<sup>®</sup> <https://www.logitech.com/en-us/promo/vr-ink.html>

的方法更适合于点的放置，而语音记录和虚拟键盘输入则比自由手写的记事方法更好的结论。

还有工作侧重研究虚拟数据的表现方式带来的感知差异。Abhari 等人<sup>[26]</sup>在其系统中考虑了头戴式设备存在的深度感知的缺陷，并增加了阴影、遮挡和网格线等深度线索辅助感知（见图 1.1 (e)），并用临床实验证明了这些考虑可以有效提高手术的成功率。



(a) 用户正在使用 Dextroscope 系统<sup>[22]</sup>；(b) Hinckley 等人提出的系统<sup>[25]</sup>；(c) Abhari 等人提出的系统<sup>[26]</sup>；(d) Allgaier 等人比较的四种工具<sup>[32]</sup>；(e) Abhari 等人提出的增加深度感知方法的示意图<sup>[26]</sup>

图1.1 用于术前规划的增强现实系统

#### 1.2.4 小结

综上所述，增强现实术前规划系统相较于传统的规划方式，观察和操作的自由度更高，更符合直觉。系统仍然在发展和不断改进中，比如仍然存在交互工具的操作精度还有待提升，交互工具和交互方式的效率还不够高等问题。

### 1.3 研究目的

本工作旨在使用头戴式增强现实设备 HoloLens 2 构建一个能够改善或解决以下四个问题的增强现实脑胶质瘤手术术前规划交互系统：

- (1) 脑胶质瘤切除手术的术前规划对有效性和准确性有着较高要求。
- (2) 为了尽可能地呈现重要区域的信息，需要结合多模态影像信息。
- (3) 目前大多增强现实术前规划交互系统交互功能仍比较基础。

(4) 结合增强现实和二维数字图像重建软件，在不同功能上分配不同的操作权限，使得增强现实设备使用者和软件使用者共同协作，以期带来更好的交互体验和交互精度，并使得使用者获取更多的信息，提高交互效率。

问题(1)、(2)是该手术场景本身提出的要求，而问题(3)、(4)则是相较于前人的工作可以改进和创新的部分。

针对上述的几个场景和领域问题，本工作提出了一个基于红外被动标记的精准多面体交互系统保证系统的精准交互，并未基于多模态术前影像的术前规划协作交互方法提供精确的交互作用点，最终实现了增强现实术前规划系统的构建。三个研究内容如图 1.2 所示，分别在第 2、3、4 章呈现。



图1.2 本工作的三个研究内容和相互关系

## 1.4 论文结构

第 2 章将围绕基于红外被动标记的精准多面体交互工具展开，阐述该工具的硬件、外观设计和与之配套的定位算法。本文利用 HoloLens 2 的红外深度传感器能识别逆向反光标记物的特性制作了十二面体状的追踪标记物，实现在使用场景下精准、高准确度定位工具尖端位置。并自主设计了电路结构，实现了除定位外的应用层功能，丰富了工具和三维医学图像数据的交互方式。

第 3 章阐述基于多模态术前影像的术前规划协作交互方法，创新性地结合了医学图像软件 3D Slicer 和以 HoloLens 2 为载体的三维增强现实系统，提出了充分结合两者优势进行以脑胶质瘤切除手术为例的术前规划的完整流程。实现了从

MRI、DTI、fMRI 等不同模态的影像中提取有效信息并有序显示于增强现实显示器中。

第 4 章为增强现实手术术前规划协作系统构建。以 Unity 3D<sup>④</sup>作为图形渲染器，3D Slicer 作为医学影像数据处理器，HoloLens 2 作为虚拟图像显示器，第 2 章提出的红外被动笔为交互工具，构建了一套高效、实时的术前规划协作系统。详细阐述了模块在系统中的功能与模块之间的连接方式。

第 5 章为对整个文章的总结，总结了本文的成果，提出了本文的创新点和未来可以进行改进的部分。

---

<sup>④</sup> <https://unity.com/cn>

## 第 2 章 基于红外被动标记的精准多面体交互工具

### 2.1 引言

为了与虚拟图像进行交互，往往需要引入一些交互工具和交互方式，但无论是什么形式都需要获取交互作用点与虚拟图像的相对位置。目前常见的人与虚拟图像交互的方式有手势和交互工具。手势交互的优势是不需要额外的工具，但存在的问题是操作指令有限，自由度小，且精度有限。而交互工具又分为两大类：

(1) 商业设备如触觉控制器 Geomagic Touch Haptic Device<sup>⑤</sup>，空中设备 VR Controller、Logitech VR Ink、Wacom VR Pen<sup>⑥</sup>等。尽管商业化工具减轻了很多技术负担，但存在的问题是适配性和可扩展性不高 (2) 自制定位工具。常用被动红外反光标记和如 ArUco、Vuforia<sup>⑦</sup>等特殊图案进行工具标记，再通过工具定义还原工具全貌。这种工具可以自由地根据场景制定策略，但对定位算法的精度和实时性也相应地提出了更高的要求。

不同的定位系统由不同的传感器和不同的定位工具组成，一般可以分为自定位系统和外部定位系统。自定位系统如 Geomagic Touch 等通过自身携带的力传感器和机械结构判定尖端位置。而外部定位系统才是更为常见的，外科领域最常用的外部定位传感器是 NDI<sup>⑧</sup>，在不同的手术场景下使用电磁或红外原理进行定位。同样 RGB 工业相机也可以被用于检测上述提到的图案标记物从而实现定位。以下将列举一些自制工具和对应的定位系统的例子。

最早的和 HoloLens 结合的被动定位工具是由 Kunz 等人<sup>[34]</sup>在 2020 年提出的。该研究利用 HoloLens 自带的近红外传感器检测红外反光标记物，从而实现还原标记物的位姿，最后实现了小于 1mm 的定位误差。

Song 等人<sup>[35]</sup>使用 Vuforia 标记制作了用于牙科手术当中的工具（见图 2.1），捕获标记画面的传感器是 HoloLens 的 PV (Photo and Video) 摄像头中的 RGB 传感器。他们还构建了一套可以直观地用该笔进行测量和标记的软件系统。

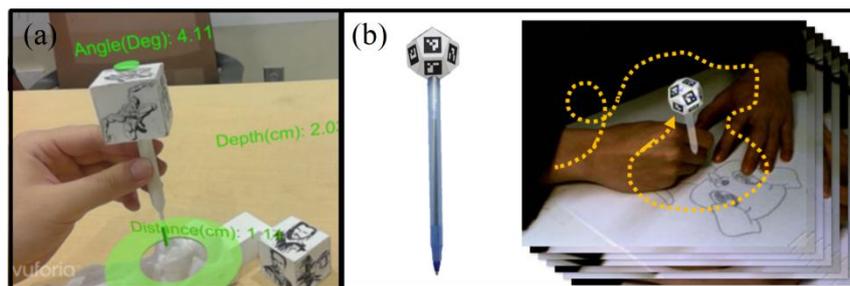
另外一种常见的图案标记是类似于二维码的 ArUco，Wu 等人<sup>[36]</sup>通过把十二面体的每个面都贴上 ArUco 标记制作了一只笔，通过一套 RGB 摄像头系统捕获画面并进行工具定位，其提出的算法达到了 0.1mm 的定位精度。

<sup>⑤</sup> <https://www.3dsystems.com/haptics-devices/touch>

<sup>⑥</sup> <https://developer.wacom.com/en-us/wacomvrpen>

<sup>⑦</sup> <https://developer.vuforia.com/>

<sup>⑧</sup> <https://www.ndigital.com/products/>



(a) Song 等人提出的工具<sup>[35]</sup>; (b) Wu 等人提出的工具<sup>[36]</sup>

图2.1 图案标记实现的定位工具

可以看出对于 HoloLens 2 而言, PV 摄像头的 RGB 传感器和深度摄像头的红外传感器都可以被用于做定位, 两者各有优劣。就分辨率而言, PV 摄像头的默认分辨率为  $1280 \times 720$ , 而深度传感器仅为  $512 \times 512$ , 仅就数值而言如果不考虑其他因素显然 PV 摄像头的高分辨率能带来更高的定位精度。但另外几个因素限制了其性能: (1) 高分辨率导致传输信道压力大, 从而可能会降低整个系统的运行效率和帧率。PV 摄像头获取图像每一帧的大小为  $1280 \times 720 \times 3 \times 8 \text{ bits}$ , 3 代表 RGB 三通道, 8 代表 uint8 数据结构的大小, 而深度摄像头的图像每一帧大小为  $512 \times 512 \times 2 \times 8$ , 2 代表强度图和深度图, 8 代表 uint8 数据结构的大小。可以得出 PV 摄像头的运行压力是深度摄像头的 5 倍 (2) PV 摄像头的默认水平视域 (Horizontal Field of View, H-FOV) 仅为 45 度, 极大地限制了工具的可活动空间。而深度传感器的 H-FOV 约为 120 度 (3) 彩色图像的边缘提取受环境光影响大, 而红外光斑提取的是连通域的中心点, 受边缘噪声影响小。

所以本工作中使用的是 HoloLens 2 的红外深度传感器, 由于该传感器集成于整个头显内, 属于内部设备, 在理想情况下和设备佩戴者具有固定的相对位置关系, 所以相比于 NDI 等外部设备, 具有无需再额外计算传感器与虚拟图像的相对位置的优势。工具的标记物相应地选择了用红外反光材料制作。

红外反光球状标记物是常用的定位标记物之一, 其具有反射面积大, 反射强度大, 易于检测的优点。但由于几何关系特殊, 在某些视角下不同深度的标记物重合导致容易造成光斑粘连现象, 影响定位的准确性。为了解决这个问题, 本工作采用了平面被动反光标记。但平面标记存在检测盲区, 由于使用场景的要求, 工具应该能被全方位定位, 于是本工作进一步对工具形状进行了改进, 使用了多面体设计, 并在每个面上都贴有一定数量的标记物, 保证工具正常视角下的任意位姿都能检测到。

此外为了使定位工具具有一定的交互性，本工作设计了能和服务器进行指令交换的控制主板，扩展并丰富了工具功能。

本章提出了一个使用外科手术中常用的被动红外反光标记作为定位标记物并具有通信功能控制主板的四面体交互工具，并提出了使用 Microsoft HoloLens 2 的深度传感器对该工具进行定义和精准定位的算法，实现了对笔尖的精准定位，以供后续的交互系统使用。

## 2.2 工具设计

### 2.2.1 概述

根据引言对前人工作的总结，我们提取出了两个主要的工具设计需求：（1）能在大部分使用情况下都被检测并识别出位姿（2）具有一定的交互功能。基于此，我们对工具的设计分为了三大部分：尾部标记物、控制电路和外壳。以下将分三部分叙述每部分的设计考量。

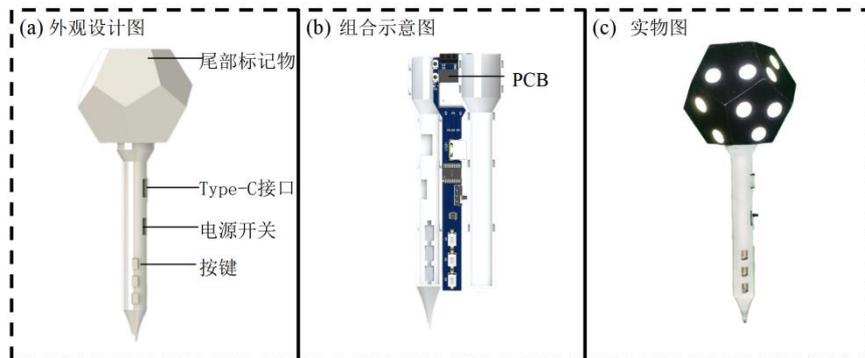


图2.2 交互工具图

### 2.2.2 尾部标记物设计

#### 预实验：反射强度和角度的关系

为了指导工具和算法的设计，我们在预实验中测量了反射强度和角度的关系。实验设置如图 2.3 (a)(b)所示，实验板是一块周围有 7 个直径为 10mm 的红外反光圆，中心为边长 10cm 的红外反光正方形的 3D 打印板。HoloLens 2 的传感器光轴和一个旋转台的轴心保持在同一直线上，两者之间相距约 30cm 至 40cm，模拟正常的交互距离。红外反光实验板中心与旋转台中心重合，并垂直于旋转台平面。实验过程中使转台匀速旋转，HoloLens 2 记录下深度传感器传回的深度图和强度

图。通过计算每一帧的平面法向量与光心之间的夹角，和中心正方形中点像素的反光强度值，得到夹角-反光强度图。

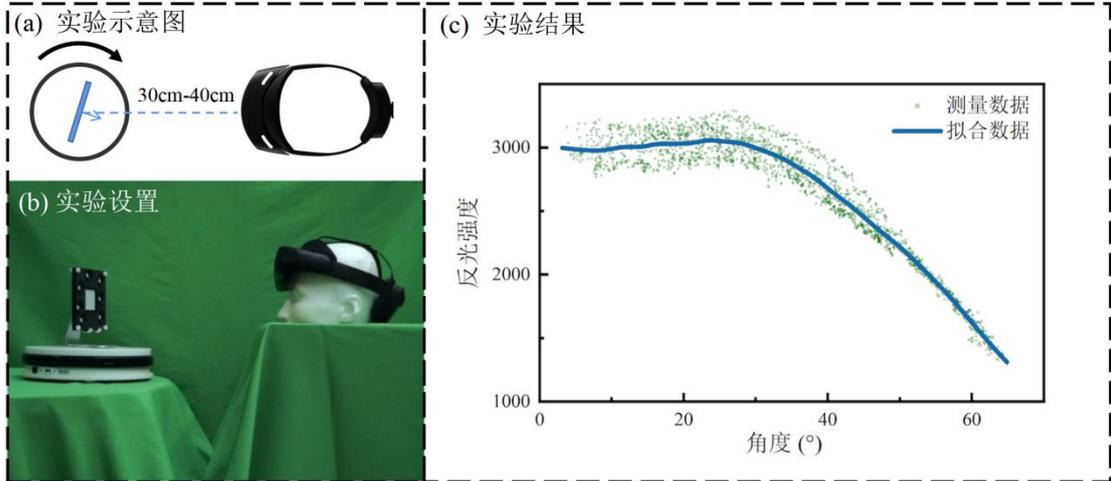


图2.3 反射强度与角度关系预实验设置与结果

首先是法向量的计算，七个圆通过连通域分析提取出七个对应的中心点坐标，记作 $\{P_i\} = \{(x_i, y_i, z_i)\}$ 。

因为七个点在同一平面上，认为每个点的坐标都符合平面方程：

$$Ax + By + Cz + D = 0$$

(2-1)

进一步简化为：

$$z = a_0x + a_1y + a_2$$

(2-2)

为方便陈述平面最小二乘计算法向量的方法，先假设有 $n$ 个点，第 $i$ 个点 $(x_i, y_i, z_i)$ 带入式 (2-1)，得到第 $i$ 个误差项：

$$v_i = a_0x_i + a_1y_i + a_2 - z_i$$

(2-3)

令误差项 $V = [v_1, v_2, \dots, v_n]^T$ ，待拟合参数 $a = [a_0, a_1, a_2]^T$ ， $L = [z_1, z_2, \dots, z_n]$ ，并令：

$$B = \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix}$$

(2-4)

代入式 (2-3), 可得:

$$V = Ba - L$$

(2-5)

欲求  $\min\{V^T V\}$ , 利用拉格朗日乘数法, 对平面参数求偏导:

$$\frac{\partial V^T V}{\partial a} = 2V^T \frac{\partial V}{\partial a} = V^T B = 0$$

(2-6)

将式 (2-5) 代入 (2-6), 得:

$$B^T B a - B^T L = 0$$

(2-7)

得到平面参数:

$$a = (B^T B)^{-1} B^T L$$

(2-8)

利用平面法向量垂直于任意一条平面上直线的定义, 设平面上有两点  $(x_1, y_1, z_1)$ ,  $(x_2, y_2, z_2)$ , 满足平面方程 (2-1), 得到方程组:

$$\begin{cases} Ax_1 + By_1 + Cz_1 + D = 0 \\ Ax_2 + By_2 + Cz_2 + D = 0 \end{cases}$$

(2-9)

两式相减, 得:

$$A(x_1 - x_2) + B(y_1 - y_2) + C(z_1 - z_2) = 0$$

(2-10)

即向量  $(A, B, C)$  垂直于平面中任意两点形成的直线, 所以平面的法向量就是  $(A, B, C)$ 。对于式 (2-3) 而言, 法向量用平面参数表示就是  $(a_1, a_2, -1)$ 。

其次是求实验板中心和光心之间的向量表示, 在相机坐标系下, 该向量即可用中心的坐标  $(x_c, y_c, z_c)$  表示, 故法向量和该向量之间的夹角表示为:

$$\theta = \arccos \left( \frac{(a_1, a_2, -1)^T (x_c, y_c, z_c)}{\|(a_1, a_2, -1)\| \|(x_c, y_c, z_c)\|} \right)$$

(2-11)

实验结果如图 2.3 (c) 所示, 从图中我们得到了红外反光标志物的反射强度会随着夹角的增大降低的结论, 且能够被传感器测量并分离出来的的极限大概在 65

度左右。如果只有一面能被检测到，算法的稳定性会降低，所以希望至少有两面能同时出现，这就决定了工具的面夹角不能太窄。在下一小节中会量化不同正多面体的面夹角，从而指导多面体面数的选择。

### 多面体的面数

如图 2.4 所示，此处对场景进行了理想化建模，假设了两种多面体和相机光心相对位置的极端情况，即光心在面夹线上和光心在面垂线上的两种情况。其中  $\alpha$  和  $\beta$  代表相机光心和面中心连线与面垂线的夹角。而  $\gamma$  代表两条相机光心和面中心连线的夹角， $\theta$  代表面夹角。容易从几何关系得到  $\gamma = \theta + \alpha + \beta - 180^\circ$ ，实际使用中  $\gamma$  取值约为  $5^\circ$  到  $20^\circ$ 。显然我们关心的是  $\alpha$  和  $\beta$  的取值范围。表 2.1 列出了常见的正多面体的面数和不同角的关系。

结论表明正四面体、六面体、八面体都不符合条件，在大部分情况下都只能检测到一面。而十二面体和二十面体都在大部分情况下都能被检测到至少两个面，但考量到体积相同的情况下面数越多，每面的面积越小，可以贴的标志物越少。所以最后选择了十二面体。

表 2.1 正多面体面数的面夹角、在两种情况下的夹角取值

面数	4	6	8	12	20
面夹角 $\theta$	$60^\circ$	$90^\circ$	$109.28^\circ$	$116.56^\circ$	$138.19^\circ$
$\alpha = \beta =$	$62.5^\circ - 70^\circ$	$47.5^\circ - 55^\circ$	$38^\circ - 45.5^\circ$	$34^\circ - 42^\circ$	$23.5^\circ - 31^\circ$
$\beta = 0^\circ \alpha =$	$125^\circ - 140^\circ$	$95^\circ - 110^\circ$	$76^\circ - 91^\circ$	$68^\circ - 83^\circ$	$47^\circ - 62^\circ$

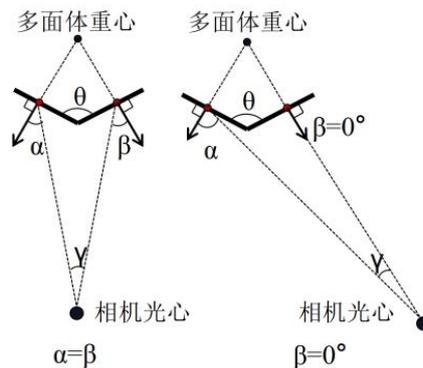


图2.4 场景的理想化建模

## 十二面体尾部的设计图和实物图

经过不同的尝试，最终选用了边长为 23mm，壳厚 3mm，底部通有孔的设计，有效面为十一面，底部的面用于和笔结构连接。在选择边长、壳厚时平衡了能被检测到和维持其合适的重量两个因素。

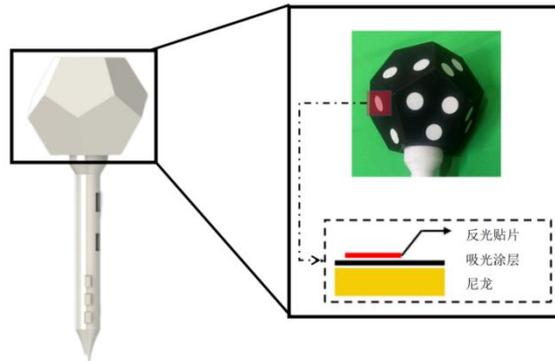


图2.5 尾部标记物表面结构示意图

最终采用尼龙材质打印，为了减少标记附近因反光造成的噪声干扰，在表面喷涂了一层均匀的红外光吸收材料。标志物是从红外反光胶带上裁剪下的 8mm 直径的圆片，间隔不等地分布于每个面。结构示意图如图 2.5 所示。

### 2.2.3 控制电路设计

本章引言提到设计需求之一为希望交互工具拥有一定的交互主动性。在本节中将需求细化为 (1) 能通过无线方式与服务器连接 (2) 可以自主编辑交互逻辑和功能 (3) 能主动向服务器发出信号 (4) 能从服务器接受信号 (5) 能独立供电。基于以上需求，我们采用了自主设计印刷电路板 (Printed Circuit Board, PCB) 的方案，电路原理图如附录 B.1 所示，PCB 预览图如图 2.6 所示。在该方案中，引入了按键作为发送信号的硬件载体，马达作为接受反馈的受体，通过调节震动模式提供触觉反馈，并设计了合理的周围电路实现开关、固件烧录、供电等功能。

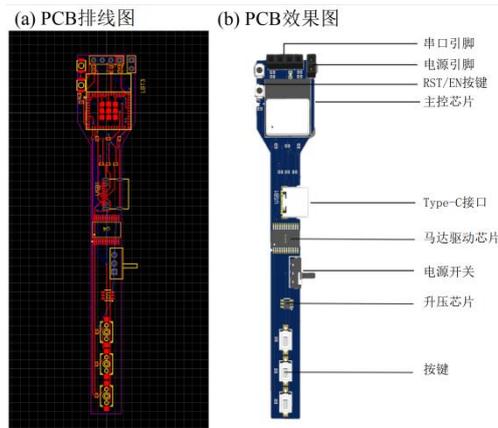


图2.6 PCB示意图

## 硬件选择与功能

主控芯片采用的是 ESP32 C3 MINI 1U，这款芯片支持 2.4GHz Wi-Fi，并且板载天线，无需安装额外的天线，可以通过局域网连接服务器并进行无线通信。以下将按图 2.6 (b)中所示从上到下的顺序介绍不同组件的功能。

最上方是串口引脚，用于烧录固件。在串口烧录时需要将 GPIO2 和 GPIO8 置为低电平，所以在设计时直接将这两个引脚接地。此外还需要使从 GPIO9 引脚置为低电平，但正常运行时 GPIO9 应为高电平，故引出一个按键在烧录时长按实现低电平控制信号的产生。最后串口开始烧录需要使能接口产生下降沿信号，故引出一个按键在烧录时短暂按下实现下降沿信号的产生。

除了串口烧录的方式外，还加入了更为便捷的 Type-C 烧录方式。采用的接口型号是有 16 个引脚的 Type-C 接口，其中最主要的两对引脚是电源对和差分数据对 D-与 D+，对应 GPIO18 和 GPIO19 引脚。再在关键的连线处加入合适大小的限流电阻与去耦电容即可。

芯片引出了 15 个 GPIO 口，在本工作中被用于按键输入和马达控制。GPIO0、GPIO1、GPIO3 用于按键检测，当按键按下时引脚电平被强制置低，按键松开状态为高电平。马达通路上主要有升压芯片 LTC1928ES6-5 和直流电机驱动芯片 TB6612FNG 以及两对马达连接引脚。升压芯片的作用是将电流转换为电压，从而实现将单片机提供的 3.7V 提升到马达驱动芯片需要的到 5V，保证了马达驱动芯片的电压、电流稳定，能维持马达震动在较高幅度。而马达驱动电路采用了三极管推挽结构实现电压转向，可以通过产生 PWM 信号调节占空比从而控制强度。

最后是电池的选择，在考量了体积和容量的取舍后采用了可以通过 Type-C 接口充电的七号锂电池，高度 43.6mm，直径 10.1mm，额定电压为 3.7V，容量为 350mAh。

经过实验测量，在空载情况下，即不驱动马达，但无线、按键等模块在工作情况下总路电流约为 30mA；在满载情况下，即驱动一个或两个马达时电流约为 120mA。这是因为由于升压芯片的稳流作用，马达数量并不显著影响电流输出。两个马达会平分电流，导致相比起一个马达的情况震动强度减半。

简单计算可得，空载续航时间约为 11.5 小时，满载续航时间约为 2.8 小时。即使考虑电荷流失后电压下降的情况，也完全符合应用场景需求。

## 控制逻辑

控制代码由 C 实现，主要实现了以下三大功能：（1）自动接入指定局域网，以工具侧作为服务端（Server），服务器侧做客户端（Client）的模式连接，并采用了 TCP 协议进行双向通信（2）以 3 比特字符编码三个按键的实时状态，并设计了防抖算法，数据打包后通过局域网发送给服务端（3）实时倾听马达控制信号的接受端口，并在马达控制信号到达时使马达震动。

### 2.2.4 外壳设计

本工作将交互工具设计为了符合人抓握习惯的笔状，外观设计图如图 2.2(a) 所示，除笔尖外的整体长度为 10.8cm，笔尖长度为 20mm。尖端设计为半球形，便于 2.3.6 中精确定位尖端中心位置。

按键摆放的位置是根据人抓握时指尖可及处设计的，而侧面两个开口分别为可以充电或烧录程序的 Type-C 接口和电源拨动开关预留。按键采用了卡扣设计，能自由组装和拆卸。

## 2.3 算法设计与验证

### 2.3.1 概述

在上一节中阐述了阐述了工具各个组件的形态及其功能，本节中将聚焦于如何获取尾部标记物的位姿并获得工具尖端的位置。

整个算法流程如图 2.7 所示，可以分为数据准备、标记点检测、工具定义和工具定位四个部分。数据准备部分包括内参标定、深度传感器数据实时获取；标

记点检测步骤从获取的图像中抽象出点的坐标与特征信息；工具定义和标记点检测共同服务于工具定位。

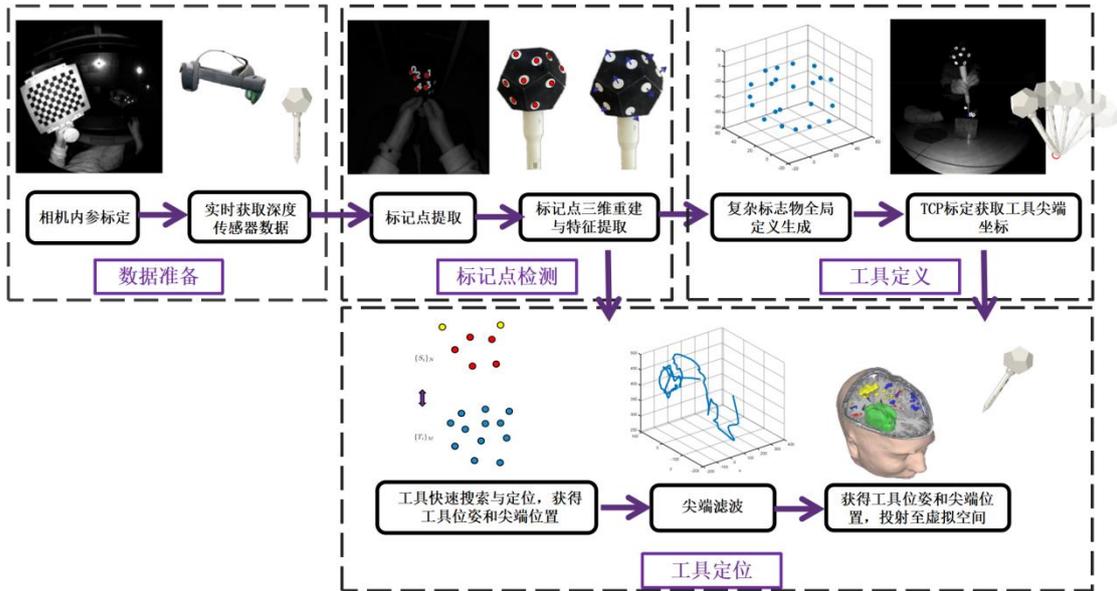


图2.7 算法流程图

### 2.3.2 数据准备

#### 相机内参标定

相机将物理三维世界中的信息投射到二维的图像平面，这个过程中蕴含着如图 2.8 所示的三次坐标系变换。而在这个物理光线透过透镜映射到传感器的过程中会产生由透镜本身的光学特性造成的畸变。标定的目的主要有两个，一是建立数学模型矫正由于镜头畸变造成的图像变形，二是获取内参矩阵后可以根据拍摄获得的二维图像重构三维场景。

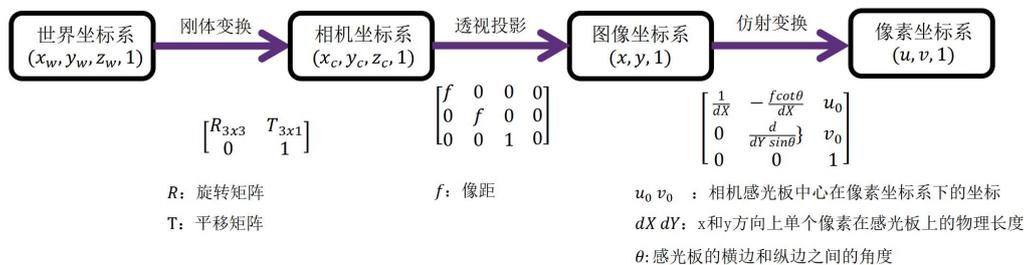


图2.8 成像过程中坐标系的转换关系

我们将内参矩阵定义为从相机坐标系到像素坐标系的转换矩阵:

$$(2-12) \quad A = \begin{bmatrix} \frac{1}{dX} & -\frac{f \cot \theta}{dX} & u_0 & f & 0 & 0 & 0 \\ 0 & \frac{d}{dY \sin \theta} & v_0 & 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

通过大于 3 张的标定板图片可以拟合出内参矩阵, MATLAB 提供了相应的工具箱。

### 深度传感器数据实时获取

HoloLens 2 为开发者提供了开发者模式 (Research Mode), 提供了对关键传感器的访问权限。深度相机使用基于相位的飞行时间 (Phase-based Time-of-flight) 主动红外照明来确定物体的深度。相机可以在两种模式下运行, 第一种模式启用高帧率 (45 fps) 近深度感测, 通常用于手部跟踪, 而另一种模式用于低帧率 (1-5 fps) 远深度感测, 用于空间映射<sup>[37]</sup>。除了深度图, 相机还提供反映红外反射强度的强度图。

### 2.3.3 标记点检测与特征提取

#### 标记点检测

在获得矫正过的深度图和强度图后, 首先要获得强度图中光斑的位置。为了实现这一步, 首先要对强度图进行二值化, 可以表达为:

$$(2-13) \quad \text{Binarize}(f(x, y)) = \begin{cases} 1, & f(x, y) \geq \text{threshold} \\ 0, & f(x, y) < \text{threshold} \end{cases}$$

其中 $\text{threshold}$ 的值根据经验决定。接着可以从二值图中获得连通域的个数、大小和具体的位置, 也就是光斑的数学抽象。第 $i$ 个连通域可以表示为一些像素点的集合 $A_i = \{(u_k, v_k)\}$ 。

我们要做的是从强度图中的众多光斑中排除明显的噪声, 筛选出标记物。为了实现这一步, 需要给连通域加以筛选条件。

(1) 连通域大小。其中 $\text{MinMarkerArea}$ 和 $\text{MaxMarkerArea}$ 均为经验数值, 这一步通常能排除一些较大的环境噪音, 如灯光光源等。

$$\text{MinMarkerArea} < \|A_i\| < \text{MaxMarkerArea} \quad (2-14)$$

(2) 理论连通域大小。这一步依据真实的标记物大小，通过相机参数计算出投射到图像坐标系后的大小理论值。其中 $r$ 为标记物半径， $s_x/f_x$ 和 $s_y/f_y$ 代表 $x$ 和 $y$ 方向上单位像素的物理大小和焦距的比值， $d_i$ 代表从深度图中提取的连通域中心处的深度值。其中 $\alpha$ 和 $\beta$ 是经验值，代表了理论大小和实际大小的偏差范围。

$$\alpha \|A_i\| < \|A_i'\| = \frac{\pi r^2}{s_x/f_x \cdot s_y/f_y \cdot d_i^2} < \beta \|A_i\| \quad (2-15)$$

由此我们获得了筛选后的像素坐标系下连通域中心坐标集合 $A_i' = \{(u_k, v_k)\}$ ，尽管其中仍可能存在未被上述步骤排除掉的噪声点，但已经排除了部分噪声，降低了后续算法的运行压力。

为获得 $A_i'$ 对应的相机坐标系中的三维位置，使用 2.3.2 中获取的内参矩阵的参数计算去畸变后的像素坐标系下的坐标 $A_i'' = \{(u_k', v_k')\}$ ，则相机坐标系下的三维位置为：

$$(x_k, y_k, z_k) = \frac{d_k}{\|(u_k', v_k', 1)\|} (u_k', v_k', 1) \quad (2-16)$$

### 预实验：深度误差评估

在上述标记点检测过程中使用了深度图提供的深度数值，但事实上该深度传感器和真实的物体距离之间存在误差，并认为这个误差和距离有相关性。为了量化这个误差，我们搭建了如图 2.9 (a)所示的实验平台。

实验板与 2.2.2 中预实验采用的打印板相同。HoloLens 2 的传感器光轴和步进电机驱动的运动轴的轴心保持在同一直线上，并让步进电机匀速运行，使得实验板匀速改变距离。首先提取周围七个圆点的像素坐标系下的二维中心位置 $\{O_i\}$ 和该中心在相机坐标系下的三维坐标 $\{O_c\}$ ，从而能根据实验板的设计模型中的相对位置关系得到正方形中心的三维坐标 $O_c$ 。使用 PnP (Perspective-n-Points) 算法将 $O_c$ 重投影回深度图中得到像素坐标 $O_i$ ，得到对应像素的深度数据 $d_c$ 。相机坐标系下的向量模长 $\|O_c\|$ 代表了 $O_c$ 到光心的距离，认为该深度数据是真实的距离。深度传感器和真实的物体距离之间的误差则为 $\delta = d_c - \|O_c\|$ 。

实验结果如图 2.9 (c) (d)所示。注意到并不是一个规则的曲线，所以为了补偿深度传感器带来的误差，我们建立了一个查找表 $LUT(d)$ 。当计算新的点的深度时先得到相机坐标系下的三维坐标 $O_C$ ，计算 $\|O_C\|$ ，最终的深度值为 $LUT(\|O_C\|)$ 。进行深度补偿后可以一定程度上减小定位误差。

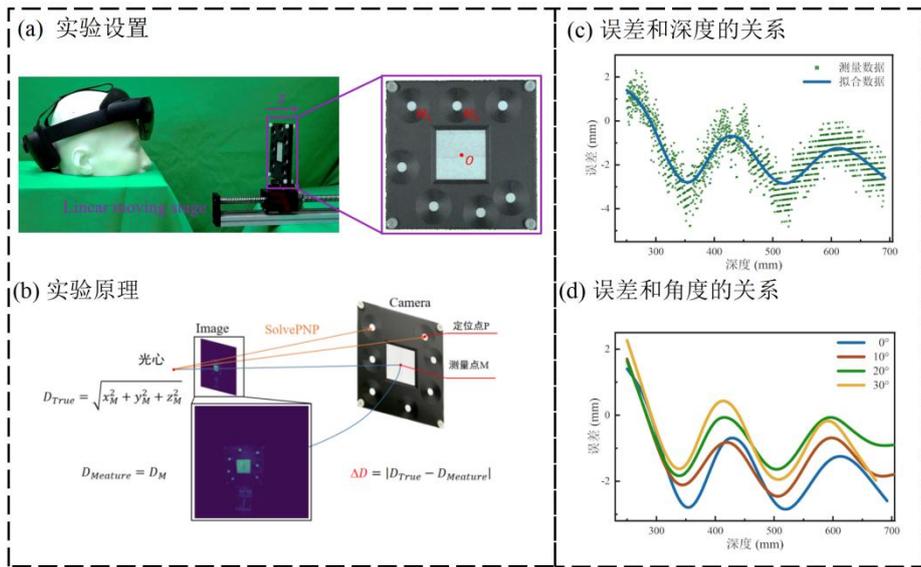


图2.9 深度误差评估实验设置和实验结果

## 特征提取

每个标记物的特征除了其在三维坐标下的位置外，由于平面拥有显著的法向特征，结合法向特征和坐标中心特征能更好地匹配相应的点，所以在这一步我们还利用 2.2.2 中用最小二乘估计法向量的方法提取了每个连通域的法向信息。

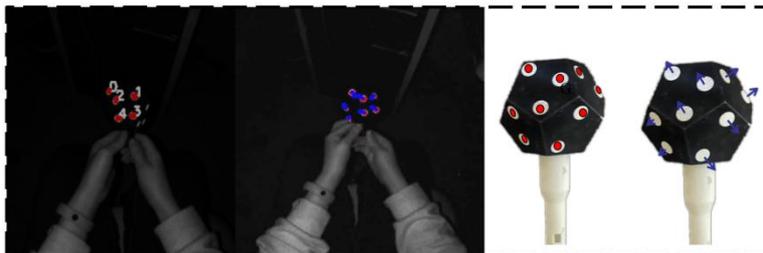


图2.10 位置特征和法向特征的示意图

### 2.3.4 定义算法设计

为了实现实时工具定位检测，需要先获得工具的形态特征，包含了每个标记点的中心位置。对于简单工具而言，所有的标记点都可以同时出现在同一个画幅中。而多面体的工具所面临的问题之一就是工具的所有点并不能同时出现。于是本节提出了一个针对多面体工具设计的时序拼接策略，利用视频流相邻帧的连续性进行多面体工具的定义。流程如图 2.11 所示。

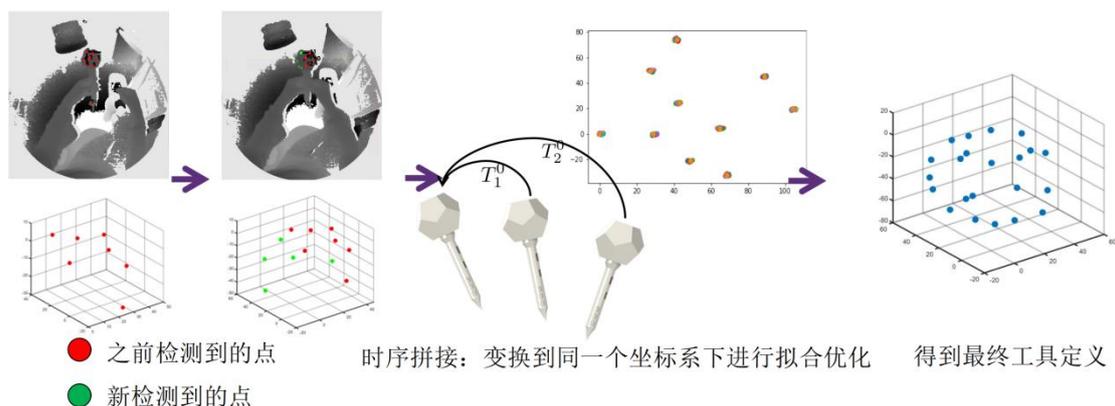


图2.11 定义算法的流程

令当前加入工具定义的点的集合为点集  $\{T_i\} = (x_i, y_i, z_i)$ ，显然初始状态下  $\|T_{i_0}\| = 0$ 。

假设在第  $n$  帧时已经维护的工具形态点集为  $\{T_i\}_{t_n}$ ，而该帧检测到的场景中所有点的点集为  $\{S_i\}_{t_n}$ ，一般而言  $\{T_i\}_{t_n}$  和  $\{S_i\}_{t_n}$  有如下几个性质：

(1)  $\|\{S_i\}_{t_n} \cap \{S_i\}_{t_{n-1}}\| \neq 0$ ，表明上一帧和当前帧中有共同的点。

(2)  $\|\{T_i\}_{t_{n-1}} \cap \{S_i\}_{t_n}\| \neq 0$ ，根据上一条性质可以推断出，因为上一帧的点被纳入了工具点集  $\{T_i\}_{t_{n-1}}$ ，故在本帧场景中出现的部分点已经被定义过。

根据性质 (2)，可以将  $\{S_i\}_{t_n}$  划分为两个子集  $\{S_m\}_{t_n}$  和  $\{S_n\}_{t_n}$ ，分别具有  $\{S_m\}_{t_n} \cap \{T_i\}_{t_{n-1}} = \emptyset$  和  $\{S_n\}_{t_n} \in \{T_i\}_{t_{n-1}}$  的性质。 $\{S_n\}_{t_n}$  的获取方法将在 2.3.7 中阐述。为了确保  $\{S_m\}_{t_n}$  中的点没有被纳入过  $\{T_i\}_{t_{n-1}}$ ，算法进行了二次检验，若：

$$\|S_m - T_i\| < ErrorTolerance \quad (2-17)$$

即  $\{S_m\}_{t_n}$  中有点和  $\{T_i\}_{t_{n-1}}$  中的点欧氏距离小于  $ErrorTolerance$ ，则执行：

$$\{S_m\}'_{t_n} = \{S_m\}_{t_n} - S_m \quad (2-18)$$

$$\{S_n\}'_{t_n} = \{S_n\}_{t_n} \cup S_m \quad (2-19)$$

最终当帧的获得的  $\{T_i\}_{t_n}$  应为:

$$\{T_i\}_{t_n} = \{T_i\}_{t_{n-1}} \cup \{S_m\}'_{t_n} \quad (2-20)$$

直到  $\|\{T_i\}_{t_n}\| = ExpectedMarkerNum$  时停止纳入新的点，并开始进行拟合，得到完整的工具定义。拟合方法在 2.3.7 中阐述。

### 2.3.5 定义算法评估

为了评估本章设计的定义算法得到的工具形态和真实的工具形态之间的误差，我们采用了一种高精度数控机床（Computer Numerical Control, CNC）加工工艺制作的平面工具（见图 2.12(a)）的设计文件作为工具形态的金标准，对不同点数的平面工具进行了误差评估。

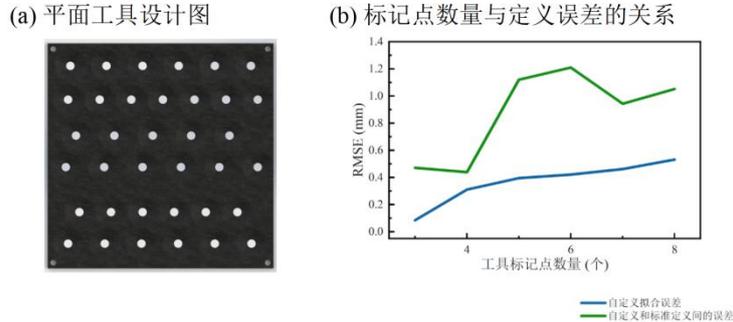


图2.12 定义算法评估实验工具与结果

将设计文件的工具定义点集记为  $\{G_i\}$ ，实验中用定义算法获得的点集记为  $\{T_i\}$ ，采用均方根误差  $RMSE(\{T_i\}, \{G_i\})$  和反映数据离散程度的定义拟合误差  $RMSE(\{T_i\}_t)$  作为衡量定义算法的误差指标。

结果如图 2.12 (b) 所示，实验表明本章提出的定义算法得到的定义误差能保持在 1mm 左右，较低时的时候接近 0.4mm。而拟合误差在 0.5mm 以下表明数据的离散程度不高。

### 2.3.6 工具尖端中心标定

工具中心点 (Tool Center Point, TCP) 标定是一种机器人领域常见的用于求解工具尖端在工具坐标系下坐标的方法。实验设置如图 2.13(a)所示, 保持 HoloLens 2 和带有球形凹槽的铁板相对静止, 将工具末端置于铁板凹槽内, 以工具末端为旋转中心摇动工具尾部, 期间记录多帧不同位姿但末端的世界坐标不变的图像。凹槽为直径略小于工具尖端半球形直径的圆柱槽, 在转动时根据图 2.13(b)右侧所示的几何关系, 半球的球心保持静止, 从而保证了所确定的尖端点为半球球心。

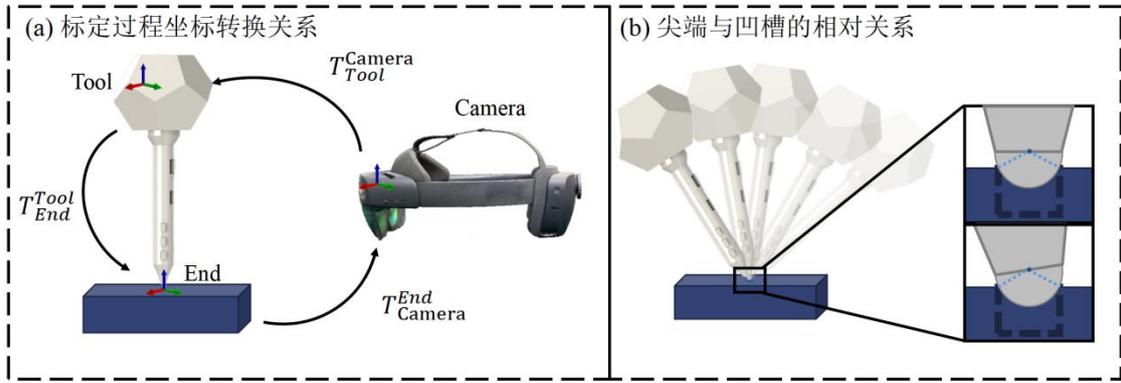


图2.13 工具尖端中心标定实验图

将工具、工具尖端和 HoloLens 2 的传感器分别记作  $Tool, End, Camera$ , 简称为  $T, E, C$ 。从 A 到 B 的变换关系由符号  $T_B^A$  表示, 且有  $(T_A^B)^{-1} = T_B^A$  的性质。TCP 标定的目的是求出  $T_{Tool}^{End}$ 。如图 2.13(a)所示, 可以得到空间变换关系:

$$T_{End}^{Camera} = T_{Tool}^{Camera} \cdot T_{End}^{Tool} \quad (2-21)$$

将  $T_B^A$  展开写为:

$$T_B^A = \begin{bmatrix} R_{AB} & t_{AB} \\ 0 & 1 \end{bmatrix} \quad (2-22)$$

其中  $R$  和  $t$  代表旋转矩阵和平移向量。代入式(2-21)有

$$\begin{bmatrix} R_{CE} & t_{CE} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_{CT} & t_{CT} \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} R_{TE} & t_{TE} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_{CT} \cdot R_{TE} & R_{CT} \cdot t_{TE} + t_{CT} \\ 0 & 1 \end{bmatrix} \quad (2-23)$$

显然有:

$$t_{CE} = R_{CT} \cdot t_{TE} + t_{CT} \quad (2-24)$$

由于相机和工具尖端保持相对静止， $t_{CE}$ 是不变的。而工具尾部位姿改变所以 $R_{CT}$ 和 $t_{CT}$ 是变化的。令第 $n$ 帧获取的 $R_{CT}$ 和 $t_{CT}$ 表示为 $R_{CT}^{(n)}$ 和 $t_{CT}^{(n)}$ ，于是有方程组：

$$\begin{aligned} t_{CE} &= R_{CT}^1 \cdot t_{TE} + t_{CT}^1 \\ t_{CE} &= R_{CT}^2 \cdot t_{TE} + t_{CT}^2 \\ &\vdots \\ t_{CE} &= R_{CT}^n \cdot t_{TE} + t_{CT}^n \end{aligned} \quad (2-25)$$

进行差分运算，得：

$$\begin{aligned} (R_{CT}^1 - R_{CT}^2) \cdot t_{TE} &= -(t_{CT}^1 - t_{CT}^2) \\ (R_{CT}^2 - R_{CT}^3) \cdot t_{TE} &= -(t_{CT}^2 - t_{CT}^3) \\ &\vdots \\ (R_{CT}^{(n-1)} - R_{CT}^n) \cdot t_{TE} &= -(t_{CT}^{(n-1)} - t_{CT}^n) \end{aligned} \quad (2-26)$$

表达为矩阵形式并进行化简则为：

$$t_{TE} = - \begin{bmatrix} (R_{CT}^1 - R_{CT}^2) \\ (R_{CT}^2 - R_{CT}^3) \\ \vdots \\ (R_{CT}^{(n-1)} - R_{CT}^n) \end{bmatrix}^{-1} \cdot \begin{bmatrix} t_{CT}^1 - t_{CT}^2 \\ t_{CT}^2 - t_{CT}^3 \\ \vdots \\ t_{CT}^{(n-1)} - t_{CT}^n \end{bmatrix} \quad (2-27)$$

$t_{TE}$ 即为工具尖端在工具坐标系下的坐标。至此我们已经获得了工具和工具尖端的定义。

### 2.3.7 定位算法设计

定位算法旨在将场景中检测到的点和工具定义中的点一对一匹配上。该问题可以抽象为如下图匹配问题。令场景中检测到的大小为 $N$ 的点集为 $\{S_i\}_N$ ，大小为 $M$ 的工具定义的点集为 $\{T_i\}_M$ ，场景中属于工具大小为 $K$ 的点集为 $\{S_i\}_K$ ，工具定义中出现在场景中大小为 $K$ 的点集为 $\{T_i\}_K$ ，欲求 $\{S_i\}_N$ 和 $\{T_i\}_M$ 的匹配关系，获得 $\{S_i\}_K$ 和 $\{T_i\}_K$ 。

在多面体工具的定位场景中，存在如下几个问题：

(1)  $\|\{T_i\}_K\| \ll \|\{T_i\}_M\|$ , 工具一定会产生遮挡, 出现在场景中的点只占整个工具点集的部分。

(2)  $\|\{S_i\}_N\| \ll \|\{T_i\}_M\|$ , 场景中出现的点数一般情况下小于 10, 甚至小于 5, 而工具点集大小往往大于 20。

(3)  $\|\{S_i\}_K\| \leq \|\{S_i\}_N\|$ , 场景中可能存在噪声。

本章中提出的匹配方案的核心思想在于:

(1) 将点匹配问题转换为边匹配问题, 将绝对位置信息转换为相对位置信息。

(2) 用工具边匹配场景边, 从而减少运算量。

定位算法的流程如下:

(1) 从点的匹配问题转换为边的匹配问题, 计算完全图的每条边信息。

(2) 通过筛选条件建立场景边与工具边的匹配图。

(3) 在匹配图中以点编号为索引遍历可能的边匹配关系, 以预设的某个边界调节计算出最优解。

(4) 获得一一匹配的工具点和场景点, 利用该匹配信息, 将工具点在工具坐标系下的坐标和 2.3.6 获得的尖端坐标刚体变换到场景点在相机坐标系下的对应位置, 获得工具当前帧的位姿与尖端位置。

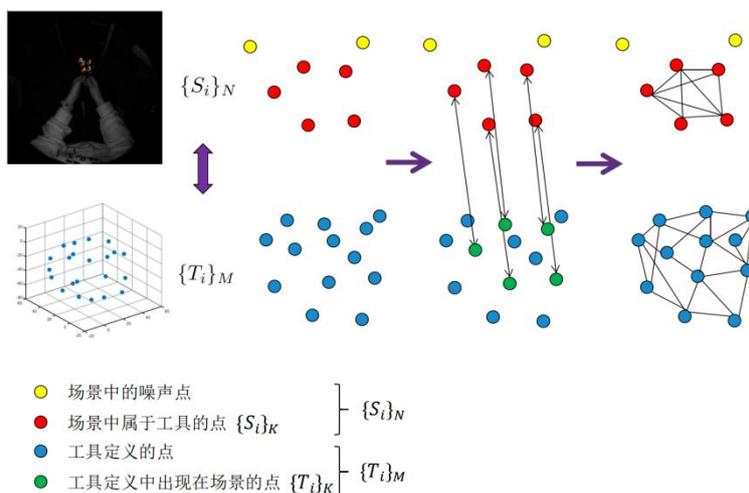


图2.14 图匹配问题示意图

### 从点的匹配问题转换为边的匹配问题

现有场景点集  $\{S_i\}_N$  和工具点集  $\{T_i\}_M$ , 场景边集  $\{E_{S_j}\}_{C_N^2}$  和工具边集  $\{E_{T_j}\}_{C_M^2}$  的计算方式为:

$$\{E_{Sj}\}_{C_N^2} = \{\|S_p - S_q\| \mid p \neq q \text{ and } 0 \leq p, q < N\}$$

(2-28)

$$\{E_{Tj}\}_{C_M^2} = \{\|T_p - T_q\| \mid p \neq q \text{ and } 0 \leq p, q < M\}$$

### 建立场景边与工具边的匹配图

对于大小为 $C_N^2$ 的场景边集 $\{E_{Sj}\}_{C_N^2}$ 中的每一条边 $E_{Sj}$ ，认为在工具边集 $\{E_{Tj}\}_{C_M^2}$ 中存在一个满足 $\|E_{Tk} - E_{Sj}\| < \delta_{tolerance}$ 的大小为 $K_j$ 的子集 $\{E_{Tk}\}_{K_j}$ ，以符号 $E_{Sj} \rightarrow \{E_{Tk}\}_{K_j}$ 表示该匹配图关系。于是可以得到：

$$\begin{aligned} E_{S1} &\rightarrow \{E_{Tk}\}_{K_1} \\ E_{S2} &\rightarrow \{E_{Tk}\}_{K_2} \\ &\vdots \end{aligned}$$

(2-29)

$$E_{S(C_N^2-1)} \rightarrow \{E_{Tk}\}_{K(C_N^2-1)}$$

其中 $K_j$ 的取值范围为 $0 - C_M^2$ 。

实际上还可以有另一种建立匹配图的方向，即对于大小为 $C_M^2$ 的工具边集 $\{E_{Tj}\}_{C_M^2}$ 中的每一条边 $E_{Tj}$ ，认为在场景边集 $\{E_{Sj}\}_{C_N^2}$ 中存在一个满足 $\|E_{Sk} - E_{Tj}\| < \delta_{tolerance}$ 的大小为 $K_j$ 的子集 $\{E_{Sk}\}_{K_j}$ ，以符号 $E_{Tj} \rightarrow \{E_{Sk}\}_{K_j}$ 表示该匹配图关系。这两种匹配方向在计算复杂度上有极大的差异。

经理论推导第一种匹配方向可以得到场景中只有一个工具情况下最好情况的复杂度为 $O(N^2)$ ，最坏情况的复杂度为 $O((M^2)^{N^2})$ ，平均情况的复杂度为 $O((\delta_{Tolerance} M^2/d)^{N^2})$ ，其中 $d$ 为工具边的平均差。第二种匹配方向同理。可以得到如下不等式：

从场景搜工具（搜索方向一）的复杂度为：

$$O(N^2) < O\left(\left(\frac{\delta_{Tolerance}}{d} M^2\right)^{N^2}\right) < O((M^2)^{N^2})$$

(2-30)

从工具搜场景（搜索方向二）的复杂度为：

$$O(M^2) < O\left(\left(\frac{\delta_{Tolerance}}{d} N^2\right)^{M^2}\right) < O((N^2)^{M^2})$$

(2-31)

显而易见的是，搜索方向一，也就是本章中选择的搜索方向，适用于 N（场景点）少，M（工具点）多，即工具复杂，场景简单的情况；而搜索方向二适用于 M（工具店）少，N（场景点）多，即工具简单，场景复杂的情况。

另外需要说明的是，尽管理论计算的平均复杂度达到了数值上较高的阶数，但实际上由于加入了优化中断算法，算法的时间消耗能保持在一个较小的值，将在 2.3.8 中进行实验验证。

### 以点编号为索引遍历可能的边匹配关系

此处采用的遍历方法是深度优先搜索（Depth First Search, DFS）算法，搜索的结果由一个工具点编号和场景点编号的映射图表征。每一次深度搜索到尽头返回本次搜索的映射图后使用 Horn 等人<sup>[38]</sup>提出的 SVD 的最小二乘刚体变换算法计算工具坐标系下的工具点和相机坐标系下的场景点的刚体变换误差。刚体变换误差原理如下：

现有已知匹配关系的场景点集  $\{S_i\}_K$  和工具点集  $\{T_i\}_K$ ，每个点对应的法向量集  $\{SN_i\}_K$  和  $\{TN_i\}_K$ ，欲求  $\{T_i\}_K$  到  $\{S_i\}_K$  的旋转矩阵 R 和平移向量 t，优化的目标是

$$Error = \min \left( \alpha \sum_{i=1}^k w_i \| (RT_i + t) - S_i \|^2 + (1-\alpha) \sum_{i=1}^k w_i \| (RTN_i + t) - SN_i \|^2 \right)$$

其中前一项是位置误差，后一项是法向量误差， $\alpha$  是误差权重，为经验数值。

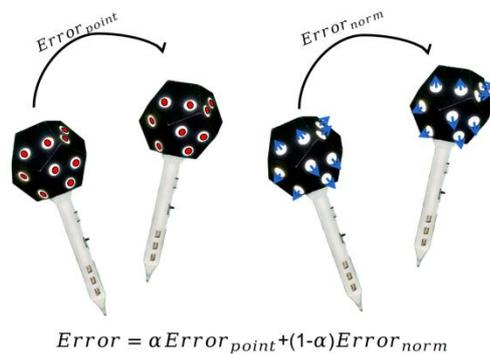


图2.15 定位算法和误差计算示意图

首先分别计算两个点集的加权中心点：

$$\bar{T} = \frac{\sum_{i=1}^k w_i T_i}{\sum_{i=1}^k w_i} \quad \bar{S} = \frac{\sum_{i=1}^k w_i S_i}{\sum_{i=1}^k w_i} \quad (2-32)$$

并将新的中心点作为原点，应用于点集：

$$x_i := T_i - \bar{T} \quad y_i := S_i - \bar{S} \quad (2-33)$$

计算大小为  $3 \times 3$  的协方差矩阵，其中  $X$  为  $x_i$  作为列的矩阵， $Y$  为  $y_i$  作为列的矩阵，大小均为  $3 \times k$ ， $W = \text{diag}(w_1, w_2, \dots, w_n)$

$$A = XWY^T$$

$$(2-34)$$

对  $A$  进行奇异值分解，

$$A = U \Sigma V^T$$

$$(2-35)$$

得到旋转矩阵  $R$  为：

$$R = V \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \det(VU^T) \quad (2-36)$$

平移向量  $t = \bar{S} - R\bar{T}$ 。

$$(2-37)$$

设置的边界条件为：

- (1)  $K \geq 3$ ，因为至少需要三点才能做刚体变换。
- (2)  $Error \leq ErrorThreshold$ ，若刚体变换误差过大说明存在错配的可能性。
- (3) ( $K_{current} = K_{previous}$  and  $Error \leq PreviousError$ ) 或  $K_{current} > K_{previous}$ ，和之前维护的最好的结果进行对比，如果匹配图的大小比之前最好的结果大，或尽管相等，但误差小于之前最好结果的误差，则将当前结果置为最好结果。

获得最终位姿和尖端位置。

上述遍历结束后，得到了最佳匹配图和相应的刚体变换矩阵，继而可以应用刚体变换矩阵求解尖端在当前帧下的坐标。

### 2.3.8 定位算法评估

实验采用了图 2.12(a)中的实验板测试帧率、定位误差和标记物数量的关系。

#### 帧率测试

离线测试环境 CPU 为 Intel Core i9-12900, 实现语言为 C++, 编译器为 MSVC。

结果如图 2.16(a)所示, 可以发现随着点数的增加运行帧率有下降趋势, 但仍然保持在 1000fps 以上。

对于有 22 个点的面体工具, 定位帧率平均为 1057fps, 完全满足场景要求。

#### 定位误差测试

随着点数增加定位误差有所增加 (见图 2.16(b)), 总体能维持在 1mm 以下。

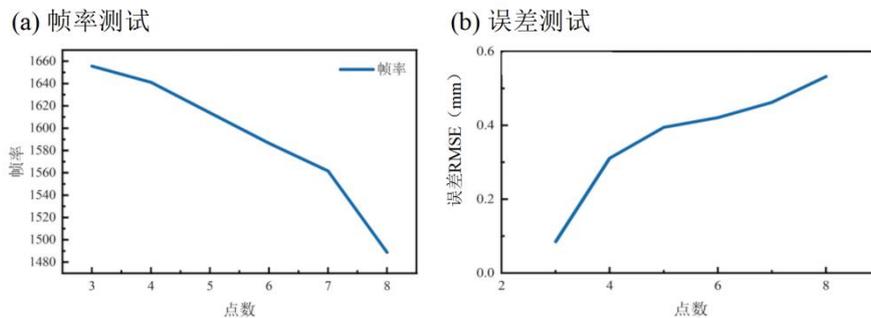


图2.16 标志点数量和检测帧率、误差的关系

### 2.3.9 尖端误差评估

在上一节中评估的误差为两个坐标变换的误差, 而使用过程中更直观的误差产生于交互作用点。本节设置了三个误差评估实验评估尖端误差。

#### 静态尖端抖动评估

该实验旨在量化尖端距原点距离对抖动的影响, 并测量当前笔长设置下尖端的抖动。将工具和 HoloLens 2 都静止置于实验台上, 录制 300 帧数据, 沿笔尖和原点连线方向取 150 个间隔为 1mm 的点, 记作  $(X, Y, Z) = \{(x_t, y_t, z_t)\}$ , 每个点的抖动通过下式进行量化:

$$jitter = \sqrt{(std(X)^2 + std(Y)^2 + std(Z)^2)/3} \quad (2-38)$$

结果如图 2.17 所示，近乎呈现线性趋势，说明减小笔长可以有效减小尖端抖动。本章设计的笔尖距原点距离为 84.4cm，抖动为 1.14188mm。

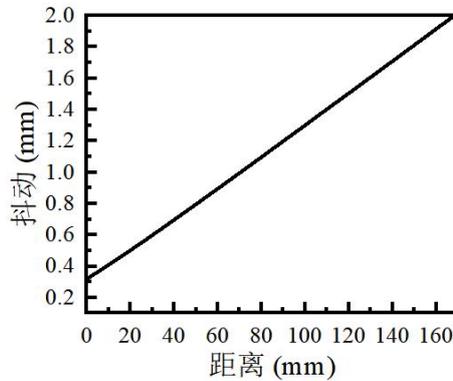


图2.17 尖端距原点距离与抖动的关系

### 相对定位误差和重复定位误差评估

实验设计如图 2.18 (a)所示，实验者佩戴 HoloLens 2 用工具尖端触碰一块带有参考架的实验板的凹槽。用本节提出的定位算法可以获得  $T_T^H$  和  $T_H^R$ ，则  $T_T^R = T_H^R T_T^H$ ，从而获得尖端相对于参考架的坐标。实验结果可视化如图 2.18 (b)所示。

重复定位误差指的是在工具尖端反复与同一个点接触时产生一系列有微小差距的坐标值，该系列的坐标值的在三个方向上的标准差即为重复定位误差。

相对定位误差指的是工具尖端与相邻两个点接触后计算出的两点距离与实际两点距离有一定的差距，实际距离和测量距离的差值即为相对定位误差。由于实验板只有 x 和 y 两个方向，故相对定位误差也只衡量了两个方向。

为了方便比较，数据进行了去均值操作，使得均值都在 0mm 处。两个误差都由 RMSE 表示，误差以小提琴图 2.18 (c)(d)的形式展示，呈现了误差结果的分布。

使用 Anderson-Darling 检验验证重复定位误差的三个方向、相对定位误差的两个方向是否符合正态分布，由表 2.2 和表 2.3 看出，除了 x 方向的相对定位误差之外其他误差分布的 p 值均远远大于 0.05，故使用期望和方差表达 x 方向的相对定位误差，以“期望±方差”的形式表达；四分位距和中位数对误差分布表达其他方向的误差分布，记作“中位数（四分位距）”。

x 方向的重复定位误差为-0.0062mm (2.0934mm)，y 方向的重复定位误差为-0.0177mm (0.5528mm)，z 方向的重复定位误差为-0.0852 mm(3.2866mm)，x 方

向的相对定位误差为  $0.8107\text{mm} \pm 2.1072\text{mm}$ ，y 方向的相对定位误差为  $-1.2452\text{mm}$  ( $5.5865\text{mm}$ )。

结果表明重复定位误差中位数均在  $0\text{mm}$  附近，但相对定位误差有方向性。此外四分位距或方差表明了分布的离散程度，根据分析结果可以得出抖动相对较大的结论。深度轴方向误差明显大于另外两个方向，这是 HoloLens 2 深度传感器自身结构带来的噪声导致的，这也是本工作提出的定位算法需要克服的问题之一。

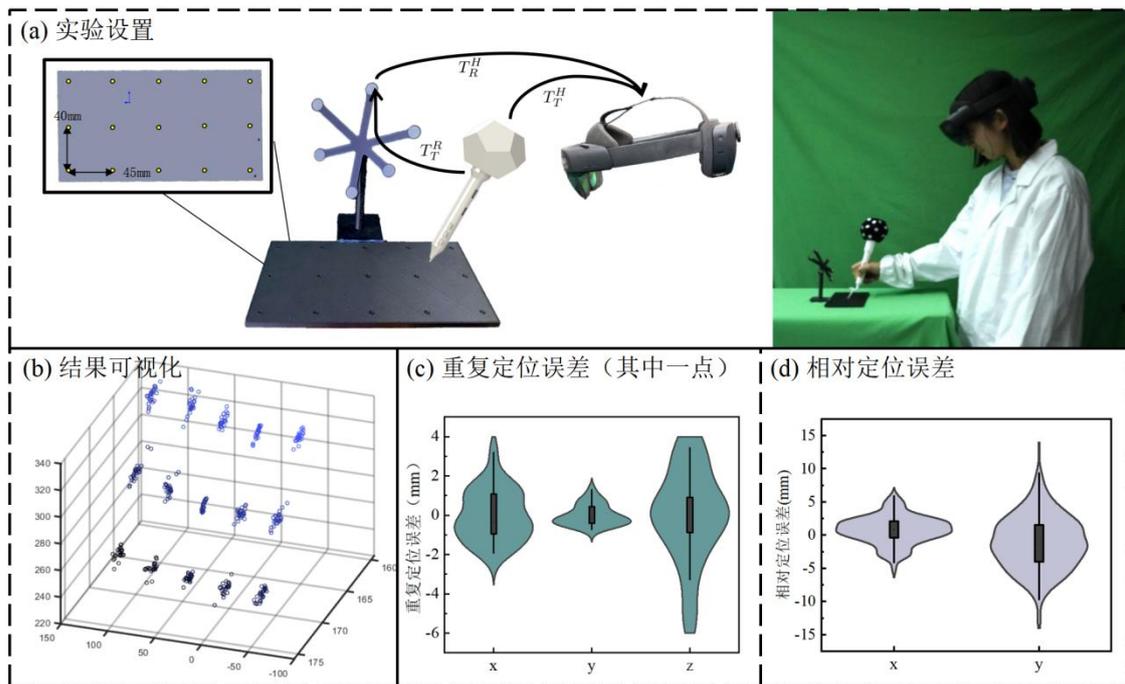


图2.18 定位误差评估实验

表2.2 实验板两个方向相对定位分布的统计学参数

	x	y
p 值	0.0076	0.7920
四分位距 (mm)	2.5274	5.5865
中位数 (mm)	0.8468	-1.2452
期望(mm)	0.8107	NA
方差(mm)	2.1072	NA

表2.3 三个方向重复定位分布的统计学参数

	x	y	z
p 值	0.5166	0.7442	0.7152
四分位距 (mm)	2.0934	0.5528	3.2866
中位数 (mm)	-0.0062	-0.0177	-0.0852

### 2.3.10 尖端位姿优化和轨迹误差评估

在上一节中我们用实验量化了尖端抖动，为了平滑该由于传感器本身的性质造成的抖动，我们用三种不同的滤波法最小二乘平滑优化、卡尔曼滤波、一阶线性差分法（原理见附录 B.2）直接对尖端的轨迹进行了平滑滤波，比较了轨迹图、和原信号的均方差 (Mean Square Error, MSE)、和原信号的协方差 (Covariance, COV)、信噪比 (Signal-to-Noise Ratio, SNR) 等指标以评估滤波效果。结果如下图所示，在信噪比和协方差指标上滤波器并没有带来太大的改善或损失，但在协方差上卡尔曼滤波后的轨迹尤其明显。这是由于卡尔曼滤波器的预测功能使得在尖端移动速度快时出现了超调现象，也就是过度预测。

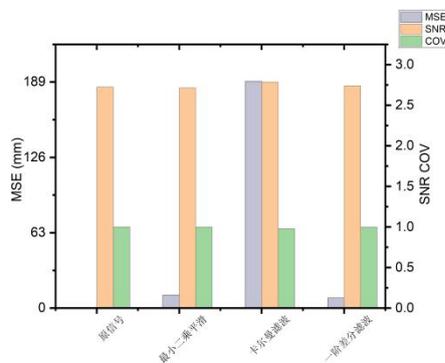


图2.19 三种滤波后信号和原信号的各个指标对比

此外实验还对比了同一种滤波器不同参数的情况，以供选择合适的参数实现不严重破坏轨迹的情况下滤除抖动，详细参数对比见附录 B.2。

为了更直观地观察尖端误差，我们采用 CNC 工艺制作了一块有规则和不规则轨迹的实验板。实验过程中笔尖紧沿着轨迹凹槽运动，记录笔尖坐标，并与设计图中的使用迭代最近点算法 (Iterative Closest Point, ICP) 进行配准，计算轨迹和设计图图案之间的均方根误差。

从直观上不同滤波器滤波后轨迹和原轨迹可以看出最小二乘平滑算法滤波后得轨迹最接近原轨迹，从数值上看，RMSE 最小的也是最小二乘平滑。

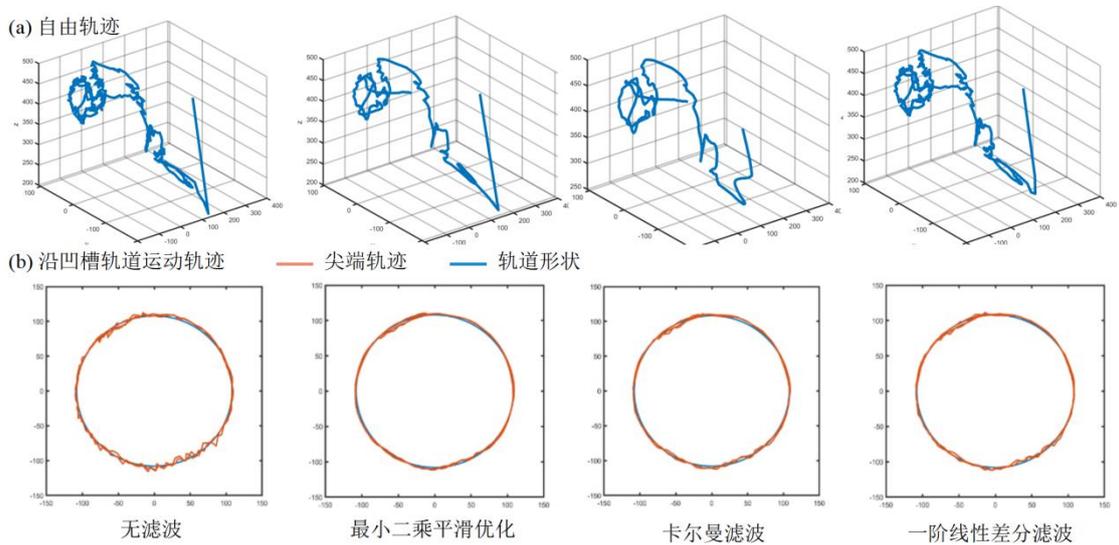


图2.20 轨迹对比

表2.4 配准后轨迹与轨道形状的RMSE

滤波器	无	最小二乘平滑	卡尔曼滤波	一阶线性差分
ICP RMSE (mm)	2.12642	1.90021	2.17271	1.999

## 2.4 总结与讨论

医疗场景应用中增强现实交互需要精准的工具提供精准的交互作用点以满足临床需求，本章提出了一个针对于 HoloLens 2 深度传感器的红外标记点工具，拥有定位和基于无线局域网的主动交互功能。

首先本章从工具的硬件构造开始介绍，依次阐述了尾部标记物的结构和设计考量、控制电路的功能与硬件实现和外壳的设计。在如多面体面数的选择等关键设计环节提供了理论和实验两方面的结果支撑。

接着开始介绍多面体工具配套的算法设计。整个算法流程包含了四大部分，数据准备、标记点检测、工具定义和工具定位。在数据准备部分详细介绍了内参标定和深度传感器数据流获取的步骤；标记点检测部分阐明了如何从场景图中抽象出关键特征；工具定义部分以数学符号形式对问题进行公式化，提出了针对多

面体工具定义时序拼接的关键概念，并对其误差进行了衡量，结论是定义误差在 1mm 及以下；获得尾部标记物定义后通过 TCP 标定算法计算出了工具尖端，也就是增强现实空间的交互作用点相对尾部的坐标；在定义算法设计部分将场景图与工具定义匹配的过程抽象为了图匹配问题，并通过理论推导证明了本算法提出的搜索方向的绝对优势；最后用帧率测试实验验证了算法的实时性和高效性，尖端误差评估以误差数值和直观轨迹图展示了滤波前后的效果，表明最终交互作用点的误差在 2mm 以内。

在本节中希望探讨误差的来源。事实上误差主要来源于两大部分：相机畸变和传感器噪声。尽管已经进行了内参标定，但由于选用的畸变模型不能完整地弥补镜片的实际畸变，所以在视野的不同方位仍然存在不同的畸变误差。而且随着时间和温度的改变，传感器不可避免地产生时间漂移和温度漂移，内参矩阵也会随之改变。此外由于传感器的画幅只有  $512 \times 512$ ，在进行连通域分析和中心点提取时会对结果取整导致半个像素的误差，该误差转换到现实世界中就会达到亚毫米量级甚至毫米级，从传感器源头的误差经过工具定义、尖端标定等步骤一步步累积放大，对最终结果的影响是显著的。

但需要说明的是，在线运行时获取的画幅都是由 HoloLens 2 自身实时依据当前状态进行矫正获得的，故可以抵消一定的畸变影响，使误差减小。

最后对本章取得的成果和创新性进行总结：

- (1) 制作了一支用于增强现实系统的拥有主动交互功能和被动定位功能的工具笔，提高了交互工具的可扩展性。
- (2) 提出了多面体拼接定义算法，实现全方位定位。
- (3) 通过优化搜索方向和设置早停机制提升了定位速率，实现了实时定位。
- (4) 提出了完善的定义、定位算法和滤波器，并进行了多角度的性能验证，实现了精准定位，为接下来即将展开的第 3 章中的交互方法提供了精准的交互作用点。

## 第3章 基于多模态术前影像的规划协作交互方法

### 3.1 引言

近二十年来不同领域的人对二维屏幕和增强现实系统协作方式进行了探索。在协作系统中二维屏幕的作用主要分为两大类：显示和标记。Gauglitz 等人<sup>[39]</sup>通过令远程用户在显示着实时场景的画面中用触摸屏勾画一个轮廓指出环境中的一个元素，并以增强现实的方式显示给手持设备的远程用户。Kim 等人<sup>[40]</sup>提出在远程会议中通过于屏幕上进行标记、指针的方式改善会议参与者的体验感。Huang 等人<sup>[41]</sup>构建了一套可以共享手势和标记的系统以提升远程协作者对任务的理解。上述提到的部分工作中搭建的协作系统是出于远程应用的场景而设置的，也有部分是为了利用二维平面上能产生更精确的标记的特质而引入的协作方式。然而这些工作分配给二维屏幕的任务和权限相对单一，且在医学领域尚未有人探索二维屏幕和三维增强现实空间的协作。

绪论中提到，当前应用于术前规划领域的增强现实系统仍然存在功能较为基础，无法实现如在影像处理软件中的包括分割、配准、自动规划等功能。由此可见在以往的工作中二维平面的优势和在增强现实空间中对医学影像的操作都尚未得到充分开发。

影像处理软件和增强现实在术前规划任务中各有优劣。二维交互的一大特点在于把复杂的信息简化使得人更容易操作，同时这样的简化也带来了信息量的损失，在二维界面进行如切割等某些三维操作时并不符合直觉。而增强现实系统提供的三维观察和操作空间自由度更大，更符合人自然的操作习惯，但由于当前硬件和算力的限制，增强现实系统难以脱离服务器自身实现复杂任务。结合影像处理软件在复杂操作上的优势和增强现实系统在观察与简单交互的优势成为一个构建术前规划交互方法的新路径。

上述提到的观察、简单操作和复杂操作具体到术前规划任务中可以分别表述和举例为（1）观察：能在三维空间中自由从多角度获取已配准的多模态影像信息，如脑结构像，脑动脉结构，神经束结构，脑功能激活区，并能选择单独观察感兴趣的结构或去除不感兴趣的结构（2）简单操作：符合人的物理直觉，不需要额外的学习成本的操作方式，如对虚拟影像进行旋转、平移、缩放，使用第2章提出的工具在影像表面或内部的感兴趣区域进行标注或擦除（3）复杂操作：由于承载了过多的信息量需要在操作层面对操作对象的复杂度进行压缩并借助

计算机辅助计算的过程，如结构分割、多模态影像配准、自动规划路径、入路风险计算等。

本章的工作结合了以 HoloLens 2 为载体的增强现实系统和二维医学影像处理软件 3D Slicer，在不同功能上分配不同的操作权限，使得增强现实设备使用者和软件使用者共同协作，以期望带来更好的交互体验和交互精度，并使得使用者获取更多的信息，提高交互效率。上一章提出并实现的可以被头戴式设备 HoloLens 2 识别并定位的精准交互工具将在本章中被应用于各类任务中的精准交互。

## 3.2 交互方式概述

### 3.2.1 交互分类

交互是一个广泛的概念，在设计领域交互定义为个体与个体之间交流或对彼此做出响应的过程。LaViola 等人<sup>[42]</sup>提出三维空间中的交互可以被分为低级别的交互，如选择 (Selection)、操纵 (Manipulation)、导航 (Navigation) 和系统控制 (System Control) 等，和由低级别交互复合而成的高级别交互，如改变一个物体的颜色包含了从菜单中选择颜色 (系统控制)、点击一个物体 (选择)、将画笔放置于三维空间中 (操纵) 几个部分。除此之外三维空间的交互还可以依据其支持的任务组、隐喻 (Metaphor)、自由度等特征细分<sup>[43]</sup>。可见交互的分类方式多种多样，在本章中，我们围绕着术前规划任务将交互分为了三个技术层次：基础交互、数据元修改和影像处理。

基础交互指的是将医学影像的整体作为交互对象，如对三维体素或面片模型进行整体的移动、缩放、旋转。此外我们还将数据展示归入了该层次，数据展示并非影像数据到用户单向的显示，而是包含了用户可以通过调整整体或局部的渲染方式改变信息传达量的过程。

数据元的修改相比于基础交互侧重的是微观的操作。体数据的由体素构成，面片数据由面元构成，当用户需要对数据局部进行标记、切割或绘制时就需要对体素或面元进行修改。

影像处理则需要综合考虑如不同解剖结构等不同模块的联系，并联系其生理意义以数值或图形的方式给出可被用于指导术前规划的结果。如分割肿瘤区域，提取血管、神经纤维束或头骨轮廓，计算肿瘤大小，计算入路方案风险数值等都被归于影像处理的范畴。

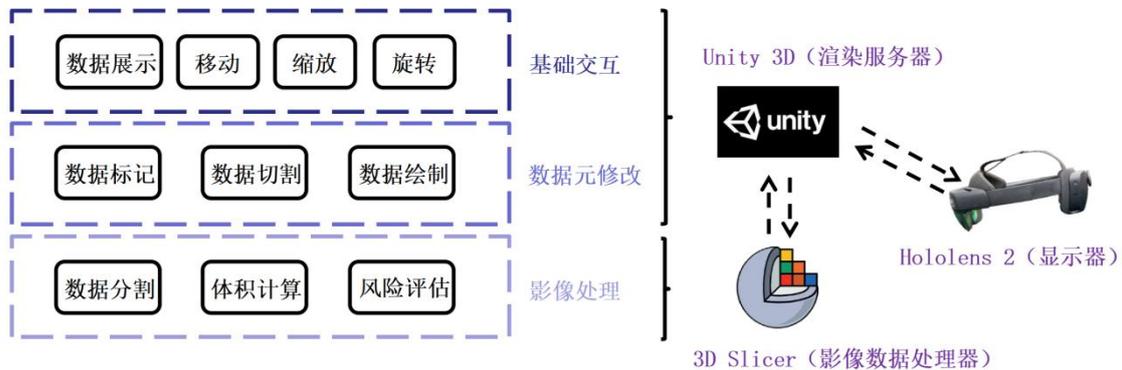


图3.1 在术前规划任务中的交互分类

### 3.2.2 交互平台与载体

鉴于当前并没有集成增强现实接口和医学影像处理接口的平台，开发一个新平台存在成本高昂和可用性低的问题，所以上述功能将分布于不同的平台并以平台间通信的方式连接。目前最常用的支持增强现实设备的渲染引擎之一是 Unity 3D，该软件和 HoloLens 2 可以一同被视为增强现实系统；被广泛应用于医学影像处理的软件之一是 3D Slicer，该开源软件为开发者提供了封装良好的应用接口，给予软件极大的可扩展性。上小节所述的基础交互和数据元修改被部署于增强现实系统，影像处理部分在 3D Slicer 上实现。两个平台的数据和指令通信过程将在第 4 章中详细阐述。

## 3.3 协作方式概述

协作是指多人或团体为完成一个共同的目标相互配合。上一节提到为了完成术前规划任务，技术被部署于两个平台。由于两个平台的任务差异较大，故需要至少两个人参与使用该系统进行术前规划流程，一人佩戴头戴式增强现实显示器，另一人操纵影像处理软件以配合任务的完成。

脑胶质瘤切除手术术前规划流程可以分为四个步骤：获取和处理术前多模态影像、观察影像建立感知、入路和切口规划和风险评估。每一个步骤又可以分为不同的子任务，依据增强现实系统和影像处理软件的各自优势被分配在相应的平台上，从而实现提高流程效率。效率流程和子任务被标注于图 3.2 中。下文将以流程中的四个步骤为主线，着重阐述增强现实系统中子任务的技术实现和两个平台任务分配的考量。

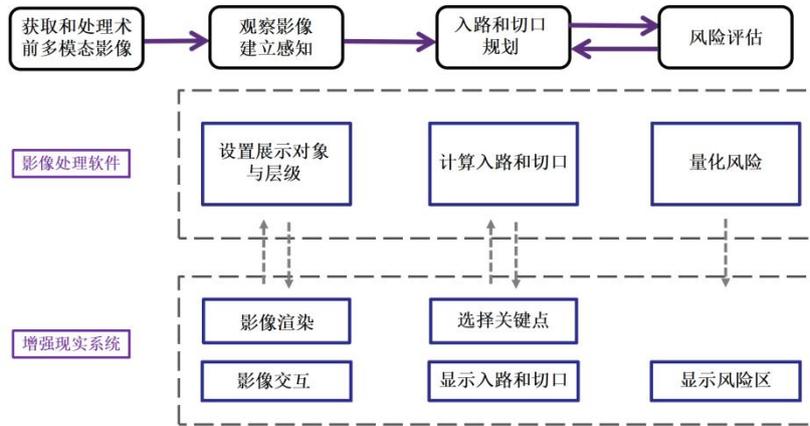


图3.2 增强现实系统和影像处理软件在术前规划流程中协作

### 3.4 获取和处理术前多模态影像

第一章中介绍了不同模态的影像和各自能表达的信息。由于设备限制原因，部分模态的数据难以获取，故本工作的部分数据来源于公开数据集。MRI 结构像、fMRI 功能像和 DTI 数据来自于脑肿瘤病人公开数据集<sup>[44]</sup>，CT 和 MRA 来自于长庚医院数据。若系统被应用于临床，影像数据应来自于同一患者的术前扫描并进行配准。需要说明的是，下文简单列举的处理方式只是可能的获取相应数据的渠道之一，处理结果可以来自于任何可用的医学影像处理软件或特定的神经网络。

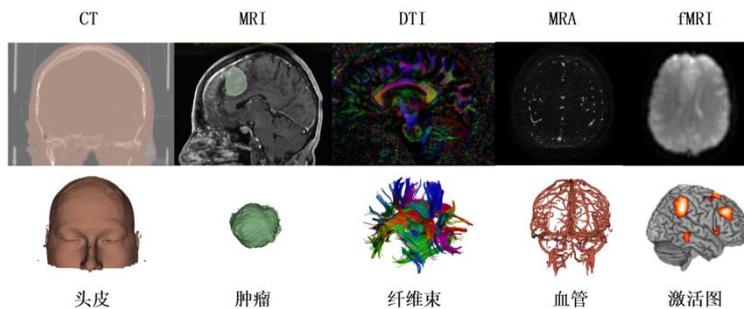


图3.3 不同模态的影像图和处理后提取出的结构

#### 3.4.2 MRI 数据处理

MRI 是一种反映解剖结构的影像，对软组织的显示较为清晰，在脑胶质瘤切除手术术前规划中被用于提取肿瘤边界等信息。肿瘤的提取使用 3D Slicer 的分割

功能 (Segmentation) 实现, 将肿瘤区域的标签信息保存并导出肿瘤边界为面数据。

### 3.4.3 DTI 数据处理

DTI 经过处理后能重建出神经纤维束, 发源于脑干的支配着全身不同感觉和运动功能的十二对颅神经也能通过 DTI 被重建。脑胶质瘤切除手术的目标之一是尽可能不损伤患者的感官和运动功能, 故 DTI 重建结果有助于在术前规划中避开十二对已知对人体有重要影响的神经束。神经束重建采用 3D Slicer 的扩散 (Diffusion) 功能实现, 重建后的神经束被导出为面数据。

### 3.4.4 CT 数据处理

在脑胶质瘤切割手术的术前规划中, 切口和入路常常标记在头皮表面, 故在增强现实显示中需要有一个可作用的表面。CT 影像可以被用于提取头皮, 采用 3D Slicer 的分割 (Segmentation) 功能粗略提取出头皮轮廓, 再用其他面数据修改软件进行精修, 得到完整的表面。

### 3.4.5 MRA 数据处理

在开颅手术中应尽可能避免大血管, MRA 影像提供了明显的血管边界。利用神经网络对血管进行分割后导出为面片数据。图 3.3 中第四列影像切面和血管重建结果。

### 3.4.6 fMRI 数据处理

在本文使用的 fMRI 数据集的刺激为运动任务和语言任务。使用 MATLAB 的 SPM 工具箱<sup>⑨</sup>进行预处理和利用时序信息及统计学方法激活图分析, 从而得到运动功能区和语言功能区的位置与分布, 以激活图的形式呈现。需要说明的是 fMRI 体数据的大小往往比结构像小, 故需要进行上采样。

## 3.5 观察影像建立感知

在使用医学影像处理软件进行脑胶质瘤切除手术术前观察影像建立感知的过程中存在以下问题: 由于观察和操作平面限制于二维, 常常无法实现精确修改三维影像。以 3D Slicer 为例, 若在三维影像上实现标记点、线功能只能先选取一个

---

<sup>⑨</sup> <https://ww2.mathworks.cn/matlabcentral/profile/authors/7905683?requestedDomain=zh>

合适的视角，鼠标在二维平面的坐标被投影到三维影像表面与该投影方向的第一个交点处；若期望实现切割功能，只能选取一个合适的视角，以鼠标在二维平面的勾画的封闭图形作为截面，以投影方向（常常为垂直于视角平面）为轴，将形成的立体区域所有与影像相交的部分去除。以切割部分头皮和颅骨模拟切口为例，由于头部的各个组织和解剖层次以包裹的形式组织在一起，这样的操作常常会将感兴趣区域也去除。于是我们期望在增强现实空间中实现一种能自由选择标记点或擦除点的精准交互方式，并能根据生理结构将操作对象分层，能对指定的体层进行操作而不影响到邻近的层。

此外增强现实空间中多模态的医学影像展示应保证信息表达的准确性。脑胶质瘤切除手术术前规划的最终目的是让医师尽可能地获取肿瘤位置大小、毗邻的如血管、神经束等重要结构与肿瘤间的相对位置关系的信息，所以组织间的遮挡关系应在任意角度都被正确表达。

在以往的神外术前规划工作中都是使用面数据居多，然而面数据只能体现结构的边界信息，并且可交互空间缺乏深度信息。在医学图像领域更能表现解剖结构间的相对关系的体数据的渲染是一个尚未被完全发掘的领域。如何组合多模态影像使每个模态在选择的渲染模式下尽可能地传达准确的信息量将为本节讨论的主要话题之一。此外我们还设计了一些辅助观察的虚拟工具，也将在下文中详细描述。

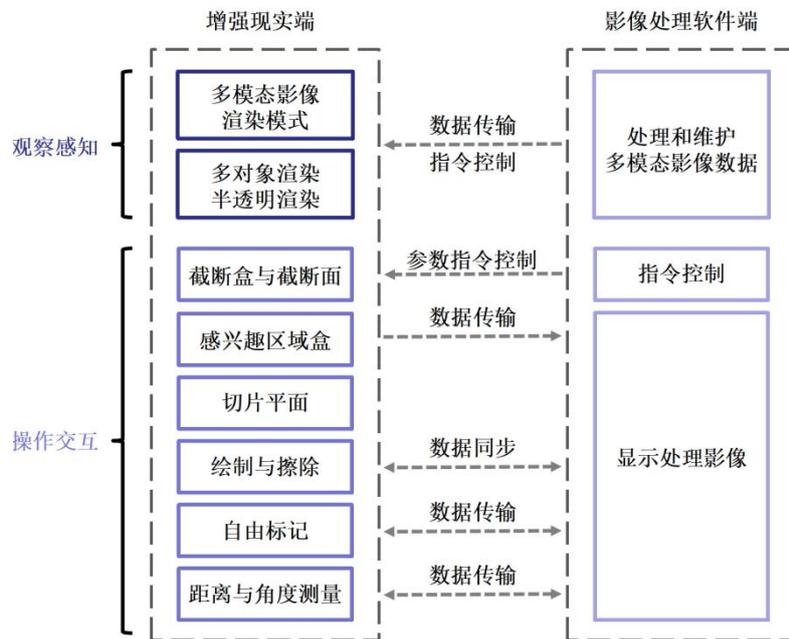


图3.4 增强现实系统和影像处理软件功能分配

本小节将针对上述存在的问题与需求提出并实现：

- (1) 相较于医学影像处理软件更精确和自由的操作方式。
- (2) 多模态渲染层次与生理结构一致的显示效果。

在方法论上，聚焦于交互层面上：

- (1) 在增强现实端实现不同模态的渲染的技术实现。
- (2) 辅助增强现实使用者进行观察的虚拟工具的技术实现。

以及协作层面上：

- (3) 影像软件端对影像展示逻辑的控制方法实现。

### 3.5.2 多模态影像渲染技术与渲染结构

#### 渲染管线

渲染管线 (Render Pipeline) 的软件概念是指一个用于描述将三维场景转为二维图像的流程与序列, 而硬件概念中描述了图形处理器 (Graphics Processing Unit, GPU) 内部的硬件架构和组件, 以及它们在生成图像过程中的协作方式。之所以渲染管线有硬件和软件的两层含义, 是因为 GPU 中有专门的处理单元与软件流程中的一些节点一一对应。GPU 是一种被设计应用于图形运算工作的硬件单元,

拥有高并行运算的能力，相比于中央处理器（Central Processing Unit, CPU）的串行处理，在实时渲染任务中有着绝对优势。

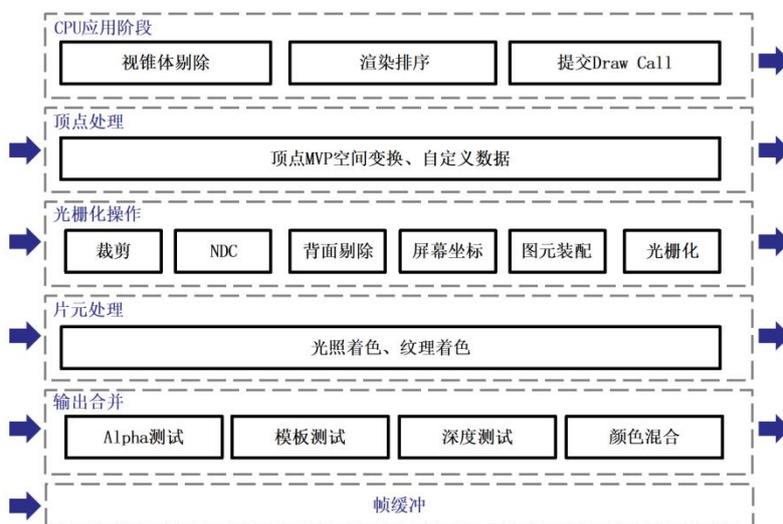


图3.5 Unity 3D的默认渲染管线

不同的渲染引擎、渲染语言有不同的渲染管线实现和配置方式。以 Unity 3D 的内置渲染管线为例，可以将该流程分为三个步骤：CPU 获取处理数据并转发给 GPU、GPU 处理数据并输送给帧缓冲区、帧缓冲区后进行后处理实现渲染目标。在 CPU 处理阶段包含了剔除视锥范围外物体、根据物体位置和渲染队列决定渲染顺序、将模型数据的参数打包发送给 GPU、调用着色器设置和执行渲染通道；GPU 处理阶段将获取的数据包通过顶点处理、光栅化操作、片元处理和输出合并的转换流程转化为屏幕画面上有指定位置和颜色信息的像素点<sup>[45]</sup>。渲染管线可以表达为图 3.5。

其中有部分结构提供了几个可以自主编程定义输入和输出和处理方式的部分，如顶点处理、片元处理和输出合并部分，被称为着色器（Shader）阶段。着色器通常可分为用于执行顶点变换和计算顶点属性的顶点着色器（Vertex Shader）、生成或修改新图元的几何着色器（Geometry Shader）、计算片段颜色、光照的片段着色器（Fragment Shader）、固定函数着色器（Fixed Function Shader）和用于管线之外进行通用计算的计算着色器（Compute Shader）。Unity 3D 封装了名为 ShaderLab 的声明性语言对着色器文件进行编写，本工作主要使用到的是顶点着色器、片段着色器和计算着色器。

## 体渲染

体渲染 (Volume Rendering) 是一种用于显示三维离散采样数据的渲染技术。体渲染的对象常常是一组以三维体数据结构存储的由 CT、MRI 等技术采集的切片数据。通过观察单个切片很难看到体数据内部的三维结构, 所以体渲染在医学影像领域的意义在于有效地可视化体数据, 而无需从数据中明确提取几何表面, 能够从不同的角度对它们进行成像, 并以一种方式遮挡不需要的部分, 从而产生表面密度或不透明度的微妙变化<sup>[46][47]</sup>。

体渲染的实现按照渲染单元可以分为图像顺序渲染 (Image Order Rendering)、对象顺序渲染 (Object Order Rendering)、基于域的渲染 (Domain-Based Rendering) 和混合渲染 (Hybrid Rendering) 技术<sup>[48]</sup>。其中图像顺序渲染指的是使用光学模型将体素值映射到颜色、不透明度等光学属性, 遍历计算每个体素对该像素的影响并获得最终的像素值, 这也是本系统中使用的渲染技术。实现该渲染技术的算法为光线步进 (Ray Marching)。

光线步进是一种简化的光线追踪 (Ray Tracing) 的算法, 不同于现实中光线从光源发出后经物体反射进入相机或视网膜, 该算法模型中光线从相机发出。光线被分为  $N$  步, 每一个样本点处的光学属性都会被纳入最终的计算。我们将采用该算法实现在三维医学影像显示中较为有意义的三种体渲染模式: 直接体渲染 (Direct Volume Rendering, DVR), 间接体渲染 (Indirect Volume Rendering, IVR) 和最大密度投影 (Maximum Intensity Projection, MIP)。

在进一步说明每一种渲染模式前, 首先要阐述数据结构。本工作中使用的体数据是一个大小为  $A \times B \times C$  的三维数据, 每一个体素存储了一个强度值或又称作密度值 (Intensity), 在不同影响模态中对应着不同的物理含义。我们希望影像能以彩色的方式显示, 所以需要将该值通过传递函数 (Transfer Function, TF) 进行升维, 表达为:

$$(Color_R, Color_G, Color_B, \alpha) = TransferFunction(Intensity) \quad (3-1)$$

经过映射后每个体素拥有了颜色和透明度的光学属性。

在 DVR 中, 考虑一条从相机发出并与体数据相交的光线  $L$  和交点  $\{P_i\}_M$ 。实际上在复杂的 DVR 中常常需要将反射、折射纳入计算, 但这些考量使得服务器计算负担重, 可能影响到帧率。而光线步进算法只考虑了一条直线上的有限采样点, 设  $P_i$  对应的光学属性为  $(C_i, \alpha_i)$ , 其中  $C$  代表 RGB 三个颜色通道,  $\alpha$  代表不

透明度。每一次步进都会进行一次光学属性的累计，也被称作透明度混合 (Alpha Blending)，从而实现半透明效果。

$$\begin{cases} C_{i+1} = \alpha_i C_i + (1 - \alpha_i) C_{i+1} \\ \alpha_{i+1} = \alpha_i + (1 - \alpha_i) \alpha_{i+1} \end{cases} \quad (3-2)$$

IVR 又被称作表面渲染 (Surface Rendering, SR) 或等值面渲染 (Iso-surface Rendering, IR)。当射线  $L$  击中密度高于设定的阈值的第一个体素时返回该像素的光学属性。数学表示为:

$$(C, \alpha) = (C_{\min(i)}, \alpha_{\min(i)}) \text{ and } Intensity_{\min(i)} > IntensityThreshold \quad (3-3)$$

在计算出等值面后再计算面上每个点的法向量，使得体数据拥有了类似于面数据的特性。该种渲染模式在计算消耗上相较而言较小，但同时也损失了部分内部的信息。

MIP 可被认为是一种简单的 DVR，最终的透明度取值为射线  $L$  上最高的密度值。数学表达为:

$$(C, \alpha) = (1, 1, 1, \max(Intensity_i)) \quad (3-3)$$

注意到 MIP 模式下并不考虑颜色，颜色被默认置为了白色。

三种渲染模式各有优劣，本工作中最常用的是 DVR，因为该渲染模式很好地结合了颜色和透明度信息，能够传达较多的信息量。

三种不同的渲染模式在渲染引擎和增强现实显示器中的显示效果如图 3.6 所示。

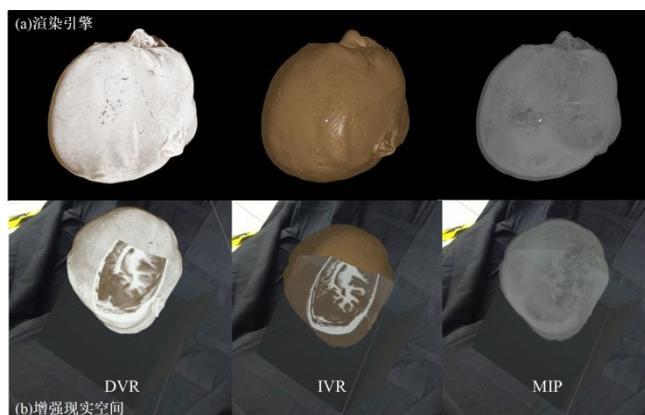


图3.6 三种不同的渲染模式在渲染引擎和增强现实显示器中的显示效果

但这仍不能满足医学影像渲染要求，如本节引言所述，生理结构是分层的，在观察和操作层面都有按解剖层次将体素分类的需求。在观察医学影像时期望能将感兴趣区域的集合以指定的颜色标注，而传递函数只能将强度值一样的体素的集合赋值为相同的颜色，故我们又引入了一个新的赋色规则：标签图 (Label Map)。

标签图以前三维和影像大小相同的  $A \times B \times C \times 4$  大小存储，4 代表  $RGB\alpha$  四通道，可以被认为是一种基于位置的传递函数。考虑了标签图的体素光学属性应为：

$$(C, \alpha) = (C, \alpha) \otimes (C_{LabelMap}, \alpha_{LabelMap}) \quad (3-4)$$

注意到标签图效果是在前述的三种渲染模式之后执行的，所以 DVR 模式并不会把标签图的数值累加。

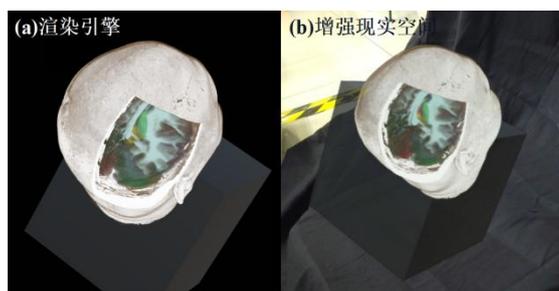


图3.7 在DVR模式下增加标签图颜色渲染

此外由于医学影像数据具有层级的生理意义，我们可以如皮肤、颅骨、脑软组织、肿瘤等具有不同生理意义的结构赋予显示优先级，使得当视角可以深入体素时保留需要的结构。由此我们引入了另一个和三维影像大小  $A \times B \times C$  相同的数据结构存储的优先级图 (Priority Map)。并设定某个显示优先级阈值，当体素的优先级大于该值时才给予渲染。

$$(C, \alpha) = \begin{cases} (C, \alpha) & Priority > PriorityThreshold \\ (C, 0) & Priority \leq PriorityThreshold \end{cases} \quad (3-5)$$

至此我们介绍了三种主要的体渲染模式和其技术原理，以及针对医学图像渲染场景加入了标签图和优先级图。上述体渲染算法均主要在片段着色器中实现。

## 面渲染

面渲染或称网格渲染 (Mesh Rendering) 是一种更常见的渲染方式, 表达的是某个表面的表面信息。面渲染的存储对象为一个表面上的顶点 (Vertex)、面片 (Fragment)、法向量 (Norm)、纹理坐标 (Texture) 和颜色等结构化信息存储。所有的面数据都由多个三角网格构成, 每一个三角网格被叫做面元。在渲染时可以对单个面元的顶点或颜色等属性进行修改, 从而获得整体的渲染结果。

面数据和体数据在增强现实空间医学影像的表达方式上两点显著差异:

- (1) 面数据缺乏深度和细节信息, 但提供了一个限制交互范围的边界。
- (2) 面数据无法提供物理同构的交互感, 只是一种物理边界的抽象。

(3) 面数据在医学影像的表达中常用于描述如头皮、血管、神经束等边界信息相较于内部细节对于观察者更重要的组织结构。在脑胶质瘤切除手术前常常需要在头皮标记切口范围, 在本系统中期望医师能在增强现实空间中于虚拟的头部上进行标记, 在规划完毕后与真实的病人头部进行配准, 实现在虚拟空间观察和操作的的结果在真实空间中呈现, 故存在对面数据进行颜色修改的需求。该技术使用计算着色器实现, 详细技术将在 3.5.9 中阐述。

在明确了体渲染和面渲染两种基本的渲染方式以及渲染管线后, 我们将进一步讨论如何用不同渲染模式表达多模态信息。

### 3.5.3 多模态影像渲染模式

在本部分讨论的是不同模态的影像数据应该由什么样的渲染模式呈现。

MRI 结构像采用了体渲染, 因为 MRI 是完整表达整个头部信息的最佳选择, 所以需要体渲染表达细节部分。同样使用体渲染的还有 fMRI 的激活图, 因为激活图是一个立体的区域, 使用面数据表达会表达错误的激活功能区的的信息。

使用面渲染的有 MRA 中提取的血管、CT 中提取的头皮、MRI 中提取的肿瘤、DTI 中提取的神经纤维束。当前体渲染精度不及面渲染, 体素边界存在不平滑的问题, 此外对于上述提到的几个生理结构最重要的信息是其边界信息, 所以面渲染可以很好地提取并平滑其边界, 向观察者传达边界所在的位置。

### 3.5.4 多对象渲染与半透明渲染

在渲染同一个场景中的多个对象时需要保证渲染顺序严格遵守生理结构层次, 尤其是在有半透明结构时遮挡关系应处理正确。着色器的设计应该保证在场景中不同的面数据和体数据以不同不透明或半透明状态组合的情境下都能正确渲染。

本部分讨论的是不同模态的影像数据在多对象渲染和半透明渲染时存在的问题以及解决方案。该问题可以被分解为两个子问题：单个对象本身渲染正确和对象间关系渲染正确。

为了实现单个对象本身渲染正确，以头皮数据为例，需要保证 (1) 从正面和反面进行观察都能看到 (2) 不透明状态下渲染顺序与离相机的远近顺序相同 (3) 透明状态下渲染顺序与离相机的远近顺序相同，并能在重合部分体现出被遮挡部分参与了颜色的混合。

需求 (1) 主要是由渲染管线中光栅化阶段中的剔除步骤决定的。为了节省渲染资源，渲染管线提供了只剔除正面 (Cull Front) 或剔除背面 (Cull Back) 的选择，从而实现只渲染一面提高渲染速率。在医学场景中需要实现双面渲染从而保证各个角度下都能正确呈现信息。尽管渲染管线提供渲染双面 (Cull Off) 的选项，但该模式下反面会被渲染为正面相对的视角画面，如图 3.8(b)所示，对于里外对称的物体这并不会带来错误，但在本文语境中显然渲染对象正反面是不一样的。

为了解决这个问题，我们引入了双通道渲染技术。渲染通道 (Rendering Pass) 的目的是将图形渲染过程划分为多个独立的步骤，以便更好地管理和控制渲染流程。每个通道独立负责实现某个效果，最终所有的通道的效果按照通道的顺序叠加。在着色器中，我们先设置了第一个渲染背面 (Cull Front) 的通道，再用其他配置相同的通道渲染正面 (Cull Back)，从而实现了双面渲染。

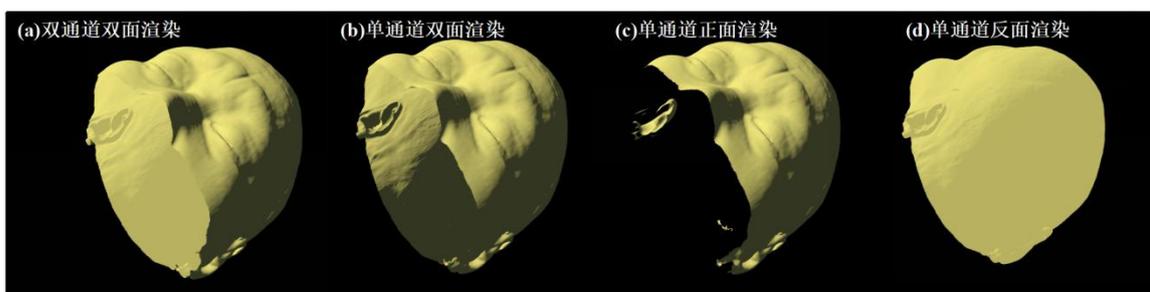


图3.8 不同的剔除模式

需求 (2) (3) 的决定性实现步骤发生在渲染管线输出合并阶段的深度测试 (ZTest) 处。在合并阶段为了存储中间结果管线设计了深度缓冲区 (ZBuffer) 和颜色缓冲区 (Color Buffer)，深度测试在逐片元比较的过程中判断每个片元的深度值和已存在于深度缓冲区的深度值的关系，如果满足通过条件并且深度写入 (ZWrite) 开启，该片元的深度值会被写入深度缓冲区。深度测试决定了片元的可见性，只有通过深度测试的片元才会被渲染，反之则被舍弃。

此外无论深度写入是否开启，只要开启了混合，通过深度测试的片元的颜色都会被写入颜色缓冲区。这一步是为了实现 (3) 而设计的，本工作中使用的混合方法是公式(3-2)呈现的透明度混合算法。通过混合存储于颜色缓冲区中的颜色值和通过深度测试的片元的颜色值就可以获得新的颜色值并更新颜色缓冲区。

本工作设计的着色器将深度测试条件设置深度值为小于或等于深度缓冲区的值时通过测试，并开启了深度写入，也就是最后渲染的是离相机最近的片元。至此我们通过深度测试、双通道渲染的方法实现了单个对象自身的正确渲染。

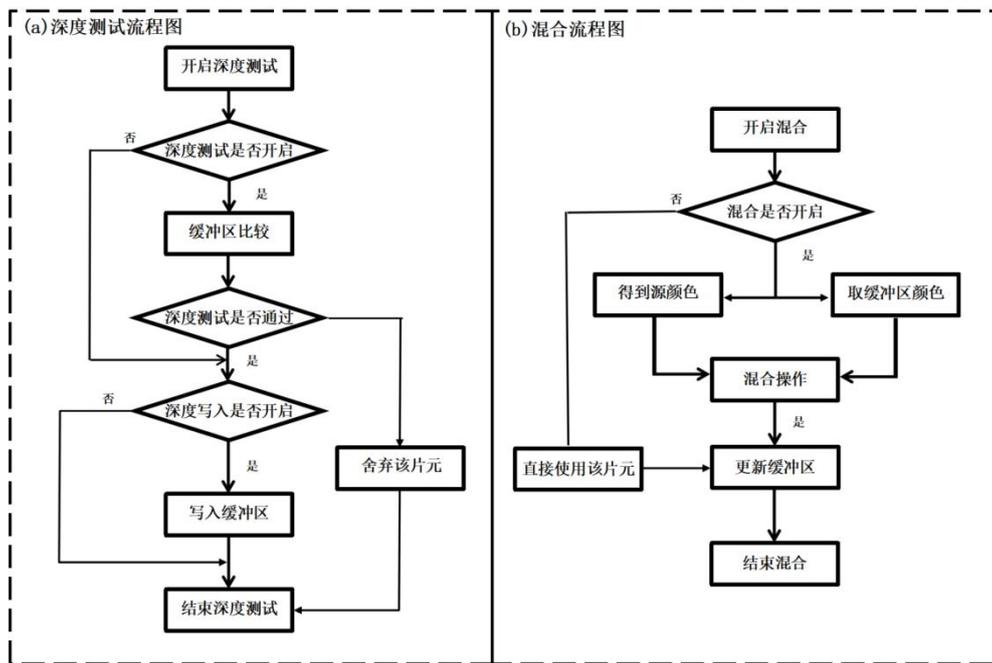


图3.9 深度测试流程图与混合流程图

为了实现多个对象间关系渲染正确，主要需要保证 (1) 不透明状态下渲染顺序与各个对象的各个部分离相机的远近顺序相同 (2) 透明状态下渲染顺序与离相机的远近顺序相同，并能在重合部分体现出被遮挡部分参与了颜色的混合。需要说明这两个需求与单个对象本身渲染不同的是存在一种特殊情况，两个对象 A 和 B 间可能存在 A 有部分位于 B 的某部分前，而 B 也存在某部分位于 A 的某部分前的情况，但在渲染管线中，深度测试使用的深度值是整个渲染对象中心点的深度值，而非单独判断每个片元的深度。所以仅使用上述方法会出现当 A 中心深度小于 B 中心深度时，B 的所有部分都会比 A 先渲染，导致 A 被 B 遮挡的部分无法正确显示。于是引入了另一个决定渲染顺序的变量：渲染队列 (Render Queue)。

渲染队列以一个整数值队列数表示，Unity 3D 提供了表 3.1 中的几个默认队列值。队列数越小越先渲染，所以队列数越大的越在视野的前方。2500 是透明和不透明的分界值。当队列数在 2500 以下时，队列数相等的对象由近到远排序；当队列数在 2500 以上时，队列数相等的对象由远到近排序。由于在医学影像渲染的场景下，被渲染对象的相对位置关系都是明确的，如头皮一定在其他所有对象的最外围，所以我们指定了血管、肿瘤等颅内组织面数据的队列数为“AlphaTest”（2450），体数据的队列数也为“AlphaTest”，而头皮的队列数为 2455，从而保证体数据和面数据同时出现时距离近的对象在视野的前方，大部分情况是体数据包住面数据，以及头皮在任何时候都位于视野的最前方。

表3.1 Unity 3D提供的默认渲染队列

渲染队列	Background	Geometry	AlphaTest	Transparent	Overlay
队列数	1000	2000	2450	3000	4000

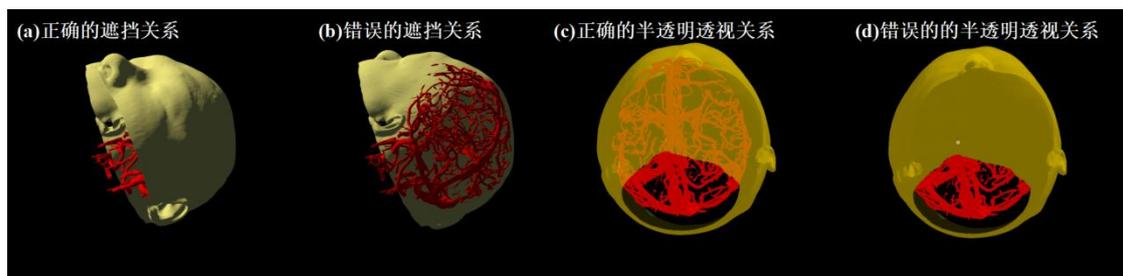


图3.10 遮挡关系与透视关系示意图

### 3.5.5 小结

本节通过设计双通道渲染、深度测试和渲染队列，实现了医学影像渲染语境下的多对象渲染、半透明渲染。并利用优先级图和标签图实现了分层渲染。3.4 中提到的多模态影像同时显示的效果的一些常用组合如图 3.11(a)(b)所示。

从下一小节开始介绍辅助增强现实设备佩戴者进行进一步观察的虚拟交互工具，为增强现实设备佩戴者赋予更多的交互主动性。

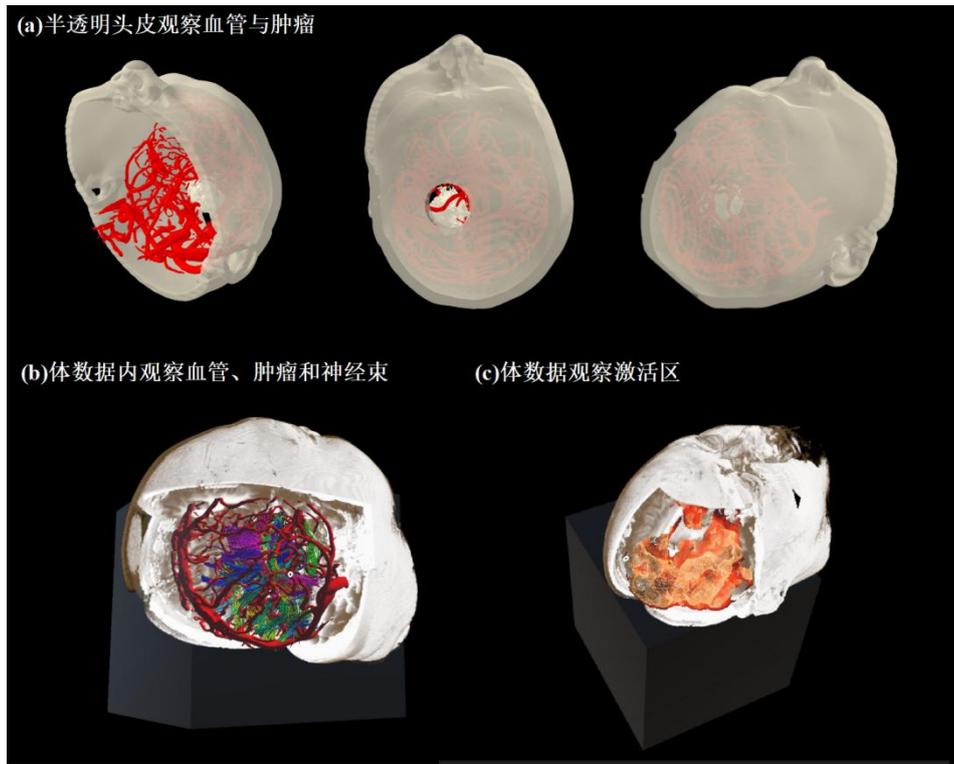


图3.11 多模态影像显示效果

### 3.5.6 截断盒和截断面

在脑胶质瘤切除手术术前规划中，存在需要自由快速观察虚拟影像内部结构而不对影像结构做修改的需求，例如透过层层组织观察肿瘤位置和形态。于是本系统引入了截断盒（Cut-off Box）和截断面工具。截断盒的视觉效果为与体数据相交或相补的部分会变为透明，从而暴露出体数据内部的细节和隐藏在体数据内部的面数据。而截断面（Cut-off Plane）的视觉效果为在截断面一侧的体数据全部透明，另一侧全部保留。

截断的原理概括来说是欲显示为透明的部分不予渲染。首先判断截断盒或截断面中心当前在世界坐标系下的位置，并通过截断盒的边长信息或截断面的  $z$  坐标和参数设定计算出一个三维区域，若当前顶点在该三维区域内则直接不予渲染。这一步抉择是在片段着色器中完成的。

该截断盒可以被单手操作自由拖动，被双手操作自由旋转，从而实现在二维影像处理软件中难以完成的自由深入观察。此外本系统还为该截断盒设置了两个参数，与 3.5.2 中介绍的体数据的优先级图配合的优先级值和选择是与体数据相交区域或相补区域展示为透明的外切或内切模式。

截断盒个数和每个截断盒的优先级值可以由 3D Slicer 端控制，截断盒个数也可以由渲染引擎当地控制，只有当体素的优先级大于截断盒的优先级值才能不被切除保留渲染。其次内切或外切模式也是由 3D Slicer 端控制的，当模式为外切时，截断盒与体数据的相交区域不予渲染，以体数据为全集的相补区域正常显示；当模式为内切时，截断盒与体数据的相交区域正常显示，以体数据为全集的相补区域被切除。

通过赋予截断盒自由拖动旋转、优先级以及内切外切的属性，观察者能利用该工具直观有效地深入了解颅内的细节和不同结构的位置关系。分配给 3D Slicer 端更多的参数调整空间是因为参数的调整在增强现实端并不方便，但同样也基于了增强现实端简化版的自主控制权。

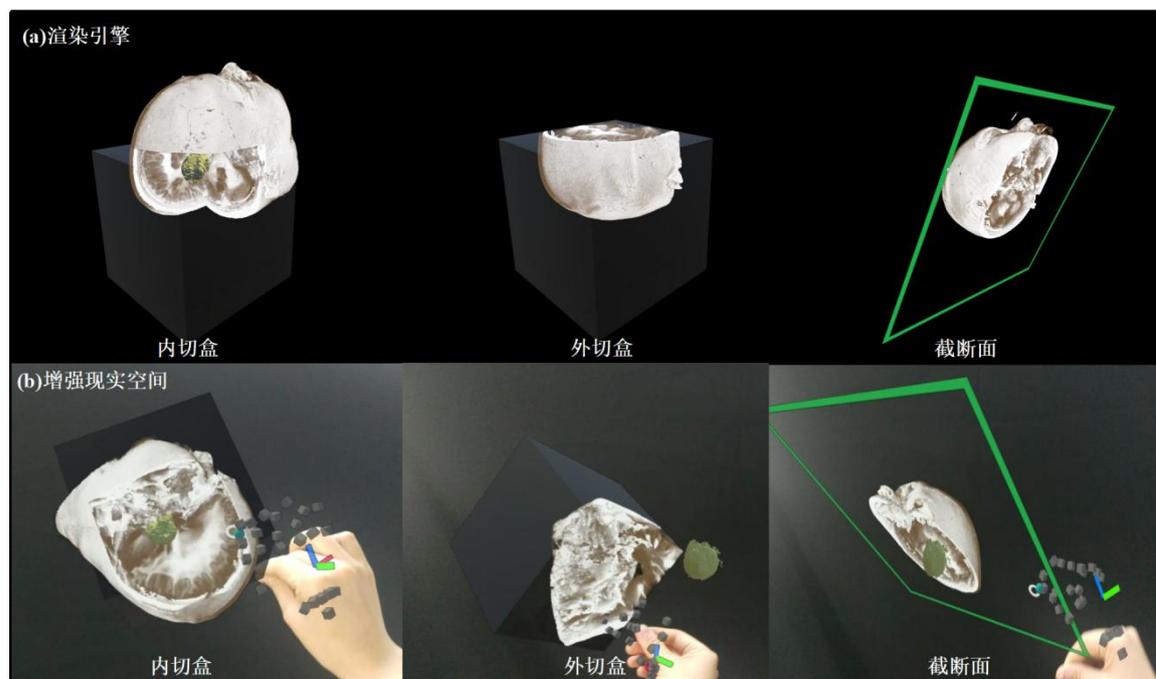


图3.12 渲染引擎和增强现实空间中的截断盒与截断面

### 3.5.7 感兴趣区域盒

感兴趣区域 (Region of Interest, ROI) 被用于选取影像中对观察者有效的部分。本系统设计的感兴趣区域盒和外切截断盒原理一致，通过只渲染框选区域内的部分实现感兴趣区域的拾取。该工具与截断盒的差异在于 (1) 可通过伸缩作用点在三个垂直方向上改变长方体区域的长宽高 (2) 在确定拾取区域后将区域内的体素传回到 3D Slicer 端进行进一步的处理。

### 3.5.8 切片平面

尽管增强现实的操作空间已经变成了三维，但有时二维的断面传达的信息同样重要。二维切面将信息压缩到一个平面上，利于去除其他干扰信息并缩减操作空间，从而提高特定操作的精度。例如当医师希望观察肿瘤离头皮的距离时，切片平面 (Slice Plane) 相比于三维显示反而能更精确地呈现该信息。

切面平面的实现原理为先取一个切面上的点，将其在世界坐标系下的坐标值转为在体数据坐标系下的值，也就是和体数据的相对坐标。再在从体数据提取的三维材质的对应坐标处采样取得对应的强度值 $v$ 。此外还设置了一个非线性映射关系，将取得的记该非线性函数中的两个参数最大值和最小值为符号 $maxv$ 和 $minv$ ，则有：

$$v' = \frac{v - minv}{maxv - minv} \quad (3-6)$$

得到最终该点的光学属性  $(v', v', v', \tau)$ 。

本系统设计了两种控制切片显示坐标的方式。第一个是通过面板上的滑动条分别控制，第二个是设计了三个相互垂直的切片平面，分别为矢状面、水平面和冠状面。在体数据上原位显示。二维面板的切面与体数据上的切面显示的内容同步，原位显示的切面可以沿垂直于切面的轴运动并随之改变二维面板的显示内容，也可以随着二维面板对应滑动条值的改变。切片平面和截断盒的结合实现了三维空间中的三维与二维混合显示，极大地提升了影像表达的信息量。该功能的控制权仅在增强现实端，因为该工具只是辅助观察的，影像软件端有类似的功能，也无需知晓增强现实端通过该工具观察到了什么中间结果，所以无需同步该工具。

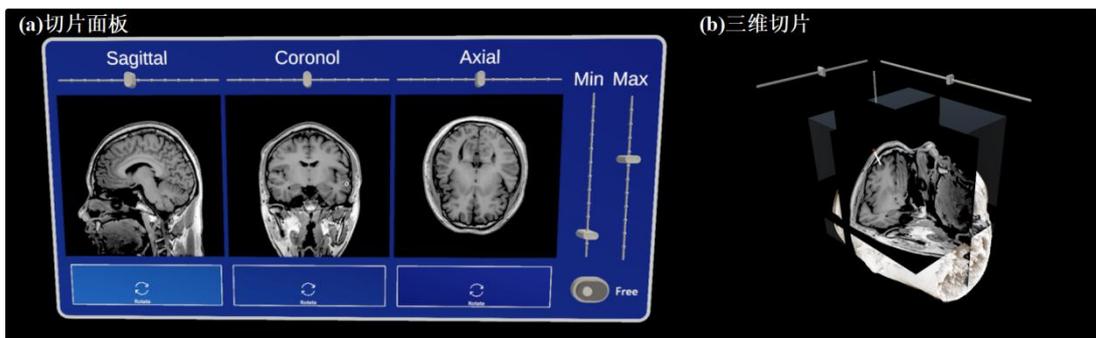


图3.13 二维切片面板和三维切片

### 3.5.9 绘制与擦除

前述的虚拟工具都是使用手势作为交互方式进行交互，从本节开始涉及对点或具体区域的精确定位，手势的交互精度不再能满足需求，故交互作用点变为第2章提出的多面体工具笔的尖端。

脑胶质瘤切除手术的术前规划的目标之一是在头皮表面标记切口，有类似于画笔的功能就尤为重要，且该画笔只应在表面作用。于是本系统提出了一种使用计算着色器控制笔尖在面数据上修改顶点颜色实现绘制与擦除的方法。计算着色器是一种用于调用 GPU 进行并行计算实现渲染管线之外的逻辑的着色器。之所以在计算着色器而非顶点着色器中进行逻辑判断的原因是计算着色器对顶点颜色的修改都会被记录并便于从 CPU 上运行的主程序中调用管理，而顶点着色器中每次修改都会被覆盖，修改记录并不能被直接获取。

我们将笔尖位置实例化为一个半径可手动调整的球体，将球体球心在世界坐标系下的坐标转换到待修改对象坐标系下的坐标，设坐标转换后球体的球心坐标为 $(x_b, y_b, z_b)$ ，待修改对象上的第 $i$ 个顶点的坐标为 $(x_i, y_i, z_i)$ ，计算两个点的欧氏距离：

$$distance = \|(x_b, y_b, z_b) - (x_i, y_i, z_i)\| \quad (3-7)$$

当两点的距离小于球体半径时，将该点的颜色和透明度更改为球体的颜色和透明度，从而实现视觉上球体给对象染色的效果。

对于擦除效果的实现，在保持上述逻辑不变的状态下，将笔刷球体的不透明度调整为 0，从而实现被擦除的部分变透明，产生该部分被删除的视觉效果（见图 3.14 (a)(b)）。

此外在体数据内部也有进行区域标记和擦除不需要的体素的需求，原理和在面数据上的逻辑一致，但修改的并不是体素本身，而是体素对应的标签图。在体素上绘制相当于在更改标签，勾画出的标签区域可以用于如粗分割等处理。

无论是修改顶点颜色或修改标签图，这个过程大部分情况下都是可逆的。对于面数据，往往只有单一的颜色，所以恢复数据只需取原色在修改过的地方进行绘制即可；对于体数据，修改的是标签图而非体素本身，故原始信息被完整保留，只需要将标签图对应的颜色和不透明度更改为 $(1, 1, 1)$ 即可恢复原样。

面数据和体数据的绘制和擦除功能提升了操作者的交互主动性和操作精准度，且绘制和擦除结果实时回传到 3D Slicer 端，使得 3D Slicer 端的使用者能观察到

在该软件中无法实现的效果（见图 3.14 (c)(d)），并能通过勾画的区域理解增强现实端使用者期望表达的意图。

对于绘制和擦除功能还有需要改变画笔大小、颜色和透明度的需求。我们设计了一个有限颜色的色盘面板，下方的滑动条为透明度调整条，右侧的滑动条调整画笔大小。从而实现对画笔的控制。

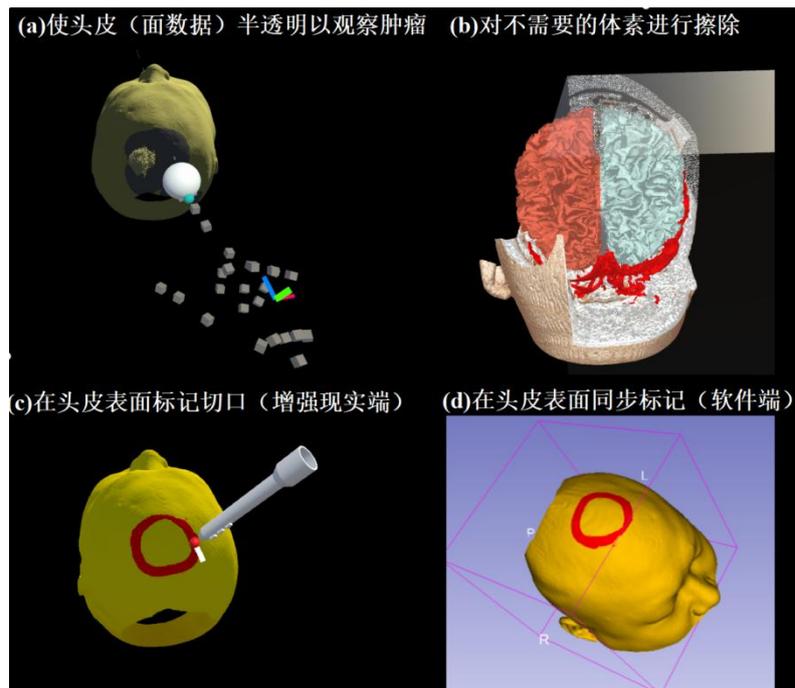


图3.14 使用画笔或手势对面数据或体数据进行修改并同步示意图

### 3.5.10 自由标记

除了在限定的面上和体内进行绘制外，有时还有在三维空间自由标记的需求。基于此我们设计了空中线绘制、点标记、基于采样点的曲线标记等三种自由标记方式。

空中线绘制的实现是通过实例化线渲染模板，设定线粗细为球直径，线颜色为球颜色，通过如按键按下为启动信号实现连续绘制和分段绘制。通过空中线绘制操作者可以在三维空间中自由标记如入路方向等记号，甚至写字。

但空中线绘制的问题在于首先 3D Slicer 中没有类似的功能，所以难以同步，此外空间线绘制是基于世界坐标系的，也就是当对象移动时线绘制仍保留在原位置。为了将三维自由标记的结果同步到 3D Slicer 且能保持和作用对象的相对位置不变，我们设计了点标记与基于点的线标记。点标记和曲线标记在 3D Slicer 中也

有对应的标记物设置,故增强现实空间中的点标记和曲线标记结果会被同步到 3D Slicer 中,以供软件进行进一步的计算。并将标记的坐标系设置为作用对象的坐标系,从而保证在对象平移旋转甚至缩放时标记的相对位置都不变。

除了把结果从 Unity 3D 传到 3D Slicer, 3D Slicer 的标记同样也可以回传到 Unity 3D。这个功能极大地弥补了影像处理软件于三维空间中的位置操作不符合空间感官的缺陷,当影像处理软件获取增强现实设备佩戴者在三维空间中的标记后还能对点的位置进行微调并回传,从而实现标记点的精准放置。

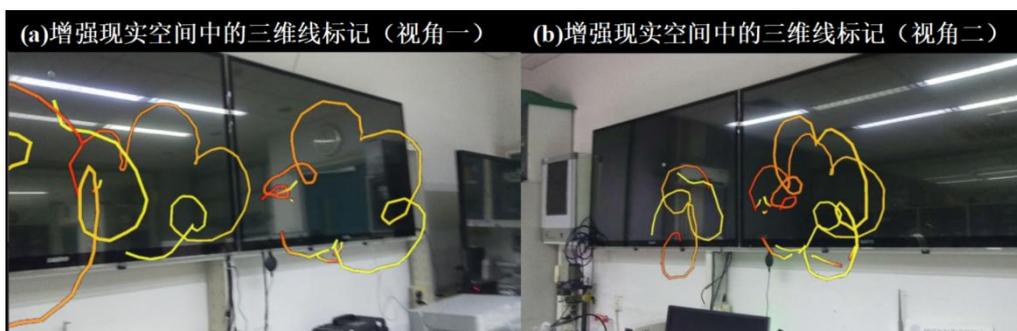


图3.15 增强现实空间中使用空中线标记

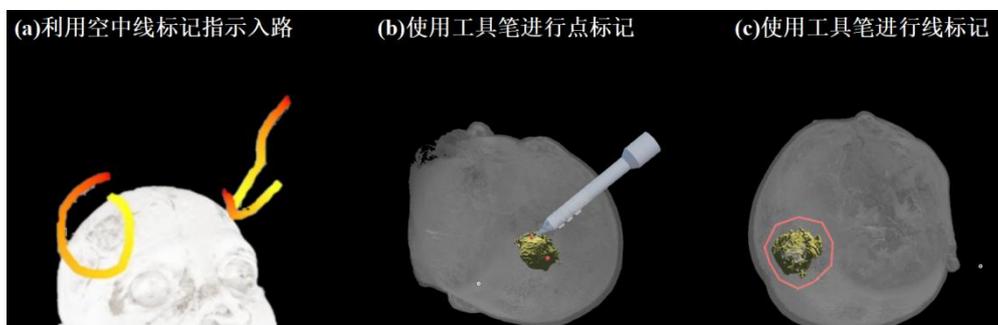


图3.16 渲染引擎中点标记和线标记示意图

### 3.5.11 距离与角度测量

尽管增强现实技术使得虚拟的医学影像能在现实空间展示,现实世界参考物的存在使得观察者评估距离和角度等物理量变得容易,但一个可以精确给出距离或角度数值的工具仍然是有必要的。我们设计了一个基于点标记的距离测试工具和角度测试工具。测量距离时需要在待测线段的两端各放置两个点标记,放置完成后两点间自动生成一条线段,并由标签给出线段以毫米为单位的长度。测量角度时依次放置三个点,放置完成后自动生成两条线段,并由标签给出两线之间夹角以度为单位的角度大小。

该功能与点标记和线标记相同，可以在增强现实端和影像软件端双向同步。

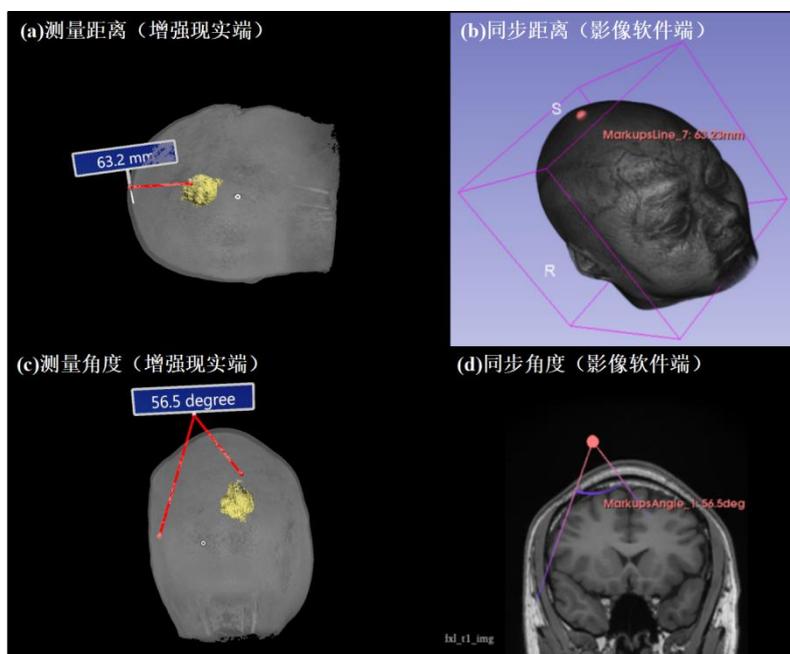


图3.17 在增强现实端测量距离和角度并在影像软件端同步

### 3.6 入路和切口规划

在医师利用本文搭建的系统对三维影像进行观察和操作后需要进行入路和切口规划。使用场景和工具设定为开放式手术下使用常规的如肿瘤镊、止血钳等手术器械将脑胶质瘤与周围组织分离并取出。入路和切口规划流程包括以下几个步骤：(1) 切口标记：使用 3.5.10 中提出的曲线标记工具绘制切口 (2) 入路生成：在头皮上切口范围内选择一点作为锚点，在肿瘤靶区内选择另一个锚点，两个锚点形成一条线段代表入路 (3) 切口模拟：以入路为轴，以切口为截面进行头皮到肿瘤点处的投影，投影体积内体素切除，保留血管、神经束等面数据。过程示意图见图 3.19 (a)-(c)。处理后的数据会从增强现实端同步到 3D Slicer 端显示，3D Slicer 也可以对标记的结果进行修改。完成这三个步骤后增强现实端和影像处理软件端的操作者都可以看到同一个显示了入路和切口标记的对象。

### 3.7 风险评估

脑胶质瘤切除手术术前规划目标为尽量避免损伤患者的感官和运动功能、尽量减少对其他包括脑部软组织、大血管在内的组织的破坏。为了衡量 3.6 中获得的规划路径对患者造成损伤的风险，本节提出了通过几项指标量化风险程度，并将该风险值以颜色表示实时可视化于入路线段上。

依据规划目标，我们提出了进行风险评估的四个评估点：

(1) 入路距十二个颅神经对的最短距离。实际上颅内存在的神经束数量庞大，但最重要的还是十二对颅神经，故只将十二对颅神经记入考量范围。颅神经对的位置和信息由 DTI 影像处理后提供。

(2) 入路距运动和语言脑功能区的最短距离。避开语言和运动区有助于降低术后失语或偏瘫的风险。脑功能区的分布由 fMRI 影像处理获得。

(3) 入路距直径大于 3mm 的血管的最短距离。开颅过程中不可避免地会破坏小血管，故只考虑直径大于 3mm 的大血管。血管模型由 MRA 模态处理得到。

(4) 头皮切口到肿瘤的距离，以入路线段的长度代表。因为期望尽可能少地破坏沿路组织，所以也期望切口到肿瘤的距离能尽量短。

受文献[50]中对术前规划进行评分的量化方式启发，本文提出了如下的风险评估量化方式。需要说明的是量化的具体方式可以由临床医师的需求进行调整和更改，本文提出的评估函数只是可能的量化方式之一，在本工作中只被用于得到和展示结果。

设路径线段的起点与终点坐标为  $P_{start} = (x_s, y_s, z_s)$  和  $P_{end} = (x_e, y_e, z_e)$ ，则该线段上可以取含端点在内等距的  $N + 1$  个采样点

$$P_i = P_{start} + \frac{(P_{end} - P_{start})}{N} i \quad 0 \leq i \leq N \quad (3-8)$$

(1) - (3) 中的风险源可以被抽象为点的集合。对于血管和神经纤维束的管状面数据，考虑到期望避开的是其表面而非中心点，故计算的是到表面的最短距离而非几何中心的距离。分别以符号  $d_{nerve}$ 、 $d_{function}$ 、 $d_{vessel}$  表示待计算入路点到神经束、功能区 and 血管的最短距离。最短距离的求解过程可以看作求解  $P_i$  到神经纤维束表面点的集合  $\{P_{nerve}\}$ ，激活图唤醒程度大于指定阈值的体素点的集合  $\{P_{function}\}$ ，血管表面点的集合  $\{P_{vessel}\}$  的最短距离。求解最短距离使用的是 KD 树 (K-Dimensional Tree) 算法，该算法能在巨大的数据量下快速求解最近

邻, 通过划分数据空间减小搜索范围, 提高搜索效率。而设计的入路长度以  $d_{scalp}$  表示。

又设 (1) - (4) 对应的风险函数为  $f(d_{nerve})$ 、 $f(d_{function})$ 、 $f(d_{vessel})$ 、 $f(d_{scalp})$ 。

对于点  $P_i$ , 由于不同的脑胶质瘤切除手术采用的手术器械、具体操作不同, 风险函数也会有所差异, 图 3.18 提供了两种可能的风险函数的表达。本质上风险函数的设计需要遵从临床事实并结合临床经验, 如对于神经、血管、功能区, 入路离这些组织和区域的距离越远, 风险越小, 故图 3.18(a) 设置了一个负线性函数; 而入路长度期望越小越好, 但通常该长度都有一个可以接受的取值范围, 故  $d_{min}$  并未取到原点, 同时也设置了  $d_{max}$ 。

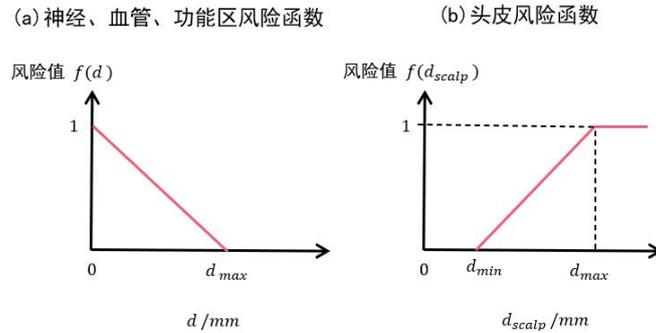


图3.18 一种可能的风险函数的表达

通过计算  $P_i$  对应的四个风险函数, 则  $P_i$  处的风险值表达为:

$$RiskValue_i = w_1 f(d_{nerve}) + w_2 f(d_{function}) + w_3 f(d_{vessel}) + w_4 f(d_{scalp}) \quad (3-9)$$

其中  $w_1 + w_2 + w_3 + w_4 = 1$ 。注意到  $d_{scalp}$  对于任意点  $P_i$  都是相同的数值, 而得到的  $RiskValue_i$  取值在 0 到 1 间。此外若有模态的数据缺失, 可以将对应的权重值置为 0, 即不参与到风险值的考虑中。

整个入路的平均风险值表达为:

$$MeanRiskValue = \frac{1}{N} \sum_{N+1}^i RiskValue_i \quad (3-10)$$

该数值会在影像处理端显示。为了让增强现实端的观察者更直观地感受该风险值, 我们用红黄颜色映射的方式表达数值。已知纯红色的 RGB 值为  $(255,0,0)$ , 纯黄色的 RGB 值为  $(255,255,0)$ , 若以黄色代表风险值低, 红色代表风险值高, 则  $P_i$  处的颜色值应为:

$$\begin{aligned}
 Color_i &= (255,255,0) - [(255,255,0) - (255,0,0)] \\
 &\quad \times RiskValue_i \\
 &= (255,255 \times (1 - RiskValue_i), 0)
 \end{aligned}
 \tag{3-11}$$

风险可视化示意图如图 3.19(d)所示，对现有的医学影像数据进行评估结果生成的部分流程如图 3.20 所示。该评估考量了血管、神经束、功能区和入路长度，设置权重相等，另一参数  $N=10$ ，故一共有 11 个点。红色的模型为血管，彩色的为神经束，半透明的为功能区，绿色的为肿瘤，图中已经绘制出一条 36.73mm 长的入路线段。生成结果的过程用时 0.592s，总风险值为 0.483，每个点的风险值见表 3.2，点序号按离肿瘤最近的点编号为 0，按距离依次编号的规则赋值。可见本案例中靠近肿瘤的点风险值越、较高。

使用者可根据实际需求更改权重、风险函数的参数甚至风险函数，以及可视化参数中生成路径的粗细（模拟工具直径）和采样点数量。

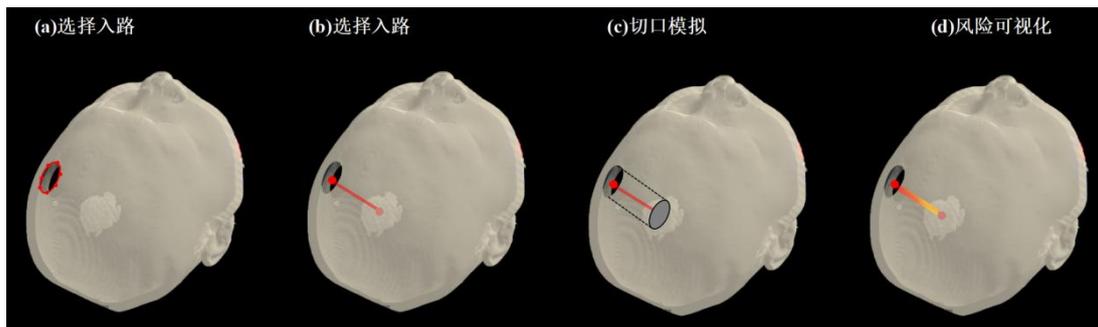


图3.19 路径规划和风险评估流程示意图

表3.2 入路上11个点对应的风险值

点序号	0	1	2	3	4	5	6	7	8	9	10
风险值	0.539	0.550	0.563	0.559	0.537	0.511	0.485	0.453	0.415	0.373	0.333

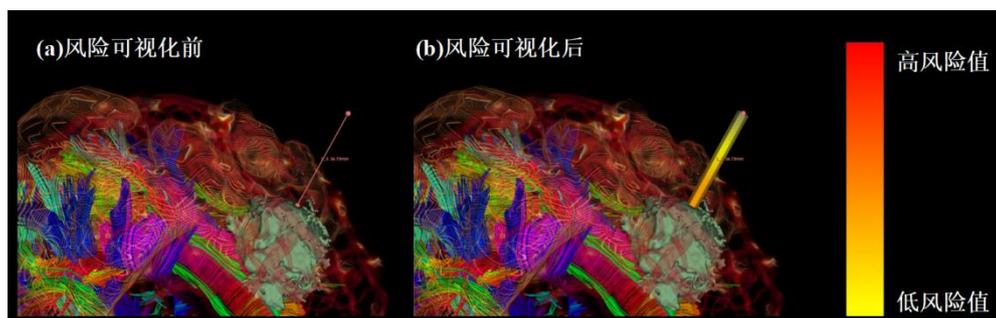


图3.20 风险可视化过程

### 3.8 影像处理软件与增强现实端的协作

3.4 至 3.7 节以开放式脑胶质瘤切除手术术前规划的手术流程为主线介绍了每个步骤中的关键功能和实现技术。3.5 节中以增强现实端的观察感知和操作交互两大功能的具体实现展开为各个子节，展示了增强现实端如何表达医学影像数据并针对不同的规划目标如何设计了不同的方式与虚拟影像进行交互，并在每小节简单提到 3D Slicer 端对该功能的控制逻辑。3.6 和 3.7 节将 3.5 节中的部分功能应用于了切口和入路设计的特定任务中，并提出了量化入路风险的算法。本节即将单独讨论影像处理软件端和增强现实端在每个任务发生时的通信过程。

在处理术前多模态影像这一步 3D Slicer 既可以依托丰富的影像处理功能从获取的影像中提取出有效信息，也可以作为已经通过其他渠道处理好的影像数据的维护器。

渲染引擎中的多模态影像是由 3D Slicer 发送的。3D Slicer 将数据转换为特定的格式后通过网络传输或内存读写传送到 Unity 3D。这一步的主动权完全在 3D Slicer 上，且 3D Slicer 可以实时添加、删除、显示、隐藏或更新面数据、体数据、标签图或优先级图。

截断盒和截断面由于有两个可调参数，不便于在增强现实空间中选择，故参数选择权也交由 3D Slicer。不同于前一个功能中两者的通信是数据传输，这一步中的通信是指令传输。一个结构化的指令中包含了参数信息，Unity 3D 通过网络传输接受到该结构化指令后进行解析，在本地实例化了相应的截断盒或截断面。

感兴趣区域盒是由增强现实端本地控制的，该工具的初始化是通过触发虚拟空间中的二维界面上的功能按钮实现的。当感兴趣区域选择完毕后，增强现实端操作者通过二维界面上的确认按钮将感兴趣区域内的体数据传到 3D Slicer 端。这

样分配控制权的原因是，在增强现实空间中有自由的观察角度，且用手势选择边界的过程相比于鼠标操作更贴合人的自然习惯，所以我们认为在该操作三维空间中执行比二维空间更有优势。

切片平面是仅部署于增强现实端的，正如前文所述，该功能无法传达使用者的意图，且 3D Slicer 中也有相应的功能，故没有传输的必要性。

绘制与擦除、自由标记和距离和角度测量等功能是在 Unity 3D 与 3D Slicer 间双向传输的，也就是两端的修改都会被同步到另一端。不同的是绘制与擦除的同步是实时的，而自由标记和距离和角度测量的同步需要手动进行。常常使用的传输方向是 Unity 3D 到 3D Slicer，这是因为在三维空间中进行绘制和擦除的精准度相较于将二维平面的操作轨迹投射到三维空间中更高。然而在一些极为细节的部位使用鼠标能定位到更精细的区域，或有对整个数据进行统一修改的需求，所以也需要传输 3D Slicer 的修改结果传输到增强现实端。

本节中探讨了引言中提出的如何将不同的操作类型分配到二维和三维空间中以发挥其最大优势的问题。我们得到的结论是有的操作如截断盒观察是更适合三维空间的，而绘制擦除、自由标记等操作在二维和三维空间中各有优劣，需要相互配合以达到最优效果。这也是本章提出的“协作”的概念核心，协作不仅是在系统层次而言增强现实用户端和影像处理软件端的配合，更是一种将指定问题分配给高效处理者的策略。

### 3.9 总结与讨论

本章提出的基于多模态术前影像的术前规划方法中关键概念有两个：交互与协作。在交互方面，本系统通过以术前规划场景中需要的技术功能为分类依据，将本任务中的交互分为了基础交互、数据元交互和影像处理三层次，并将基础交互、数据元交互的处理核心指定为渲染服务器，而影像处理交由影像数据处理器执行。而协作在本文的语境下指增强现实头戴式设备佩戴者和影像数据处理器的操作者之间为了达到完成术前规划流程在同一任务上依据增强现实和影像处理软件的优势被分配了不同的操作权限，提高了操作的效率。

本章将脑胶质瘤切除手术规划流程划分为获取和处理术前多模态影像、观察影像建立感知、入路和切口规划与风险评估四部分。在观察影像建立感知部分介绍了如何使用双通道渲染和深度测试正确显示虚拟的三维医学影像，并介绍了六种辅助观察的虚拟工具及其技术原理，结合手势交互和第 2 章提供精准交互点的

定位工具交互方式在三维空间中对影像进行高自由度精准的观察和修改。在入路和切口规划部分利用虚拟工具进行切口和入路的可视化。在风险评估部分将风险因素量化，并将风险值可视化，以提供手术方案的评估反馈。

本章的主要成果和创新点为：

- (1) 提出并实现了多模态医学影像渲染语境下的渲染方式。
- (2) 实现了针对多模态医学影像的虚拟交互工具。
- (3) 提出了完整的基于本工作提出的交互方式的脑胶质瘤手术术前规划流程。
- (4) 探讨了不同操作和交互方式在二维空间和三维空间中的优势与劣势，并依据此建立了影像处理软件和增强现实系统的协作权限的分配关系。创新性地结合 3D Slicer 和 Unity 3D 构建了用于医学场景的协作交互系统，提高了术前交互的效率，并相较于影像处理软件能更好地增强医师对手术案例的感知与认知。

## 第 4 章 增强现实术前规划系统构建

### 4.1 引言

本文第 2 章提出了一个基于红外被动标记的精准多面体交互工具及算法，实现了使用增强现实头戴式设备 HoloLens 2 的深度相机实时追踪交互工具的位姿并定位尖端坐标，为系统提供了精准交互作用点。第 3 章提出了一套完整的基于多模态术前影像以脑胶质瘤切除手术为应用场景的术前规划协作和交互流程，针对手术场景和临床需求设计了针对多模态医学影像的渲染模式及辅助观察的虚拟交互工具，构建了完善的影像处理软件和增强现实系统的协作流程，为系统提供了高效率的交互方式。本章将叙述本文提出的增强现实术前规划系统中包括增强现实头戴式设备 HoloLens 2、第 2 章制作的交互工具和第 3 章渲染服务器、影像处理器在内的模块各自的结构与功能以及模块之间的数据和指令通信过程。

### 4.2 系统框架

如图 4.1 所示，系统可以分为增强现实显示器和图形计算服务器两部分。增强现实显示器部分的主要功能是形成虚拟影像和与虚拟影像进行交互，包含了第 2 章设计的交互工具和既有定位功能又有虚拟显示功能的头戴式设备 HoloLens 2 两个模块。图形计算服务器承载的功能是对渲染、交互和协作逻辑进行运算与执行，包含了渲染服务器和影像数据处理器。

下文将按照模块顺序展开叙述每个部分的功能和模块与模块间的连接方式。

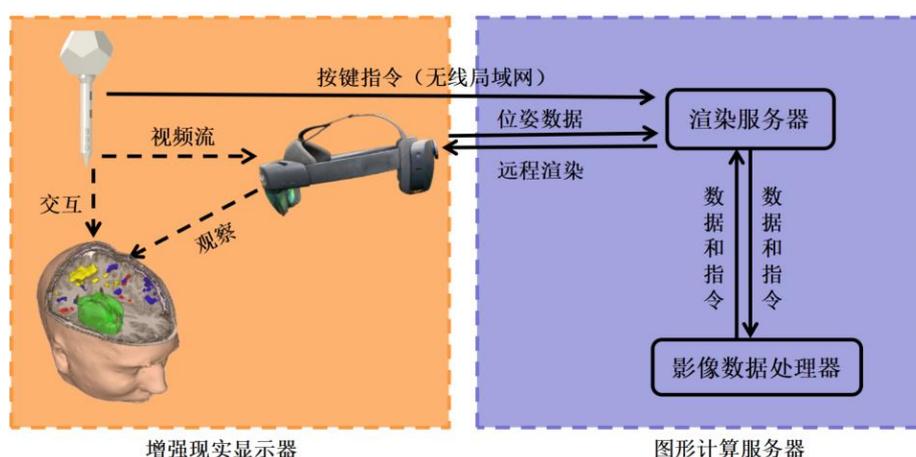


图4.1 系统框图

## 4.3 交互工具

第 2 章提出的交互工具除了具备可被定位的尾部标记物，还拥有一块可以连接局域网并配备三个按压式按键的电路板，作为主动交互的硬件支持。本节将探讨的话题主要有四个（1）向渲染服务器发送按键状态（2）按键的功能分配（3）功能界面的设计（4）虚像工具的必要性。

### 4.3.1 按键状态传输

向渲染服务器发送按键状态的过程包含三个步骤（1）读取当前按键状态（2）将三个按键的状态编码为数据包（3）发送数据包。

我们采用 Arduino 编译器对 C 语言程序进行编译烧录。首先通过读取按键连接的 GPIO 引脚的电平获取按键状态，再将三个按键状态编码为一个大小为 3 比特的比特串，以笔端为服务端，通过 TCP 协议向指定接口发布数据。以交互工具作为服务端而非客户端的优势在于若使用过程中断连，渲染服务器作为客户端有足够灵活的应对机制重连。且渲染服务器不同会导致连接的 IP 不同，若以交互工具作为客户端，需要重新烧录程序更改请求连接的 IP 地址。而以交互工具作为服务端则只需在渲染服务器中进行灵活修改即可。

### 4.3.2 按键功能分配

在 3.5 中我们提出了六个辅助观察的交互工具，其中前三个由于主要是对平面操作，主要由手势进行交互。而后三个涉及到点操作，需要确定一个小范围进行精准操作，故由笔尖作为交互点。我们期望常用的功能可以通过按键唤起，常常是以按下的状态判断，且操作容易记忆。所以需要将绘制、擦除、标记点、标记线、测量距离、测量角度六个功能分配给不同的按键状态。我们设计了一个状态表表示在何种状态下功能会被唤起。由于单个按键按下只有三个状态，但有六个功能，且单个功能的实现往往需要多个按键状态的配合，我们引入了功能面板扩展状态。

每一个状态都由按键状态和面板状态构成，按键状态包括按下和松开两种，面板状态包括无状态、标记点状态、标记线状态、测量距离状态和测量角度状态。可以法向面板状态尽管是不同的功能，但基础操作都是放置点，但连接点的逻辑有所不同。

表4.1 面板状态为无状态时按键状态改变触发的功能

按键状态	按键 1 按下	按键 2 按下	按键 3 按下	按键松开
触发功能	绘制	擦除	空中线绘制	无

表4.2 面板状态为其他状态时按键状态改变触发的功能

按键状态	按键 1 按下	按键 2 按下	按键 3 按下	按键松开
触发功能	标记点跟随	标记点跟随	标记点跟随	标记点放置

### 4.3.3 虚像工具的必要性

最后我们选择了在虚拟空间叠加笔的模型，并使其与真实的笔重合，并有一个球体指示笔尖位置。尽管增强现实空间中有真实物体作为参考，但与虚拟物体交互的作用点仍在虚拟空间下，故需要有虚拟模型作为视觉参考。真实物体的意义在于产生了真实的体感，虚拟模型的移动与旋转和抓握的笔一致，这样就能建立起虚实空间的心理映射。此外在物理按键按下时，虚拟模型相应的按键处会以颜色变化提示按压操作有效。这些虚拟模型上的设计都能极大地提升使用者的交互感。



图4.2 虚像工具

## 4.4 增强现实显示器

增强现实显示器 HoloLens 2 是整个系统的核心部分之一，在本系统中 HoloLens 2 的主要功能有 (1) 获取含有交互工具标记物的深度图和强度图并进行位姿估算并发送给渲染服务器 (2) 获取渲染服务器形成的虚拟图像并显示。

#### 4.4.1 位姿数据的获取与传输

通过调用 HoloLens 2 平台提供的深度传感器接口获取深度图和强度图，并将 2.3.7 中提出的定位算法以 C++ 程序实现，在 HoloLens 2 本地获得每一帧的工具位姿。位姿矩阵通过局域网向渲染服务器的指定端口发送。

#### 4.4.2 远程渲染

由于 HoloLens 2 的内存只有 4GB，难以承载本系统大量的数据交换过程，而 Azure 远程渲染技术可以被用于渲染 GPU 密集型交互式三维内容，低延迟将渲染画面流式传输到混合现实设备<sup>[49]</sup>。故我们采用了将渲染计算任务交给渲染服务器，而 HoloLens 2 只负责接受视频流并显示及交互输入，从而减小显示器的计算负担。

### 4.5 影像数据处理

影像数据处理器在本系统中起到的作用是逻辑计算和数据处理的核心组件，也是维护和其他各个组件连接的总服务器。本节将讨论影像数据处理器和增强现实显示器的控制关系，和影像处理器间建立数据和指令通道与交互界面的设计。

#### 4.5.1 增强现实显示器指令通道

3D Slicer 与 Unity 3D 之间传输面数据、体数据和指令的通信方式已在 4.5.2 中详述。通过 ZeroMQ 协议向 Unity 3D 发送工具定义和开始追踪定位的指令并转发给 HoloLens 2，总而保证了在更换工具时无需再向 HoloLens 2 重新烧录程序，提高了应用的可扩展性，且不需要追踪时关闭深度传感器，节省了运算资源。

#### 4.5.2 数据与指令通道

和影像处理器间建立的通信通道要满足延时低的需求。对于以字典格式存在的指令和面数据等轻量级的数据，我们采用了轻量级的通信协议框架 ZeroMQ<sup>®</sup> 进行传输。实验环境所使用的无线路由器型号为华硕 ROG GT-AX11000，最大带宽为 11000Mbps，通常场景中使用的最大的面数据不超过 100 MB，在不考虑其他设备分带宽、网络传输协议有延时等理想条件下传输该面数据最快只需要 20ms 不到。人眼能感受到连续的最低帧率为 24-30 fps 左右，显然该带宽下是完全可以满足实时更新的要求的。

---

<sup>®</sup> <https://zeromq.org/>

然而对于体数据，通常大小在  $200 \times 200 \times 200$  到  $500 \times 500 \times 500$  的范围内，每个体素由一个 32 位的 float 值存储，则体数据的大小范围在 30MB-500MB 内。除了体素本身需要被传输之外，我们还需要传输大小为体素四倍的标签图和同等大小的优先级图，故最终的传输大小在 150MB-2500MB 左右。显然此时即使在理想条件下也难以达到高于人眼刷新率的帧率，且可能由于数据量过大导致传输不稳定。于是我们采用了直接通过内存读写的方式传输数据和手动更新两种优化方法。大部分内存的读写速度都是 GB 量级，相比于网络传输而言更可靠稳定，但仍然会产生些许可以感知的延时，所以体数据不再采用实时更新的策略，而是在需要的时候进行请求获取最新的更改数据。内存读写具体使用的技术是内存映射，当数据写入内存后内存映射函数返回内存指针，只需再发送一条包含内存指针的指令告知数据存在的地址即可在另一端将数据读取出来。

### 4.5.3 用户界面

交互逻辑还应以用户界面的形式进行展示，本节着重描述我们设计的 3D Slicer 扩展界面。3D Slicer 作为一款开源软件支持用 Python 和 Qt 编写插件逻辑和界面，封装了现有功能的接口并支持用户实现定制的功能。3.8 中我们总结了影像处理软件与增强现实端的协作方式，为了实现 3D Slicer 对 Unity 3D 的控制，我们设计了如下面板。



图4.3 体数据管理界面和面数据管理界面

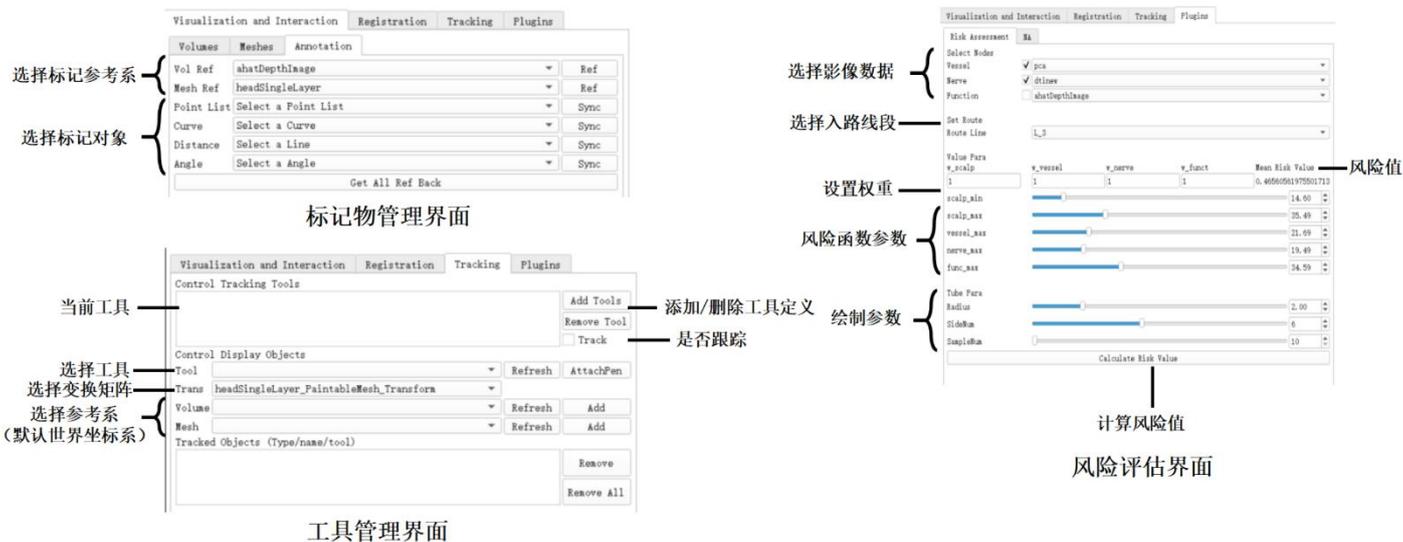


图4.4 标记管理界面、工具管理界面和风险评估界面

## 4.6 渲染服务器

Unity 3D 作为一款游戏开发引擎，提供了丰富的交互功能和渲染模型。在本系统中通过利用 MRTK (Mixed Reality Toolkit) 混合显示开发工具包<sup>①</sup>我们构建了一个增强现实开发环境，并利用 C#脚本控制交互逻辑，ShaderLab 脚本控制渲染逻辑，实现场景的构建。

渲染服务器按照 4.5.2 中的描述建立相应的配置的传输接口发送和获取数据。

在 Unity 3D 端我们需要解决的问题还有 (1) 接受交互工具发送的按键状态 (2) 虚拟二维用户界面的设计。

### 4.6.1 按键状态接收

2.2.3 中提到交互工具在通信中作为服务端，故 Unity 3D 作为客户端监听交互工具 IP 指定的端口，从而获取打包数据并解析，供后续逻辑分析使用。

### 4.6.2 用户界面

增强现实系统按照功能分类设置了三个面板：体数据控制面板、标记面板和交互控制面板。分别用于体数据控制、虚拟标记物控制和交互控制 (见图 4.5)。

<sup>①</sup> <https://learn.microsoft.com/zh/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05>

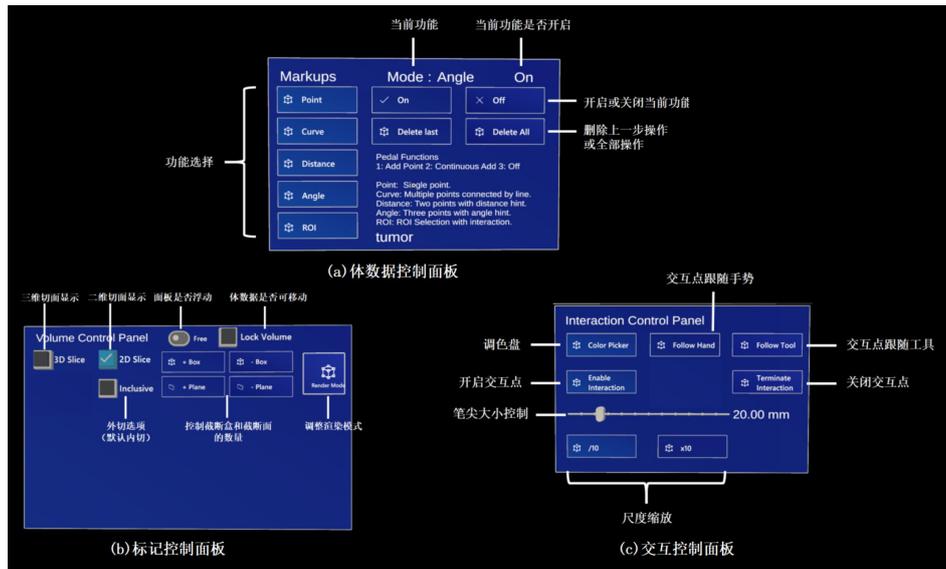


图4.5 三个控制面板示意图

## 4.7 系统验证

### 4.7.1 交互工具精度验证

为比较第 2 章提出的交互工具和手势交互方法在本系统中的精度，以验证交互工具更能胜任标记等需要精细操作的任务，我们设计了以下实验。

#### 实验设置

实验在增强现实空间中设置了三列在空间中静止且相互垂直、间隔均为 3cm 的白色小球，每个轴上小球的数量为 11 个。实验分为手势组和工具组两组，手势组通过捏合食指与拇指控制离捏合点 2cm 的红色小球，工具组通过追踪第 2 章中提出的工具控制工具尖端的红色小球，被试分别通过两种方式控制红色小球与白色小球尽量重合，记录重合时的目标白色小球和控制红色小球的位置坐标。每个实验组做两次重复实验。实验结束后被试填写性别、年龄等个人信息与对手势控制和交互笔控制的准确度体验与偏好度的五点李克特评分量表。

需要额外说明的三点为上述球的半径均为 1cm；在实验前被试均做过眼动矫正；实验中被试均使用优势手进行控制。

我们提出的假设是：

$H_1$ : 在该任务中交互工具控制的精度优于手势控制精度

$H_2$ : 在该任务中受试对交互工具的偏好性大于手势

## 实验结果

一共有 10 名受试参与了本实验，受试特征如表 4.3 所示。

表4.3 受试特征

特征	值	平均值±标准差	百分比
年龄 (岁)	[22-32]	24.8±3.39	NA
性别			
男	6	NA	60%
女	4	NA	40%
使用 AR 的经验			
从未使用过	4	NA	40%
使用过几次	5	NA	50%
经常使用	1	NA	10%
优势手			
右手	10	NA	100%
左手	0	NA	0%

实验结果分为两大部分：任务的客观指标分析和量表的主观分析。任务客观主要评价指标为目标距离与实际距离间的误差，所有个体数据都经过了数据筛选，即绝对值在三倍标准差外的数值被认为是异常点。

表 4.4 分析了每个被试两种不同交互方式的误差、p 值和误差降低率，工具交互误差总是小于收手势交互误差，且 p 值均小于或远远小于 0.05，可见两种交互方式间有显著差异。图 4.6 比较了组间误差，受试 2 作为异常数据被抛出。手势控制的误差为  $1.979 \pm 0.604\text{mm}$ ，工具控制误差为  $1.397 \pm 0.241\text{mm}$ 。ANOVA 分析的结果为  $F=1008.05$ ， $p \sim 0 < 0.05$ 。“~0”表示 p 值接近 0 且小于计算软件精度，被认为近乎为 0。可见组间差异显著性大，认为手势交互与工具交互在精度上存在显著差异。而综上认为工具交互带来了显著的精度和稳定度的提升。

表 4.5 为偏好性统计，受试从抓握舒适程度、可控制程度等方面进行选择，80%的受试选择了交互工具。此外用户对两者的精准度通过五点量表进行评分，手势控制分值为  $3.3 \pm 0.6749$ ，工具控制分值为  $4 \pm 0.6667$ 。

可见需要进行精细操作的任务中在精度、抖动和舒适程度等方面交互工具控制优于手势控制，假设 $H_1$ 、 $H_2$ 均成立。

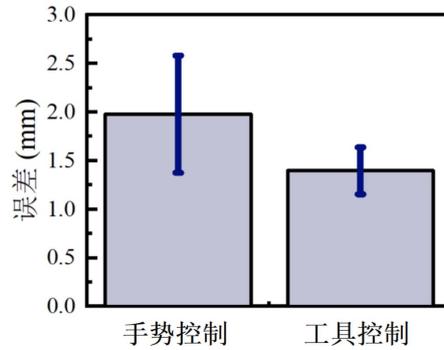


图4.6 两种控制方式组间分析的均值-方差图

表4.4 个体交互误差分析

受试编号	手势交互误差(mm)	工具交互误差(mm)	p	误差降低率
1	2.3±1.00	1.2 ±0.37	~0	49.39%
2	9.5±1.21	7.2±0.82	9.8039e-59	23.64%
3	1.4±0.50	1.3±0.61	1.5951e-06	5.34%
4	1.9±0.82	1.3±0.52	~0	29.29%
5	2.0±0.77	1.8±0.85	4.8839e-147	11.67%
6	1.4±0.45	1.1±0.33	~0	22.32%
7	2.6±1.50	1.5±0.63	~0	42.13%
8	1.5±0.53	1.5±0.67	0.0373	1.01%
9	1.6±0.58	1.2±0.33	7.9691e-174	23.5%
10	3.1±1.50	1.7±0.39	5.5973e-200	45.93%

表4.5 偏好性统计

特征	值	百分比
手势控制	2	20%
工具控制	8	80%

## 4.8 总结与讨论

本章介绍了如何将第2章提出的交互工具、第3章提到的渲染服务器 Unity 3D 与影像处理器 3D Slicer 整合到一个系统内, 并介绍了从系统层面每个部分的设计考量。最后对交互工具的精度进行了验证评估, 证明了本系统能提供精准交互。

本章的主要成果与创新点为:

(1) 首次做出结合 3D Slicer 和以 Unity 3D 为渲染引擎的增强现实系统的尝试, 结合了两者的优势。实现了主要功能操作结果的实时双向或单向传输, 使得增强现实设备的佩戴者和影像处理软件的使用者对同一个医学影像对象的操作被记录和共享, 每个操作都得以被分配到优势端。期望相较于医学影像处理软件和以往应用于医疗领域的增强现实系统能够提升操作效率。

(2) 利用多面体交互工具进行精细操作的控制, 提升了交互精度。

至此一个完整的基于增强现实的术前规划交互系统已经构建完毕。

## 第 5 章 总结与展望

### 5.1 论文成果与创新点

第 1 章绪论阐述了脑胶质瘤手术术前规划的相关临床背景，提出脑胶质瘤切除手术是一种对精度要求高，需要综合考量因素复杂的手术。在此章中还介绍了常用于术前规划的数字三维重建图像软件的优势与劣势，并紧接着提出增强现实系统可以弥补医学图像软件被局限于二维空间观看和操作的劣势。但目前用于医疗应用场景的增强现实系统在精度和效率上仍有可以提升的空间。

第 2 章提出并实现了一个基于红外被动标记的精准多面体交互工具，首先介绍工具的硬件构造，依次阐述了尾部标记物的结构和设计考量、控制电路功能与硬件实现和外壳设计。在如多面体面数的选择等关键设计环节提供了理论和实验两方面的结果支撑。接着开始介绍多面体工具配套的算法设计，结合理论推导和实验评估了该算法的误差、运行速率等指标。研究内容一的主要成果和创新点为：

- (1) 制作了一支用于增强现实系统的拥有主动交互功能和被动定位功能的工具笔，提高了交互工具的可扩展性。
- (2) 提出了多面体拼接定义算法，实现全方位定位。
- (3) 通过优化搜索方向和设置早停机制提升了定位速率，实现了实时定位。
- (4) 提出了完善的定义、定位算法和滤波器，并进行了多角度的性能验证，实现了精准定位，第 3 章中的交互方法提供了精准的交互作用点。

第 3 章聚焦于交互和协作方法，将脑胶质瘤切除手术规划流程划分为获取和处理术前多模态影像、观察影像建立感知、入路和切口规划与风险评估四部分。简要介绍了数据来源和处理方式，着重展开如何通过渲染技术与交互技术提升观察影像建立感知过程中效率，并针对脑胶质瘤切除手术设计了入路和切口规划以及风险评估的增强现实表达。最后总结了该流程中影像处理软件和渲染服务器的协作关系。研究内容二的主要成果和创新点为：

- (1) 提出并实现了多模态医学影像渲染语境下的渲染方式。
- (2) 实现了针对多模态医学影像的虚拟交互工具。
- (3) 提出了完整的基于本工作提出的交互方式的脑胶质瘤手术术前规划流程。
- (4) 探讨了不同操作和交互方式在二维空间和三维空间中的优势与劣势，并依据此建立了影像处理软件和增强现实系统的协作权限的分配关系。创新性地结

合 3D Slicer 和 Unity 3D 构建了用于医学场景的协作交互系统，提高了术前交互的效率，并相较于影像处理软件能更好地增强医师对手术案例的感知与认知。

第 4 章以前两个研究内容为基础，以系统的视角展开了每个子系统的功能结构与子系统间的连接方式。并在每个部分的技术难点提出了经验累计后进行的理论分析与相应的解决方案。研究内容三的主要成果和创新点为：

(1) 首次做出结合 3D Slicer 和以 Unity 3D 为渲染引擎的增强现实系统的尝试，结合了两者的优势，相较于医学影像处理软件和以往应用于医疗领域的增强现实系统能够提升操作效率。

(2) 实现了主要功能操作结果的实时双向或单向传输，使得增强现实设备的佩戴者和影像处理软件的使用者对同一个医学影像对象的操作被记录和共享，每个操作都得以被分配到优势端，使得术前规划的结果精度获得了提高。

## 5.2 论文未来工作展望

尽管系统原型已经构建完毕，但尚未步入用户实验的阶段。在未来期望能够设计合理的实验测试任务，通过量表和任务结果的统计对比证明系统精准与高效率的性质。在临床应用领域，术前规划的路径和切口的可视化结果期望能与真实的患者头部进行精准的配准，将虚拟影像上的标记转移到真实的头部。

## 插图索引

图 1.1	用于术前规划的增强现实系统 .....	7
图 1.2	本工作的三个研究内容和相互关系 .....	8
图 2.1	图案标记实现的定位工具 .....	11
图 2.2	交互工具图 .....	12
图 2.3	反射强度与角度关系预实验设置与结果 .....	13
图 2.4	场景的理想化建模 .....	15
图 2.5	尾部标记物表面结构示意图 .....	16
图 2.6	PCB 示意图 .....	17
图 2.7	算法流程图 .....	19
图 2.8	成像过程中坐标系的转换关系 .....	19
图 2.9	深度误差评估实验设置和实验结果 .....	22
图 2.10	位置特征和法向特征的示意图 .....	22
图 2.11	定义算法的流程 .....	23
图 2.12	定义算法评估实验工具与结果 .....	24
图 2.13	工具尖端中心标定实验图 .....	25
图 2.14	图匹配问题示意图 .....	27
图 2.15	定位算法和误差计算示意图 .....	29
图 2.16	标志点数量和检测帧率、误差的关系 .....	31
图 2.17	尖端距原点距离与抖动的关系 .....	32
图 2.18	定位误差评估实验 .....	33
图 2.19	三种滤波后信号和原信号的各个指标对比 .....	34

图 2.20	轨迹对比 .....	35
图 3.1	在术前规划任务中的交互分类 .....	39
图 3.2	增强现实系统和影像处理软件在术前规划流程中协作 .....	40
图 3.3	不同模态的影像图和处理后提取出的结构 .....	40
图 3.4	增强现实系统和影像处理软件功能分配 .....	43
图 3.5	Unity 3D 的默认渲染管线 .....	44
图 3.6	三种不同的渲染模式的显示效果 .....	46
图 3.7	在 DVR 模式下增加标签图颜色渲染 .....	47
图 3.8	不同的剔除模式 .....	49
图 3.9	深度测试流程图与混合流程图 .....	50
图 3.10	遮挡关系与透视关系示意图 .....	51
图 3.11	多模态影像显示效果 .....	52
图 3.12	渲染引擎和增强现实空间中的截断盒与截断面 .....	53
图 3.13	二维切片面板和三维切片 .....	54
图 3.14	使用画笔或手势对面数据或体数据进行修改并同步示意图 ...	56
图 3.15	增强现实空间中使用空中线标记 .....	57
图 3.16	渲染引擎中点标记和线标记示意图 .....	57
图 3.17	在增强现实端测量距离和角度并在影像软件端同步 .....	58
图 3.18	一种可能的风险函数的表达 .....	60
图 3.19	路径规划和风险评估流程示意图 .....	61
图 3.20	风险可视化过程 .....	62
图 4.1	系统框图 .....	65
图 4.2	虚像工具 .....	67
图 4.3	体数据管理界面和面数据管理界面 .....	69

图 4.4	标记管理界面、工具管理界面和风险评估界面 .....	70
图 4.5	三个控制面板示意图 .....	71
图 4.6	两种控制方式组间分析的均值-方差图 .....	73

## 表格索引

表 2.1 正多面体面数的面夹角、在两种情况下的夹角取值.....	15
表 2.2 实验板两个方向相对定位分布的统计学参数.....	33
表 2.3 三个方向重复定位分布的统计学参数.....	34
表 2.4 配准后轨迹与轨道形状的 RMSE.....	35
表 3.1 Unity 3D 提供的默认渲染队列.....	51
表 3.2 入路上 11 个点对应的风险值.....	61
表 4.1 面板状态为无状态时按键状态改变触发的功能.....	67
表 4.2 面板状态为其他状态时按键状态改变触发的功能.....	67
表 4.3 受试特征.....	72
表 4.4 个体交互误差分析.....	73
表 4.5 偏好性统计.....	93

## 参考文献

- [1] Goodenberger M K L, Jenkins R B. Genetics of adult glioma[J]. *Cancer genetics*, 2012, 205(12): 613-621.
- [2] 中国抗癌协会. 《中国恶性肿瘤学科发展报告 (2021) 》——神经肿瘤研究进展篇 [DB/OL]. (2022-05-16)[2023-02-11]. [http://app.weiyilewen.com/webtopic17?mpg=topic\\_news\\_detail&newsid=769](http://app.weiyilewen.com/webtopic17?mpg=topic_news_detail&newsid=769)
- [3] Sung H, Ferlay J, Siegel R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA: a cancer journal for clinicians*, 2021, 71(3): 209-249.
- [4] Chen F, Wendl M C, Wyczalkowski M A, et al. Moving pan-cancer studies from basic research toward the clinic[J]. *Nature Cancer*, 2021, 2(9): 879-890.
- [5] Louis D N, Perry A, Wesseling P, et al. The 2021 WHO classification of tumors of the central nervous system: a summary[J]. *Neuro-oncology*, 2021, 23(8): 1231-1251.
- [6] Chiacchiarretta P, Perrucci M G, Caulo M, et al. A Dedicated Tool for Presurgical Mapping of Brain Tumors and Mixed-Reality Navigation During Neurosurgery[J]. *Journal of digital imaging*, 2022, 35(3): 704-713.
- [7] 中华人民共和国国家卫生健康委员会. 脑胶质瘤诊疗指南 (2022 年版) [DB/OL]. (2022-04-11)[2023-02-11]. <http://www.nhc.gov.cn/yzygj/s2911/202204/a0e67177df1f439898683e1333957c74/files/2888d8e5c72c48ca8a9844000f55be58.pdf>
- [8] National Comprehensive Cancer Network. NCCN Guidelines Version 1.2019 Central Nervous System Cancers[DB/OL]. (2019-03)[2023-02-11]. [https://www.nccn.org/guidelines/category\\_1](https://www.nccn.org/guidelines/category_1)
- [9] Weller M, Van Den Bent M, Hopkins K, et al. EANO guideline for the diagnosis and treatment of anaplastic gliomas and glioblastoma[J]. *The lancet oncology*, 2014, 15(9): e395-e403.
- [10] Krivosheya D, Prabhu S S, Weinberg J S, et al. Technical principles in glioma surgery and preoperative considerations[J]. *Journal of neuro-oncology*, 2016, 130: 243-252.
- [11] Katsevman G A, Greenleaf W, García-García R, et al. Virtual reality during brain mapping for awake-patient brain tumor surgery: proposed tasks and domains to test[J]. *World Neurosurgery*, 2021, 152: e462-e466.
- [12] Trinh V T, Fahim D K, Maldaun M V C, et al. Impact of preoperative functional magnetic resonance imaging during awake craniotomy procedures for intraoperative guidance and complication avoidance[J]. *Stereotactic and functional neurosurgery*, 2014, 92(5): 315-322.

- [13] Kuchcinski G, Mellerio C, Pallud J, et al. Three-tesla functional MR language mapping: comparison with direct cortical stimulation in gliomas[J]. *Neurology*, 2015, 84(6): 560-568.
- [14] Azad T D, Duffau H. Limitations of functional neuroimaging for patient selection and surgical planning in glioma surgery[J]. *Neurosurgical Focus*, 2020, 48(2): E12.
- [15] 张昊. 神经外科开颅手术病灶增强现实显示技术研究[D]. 天津理工大学, 2021. DOI:10.27360/d.cnki.gtlgy.2021.000661.
- [16] Kockro R A, Serra L, Tseng-Tsai Y, et al. Planning and simulation of neurosurgery in a virtual reality environment[J]. *Neurosurgery*, 2000, 46(1): 118-35; discussion 135.
- [17] Cipresso P, Giglioli I A C, Raya M A, et al. The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature[J]. *Frontiers in psychology*, 2018: 2086.
- [18] Wilson M. Six views of embodied cognition[J]. *Psychonomic Bulletin & Review*, 2002, 9(4):625.
- [19] 杜云霄, 王殊轶, 于德旺, 陈俊. 增强现实技术在医疗领域的应用[J]. *软件导刊*, 2019, 18(09):139-142.
- [20] Albrecht U V, Folta-Schoofs K, Behrends M, et al. Effects of mobile augmented reality learning compared to textbook learning on medical students: randomized controlled pilot study[J]. *Journal of medical Internet research*, 2013, 15(8): e2497.
- [21] Poston T, Serra L. Dextrous virtual work[J]. *Communications of the ACM*, 1996, 39(5): 37-45.
- [22] Yang D L, Xu Q W, Che X M, et al. Clinical evaluation and follow-up outcome of presurgical plan by Dextroscope: a prospective controlled study in patients with skull base tumors[J]. *Surgical neurology*, 2009, 72(6): 682-689.
- [23] Wang S, Zhang S, Jing J. Stereoscopic virtual reality models for planning tumor resection in the sellar region[J]. *BMC neurology*, 2012, 12(1): 1-8.
- [24] Qiu T, Zhang Y, Wu J S, et al. Virtual reality presurgical planning for cerebral gliomas adjacent to motor pathways in an integrated 3-D stereoscopic visualization of structural MRI and DTI tractography[J]. *Acta neurochirurgica*, 2010, 152: 1847-1857.
- [25] Hinckley K, Pausch R, Downs J H, et al. The props-based interface for neurosurgical visualization[C]. *Medicine Meets Virtual Reality*. IOS Press, 1997: 552-562.
- [26] Abhari K, Baxter J S H, Chen E C S, et al. Training for planning tumour resection: augmented reality and human factors[J]. *IEEE Transactions on Biomedical Engineering*, 2014, 62(6): 1466-1477.

- [27] Wesarg S, Firl E A, Schwald B, et al. Accuracy of needle implantation in brachytherapy using a medical AR system: a phantom study[C]. Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display. SPIE, 2004, 5367: 341-352.
- [28] Liao H, Inomata T, Sakuma I, et al. 3-D augmented reality for MRI-guided surgery using integral videography autostereoscopic image overlay[J]. IEEE transactions on biomedical engineering, 2010, 57(6): 1476-1486.
- [29] Liu J, Li X, Leng X, et al. Effect of 3D slicer preoperative planning and intraoperative guidance with mobile phone virtual reality technology on brain glioma surgery[J]. Contrast Media & Molecular Imaging, 2022, 2022.
- [30] Zhou L, Wang W, Wei H, et al. Clinical application of 3D Slicer combined with Sina/MosoCam multimodal system in preoperative planning of brain lesions surgery[J]. Scientific Reports, 2022, 12(1): 19258
- [31] Deng W, Li F, Wang M, et al. Easy-to-use augmented reality neuronavigation using a wireless tablet PC[J]. Stereotactic and functional neurosurgery, 2014, 92(1): 17-24.
- [32] Allgaier M, Chheang V, Saalfeld P, et al. A comparison of input devices for precise interaction tasks in VR-based surgical planning and training[J]. Computers in Biology and Medicine, 2022, 145: 105429.
- [33] Hellum O, Kersten-Oertel M, Xiao Y. Assessment of user-interaction strategies for neurosurgical data navigation and annotation in virtual reality[J]. Virtual Reality, 2022: 1-11.
- [34] Kunz C, Maurer P, Kees F, et al. Infrared marker tracking with the HoloLens for neurosurgical interventions[J]. Current Directions in Biomedical Engineering, 2020, 6(1).
- [35] Song T, Yang C, Dianat O, et al. Endodontic guided treatment using augmented reality on a head - mounted display system[J]. Healthcare Technology Letters, 2018, 5(5): 201-207.
- [36] Wu P C, Wang R, Kin K, et al. DodecaPen: Accurate 6DoF tracking of a passive stylus[C]. Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. 2017: 365-374.
- [37] Microsoft HoloLens 2: Improved Research Mode to facilitate computer vision research [DB/OL].(2020-08-28)[2023-06-01].<https://www.microsoft.com/en-us/research/blog/microsoft-holoLens-2-improved-research-mode-to-facilitate-computer-vision-research/>
- [38] Horn B K P. Closed-form solution of absolute orientation using unit quaternions[J]. Josa a, 1987, 4(4): 629-642.
- [39] Gauglitz S, Nuernberger B, Turk M, et al. In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation[C]//Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology. 2014: 197-205.

- [40] Kim S, Lee G, Sakata N, et al. Improving co-presence with augmented visual communication cues for sharing experience through video conference[C]//2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, 2014: 83-92.
- [41] Huang W, Billingham M, Alem L, et al. HandsInTouch: Sharing gestures in remote collaboration[C]//Proceedings of the 30th Australian Conference on Computer-Human Interaction. 2018: 396-400.
- [42] LaViola Jr J J, Kruijff E, McMahan R P, et al. 3D user interfaces: theory and practice[M]. Addison-Wesley Professional, 2017.
- [43] Weise M, Zender R, Lucke U. A comprehensive classification of 3D selection and manipulation techniques[M]. Proceedings of Mensch und Computer 2019. 2019: 321-332.
- [44] Pernet C R, Gorgolewski K J, Job D, et al. A structural and functional magnetic resonance imaging dataset of brain tumour patients[J]. Scientific data, 2016, 3(1): 1-6.
- [45] 冯乐乐. Unity Shader 入门精要[M]. 人民邮电出版社, 2016.
- [46] Drebin R A, Carpenter L, Hanrahan P. Volume rendering[J]. ACM Siggraph Computer Graphics, 1988, 22(4): 65-74.
- [47] Levoy M. Display of surfaces from volume data: IEEE Comput. Graph. & Appl. Vol 8 No 3 (1988) pp 29–37[J]. 1988.
- [48] Hansen C D, Johnson C R. Visualization handbook[M]. Elsevier, 2011.
- [49] Introduction to Azure Remote Rendering ? [DB/OL].(2023-05-09)[2023-06-01]. <https://learn.microsoft.com/zh-cn/training/modules/intro-to-azure-remote-rendering/>
- [50] 梁利斌. 经皮热消融自动规划方法研究与规划系统构建 [D].清华大学, 2021.

## 致 谢

行文至此，百感交集。

首先感谢王广志老师和丁辉老师的指导，从我大三暑假刚进入实验室到选题开展毕业设计的过程中提供了中肯的建议与丰富的资源，让我逐渐对增强现实领域有了宏观的把握。也感谢实验室的师兄师姐对我的工作提出的建议与帮助。

感谢我的父母，在我做出选择时他们总是无条件支持我，也在我因为冲动和感性脱离轨道时一次次将我拉回正轨。是他们教我人活一辈子应该做自己想做的事，成为想成为的人。二十载养育之恩无以为报，但愿父母万事胜意，健康顺遂。

大学四年遇到很多善良可爱的人，他们出现在我人生路上不同的阶段，或让我懂得了一些道理，或教导我知识，或借了我一把力让我渡过坎坷。我也在四年体验了近乎我想体验的所有生活后逐渐明白自己在追求的是什么，感谢所有遇见。

而我的大四被持续了一年的毕设为主线贯穿。一切的开始是李浩伟学长一年前的此时在独峰书院与一心想做工程的我畅谈增强现实领域，他说他那里可以让我接触到近乎所有想学的技术，于是我来了，至今仍记得第一次戴上 HoloLens 时被那只彩色的蜂鸟深深吸引住的震撼。在过去的一年里，学长为我提供了尽可能多的资源与平台，也经常给我上压力，但我也正是在这种压力下不断成长。我开始挑战那些我曾经以为不擅长的事情，比如算法和数学推导，而我擅长的工程能力也在整个过程中得到飞速提升。感谢一次次他和我小到技术实现大到人生方向的指点与讨论，是他让我看到了未来更多的可能性，也让我在四年画下句号前实现了构建一个完整系统的梦想。

感谢我最好的朋友黄煜展同学的陪伴、关怀与包容。

最后感谢在每次基于内心的选择中勇敢而热烈的自己。落幕的是我的大学四年，而不是我依然拥有着千万种可能性的人生。谨以此文，纪念我一步一个选择，一步一次成长的大学四年。

## 声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名： \_\_\_\_\_ 日 期： \_\_\_\_\_

## 附录 A 外文资料的书面翻译

### A Dedicated Tool for Presurgical Mapping of Brain Tumors and Mixed-Reality Navigation During Neurosurgery

神经外科手术期间脑肿瘤术前映射和混合现实导航的专用工具

#### A.1 摘要

脑肿瘤手术需要在完全切除肿瘤组织和尽量减少脑功能损失之间进行微妙的权衡。功能性磁共振成像 (fMRI) 和弥散张量成像 (DTI) 已经成为无创评估人类大脑功能的宝贵工具, 被用于确定应该保留的大脑区域以防止手术后的功能损害。图像分析需要不同的软件包, 然而这些软件包主要是为研究目的而开发的, 通常很难在临床环境中使用, 这阻碍了术前绘图的大规模推广。我们开发了一个能够实现多模态 MRI 术前绘图的自动分析, 并将结果转移到神经导航仪上的专门的软件。此外, 成像结果被整合到一个市售的可穿戴设备中, 使用优化的混合现实方法自动将从 MRI 获得的三维全息图锚定到病人的物理头部上。这将使外科医生能够虚拟地探索更深的组织层, 突出需要保留的关键大脑结构, 同时保留手眼协调的自然特性。这一应用将显著提高手术的准确性和安全性, 预计将为医疗保健系统和相关的工业投资者带来巨大的利益。

#### A.2 引言

由于要在尽可能多地切除肿瘤组织和尽量减少脑功能损失之间进行微妙的权衡, 脑肿瘤手术面临着复杂的挑战。这种权衡的结果通常是通过术中直接皮质刺激 (DCS) 获得的, 这被认为是皮质功能图谱的金标准技术<sup>[1][2]</sup>。然而 DCS 可能会引发癫痫发作而且只能在术中进行, 因此延长了手术的时间, 还无法在术前对干预措施进行详细规划。在过去的二十年里, 功能磁共振成像 (fMRI) 已经成为非侵入性评估人类大脑活动的宝贵工具, 现在被用来确定应该保留的大脑区域, 以防止手术后的功能损伤<sup>[3][4][5][6][7]</sup>。在这方面, fMRI 和其他如飞行时间 (TOF) 或对比度增强的核磁共振血管造影和扩散张量成像 (DTI) 白质束造影等 MRI 模态相辅相成, 提供了关于运动型语言中枢和肿瘤附近或内部的皮层或皮层下连接

的解剖和功能信息<sup>[4]</sup>。大量的数据为优化治疗计划并显著减少手术时间提供了契机。成像结果通常被发送到一个神经导航系统，该系统计算出手术空间和术前图像空间之间的转换矩阵。通过这种转换外科医生可以在工作站屏幕上看到指针尖端处放置在特定解剖结构上对应的虚拟图像，虚拟图像与预先获得的 MRI（或 CT）脑扫描中的相应区域重叠。通过这种方式，神经导航在手术中为外科医生提供支持，允许提高识别目标区域和需要保留的关键功能结构的准确性。

然而，尽管这些成像技术和随之而来的辅助手术的科学基础已经确立，但其目前的实施有两个主要的技术限制。

首先，仍然缺乏一个专门的软件来计算、整合和及时把成像结果输出到手术导航系统。事实上通常需要一连串复杂的操作来传输、分析、重新格式化和输出数据到神经导航仪。在大多数情况下，分析过程是使用为神经影像领域的研究目的而开发的软件包（AFNI、FSL、SPM、BrainVoyager）进行的。虽然这些程序的复杂性和性能已经达到了相当高的水平，但在临床语境中并不明了。此外，每个软件一般都是针对特定的成像模式而优化的，因此需要掌握不同的软件包来完成整个程序。事实上，术前测绘的应用仍然局限在那些除了临床团队外还能依靠多学科研究团队的影像中心，这阻碍了术前测绘的大规模推广。

其次，尽管神经导航系统在协助外科医生方面具有公认的优势，但其目前的实施方式在人机工程学上是有限的，它迫使神经外科医生不断地将目光从手术区域转移到工作站屏幕上，那里显示的是术前图像和虚拟手术器械的二维表示。这就要求外科医生付出巨大的努力在头脑中整合来自不同模式的信息，重建病人感兴趣的三维解剖特征，这会延长干预时间，增加失误风险。

在这方面，诸如增强现实（AR）和混合现实（MR）等新兴技术在神经外科应用中显示出良好的效果<sup>[8][9][10][11][12][13]</sup>。特别是 MR 方法在提高人体工程学方面具有最大的潜力，因为它混合了真实和虚拟对象，产生了一个物理和数字对象共存并实时互动的可视化环境。然而，以前的研究大多使用解剖图像，而结合 AR 或 MR 的功能性术前绘图数据的全部潜力还没有被开发出来。

这些问题激发了这项工作的主要目标，如下所述：

(1) 开发一个专门的软件，能够在应用程序中实现多模态 MRI 的分析，并将结果转移到神经导航仪上，最大限度地减少手工操作。这将扩大手术前绘图程序的推广范围，使其也能在学术界以外的医院使用。

(2) 使用新兴的混合现实成像技术，开发一种专门的程序，将术前映射 MRI 数据的全部潜力集成到商用可穿戴设备中。特别是从 MRI 获得的 3D 全

息图将与患者的身体头部锚定，使外科医生能够轻松定位开颅手术的最佳颅骨区域，同时虚拟探索更深的组织层，突出显示需要被保留的关键大脑结构。预计这将显著提高外科医生的准确性并减少错误。

### A.3 方法

#### A.3.1 分析平台

针对这项工作的第一个目标，我们搭建了一个软件平台用于分析和集成来自不同模式的成像数据，这些模式可用于脑肿瘤的术前映射，包括多模式磁共振成像（MRI）和计算机断层扫描（CT）图像。图像分析基于不同的开源软件工具，这些工具在神经影像科学界享有盛誉。这些项目包括 AFNI、FreeSurfer、FSL，以及用 Python 实现的定制软件。

该平台已在专用工作站中实施，并已配置为 DICOM（医学数字成像和通信）节点，因此在 MRI 检查结束时，图像可以直接以 DICOM 格式发送到平台。到达后，DICOM 数据根据 BIDS 结构进行重构，以标准化预处理程序。

我们平台目前处理的脑术前成像技术包括结构 MRI、fMRI、DTI 和 TOF，下面将与相关分析程序一起描述。

结构 MRI：这些图像可以通过不同的采集方案获得，从而在脑组织之间产生不同类型的对比。用于此目的的最常见 MRI 采集是 T1 加权序列，可以很好地描述灰质和白质结构的形状、大小和完整性。此外，在确定模态之间的配准矩阵后，结构 MRI 为从 fMRI 获得的激活图案或从 DTI 获得的白质纤维束的可视化提供了解剖学参考。T1 图像的处理是用 FreeSurfer 执行的，包括去除非脑组织（脂肪、头骨或颈部）和灰质分割两个步骤。此外，头部皮肤和皮质表面的三维重建是使用 3D Slicer 和 AFNI 例程创建以 .stl 格式存储的网格完成的。肿瘤分割也是通过使用 ITK-SNAP 的主动轮廓分割半自动进行的，相应的网格也保存在 .stl 文件中。头部皮肤网格最终将导入混合现实应用程序，并与来自 HoloLens 相机的图像一起使用，以实现虚拟世界和现实世界之间的锚定。肿瘤网格也将被导入，以可视化正确叠加在患者头部的相应全息图，从而突出显示病变边界和皮肤下方的位置。这将有助于外科医生决定开颅手术的最佳区域（图 1）。



图1: 肿瘤全息图的一个例子, 它来自核磁共振数据, 在锚定步骤后在HoloLens中与 "真实世界" 的头皮一起可视化。混合现实可以让人立即清楚地了解肿瘤在颅骨下的位置, 帮助外科医生划定开颅手术的最佳区域。

功能性核磁共振 (fMRI) : 无创的大脑功能图谱通常使用血氧水平依赖性 (BOLD) 技术进行<sup>[14][15]</sup>。简而言之, 当病人参与执行旨在激活运动型语言区域 (如运动、感觉或语言区域) 的范式时, 获得对随着神经元激活发生的脱氧血红蛋白变化敏感的时间系列图像。在我们的演示中, 我们使用了一系列的 T2 加权超声平面图像 (EPI), 参数如下: TR/TE=1000/30 ms; 体素大小=3×3×3 mm<sup>3</sup>; 48 层; 帧数=290。在 fMRI 过程中病人根据一个块范式执行运动、视觉和语言任务。fMRI 数据的预处理步骤包括运动校正、切片时间校正和时间过滤, 而时间序列的统计分析则使用带有双伽马血流动力学反应函数的一般线性模型 (GLM) 进行<sup>[16]</sup>。然后用仿射变换对功能和结构数据集进行注册。

所有这些分析步骤都是通过 AFNI 例程在平台上实现的。最后的结果是一个三维 "激活图", 代表了在手术中应予保留的大脑皮层区域 (图 2) 。

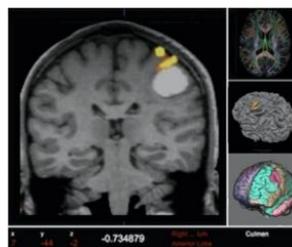


图2: 手部运动区的fMRI激活图 (红黄色) 叠加在脑瘤患者的结构扫描 (灰色) 上的例子, 该图像由我们的平台计算得出。

纤维追踪 (纤维束成像) : 使用弥散张量成像 (DTI) 数据重建连接不同功能区域的白质纤维束<sup>[17]</sup>。该 MRI 应用通过在不同方向应用扩散编码梯度获取图像, 从而可以计算图像中每个体素的脑组织中水分子的扩散张量。纤维方向由张量主特征向量表示。然后对该向量进行颜色编码以生成纤维方向图, 该图与结构图像共同配准, 以正确表示相对于肿瘤的纤维结构。在决定去除的病变而不损坏这些连接的最安全路径时, 这是一个有价值的信息<sup>[18][19]</sup>。使用 FSL/Freesurfer 模块实施 DTI 分析以创建概率纤维束成像 (图 3) 。

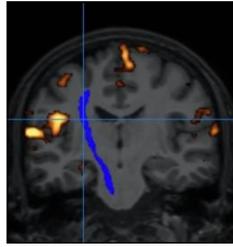


图3: 从我们平台的DTI 模块获得的结果示例, 显示皮质脊髓束

MR 血管成像 (MRA) 。代表大脑中动脉和静脉血管网络的图像是通过血液运动敏感的采集序列获得的, 如三维飞行时间 (TOF) 和三维相位对比血管成像 (PCA) 。在这个项目中我们使用了 TOF 采集, 并通过源自迭代最接近点方法的算法实现了这些图像与结构参考的核心关联。这些图像并不常规用于脑部术前绘图程序; 然而我们为未来的发展实现了这种可能性。

重要的是, 我们还设计了一个用户友好的图形界面, 专门面向需要直接工具来完成该程序的临床用户。

这个环境实际上只需要操作者的几个步骤, 包括 (1) 用一个命令开始分析, (2) 对不同图像模式之间的正确核心配对进行视觉检查, 以及 (3) 完成半自动的肿瘤分割程序。

具体来说, 随着从图片存档和通信系统 (PACS) 加载核磁共振检查, 分析术前绘图数据所需的复杂级联操作被自动启动, 并以上下文方式进行。结果也会自动准备好发送到神经导航仪系统和微软 HoloLens。

肿瘤分割是唯一需要放射科医生进行少量手工操作的程序。

一旦经过预处理, 图像就会通过网络浏览器 (Mango) 显示在平台上, 用户可以调查映射结果, 检查不同模式之间的注册, 并对检查进行全面评估。一旦控制过程结束, 就会创建一个包含每个模态的网格的 STL 文件。

该平台在富士通 CELSIUS 工作站 (R970 2 10 Core Xeon Silver 4114, 96 GB DDR4 RAM, ASUS TURBO-GTX1070TI-8G 显卡) 上运行。软件架构的开发很容易适应云计算应用, 这对最终用户来说有很多好处, 避免了硬件上的花费和安装和维护所有需要的 (开源) 软件工具的必要性。这也大大简化了升级操作, 不会对用户提出要求。

该原型已在 3 层基础架构中实现 (图 4) :

- 公共层: Frontend Angular。
- 应用层: Backend Ruby on Rails。
- 数据库层: MongoDB。

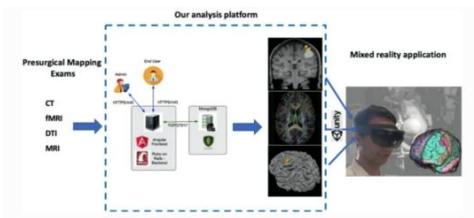


图 4: 我们的原型分析平台和混合现实应用程序的示意图

前后端之间的认证是通过 `JavaWebToken` 来完成的。前端可以使用有效证书通过 `HTTPS` 发布。该解决方案的三层结构允许网络隔离，确保数据可以存储在外部网络无法访问的最安全的环境中。前端/后端的实现是通过交换 `JSON` 文档来完成的。该解决方案提高了技术的可扩展性，优化了移动应用程序等未来功能的实施，因为后端可以与移动和桌面客户端共享。

由于不同的开源工具通常以专有格式保存结果，因此专门的模块在内部处理不同的图像格式，通常从扫描仪输出的 `DICOM` 标准转换为目标各种开源软件程序最常见的输入，可用于特定分析的 `NIFTI` 格式。

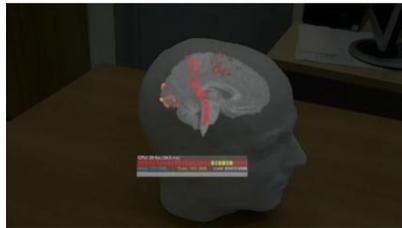


图5: 混合现实应用程序中使用体素查看器的例子。当外科医生通过HoloLens观察头部时，他可以使用半空中的手势激活一个重新切割选项，以便将从结构体素（此处为灰色）和功能或DTI数据（此处为彩色）中获得的内部切片可视化。

在图像处理，映射结果再次转化为 `DICOM` 格式，这是目前神经导航系统所接受的标准，或者将其留在 `NIFTI` 中，用于为 `HoloLens` 开发的体积查看器。

### A.3.2 在 `HoloLens` 中整合术前绘图

针对这项工作的第二个目标，我们开发了一个应用程序能够利用我们的平台获得的术前测绘结果进行导航，使用市面上的可穿戴设备 (`Microsoft HoloLens 1.0`) 实现的 `MR` 方法。该应用程序使用 `Unity` 图形包开发，能够将从例如 `MRI` 和 `CT` 扫描中获得的三维体积数据可视化为与患者真实头部锚定的三维全息图 (图 1)。此外，使用简单的空中手柄手势作为指令，外科医生可以根据偏好的角度对容积图像进行重新切片，从而通过二维切片滚动来查看突出深层关键大脑结构的内部切片 (图 5)。以下是对该应用及其开发过程的详细描述。

### A.3.3 锚定到真实头部

首先，在头部和不同的手术前映射图像的虚拟信息之间构建了一个自动锚定程序。由于实际原因，锚定程序是在一个头部模型上开发的，这个头部模型是用3D打印机从志愿者身上获得的高分辨率解剖核磁共振成像扫描的结果打印出来的。具体来说，用3D Slicer 创建了一个头部的三维虚拟模型，并保存在一个.stl文件中。从使用3D打印机用这个文件打印了一个由丙烯腈-丁二烯-苯乙烯 (ABS) 制成的模型 (图6)。



图6: 从结构性MRI数据集中获得的3D打印头部模型，用于开发锚定程序。

然后，我们使用 HoloLens 的软件开发工具包 (SDK) 来访问从 HoloLens 的正面摄像头获得的真实世界图像。利用 HoloLens 的自我定位和空间映射能力，这些图像被转换为网格，随着用户在假体上的移动而不断更新。然后，这些网格和从核磁共振得到的头部三维虚拟模型导入的网格可作为计算机视觉算法的输入，以完成虚拟物体和现实世界之间的锚定。具体来说，我们使用 Vuforia 库中的算法在 Unity 环境中工作，以实现我们的锚定程序。

请注意，使用头部皮肤获得的锚定也适用于所有其他核磁共振网格和三维体素 (特别是功能图和神经束图)，因为分析平台已经完成了不同模式之间的注册。

### A.3.4 叠加体素查看器

在开发的应用程序中，我们还实现了一个体素查看器，允许根据偏好方向重新切分体积 MRI 数据 (图5)。这为外科医生提供了有关头部全息渲染的额外信息，允许精确了解位于大脑深处的解剖和功能结构。可视化选项包括以下可能性：

- (1) 选择要重新切片/可视化的 MRI 体积 (结构 MRI、fMRI、DTI 等)；
- (2) 选择重新切片的平面；
- (3) 在结构扫描上叠加 fMRI 和 DTI 数据；以及
- (4) 改变 fMRI 地图的统计阈值 (提高阈值可以更好地定位激活的峰值)。

这些功能已经在 Unity 中使用光线投射（着色器）程序实现。此外，还开发了一个允许多重叠的自定义功能。重要的是，所有的可视化选项都可以由用户通过半空中的手势来选择，这有助于保持整个过程更加无菌，使神经外科医生能够以简单和快速的方式查询数据。

#### A.4 结果

参与该项目的临床人员对分析平台和 MR 应用都进行了评估。放射科医生表示，与现有的软件包相比，该平台在自动化和图形界面的易用性方面为整个分析程序提供了重大改进。另外，没有安装和维护问题也是非常值得赞赏的。神经外科医生表示，所实施的 MR 解决方案对手术计划非常有用。特别是叠加的体素查看器，允许滚动浏览并同时观察与 fMRI、DTI 和结构图像相对应的全息图，得到了非常积极的反馈。此外，自动锚定程序的速度和简易性也非常值得赞赏。最后，外科医生表示，该系统不仅在手术规划方面有很大的潜力，考虑到半空中手势指令所提供的无菌设置，其在术中手术指导方面也有很大的潜力。

#### A.5 讨论

在本文中，我们描述了 (1) 一个“临床”软件平台的实施，以分析从多模态脑瘤术前绘图中获得的数据，以及 (2) 一个应用程序，该应用程序使用优化的混合现实方法将绘图结果整合到一个市售的可穿戴设备中，提供身临其境和增加可视化能力。

尽管先进的结构和包括 fMRI 和 DTI 图的功能无创 MRI 技术已经越来越多地被用于高度专业化的中心作为手术前的绘图程序，帮助外科医生在获得最大的肿瘤切除的同时保留大脑功能，但由于分析工具的复杂性，它们在学术界以外的传播仍然有限。我们强调，据我们所知，目前仍缺乏一个专门的综合软件来分析术前绘图数据。

在这方面，我们的解决方案提供了更高的自动化水平和用户友好的图形界面，专门用于将临床用户从典型的复杂的研究工具中释放出来，从而使非专业的图像处理人员也能完成这一程序。此外，我们以云计算方法为基础，以提高平台的可扩展性，并帮助用户摆脱有关安装和维护操作的进一步问题。虽然分析步骤是基于不同的开源神经影像工具，但用户不需要掌握任何这些软件包，可以继续专注于结果的解释。该平台需要最少的手动操作，以便在临床环境中获得极大的方便。

特别是，我们的平台除了半自动的肿瘤分割外在向现有的神经导航系统输出前可仅由一个操作者执行，他只需进行最后的视觉检查（例如，功能和解剖图像之间的正确注册）。

我们还强调了标准神经导航设备的局限性。在这些系统中，一个二维虚拟环境被创建到工作站屏幕上，包括虚拟手术器械和从术前 CT 或 MRI 扫描中获得的病人特定虚拟解剖细节。外科医生将指针放在真实的解剖目标上，并观察其与包含额外信息的虚拟目标的对应关系（例如，在暴露的大脑上明显不可见的皮质区域和白质纤维）。然后，他必须在思维上整合二维信息，以获得肿瘤和周围关键结构之间的三维空间关系，从而选择最佳的手术策略。即使是最近使用增强现实技术，用虚拟信息丰富真实环境的数据的方法，也显示出对手术神经导航的有限改进<sup>[6]</sup>。事实上，尽管这类系统可以将现实世界的图像与虚拟对象（如术前图像或病人的其他临床数据）合并，但可视化仍然是在显示器和平板电脑等 2D 设备上进行的。相反，在新兴的混合现实技术中，虚拟信息被 "锚定" 在真实物体上，而不是简单地一起显示。在我们的应用中，一旦正确地进行了锚定，当外科医生改变位置并从不同方向观察病人的头部时，真实物体和相应的三维全息图之间的叠加就会自动更新，从而能够立即改善对肿瘤、显露区、白质纤维、血管和颅骨之间空间关系的理解。这一功能极大地促进了开颅手术和手术切除的规划，特别是在涉及复杂的解剖结构时，为神经外科领域提供了更大的改进空间。

与标准的神经导航系统相比，头戴式混合现实设备在术前神经外科规划中的工效提高和临床适用性已经在以前的研究中的现场和模拟手术中得到了证明。

特别的是，一项前瞻性的临床研究为 HoloLens 在脑瘤术前规划中的临床应用提供了概念证明，显示了在保持注意力和注意力方面的显著优势，以及由于三维全息可视化而对肿瘤-脑/颅骨的空间关系有了更好的理解<sup>[10]</sup>。此外，作者获得的定量结果测量显示，在标准神经导航系统和 HoloLens 之间，肿瘤定位的中位数偏差为 4 毫米，在 25 名患者中，有 9 名患者的定位没有差异，这表明这种 MR 系统在手术室的潜力。在该研究中，从对比度增强的 T1 加权 MRI 结构图像的分割中获得的肿瘤的三维全息图重建被用作术前虚拟对象。

另一项研究增加了 HoloLens 中的可视化组件，包括可以滚动浏览 3D MRI/CT 术前体素数据以及从共焦激光内窥镜和超声波获得的术中图像，这些都可以在 HMD 内的虚拟 2D 屏幕上看到<sup>[11]</sup>。作者还增加了新的可视化选项，例如可以调整代表三维器官模型的虚拟对象的比例、方向和透明度，如术前已经分割的肿瘤和血管结构。此外，他们还进行了一项试点用户研究，要求一组具有不同医学专业

的外科医生（9 人中有 4 人是神经外科医生）回答关于拟议系统在手术室的可用性的具体调查问卷。结果显示了对该系统的高度赞赏，特别是将选定的三维器官模型作为全息图进行可视化的可能性，以及对在术前计划阶段之外使用带有 MR 方法的 HMD 进行术中手术指导的普遍强烈兴趣。然而，在这些先前的工作中，只使用了特定模式的术前绘图数据（大多只是显示脑肿瘤的结构图像）。在这方面，我们的平台和磁共振应用扩大了技术水平，将先进的功能性术前绘图模式的全部潜力纳入脑肿瘤手术。特别是，表达区的 fMRI 和 DTI 束图对神经外科至关重要，并受到外科医生的高度重视。这些先进的 MRI 技术决定了手术中的一些优势，如量身定做的开颅手术、知情选择 DCS 部位导致刺激次数减少、术中癫痫发作的可能性降低、手术时间加快并减少病人的疲劳<sup>[4]</sup>。

然而，这些先进的功能成像技术到目前为止只在标准的神经导航系统中得到利用。据我们所知，这是第一项将所有可用的术前映射 MRI 模式整合到 MR HMD 中的工作。我们的原型系统允许可视化全息图的选择组合，代表大脑解剖结构、肿瘤边界、能动皮质区域、白质纤维束和血管网络，并有可能改变不同对象的透明度。此外，外科医生可以根据偏好的角度对体积图像进行重新切片，从而通过二维切片滚动查看内部切片，突出显示深层关键的大脑结构，如白质纤维束以及它们与肿瘤的空间关系。重要的是，这些功能可以用简单的半空中手柄手势作为命令来激活，而不需要实际接触任何物体，这是保持手术室一切无菌的重要资产。此外，作为一项重大改进，我们实施了一个自动的从真实到虚拟的锚定程序，这将导致减少干预时间。

## A.6 研究局限性

在这项工作中，我们没有对我们的自动注册与以前研究中使用的手动注册方法的准确性进行定量结果测量，因为在这个阶段，我们主要关注的是提高一般工作流程的自动化和人机工程学，从分析手术前的映射数据到它们在 HoloLens 中的可视化。虽然速度在手术室里当然很重要，但我们明白，为了将这些系统不仅用于手术规划，还用于指导干预，需要对准确性进行评估，在未来的发展中，我们计划使用与[10]类似的方法进行评估。然而，我们也指出，手术中神经导航的一个主要问题是开颅手术后的大脑变形。这种现象被称为“脑移位”，涉及由于脑脊液流失、颅内压力变化和其他原因造成的大脑结构的非刚性变形。其结果是，当脑移位显著时，术前图像与通过手术显微镜或 HMD 看到的病人图像不一致。这

个话题是激烈研究的对象<sup>[20]</sup>，我们计划在未来的研究中也朝这个方向努力。此外，未来用于手术的 MR 技术的一个重大改进是开发能够提供某种放大系数的专用 HMD，最好是将神经导航仪和手术显微镜整合在一个仪器中，并增加全息可视化的潜力。

最后，我们没有使用结构化调查问卷对广大的神经外科医生进行用户研究。然而，我们单位的临床工作人员对分析平台和 MR 应用程序的可用性的反馈是非常积极的。特别是，放射科工作人员非常赞赏我们的解决方案所提供的分析程序的简单工作流程，而外科医生则高度评价了对大脑结构之间的空间关系和 MR 应用的一般人体工程学的理解。

总之，尽管核磁共振在手术室中的应用仍处于早期阶段，但我们的工作和以前的研究结果证明了在术中协助外科医生的巨大潜力。

#### 参考文献

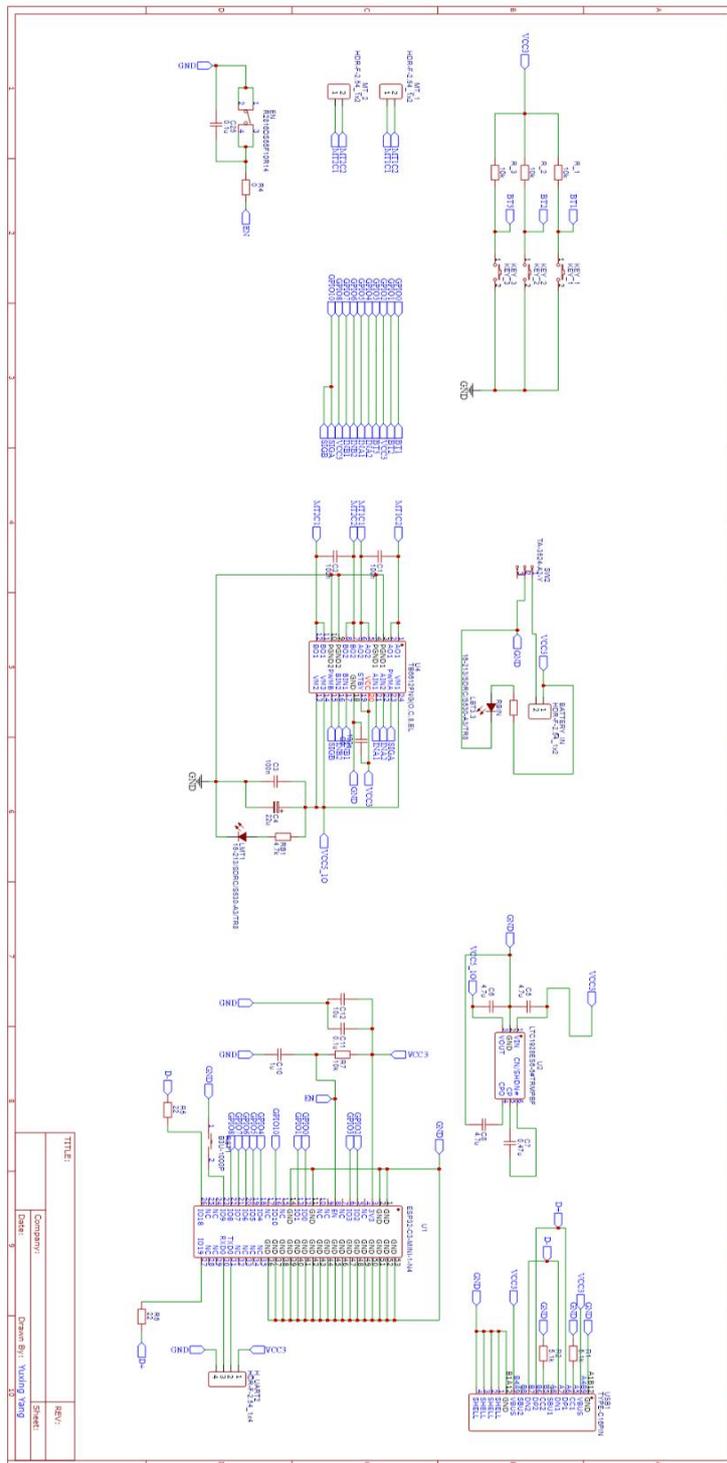
- [1] Desmurget M, Song Z, Mottolese C, Sirigu A. Re-establishing the merits of electrical brain stimulation. *Trends Cogn Sci*. 2013;17(9):442–449. doi: 10.1016/j.tics.2013.07.002. - DOI - PubMed
- [2] P. C. De Witt Hamer, S. G. Robles, A. H. Zwinderman, H. Duffau, and M. S. Berger, “Impact of intraoperative stimulation brain mapping on glioma surgery outcome: a meta-analysis.” *J Clin Oncol*, vol. 30, no. 20, pp. 2559–2565, Jul. 2012. - PubMed
- [3] Agarwal S, Sair HI, Gujar S, Pillai JJ. Language mapping with fMRI: current standards and reproducibility. *Top Magn Reson Imaging*. 2019;28(4):225–233. doi: 10.1097/RMR.0000000000000216. - DOI - PubMed
- [4] A. Castellano, S. Cirillo, L. Bello, M. Riva, and A. Falini, “Functional MRI for surgery of gliomas.” *Curr Treat Options Neurol*, vol. 19, no. 10, p. 34, Aug. 2017. - PubMed
- [5] Chaudhry AA, Naim S, Gul M, Chaudhry A, Chen M, Jandial R, Badie B. Utility of preoperative blood-oxygen-level-dependent functional MR imaging in patients with a central nervous system neoplasm. *Neuroimaging Clin N Am*. 2021;31(1):93–102. doi: 10.1016/j.nic.2020.09.009. - DOI - PubMed
- [6] J.-M. Lemée, D. H. Berro, F. Bernard, E. Chinier, L.-M. Leiber, P. Menei, and A. Ter Minassian, “Resting-state functional magnetic resonance imaging versus task-based activity for language mapping and correlation with perioperative cortical mapping.” *Brain Behav*, vol. 9, no. 10, p. e01362, Oct. 2019. - PMC - PubMed

- [7] Unadkat P, Fumagalli L, Rigolo L, Vangel MG, Young GS, Huang R, Mukundan S, Golby A, Tie Y. Functional MRI task comparison for language mapping in neurosurgical patients. *J Neuroimaging*. 2019;29(3):348–356. doi: 10.1111/jon.12597. - DOI - PMC - PubMed
- [8] J. W. Yoon, R. E. Chen, E. J. Kim, O. O. Akinduro, P. Kerezoudis, P. K. Han, P. Si, W. D. Freeman, R. J. Diaz, R. J. Komotar, S. M. Pirris, B. L. Brown, M. Bydon, M. Y. Wang, R. E. Wharen, and A. Quinones-Hinojosa, “Augmented reality for the surgeon: systematic review.,” *Int J Med Robot*, vol. 14, no. 4, p. e1914, Aug. 2018. - PubMed
- [9] Meola A, Cutolo F, Carbone M, Cagnazzo F, Ferrari M, Ferrari V. Augmented reality in neurosurgery: a systematic review. *Neurosurg Rev*. 2017;40(4):537–548. doi: 10.1007/s10143-016-0732-9. - DOI - PMC - PubMed
- [10] Incekara F, Smits M, Dirven C, Vincent A. Clinical feasibility of a wearable mixed-reality device in neurosurgery. *World Neurosurg*. 2018;118:e422–e427. doi: 10.1016/j.wneu.2018.06.208. - DOI - PubMed
- [11] Cartucho J, Shapira D, Ashrafian H, Giannarou S. Multimodal mixed reality visualisation for intraoperative surgical guidance. *Int J Comput Assist Radiol Surg*. 2020;15(5):819–826. doi: 10.1007/s11548-020-02165-4. - DOI - PMC - PubMed
- [12] J. Huang, M. Halicek, M. Shahedi, and B. Fei, “Augmented reality visualization of hyperspectral imaging classifications for image-guided brain tumor phantom resection.,” *Proc SPIE Int Soc Opt Eng*, vol. 11315, Feb. 2020. - PMC - PubMed
- [13] C. M. Morales Mojica, J. D. Velazco-Garcia, E. P. Pappas, T. A. Birbilis, A. Becker, E. L. Leiss, A. Webb, I. Seimenis, and N. V. Tsekos, “A holographic augmented reality interface for visualizing of MRI data and planning of neurosurgical procedures.,” *J Digit Imaging*, May 2021. - PMC - PubMed
- [14] Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Uğurbil K. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*. 1992;89(13):5951–5955. doi: 10.1073/pnas.89.13.5951. - DOI - PMC - PubMed
- [15] Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America*. 1992;89(12):5675–5679. doi: 10.1073/pnas.89.12.5675. - DOI - PMC - PubMed
- [16] Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp*. 1994;2(4):189–210. doi: 10.1002/hbm.460020402. - DOI

- [17] Jellison BJ, Field AS, Medow J, Lazar M, Salamat MS, Alexander AL. Diffusion tensor imaging of cerebral white matter: a pictorial review of physics, fiber tract anatomy, and tumor imaging patterns. *American Journal of Neuroradiology*. 2004;25(3):356–369. - PMC - PubMed
- [18] Romano A, D'andrea G, Minniti G, Mastronardi L, Ferrante L, Fantozzi LM, Bozzao A. Pre-surgical planning and MR-tractography utility in brain tumour resection. *Eur Radiol*. 2009;19(12):2798–2808. doi: 10.1007/s00330-009-1483-6. - DOI - PubMed
- [19] Ulmer JL, Klein AP, Mueller WM, DeYoe EA, Mark LP. Preoperative diffusion tensor imaging: improving neurosurgical outcomes in brain tumor patients. *Neuroimaging Clin N Am*. 2014;24(4):599–617. doi: 10.1016/j.nic.2014.08.002. - DOI - PubMed
- [20] I. J. Gerard, M. Kersten-Oertel, J. A. Hall, D. Sirhan, and D. L. Collins, “Brain shift in neuronavigation of brain tumors: an updated review of intra-operative ultrasound applications,” *Front Oncol*, vol. 10, p. 618837, 2020. - PMC - PubMed

## 附录 B 其他内容

### B.1 交互工具电路设计图



42

## B.2 三种滤波算法原理

### B.2.1 最小二乘平滑优化算法原理

设平滑窗的大小为  $2m+1$ ，多项式阶数为  $n$ 。平滑窗内的每一项可以表示为：

$$\begin{aligned}
 X_{-m} &= b_0 + b_1 \cdot (-m)^1 + \dots + b_n \cdot (-m)^n \\
 &\vdots \\
 X_{-1} &= b_0 + b_1 \cdot (-1)^1 + \dots + b_n \cdot (-1)^n \\
 X_0 &= b_0 + b_1 \cdot (0)^1 + \dots + b_n \cdot (0)^n \\
 \text{(B-1)} \quad X_1 &= b_0 + b_1 \cdot (1)^1 + \dots + b_n \cdot (1)^n \\
 &\vdots \\
 X_m &= b_0 + b_1 \cdot (m)^1 + \dots + b_n \cdot (m)^n
 \end{aligned}$$

又可以写作：

$$\begin{aligned}
 &\begin{bmatrix} 1 & (-m)^1 & \dots & (-m)^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (-1)^1 & \dots & (-1)^n \\ [1 & (0)^1 & \dots & (0)^n ] \\ \text{(B-2)} \quad & \begin{bmatrix} 1 & (1)^1 & \dots & (1)^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (m)^1 & \dots & (m)^n \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} X_{-m} \\ \vdots \\ X_{-1} \\ X_0 \\ X_1 \\ \vdots \\ X_n \end{bmatrix}
 \end{aligned}$$

记作  $Ab = X$ ，由最小二乘可以得到由  $X$  表征的权重系数  $b = (A^T A)^{-1} A^T X$ ，代入得  $A(A^T A)^{-1} A^T X^{(t-n)} = X^{(t)}$ 。

(B-3)

也就是以  $A(A^T A)^{-1} A^T$  作为权重矩阵更新平滑窗内  $X$  的值。

### B.2.2 卡尔曼滤波原理

卡尔曼滤波是一种常见的用于实时轨迹滤波的算法。该算法将系统的状态方程为  $x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$ ，根据上一时刻的状态和控制变量预测当前时刻的状态。由于该算法过于经典且篇幅较长，在此不做展开。在一阶卡尔曼滤波器中控制变量考量了速度，二阶中考量了速度和加速度。

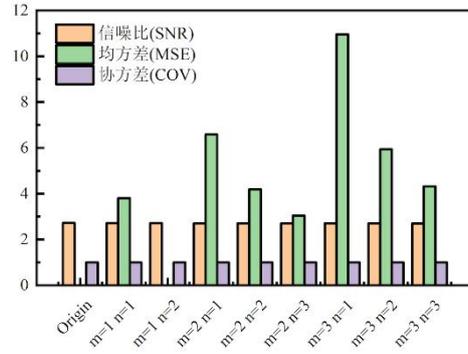
### B.2.3 一阶线性差分滤波原理

原理为用上一刻的输出和当前的输入进行加权求和得到当前的输出：

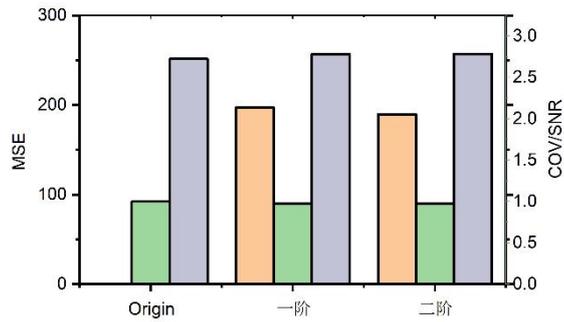
$$y[n] = ay[n-1] + bx[n] \quad \text{(B-4)}$$

## B.2.4 滤波器参数对比

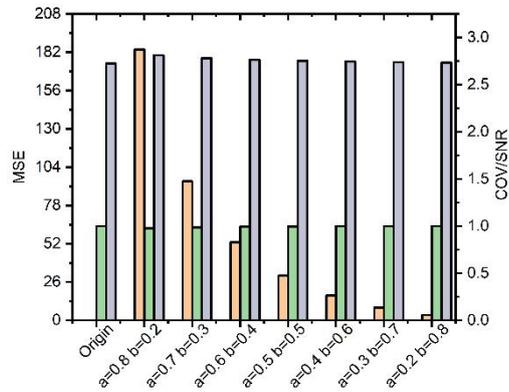
### 最小二乘平滑优化算法参数对比



### 卡尔曼滤波参数对比



### 一阶线性差分参数对比



## 在学期间参加课题的研究成果

[1] Li H, **Yang Y**, Ji Y, Peng W, Martin-Gomez A, Yan W, Qian L, Ding H, Zhao Z, Wang G. 3D Slicer-AR-Bridge: 3D Slicer AR Connection for Medical Image Visualization and Interaction with AR-HMD. [C] 22nd IEEE/ACM International Symposium on Mixed and Augmented Reality. IEEE., 2023. (共同一作在投)

[2] Li H, Yan W, **Yang Y**, Zhao Z, Ding H, Wang G. Inside-out Accurate Head Tracking with Head-Mounted Augmented Reality Device. [C] The 12th IFMBE Asian Pacific Conference on Medical and Biological Engineering, IFBME., 2023.



主要内容:

本文以提升术前规划效率和精度为研究目标, 具体研究内容:

1. 基于红外被动标记物的精准定位工具。利用 HoloLens 2 的红外深度传感器识别逆向反光标记物的特性制作十二面体状的追踪标记物, 实现在使用场景下精准定位工具尖端位置, 自主设计电路结构, 实现除定位外的应用层功能, 丰富工具和三维医学图像数据的交互方式。

<p>学生姓名</p>	<p>杨焯行 学号 2019011051 班级 生医 91</p>
<p>论文题目</p>	<p>基于增强现实的术前规划协作交互方法。创新性地结合二维医学图像软件和三维增强现实系统, 提出充分结合两者优势进行术前规划的完整流程。从不同模态的影像中提取有效信息并有序显示于增强现实显示器中。</p> <p>3. 增强现实术前规划协作系统构建和测试。以 Unity 3D 作为图形渲染器, 3D Slicer 作为医学影像数据处理器, HoloLens 2 作为虚拟图像显示器, 构建一套高效、实时的术前规划协作系统。</p> <p><b>进度规划:</b></p> <p>2023 年 1-2 月 完成定位算法</p> <p>2023 年 3 月 完成工具设计和算法测试</p> <p>2023 年 4-5 月 设计交互逻辑和构建交互系统, 中期答辩</p> <p>2023 年 5 月 系统测试与集成</p> <p>2023 年 6 月 撰写论文和答辩</p> <p style="text-align: right;">指导教师签字: _____</p> <p style="text-align: right;">考核组组长签字: _____</p> <p style="text-align: right;">年 月 日</p> <p>通过中期考核。</p>
<p>主要内容以及进度安排</p>	<p>中期考核意见</p> <p style="text-align: right;">考核组组长签字: _____</p> <p style="text-align: right;">年 月 日</p>

<p style="text-align: center;">指导教师评语</p>	<p style="text-align: right;">指导教师签字: _____ 年 月 日</p>
<p style="text-align: center;">评阅教师评语</p>	<p style="text-align: right;">评阅教师签字: _____ 年 月 日</p>
<p style="text-align: center;">答辩小组评语</p>	<p style="text-align: right;">答辩小组组长签字: _____ 年 月 日</p>

总成绩: \_\_\_\_\_

教学负责人签字: \_\_\_\_\_

年 月 日