

光学学报

多模态图像引导手术导航进展

杨健^{1,2}, 王媛媛^{1,2}, 艾丹妮^{1,2}, 宋红³, 范敬凡^{1,2}, 付天宇⁴, 肖德强^{1,2}, 邵龙³, 顾瑛^{5**}, 王涌天^{1,2*}

¹北京理工大学光电学院, 北京 100081;

²北京市混合现实与新型显示工程技术研究中心, 北京 100081;

³北京理工大学计算机学院, 北京 100081;

⁴北京理工大学医学技术学院, 北京 100081;

⁵中国人民解放军总医院激光医学科, 北京 100853

摘要 手术导航综合运用器官分割建模与手术规划、位姿标定与跟踪定位、多模态图像配准与融合显示等技术,使医生精确定位病灶与手术工具的位置,透过组织表面对内部组织进行观测,可大幅提升手术的安全性,缩短手术时间并提高手术效率。常规手术通常使用超声、内窥镜或X光等单模态影像进行手术过程引导,信息单一且均为二维影像,空间立体信息缺失,手术过程严重依赖医生经验;而多模态图像引导的手术导航技术通过融合多模态图像的优势,在三维空间提供病灶的结构或功能信息,大幅提升医生对血管、神经以及重要组织结构的空间辨识力。由此,本文针对多模态图像分割建模、手术方案决策、手术空间位姿标定与跟踪、多模态图像配准、图像融合与显示等多模态图像引导手术导航的关键技术进行总结和分析,提出其进一步发展面临的挑战并展望其未来发展趋势。多模态图像引导手术导航技术已成为神经外科、颅颌面、骨科、经皮穿刺、血管介入等临床科室精准治疗的新兴手段,具有重要的应用前景。

关键词 手术导航; 手术机器人; 多模态医学影像; 医学图像处理; 定位跟踪; 配准融合

中图分类号 TP249

文献标志码 A

DOI: 10.3788/AOS230742

1 引言

精准微创介入治疗在效率、创伤、疗效、费用等方面显著优于传统外科手术。手术导航是实现精准、安全、微创手术的核心手段。纵观外科手术的发展历程,从最初创伤大、疼痛重的开放式手术,逐步发展为超声、内窥镜和X光等单一模态图像引导的微创治疗。而今,由于国家精准诊疗发展的迫切需求,联合二维图像与三维图像的多模态图像引导手术导航应运而生。微创手术也从过去的依赖经验、精度低,向精准量化、安全可控的方向发展。多模态影像引导手术导航具有创伤小、疗效好、恢复快、费用低、可重复等优点,可实现微小切口下的良好疗效,并降低手术的并发症,因而成为21世纪极具前景的微创治疗方法之一,是外科手术发展的必然趋势^[1],已成为国际精准诊疗的前沿热点^[2-3]。

手术导航系统由成像设备、跟踪定位核心部件、末端执行机构、手术工具等硬件组成,并将现代影像技术、立体定向技术、计算机技术和人工智能技术相结合,使患者获得安全、精确、微创的手术治疗。手术导航系统涉及多模态图像分割建模、手术方案决策、手术

空间位姿标定与跟踪、多模态图像配准、图像融合显示等各个方面的核心理论和方法,进而解决三个关键任务:一是通过路径规划,告诉医生“手术怎么做”;二是通过空间定位,让医生知道“病灶在哪里”;三是通过融合显示,告诉医生“皮下有什么”。

医学图像作为手术导航的主要信息来源,其成像技术和图像质量的不断提升为手术导航的革新提供了基础,成为医生进行手术操作的必备工具。超声(US)、X光、内窥镜、电子计算机断层扫描(CT)和磁共振成像(MRI)等是常用的医学影像模态,各类影像的成像特点不同,在术中对于病灶的呈现能力也各不相同。US因具有非侵入、无辐射、低成本、便携、实时等特性,成为经皮穿刺手术最常用的术中成像方式之一^[4]。然而与CT和MRI相比,US的图像对比度和分辨率相对较低,难以准确分辨微小病灶。内窥镜成像因其无创、安全、清晰的术中成像特点,已广泛应用于临床医学领域。内窥镜可通过口腔、鼻腔、呼吸道、直肠等进入体内,辅助医生观察内脏器官的实时状况,并进行病变切除、活检等治疗操作^[5]。然而其成像视野有限,无法显示组织深度信息,图像纹理特征弱,易受高光反射影响。X光成像是透视投影成像,其基于人

收稿日期: 2023-03-29; 修回日期: 2023-06-10; 录用日期: 2023-06-30; 网络首发日期: 2023-07-10

基金项目: 国家自然科学基金(62025104)

通信作者: *wyt@bit.edu.cn; **guyinglaser301@163.com

体不同组织结构的密度不同,通过 X 光穿透人体后所能接收到的信号强度差异而进行成像,目前已被应用于骨科、介入治疗和心血管等手术引导中。然而,X 线透视受深浅组织影像相互重叠和隐藏的影响,通常需要多次多角度拍摄,并且其密度分辨率有限,难以分辨密度差异较小的组织器官和病变^[6]。CT 和 MRI 分别是对骨骼结构和软组织结构显像效果最佳的医学影像,而 CT 具有电离辐射,在术中照射应用时对患者和医生的损伤较大。MRI 影像扫描时间长,难以实现术中实时成像,且 MRI 在应用于术中引导时需要采用低磁场运动模式,从而降低了图像质量,限制了导航的精准度。

多模态图像引导手术导航有效利用多模态图像互补信息,克服单一影像引导所面临的技术挑战^[7]。患者需要在术前进行医学影像扫描,并在计算机上对感兴趣区域进行分割、建模,并选择最佳手术路径,制定最佳手术方案;术中将术前信息与术中影像进行信息配准并实时渲染,同时利用外部跟踪设备实时定位固定在手术器械上的标志物,实现手术器械定位,并将其位置在术中影像上实时融合显示,进而直观、精准地引导手术操作。多模态图像引导手术导航系统可为外科医生提供器械实时定位信息,帮助医生在切除病变组织的同时避开重要组织结构直达目标位置,有效地保护病灶周围的重要组织、血管与器官,避免不必要的损伤并降低手术并发症发生的概率。

手术导航因其巨大的临床应用潜力获得了广泛研究,技术上已经取得系列突破并成功应用于临床,辅助医生高效地完成手术^[8-9]。自 1994 年加拿大 ISG 公司开发的第一台手术导航系统 ISG Viewing Wand 问世以来,诸多商用手术导航系统被成功推入市场,如美国美敦力、德国 Brainlab、加拿大 7D Surgical 等知名公司开发的导航产品。与此同时,手术导航机器人发展迅速,如美国直觉外科、奥地利 Interventional Systems、印度 Perfint Healthcare、美国 Medical Robotics Technology、新加坡 NDR Medical Technology 等公司研发的导航机器人相继用于神经外科、颅颌面外科、骨科、经皮穿刺、血管介入等领域。近年来,国内导航系统发展势头迅猛,如天智航、微创、艾瑞迈迪、华科精准、柏慧维康、威高、深圳安科、迪凯尔等都取得了不错的成绩。手术导航系统从仅根据术前影像数据为外科医生提供结构信息,逐渐演变为具有重塑三维病灶结构、显示术野周围特征、规划最优手术路径、辅助医生在操作器械过程中规避重要组织结构、精准到达病灶部位等多功能的综合化手术引导系统,已成为临床多种术式中的重要工具。

本文围绕多模态图像引导手术导航,概括性地介绍了手术导航的硬件组成,并针对器官分割建模与手术规划、位姿标定与跟踪定位、多模态图像配准与融合显示等多模态图像引导手术导航的关键技术开展调研

分析。同时,对多模态影像手术导航系统的临床应用现状进行了总结概括。最后,对多模态图像引导手术导航面临的挑战和未来发展方向进行了讨论和展望。

2 多模态图像引导手术导航的硬件组成

2.1 术中成像设备

临床上用于术中成像的设备一般包括 C 型臂 X 光机、超声扫描仪和内窥镜系统。基于 X 射线图像引导的手术导航常应用于肝血管、冠脉等介入治疗中^[10],术中需借助 C 型臂对病灶区域进行透视成像。X 射线图像可实时呈现介入器械的位置及病灶血管的形态^[11],但持续的 X 射线及造影剂的使用将会对患者及医生造成额外的辐射伤害,且其透视成像的特点造成血管立体空间结构信息缺失,术中对医生的专业知识、临床经验及三维空间想象力要求较高^[12]。US 成像由于其成像具有非侵入性、实时性、无辐射和低成本等优点,被广泛应用于肝脏、前列腺、脑部等多个临床场景中。与此同时,US 图像灰度值代表声学阻抗的变化,由明显的斑点噪声和各种方向伪影所覆盖,US 图像存在图像信噪比、软组织对比度和分辨率均较低的问题。临床上,往往将三维术前图像的高病变可见性和丰富的空间信息与术中二维超声图像的实时成像能力相结合,从而提高器官病变的可识别性和定位精度,目前已被应用于肝脏、颅脑和前列腺等器官肿瘤治疗中。基于内窥镜成像的微创手术具有创口小、成像直观、患者疼痛更小和住院恢复时间更短等特点。在内窥镜引导的微创手术中,外科医生仅需要通过自然孔道或单切口置入内窥镜就可以获得对病灶较好的观察。内窥镜技术已在外科手术的各个领域得到普及。然而,由于内窥镜成像的对象通常在狭窄空间的软组织表面,存在较大的均匀区域、镜面反射和光照不均等问题。此外,由于内窥镜成像结果为三维结构的二维投影图像,故丢失三维信息,无法提供有关器官表面下的组织结构信息^[13],因此需借助导航技术引入术区三维先验,更好地辅助外科医生进行安全、精准的手术。近年来,荧光成像也逐渐应用于图像引导肿瘤手术治疗中。吖啶菁绿(ICG)作为一种安全无毒的菁染料已被尝试用于引导肝脏肿瘤切除手术,因其荧光特性可发射波长 840 nm 左右的近红外光,组织穿透范围在 0.5~1.0 cm 之间。ICG 的荧光成像特性可用于实时可视化生物结构,通过观察肿瘤组织和非癌性肝组织中 ICG 的胆道排泄紊乱来检测肝癌^[14]。同时,注射少量 ICG (0.25 mg) 的靛蓝胭脂红溶液可以提高肝段识别的成功率,特别是对于肝硬化和因既往手术导致厚结缔组织覆盖的患者^[15]。但 ICG 也有着显著的缺点,例如 ICG 荧光成像仅限于 ICG 染色的生物结构,难以观察未染色的组织,且容易受到环境条件等因素的干扰,严重影响成像质量。另外,光动力治疗(PDT)是一种符

合 21 世纪肿瘤治疗发展方向(靶向、局部、个体化)的新技术^[16],其作用机理是光敏剂经静脉给药后富集于肿瘤组织中,经激光照射激发产生大量活性氧物质,选择性地清除病变组织,实现靶向治疗目的。光动力治疗与手术、化疗、放疗并列为肿瘤治疗的四种主要方法^[17]。光动力治疗在腔内照射时所用的激光光纤(管径仅 100~600 μm)可顺畅通过内窥镜工作通道抵达腔道病变部位。目前,凡能耐受常规内窥镜检查又能够清晰窥见的部位,如喉/鼻咽、肺、食管、胃肠、胰胆管、肠、膀胱的恶性肿瘤都可进行光动力治疗。

2.2 跟踪定位核心部件

应用于手术导航的跟踪定位核心部件可分为光学、电磁、激光和结构光定位系统。光学定位系统通过由间隔一定距离的多个摄像机组成的立体视觉图像采集系统,从不同方位同时获取目标及其背景的二维图像,随后利用特征识别算法提取二维图像中的目标特征信息,并根据视差原理从二维特征中重建目标的三维坐标。术中,定位标识器被固定在手术器械、成像设备、患者、参考框架等需要被跟踪的刚性物体上。为实现实时鲁棒跟踪,首先利用光学定位系统对标识器进行定位,而

后根据标识器与目标的相对位姿关系计算出目标物体的位置和方向^[18]。根据标识器与定位系统光照模式的不同,可将光学定位系统分为被动式和主动式两类。

被动式光学定位系统通过捕获标识器反射的光能量进行目标图像的捕获,从而实现对标识器的定位和跟踪。此类系统要求标识器在图像中有区别于其他物体的纹理信息,商用系统中常用的被动标识器包括反光球、反光贴,以及玻璃球等^[19-20]。在神经外科、颅颌面外科手术中,反光球使用较广泛,在骨科领域,反光贴和玻璃球因其便于清洗和重复利用而使用更多。被动式系统的定位精度会受到双目标定及图像分辨率的直接影响,同时环境背景的高光反射会在图像上形成错误的斑点,导致定位数据出错。因此被动式系统在定位精度和稳定性方面有待提高。主动式光学定位系统采用红外发光二极管作为标识器,并将其安装于目标物体上。立体视觉系统对标识器进行红外成像,并重建出中心点三维坐标,从而获得标识器和目标物体相对于参考坐标系的空间姿态。主动式光学定位系统由于抗环境光和背景干扰的能力强,因而在定位精度和稳定性方面均优于被动式光学定位系统。

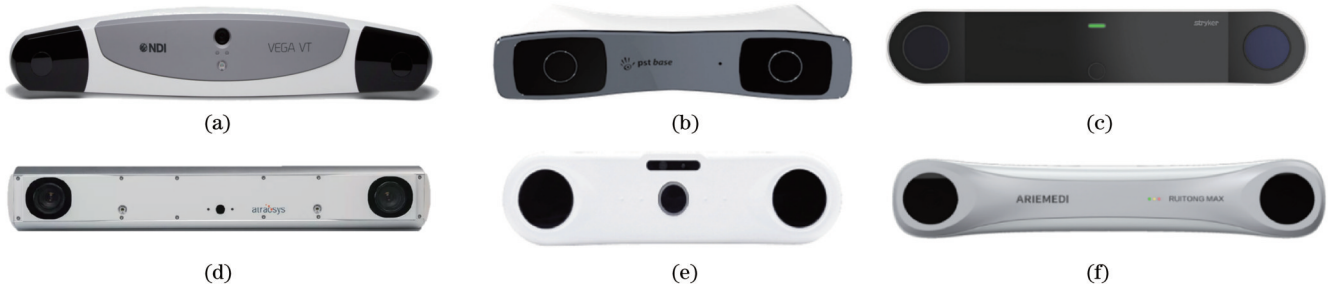


图1 商用光学定位系统。(a)NDI公司的Polaris Vega光学定位系统;(b)PS-Tech公司的PST Base定位系统;(c)Stryker公司的FP8000定位系统;(d)Atracsys公司的FusionTrack 500定位系统;(e)艾目易公司的AimPosition定位系统;(f)艾瑞迈迪公司的瑞瞳MAX定位系统

Fig. 1 Commercial optical positioning systems. (a) Polaris Vega optical positioning system (NDI©); (b) PST Base positioning system (PS-Tech©); (c) FP8000 positioning system (Stryker©); (d) FusionTrack 500 positioning system (Atracsys©); (e) AimPosition positioning system (Aimooe©); (f) RT MAX positioning system (Ariemedi©)

在商用光学定位系统中,加拿大NDI公司推出的Polaris系列占据了大部分市场,被业界公认为光学定位系统的标杆,该公司的多款定位系统的精度均可达到亚毫米量级,且稳定性极佳。荷兰PS-Tech公司开发的PST系列光学定位系统同样占有较大市场,旗下的PST Pico、Base、Iris系列定位系统可适应不同场景的应用需求,跟踪距离涵盖5~700 cm的深度空间。美国Stryker公司近期推出了一款颠覆性光学定位系统,命名为FP8000专利摄像头,该定位系统拥有无与伦比的刷新速度,可灵活选用多种光学跟踪模式,包括全光谱主动与被动混合光学跟踪。瑞士Atracsys公司开发的Atracsys FusionTrack系列光学定位系统拥有超高速刷新帧率,达335 Hz,支持同时追踪主动和被动标记,该系统在2 m跟踪范围内的定位误差小于

0.1 mm。然而,上述国外光学定位系列产品价格昂贵,且供货及返修周期长,存在“卡脖子”的风险,因此,国产化替代的需求日趋急迫。国内艾目易公司主推的AimPosition系统以及由北京理工大学研发团队自主研发,并转化至艾瑞迈迪公司的瑞瞳MAX系列光学定位系统成为打破技术垄断的试金石。瑞瞳MAX光学定位系统将视觉定位算法集成于可编程门阵列(FPGA)板卡,提升了系统刷新率,可实时精准地追踪标记物的位置和方向,同步追踪多个标记点,为医疗、工业等领域提供核心跟踪定位解决方案。图1分别展示了Polaris Vega、PST Base、FP8000、FusionTrack 500、AimPosition、瑞瞳MAX等光学定位系统。

尽管光学定位系统具有高精度定位性能,但此类系统易受遮挡影响。而电磁定位系统可有效解决遮挡

问题,主要由磁场发射器和磁传感器组成,该系统利用三轴线圈发射低频磁场,并通过固定在目标物体上的三轴磁传感器探测磁场变化,而后根据发射信号和感应磁信号间的耦合关系计算目标物体相对于磁场线圈的位置和方向。Sheinker等^[21]提出了针对配对磁信标安置方式的定位方法,该方法无须输入每个信标的精确磁矩即可实现对角线位置的定位,但由于信标不在同一平面会产生较大定位误差。

国外商用电磁定位系统主要由加拿大NDI公司垄断,该公司研发的电磁定位系统Aurora具有精度高、抗干扰强和易于集成等特点,已成功应用于腹腔软组织穿刺、介入治疗及骨科手术的辅助治疗中。



图2 商用电磁定位系统。(a) NDI公司的Aurora定位系统;(b) Polhemus公司的VIPER定位系统;(c) Polhemus公司的LIBERTY定位系统

Fig. 2 Commercial electromagnetic positioning systems. (a) Aurora positioning system (NDI©); (b) VIPER positioning system (Polhemus©); (c) LIBERTY positioning system (Polhemus©)

激光定位系统具有测量精度高、响应速度快的优点,该系统依据三角测量法则,由线激光器发出单束或多束激光投射到目标表面,利用双目相机捕获带激光线的目标图像,而后基于系统标定得出参数建立数学模型,重建目标的三维信息,最后执行术前医学影像与重建点云的注册^[22]。Ledderose等^[23]验证了表面激光注册后鼻窦和侧颅底的导航准确性,对称双侧表面激光配准的总体误差小于2 mm。Marmulla等^[24]将激光注册技术用于患者耳廓,在CT成像和激光扫描过程中,在耳廓不变形时激光注册的精度较高。由于激光注册的耗时相对较长、激光扫描为稀疏点云且有辐射,用于手术导航的激光注册技术的研究处于下降趋势。

结构光定位系统通过特定参数配合手术灯对患者暴露的解剖部位进行投射,而后通过双目相机进行捕获,并创建一个高精度的三维表面。国外7D Surgical公司的Envision 3DTM图像导航系统[图3(a)]可在几秒钟将术中图像与患者术前CT或MRI影像相结合^[25]。通过将结构光图案投射到患者解剖结构上,利用可见光相机捕获图案并进行相位解包,从而可获得密集三维点云,而后通过点云配准算法将术前医学影像与重建点云进行匹配。北京理工大学研究团队提出一种基于手持式三维扫描仪的结构光配准技术^[26],该方法通过手持式三维扫描仪获取患者面部点云,进而将术前CT影像配准至扫描点云以获取术前CT影像与患者空间的位姿映射关系。华科精准公司推出的SR1-3D神经外科手术导航机器人首次搭载了结构光

Aurora的小型线圈可与多种介入工具结合,为医生提供患者体内的实时导航影像。Polhemus公司最新款电磁跟踪系统VIPER具有体积小、采样速率高等优点,该产品具有多通道采样的扩展能力,无视线遮挡。该系统采用了飞真技术,可在需要时及时补偿失真,确保高保真位置追踪。Polhemus LIBERTY电磁跟踪系统配备了优质数据保证和失真检测标准,该系统利用交流电磁技术可实现对物体位置的连续跟踪。图2分别展示了Aurora、Polhemus VIPER和Polhemus LIBERTY等系统。然而,电磁定位系统对工作环境中的电磁干扰较为敏感,对手术环境有较高要求。

相机,作为手术导航系统的“眼”,3D结构光技术需数十秒即可采集头颅的数百万个特征深度和位置信息,与人工注册相比注册效率和精度都得到大幅提升。目前,用于手术导航的商用结构光相机普遍适用于神经外科、脊柱穿刺等术式,对关节置换、腹腔穿刺等术式的应用仍须进一步验证。

2.3 导航执行部件

手术导航系统中,机械臂是最为常用的导航执行部件。目前适用于手术导航系统的医用机械臂主要分为主从式机械臂和协作式机械臂。相较于主从式机械臂,更加小型化的协作式机械臂通过与力学、视觉及触觉传感的集成,结合自动控制、智能避障等算法,可实现更好的人机交互,完成更复杂的手术任务,减轻医生的手术操作压力。目前手术导航机器人多采用协作式机械臂,由导航平台控制,自动控制末端手术器械移动至目标位置,同时在治疗过程中始终保持稳定,减少扰动对治疗效果的影响,从而提高手术的精确度和安全性。例如在脊椎手术导航中,机械臂能够在导航系统引导下执行脊柱椎弓根螺钉的精确定位和椎弓根切除等操作^[27]。关节手术机器人借助机械臂辅助医生快速准确截骨、安装假体^[28]。在图像引导腹腔穿刺治疗中,机械臂辅助医生准确定位穿刺点,稳定精准地执行穿刺操作^[29]。

导丝是微创血管介入治疗中最常用的器械,用于建立经皮至血管的入路。血管造影剂注入、支架植入以及狭窄管腔的疏通均依赖导丝的侵入操作。在传统

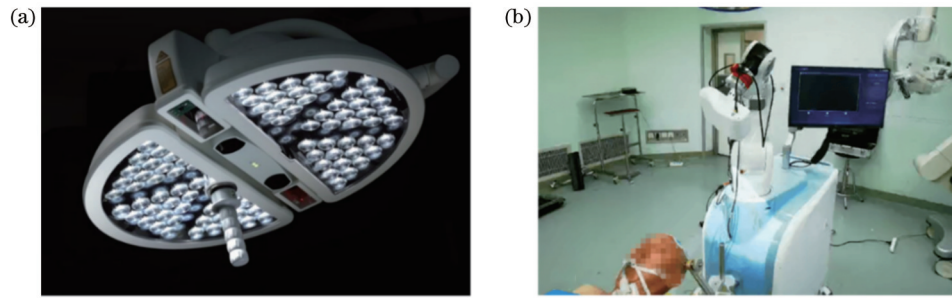


图 3 用于手术导航的 3D 结构光技术。(a)Envision 3DTM 图像导向系统^[25];(b)华科精准的结构光相机

Fig. 3 3D structured light technology for surgical navigation. (a) Envision 3DTM image guidance system^[25]; (b) structured light camera (Sinovation©)

血管介入治疗中,医生通过推拉和底部旋转来操作导丝,这对医生的操作技巧和注意力具有较大挑战^[30]。此外,大多数微创手术中,外科医生经常使用 X 射线透视获得血管的实时视觉反馈,增加了医生和患者的辐射暴露风险^[31]。Zhao 等^[32]提出了一种亚毫米量级、自润滑软连续体机器人,具有基于磁驱动的全方位转向和导航能力,该功能通过在其软体中编程铁磁畴并在其表面上生长水凝胶皮肤来实现,该系统在复杂和受限的环境中具有良好的导航性能。Jeong 等^[33]提出了一种采用两对亥姆霍兹线圈和控制算法的磁导丝装置,可进行导丝精确导航。Zhang 等^[34]提出了一种磁导向导丝,并设计了一个具有大工作空间的三维磁制动系统。所提磁导向导丝由商用导丝连接到由脂肪族芳香族无规共聚酯和磁粉制成的软执行器组成。制动系统包括两个对称部分,每个部分由三个直线电机组成。该系统的磁导丝可以在三维空间中轻松实现急转弯,并通过编程驱动磁场实现了在一个复杂而受限的

血管中的高精度导航。

2.4 跟踪定位工具

手术导航的跟踪定位工具主要包括定位探针和标定工装。定位探针按被跟踪模式的不同可分为光学导航探针、电磁探针等。光学导航探针主要包含如下几种:基于近红外识别的反光球探针、玻璃球探针和反光贴探针;基于可见光识别的激光探针和图形编码探针^[35]。近红外反光球、玻璃球和反光贴探针的识别原理一致,均通过立体视觉系统识别双目图像中标志球的中心,进而基于视差图重建标志球中心三维坐标。激光探针通过投射激光斑点至患者体表,而后利用立体视觉系统捕获激光斑点信息,进而重建斑点对应的三维坐标。图形编码探针的识别依靠可见光相机对其关键图形特征点的检测而实现,该探针往往应用于口腔颌面外科手术。图 4 展示了上述 5 种光学探针的示意图。电磁探针通过在手术器械上刚性固定磁传感器以跟踪手术器械的实时位姿。

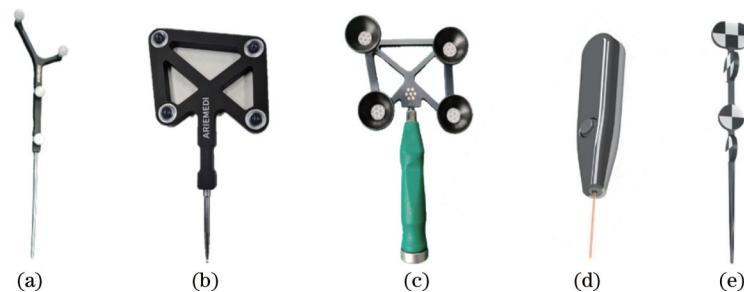


图 4 光学探针示意图。(a)反光球探针;(b)玻璃球探针;(c)反光贴探针;(d)激光探针;(e)图形编码探针

Fig. 4 Schematic diagram of optical probes. (a) Reflective sphere probe; (b) glass sphere probe; (c) reflective paste probe; (d) laser probe; (e) graphic coding probe

3 多模态图像引导手术导航的关键技术

3.1 多模态图像分割建模

基于术前多模态医学图像数据的分割建模技术可获得目标组织和器官的空间结构和位置信息,为术前手术规划和术中实时引导提供重要的数据基础。在术前常用的 CT/MRI 图像上,图像分割建模技术可实现

骨骼、实体器官、血管、神经等各类组织的信息提取与三维模型建立,在神经外科手术中,辅助医生进行患者头部手术入路及切口定位设计,确保手术精准安全^[36];在血管介入治疗中,实现路径实时规划与全局结构感知^[37];在穿刺介入治疗中,规划最优穿刺路径,为术者提供可视化路径辅助,指导临床医生精准实施介入操作^[38]。

然而,多模态医学图像不同模态间存在较大的物

理特性和图像特征差异,各类组织具有复杂的结构和多样的特征,且不同组织或器官之间相互交叠、灰度相似,且存在噪声、伪影等干扰因素。因此,多模态医学图像分割技术面临着诸多难点和挑战。近年来,围绕医学影像分割问题,国内外学者开展了大量研究工作,提出了多种不同分割方法^[39-40],主要包括阈值分割、匹配滤波^[41]、区域增长^[42]、图割^[43]、主动轮廓分割^[44]等,这些方法不依赖于大量训练样本,但当应用场景复杂时,其分割效果较差、分割速度较慢。深度学习作为一种高度非线性拟合算法,可自动从大样本数据中学习获得优良的高阶特征表达,具有较强的学习和泛化能力,已广泛应用于多种医学影像分割任务并成为当前医学图像分割领域的主流方法^[45]。

全卷积神经网络(FCN)是用于图像语义分割任务的代表性深度学习网络之一,该网络用全卷积层替代了卷积神经网络(CNN)末端的全连接层,对图像尺寸适应性强,避免了由于使用图像像素差异导致的重复计算和存储的问题。然而,FCN上采样分割结果模糊,对图像细节成分不敏感,分割效果不够精细。针对此问题,在FCN的基础上发展出了DeepLab系列^[46-47]、SegNet^[48]和级联FCN(CFCN)^[49]。Chen等^[46]提出了DeepLab-V1网络,通过在FCN之后串联一个全连接的条件随机场模型,提升了FCN的分割精度。DeepLab-V2将传统网络的VGG-16模块替换为ResNet模块,以进一步提升图像分割效果^[50]。另外,DeepLab-V2在DeepLab-V1的基础上增加了空洞空间卷积池化金字塔模块,利用多采样率的空洞卷积操作实现对特征图的并行采样,并融合输出结果获取更多空间信息。Badrinarayanan等^[51]以FCN的语义分割任务为基础,构建编解码器对称结构SegNet,该网络可利用较低分辨率对其输入特征映射进行上采样,从而显著降低模型的内存占用和计算负担。Christ等^[49]提出CFCN结构,通过叠加一系列FCN,利用从前一级模型的预测图中提取的上下文特征来提高图像分割性能。

在FCN的基础上,充分考虑多模态医学图像模态

差异大、组织边界模糊、噪声干扰大等特点,U-Net^[52]架构在保证全卷积结构的同时增加了编码与解码网络的跳连接,使网络具有完全对称性,且改变了FCN中较为简单的逐点相加的特征融合方式,借助残差连接将编码器与解码器各分辨率层级之间的特征信息进行通道维度的拼接,充分融合高维与低维特征,使得网络能够将上下文信息向更高层分辨率传播。该特性使得U-Net更加适应多模态医学图像的特点,因此U-Net及其变体网络已广泛应用于医学影像分割任务中^[53-54]。

另一方面,医学图像标注耗时耗力,且依赖专业知识和经验。因此,基于弱监督与无监督的深度学习成为当前研究的热点和难点。弱监督方法主要利用部分标注或标签不完整的数据进行图像分割,例如使用边缘信息或区域级别标注。而无监督方法则是利用未标注的数据进行分割,例如聚类、自编码器等。利用对抗生成网络(GAN)^[55]进行图像语义分割是近几年出现的弱监督或无监督代表性方法之一,通过训练使两个网络相互竞争,其中生成网络负责生成合成数据,而判别网络负责分辨数据真实性,可发现潜在的高维数据分布。Xue等^[56]基于GAN思想设计了一种分割对抗网络,以U-Net结构作为GAN的生成网络,并提出了一种多尺度损失函数来优化分割网络,成功克服了U-Net无法有效解决医学图像中像素类别不平衡的缺陷。Dai等^[57]提出结构校正对抗网络,简化了下采样模块,实现了基于胸部X光图像的肺和心脏区域的快速精准分割。

基于深度学习的医学影像分割研究覆盖面广,大量研究进展得到关注。然而,基于深度学习的多模态医学图像分割性能受医学影像训练数据量不足和深度学习模型自身缺陷的限制。进一步加强数据增强和基于元学习模型的小样本图像分割研究,引入大模型与泛化迁移训练理念,优化网络结构设计、三维数据分割模型设计和损失函数设计,是未来提高医学影像分割性能的重要研究方向。图5展示了EDDA科技公司的IQQA®-3D系统对不同组织器官的分割建模效果。

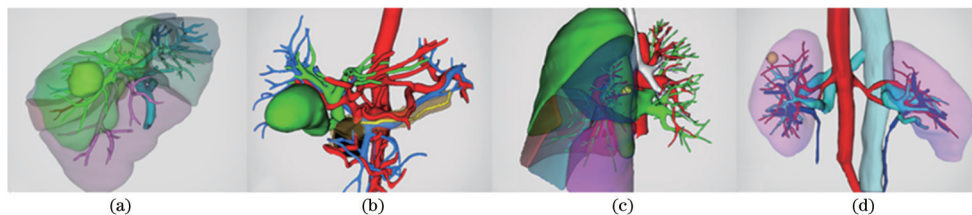


图5 组织器官分割建模效果图(EDDA IQQA®-3D)。(a)肝脏3D渲染图;(b)主动脉3D渲染图;(c)肺部3D渲染图;(d)肾脏3D渲染图

Fig. 5 Organ segmentation modeling results (EDDA IQQA®-3D). (a) 3D rendering of liver; (b) 3D rendering of aorta; (c) 3D rendering of lungs; (d) 3D rendering of kidneys

3.2 手术方案决策

手术方案决策用于引导术中手术操作,可通过组

织、器官、病灶三维模型位置关系制定手术方案。面向不同术式的术前方案规划存在较大的差异性,可分为

手术路径规划和手术方案规划两类。

在神经外科领域,医生根据术前患者脑部影像分割重建结果,手动勾勒出肿瘤轮廓,从而合理规划开颅切口。Dai 等^[58]提出一种新型交互式手术规划方法确定开颅切口位置,该方法基于病灶建模结果生成虚拟手术切口,并根据虚拟切口将实际的开颅手术路径规划到患者头部。在经皮穿刺治疗中,术前路径规划用于保障穿刺路径避开重要组织、器官等,基于最优化数值求解计算最佳穿刺路径^[59-60],一般可分为概率计算法和目标函数法。基于概率计算的方法可计算出穿刺针最佳转折点,但生成路径不平滑;基于目标函数的方法可生成较为平滑的穿刺路径,但计算量大、规划效率

低。研究如何快速精准地生成平滑的穿刺路径仍是目前手术规划领域亟待解决的关键问题之一。

手术方案规划以骨科与颌面外科为代表。在骨科手术中,医生通过术前规划系统设计骨肿瘤切除保肢重建手术的截骨平面^[61-62],避开血管、神经、软组织等重要的解剖结构,并进行个性化模拟假体的植入规划,计算假体植入后的力学参数。在颅颌面整形和下颌骨重建等颌面外科手术中,手术前须精确测量颌面结构的位置关系,计算截骨位置与植入物放置位置,规划和模拟最佳手术方案,以保证术后的治疗效果兼顾美观性和功能性^[63-65]。图 6 展示了美国 BrainLab 公司产品的



图 6 手术方案规划效果图(BrainLab®)。(a)放疗规划;(b)颅颌面手术规划;(c)脊柱手术规划

Fig. 6 Surgical planning results (BrainLab®). (a) Radiotherapy planning; (b) craniomaxillofacial surgical planning; (c) spinal surgery planning

3.3 手术空间位姿标定与跟踪

手术导航系统基于跟踪定位核心部件追踪术中手术器械的实时位姿,获取术前重建模型^[66]、术中患者和手术器械间的相对位置关系。该过程包括工具配准和患者配准。工具配准将手术器械坐标系转换到患者的

空间坐标系中;患者配准实现从患者的空间坐标系到 CT/MRI 等图像坐标系的三维刚性变换,从而将手术器械尖端的坐标信息通过一系列坐标系转换映射到术前图像坐标系中,进而在图像坐标系中精确显示患者与手术器械的实时位置关系。



图 7 体外生物标志物固定方法示意图。(a)立体定向框架(Leksell Stereotactic System®);(b)植入标志物;(c)模具配准;(d)体外定位标志架(Mako SmartRobotics™)

Fig. 7 Illustration of *in vitro* biomarker fixation method. (a) Stereotactic frame (Leksell Stereotactic System®); (b) surgical implantation marker; (c) mold registration; (d) *in vitro* positioning marker frame (Mako SmartRobotics™)

位姿标定方法根据配准点选取方式不同,可分为基于体外特征的配准和基于体内特征的配准。体外特征固定在患者体表的跟踪标志物,固定方式包括模具嵌套、立体定向框架、体表粘贴与螺钉植入骨质^[67]等(如图 7 所示),有创侵入式配准方式精度更高但会给患者带来损伤,无创标志方式因固定不稳导致精度略低。基于体内特征的配准可分为基于标志^[68]、面形^[69]和图像^[70]的三类配准方法。基于标志的配准方法使用

解剖及几何特征标志点作为配准基准,基于面形的配准方法使用图像分割得到的三维面形数据进行表面配准。这两种方法的计算速度快,但是依赖于标志点选取或分割的精度,操作复杂度高。基于图像的配准方法直接对图像像素灰度进行匹配操作,缺点是对图像质量及待配准图像间的相似度有较高的要求,配准稳定性较差。

基于标志物的选取与术前标定,术中通过光学或

电磁定位设备实现对术前医学图像、重建模型、术中患者和手术器具的空间位置坐标进行实时捕捉、更新和同步关联。术中患者的位姿和病灶结构在手术过程中发生变化,可精确校正患者移位并追踪手术工具,进而完成实时精准手术导航。除此之外,手术跟踪还可对重要组织器官进行安全预警^[71],医生通过实时监测手术器械与患者重要组织器官的实时距离,及时发现潜在的风险因素并采取相应的措施,确保手术的安全性。

3.4 多模态图像配准

在手术导航中,通过多模态影像配准,实现不同源的组织结构或功能信息互补。配准的实质是寻求多模态医学影像间坐标变换,使得它们在空间坐标系上对齐统一。图像配准流程包括:图像特征提取、特征相似性描述、几何变换求解。根据维度不同,多模态影像配准可以分为三维-三维(3D-3D)与三维-二维(3D-2D)配准两类,其中 3D-3D 配准以 CT、MRI、正电子发射断层扫描(PET)图像间的配准为代表,3D-2D 则以 CT/MRI-US 图像配准为代表。

对于 3D-3D 配准,由于图像维度一致,可以通过在线迭代相似性测度,优化刚性或非刚性变换模型参数,实现多模态图像的匹配对齐。Giesel 等^[72]选取 CT、MRI 和 PET 三种模态影像数据的对应特征点,将其作为控制点驱动样条模型求解最优非刚性变换。Sundar 等^[73]提出了一种 CT 和单光子发射计算机断层扫描(SPECT)图像配准融合的 Octree 方法,该方法通过估计图像熵的非均匀自适应采样,提高算法对图像灰度变化的鲁棒性,从而减少局部最大值对模态配准结果的影响。

传统 3D-3D 在线图像配准算法在迭代优化最佳变换参数时,易受图像间较大差异影响,导致出现计算复杂度高、配准精度有限的问题。为此,学者们提出基于深度学习的图像配准方法,通过网络训练,离线学习图像对之间的线性映射关系或非线性映射关系,弥补传统方法适应性差和复杂度高的缺陷。目前,基于深度学习的医学图像配准主要包括全监督、弱监督和无监督(自监督)学习三类。全监督学习的配准方法需要稠密形变场标签信息来学习网络参数^[74]。常用的形变标签包括随机生成的形变、由传统配准方法获得的形变^[75],以及基于模型生成的形变^[76]。Cheng 等^[77]提出一种堆叠降噪自编码器全监督网络,以 CT 和 MRI 的图像块(patch)之间匹配关系为形变场标签,学习图像之间的相似性度量,完成多模态刚性配准融合。然而真实的形变标签不易获得,且模型配准的性能受到标签质量的限制。为了降低对标签质量的限制,弱监督应运而生。弱监督学习是指在网络训练过程中同时利用图像间的相似性度量和样本的标签信息^[78]。Fan 等^[79]提出基于弱监督学习的图像配准网络,在设计目标函数时,同时度量变形后的待配准图像与模板图像之间的相似性,以及真实形变场与预测形变场之间的

相似性。Blendowski 等^[80]提出一种基于形状编码器-解码器网络框架的 MRI-CT 多模态弱监督配准方法。该方法借助图像的分割标注训练分割网络,根据分割结果迭代能量函数来估计图像对之间的形变,可降低多模态图像在大形变条件下的配准难度。弱监督学习可降低配准任务对于稠密真实形变场的依赖,更适用于临床应用。与无监督学习相比,弱监督学习更加充分地利用了实际采集的数据信息,为模型训练提供进一步指导。由于弱监督方法依然需要形变场标签,其训练效果本质上仍受限于标注质量,因此学者们提出无监督学习或自监督学习。无监督学习利用配准后图像与模板图像之间的相似性来驱动网络的学习^[81-82]。Lin 等^[83]构建了基于自适应通道注意力模块的无监督模型,不需要形变场标签,而是以配准后图像间的局部互相关作为损失函数,调整神经网络对不同形变特征的偏重,以解决模型偏重于较大形变而对小形变配准效果不佳的问题,并在损失函数中加入平滑性约束来提高形变场的平滑性,有效提升配准网络的泛化能力和稳定性。

对于 3D-2D 配准,由于术中二维图像缺乏立体结构信息,与术前三维图像匹配的目的在于拓展手术视野、定位手术空间,因此 3D-2D 图像以刚性配准为主,用于求解二维图像在三维图像空间中的位置、姿态^[10, 84]。临床上常用的二维图像包括二维投影图像(数字减影血管造影、数字重建放射图像)与二维切片图像(US 图像)。现有二维投影图像与三维影像的配准主要以灰度的相似性衡量图像间的差异^[85-86]。在这些传统方法中,通过模拟虚拟 X 射线的衰减,从三维 X 射线衰减图中得到模拟的 X 射线数字重建放射图像。通过优化器最大化数字重建放射和 X 射线图像之间基于灰度的相似性度量,在小范围内搜索最优变换参数,实现二维图像到三维图像的配准。尽管基于灰度的方法能够在空间中小范围内搜索变换参数达到较高的精度,但是需要的配准时间较长,且匹配范围受限^[87-89]。李文举等^[90]针对现有 2D-3D 配准方法大范围参数搜索精度不高、耗时较长等问题,提出一种基于训练-推理解耦架构的 2D-3D 医学图像全监督配准方法,通过生成模拟数据获取稠密形变场标签,以此作为金标准在提高 2D-3D 医学图像配准精度的同时,更好地满足了临床实时性需求。Wohlhart 等^[91]提出使用 CNN,从图像中学习二维图像位姿描述符,然后使用 k-Nearest Neighbor 进行位姿估计。Miao 等^[86]提出利用图像信息,引入自动特征提取来计算二维图像在空间中的位姿特征,并采用 CNN 直接估计变换参数。

在二维切片图像与三维影像配准的现有传统方法中,主要关注于多模态多维度相似性测度构建,如模态独立领域描述符^[92](MIND)、自相似上下文^[93](SSC)、自相似结构矢量^[94](WSSV)等。Wang 等^[94]采用不同模态图像的对应解剖结构轮廓相似特性,提出一种自

相似结构矢量加权的 US-MRI 图像配准方法(图 8)。该方法利用 US 探头标定信息初始化位姿配准矩阵,通过提取 US 图像和术前图像的归一化自相似结构矢量描述符,结合 US 图像纹理权重图,计算两个结构描述符间的加权相似性测度,利用非线性迭代优化方法实现配准。然而传统方法与初始变换参数极为敏感,为了达到最佳配准效果需要较为精确的参数初始化结

果。针对以上问题,Wei 等^[95]提出通过使用分类网络 ResNet 初始估计 US 探头的角度,在此基础上将配准问题表述为分割任务,通过分割任务学习估计三维 CT/MRI 中的 US 平面位姿。Markova 等^[96]提出使用 CNN 学习稠密的关键点描述符,然后根据描述符估计配准位姿,实现初始化配准。

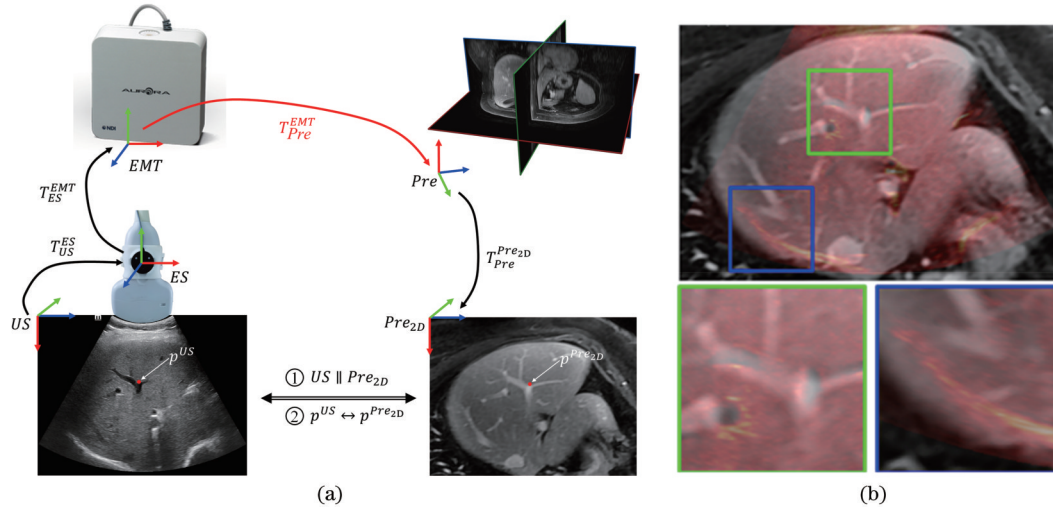


图 8 基于自相似结构矢量加权(WSSV)的MRI-US图像配准^[94]。(a)配准框架;(b)配准效果

Fig. 8 Weighted self-similarity structure vector (WSSV)-based MRI-US image registration^[94]. (a) Registration framework; (b) registration result

3.5 多源信息融合显示

多源信息融合显示是将不同图像、组织模型、手术方案、跟踪位姿等多种信息整合在同一坐标系下进行二维或三维显示,该技术克服了单一数据源在信息呈现上的局限性,辅助医生进行精准诊疗。

Dey 等^[97]通过对内窥镜的光学定位跟踪,结合内窥镜的内参,计算内窥镜图像在三维空间的坐标和姿态,将内窥镜图像中提取的组织表面结构模型,运用纹理映射与术前获取的 CT/MRI 进行融合。北京理工大学李文杰等^[98]利用内窥镜的内外参,将二维图像放置于三维空间,并与 CT 影像上分割的血管模型相融合,有效扩展狭小的内窥镜视野。Suciadi 等^[99]提出支气管镜图像与三维 CT 影像的融合算法,将畸变校正后的支气管镜视频序列图像与虚拟相机视图进行匹配,在此基础上,将病灶区域融合渲染到支气管镜视频图像空间。Behrens 等^[100]提出一种基于非线性多尺度融合的内镜图像增强渲染方法,利用内镜照明特性对图像掩膜进行建模,并对图像进行拉普拉斯金字塔分解与合成,该方法能够保留高光区域和细小血管的结构信息,解决图像融合过程的边缘结构模糊问题。Tan 等^[101]提出一种新的多模态医学图像融合方法,该方法基于边界测量脉冲耦合神经网络融合策略和能量属性融合策略,实现了 MRI 和 PET 图像,以及 MRI 和 SPECT 图像的高质量融合(图 9)。Higgins 等^[102]提出

了胸部 CT 影像和内镜视频图像的多模态融合模型,利用视频图像生成稀疏视频帧,并构建呼吸道和视频帧图像的跟踪匹配关系,进而实现内镜图像颜色黏膜信息与三维腔道内表面的融合映射。

三维的影像数据和面片数据的渲染是多源信息融合显示的基础。影像数据的体渲染(DVR)可实现对不同组织的快速预览,但是为了突出目标或者其重要部分结构,需要设置多维度的传递函数,并调节大量参数。不同于体渲染,面片渲染通过将组织稠密结构化体数据转换为表面非结构化稀疏的多个面片数据进行显示,通过调整颜色和透明度可灵活地展示组织拓扑结构,但当存在多个组织面片数据距离较近时,它们透明度的叠加容易出现错误进而造成视觉混乱,影响医生对进深的判断。

在基于增强现实的多模态图像引导手术导航中,面片和体数据的混合渲染可视化更为常用,因为常需要在同一场景中同时呈现体数据和多个面片数据。其中体数据能够用于表示软组织、骨骼和体腔等,经过渲染能够提高解剖结构周围环境信息,而面片数据常用于指示重点目标、手术工具等的边界信息。两种数据由于交错的空间关系,会出现相互嵌入的情况,因而导致其渲染结果存在一些干扰和错误。为了优化增强虚实融合渲染的结构信息、提高空间立体感并提高融合渲染效率,北京理工大学楚亚奎等^[103]对增强现实融合

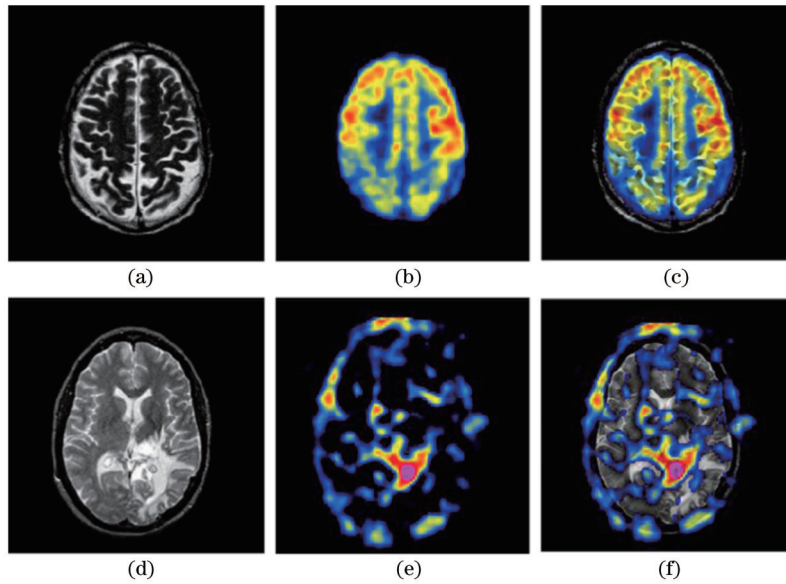


图 9 多模态医学图像融合^[101]。(a)患者 1 的 MRI 图像;(b)患者 1 的 PET 图像;(c)患者 1 的 MRI 与 PET 融合图;(d)患者 2 的 MRI 图像;(e)患者 2 的 SPECT 图像;(f)患者 2 的 MRI 与 SPECT 融合图

Fig. 9 Multimodal medical image fusion^[101]. (a) MRI image of patient 1; (b) PET image of patient 1; (c) MRI and PET fusion image of patient 1; (d) MRI image of patient 2; (e) SPECT image of patient 2; (f) MRI and SPECT fusion image of patient 2

渲染效果进行改进,提出了一种重要性驱动的数据融合渲染方法,通过着色权重映射、面片重要性驱动因子计算及深度缓存立方体剖切渲染,实现对病灶结构渲

染结果的颜色优化和手术规划路径进深感增强,并在此基础上研发了一套鼻内镜增强显示手术导航系统(图 10)。

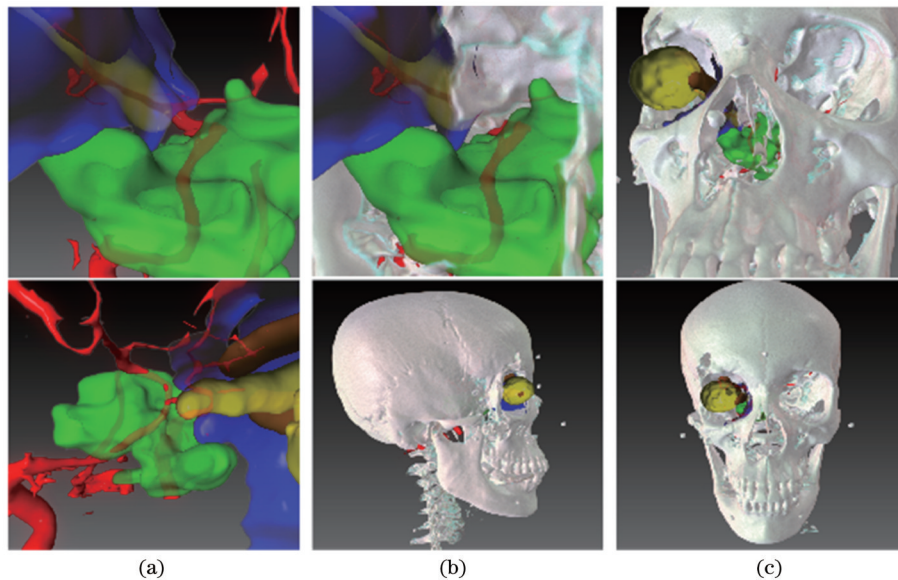


图 10 重要性驱动的数据融合渲染^[103]。(a)重要性排序;(b)侧面混合渲染效果图;(c)正面混合渲染效果图

Fig. 10 Importance-driven data fusion rendering^[103]. (a) Importance ranking; (b) hybrid rendering image on side view; (c) hybrid rendering image on front view

4 多模态图像引导手术导航的临床应用

4.1 神经外科手术导航

颅脑神经结构复杂、功能重要,神经外科手术操作空间狭小、难度高、风险大,要求术者有较高的手术操

作技能和经验。神经外科技术包括经颅骨介入与经鼻腔介入操作。在经颅骨介入治疗中,多模态图像引导手术导航可基于术前 CT 或 MRI 等三维医学图像对病灶进行分割建模,准确规划头皮上的开颅切口位置。另外,利用手术导航的术中器械定位跟踪和图像配准融合技术可为开颅缺损修复提供精准引导,提高患者

预后效果。在经鼻腔介入治疗中,针对传统内镜引导手术中内镜图像缺乏立体感、全局感及存在后端盲区等缺陷,多模态图像引导手术导航可将感兴趣区域虚拟叠加在鼻内镜图像视野中,让医生具备“透视眼”,实现对皮下组织结构的实时观测,提升手术视野的立体感,实现重要组织的预警,从而大幅提升手术的精确性和安全性^[44, 98]。目前,美国 Stryker、Medtronic, 法国

Medtech S. A. S.、德国 Brainlab 以及中国的柏慧维康、华科精准、上海复旦和深圳安科等公司均相继研发了神经外科手术导航系统,并已推广到临床应用。艾瑞迈迪公司研制的颅底手术机器人导航系统,具备虚实融合、安全预警等独有功能。图 11 展示了部分商业化神经外科手术导航系统。

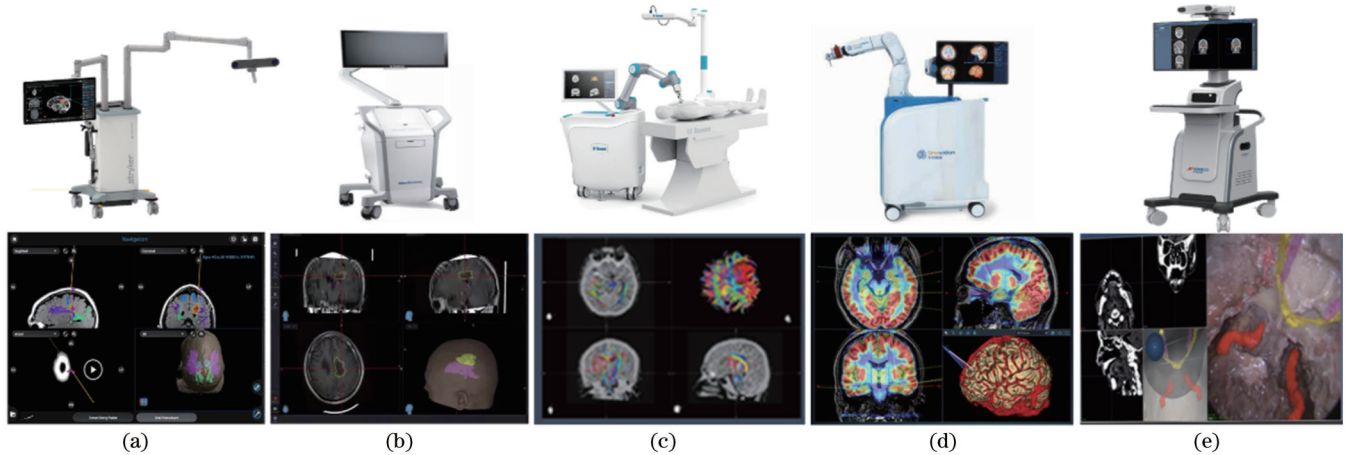


图 11 商用神经外科手术导航系统。(a)Stryker公司的Q导航系统;(b)Medtronic公司的 StealthStation™ S8手术导航系统;(c)柏慧维康公司的神经外科手术机器人系统;(d)华科精准公司的神经外科手术机器人系统;(e)艾瑞迈迪公司的颅底手术机器人导航系统
Fig. 11 Commercial neurosurgery surgical navigation systems. (a) Q navigation system (Stryker©); (b) StealthStation™ S8 surgical navigation system (Medtronic©); (c) neurosurgery surgical robot system (Remebot©); (d) neurosurgery surgical robot system (Sinovation©); (e) skull base surgical robot navigation system (Ariemedi©)

4.2 颅颌面外科手术导航

颅颌面外科手术局部解剖结构复杂,切口暴露有限,出血、神经损伤和医源性骨折的现象时有发生,是目前最具挑战性的手术之一。构建精准的颅颌面外科手术导航系统,术前合理规划手术路径,并在术中进行精确的引导定位,可有效减少手术过程中周围解剖结构不必要的损伤,已成为颅颌面外科手术领域的重要研究方向。在颌面肿瘤切除手术中,Tarsitano等^[104]利用术前三维重建技术确定肿瘤切除范围,并在术中叠加实时显示,帮助医生实时控制肿瘤切除范围。在正畸手术中,导航系统帮助术者在术中准确定位髁突位置,有效解决由于上颌骨与下颌骨相对移动造成的下颌骨髁突位置改变的问题,显著提升正畸疗效。在正颌外科手术中,手术导航系统有助于将上颌骨和下颌骨段放置在理想位置,提高手术的治疗效果和安全性。在商业化应用方面,美国医疗公司 Neocis 于 2016 年推出了世界上第一台商用牙科种植手术导航系统^[105]。国内,空军军医大学研发的“雅客智慧”种植牙手术导航系统是国内首款“自主式”种植牙手术导航系统^[106]。此外,迪凯尔公司研发的“易植美”牙种植手术导航系统和柏慧维康公司研发的“瑞医博”口腔手术机器人系统均已开始临床推广应用。上海优医基公司推出的 AccuNavi-A 颌面修复手术导航系统已在多例临床手

术中验证其性能。图 12 展示了部分商业化颅颌面外科手术导航系统。

4.3 骨科手术导航

多模态引导手术导航在骨科领域的应用主要包括关节置换、椎弓根置钉、椎体成形和全椎间盘置换等手术操作。关节置换术是治疗终末期关节疾患的有效手段。传统关节置换手术术后常因假体安放位置不妥、下肢力线不准确及软组织失衡等导致患侧关节疼痛及假体早期松动,手术效果欠佳。多模态图像引导骨科手术导航可根据患者术前关节 CT 影像三维重建结果,对关节截骨的位置、假体大小、接入方向及位置进行详细的术前规划,并在术中通过虚实融合将术前、术中影像叠加显示,准确引导术者进行每一项手术操作,精确安放符合肢体力线的手术假体^[107-108]。在椎弓根置钉手术中,手术导航有助于术者医生实时获取椎弓根和螺钉位置,以确定进针方向及置钉深度,实现安全精准的手术操作^[109]。在椎体成形手术中,手术导航可根据术前影像三维建模结果来模拟手术操作,规划进针位置、深度和角度等手术方案,提高椎弓根穿刺安全性和成功率^[110]。在全椎间盘置换手术中,手术导航的应用可更好地评估椎间盘置入物的大小、优化置入位置、减少手术并发症^[111]。目前,美国 Surgalign 公司研发的 Holo Portal 系统、Stryker 公司研发的 Q 导航系

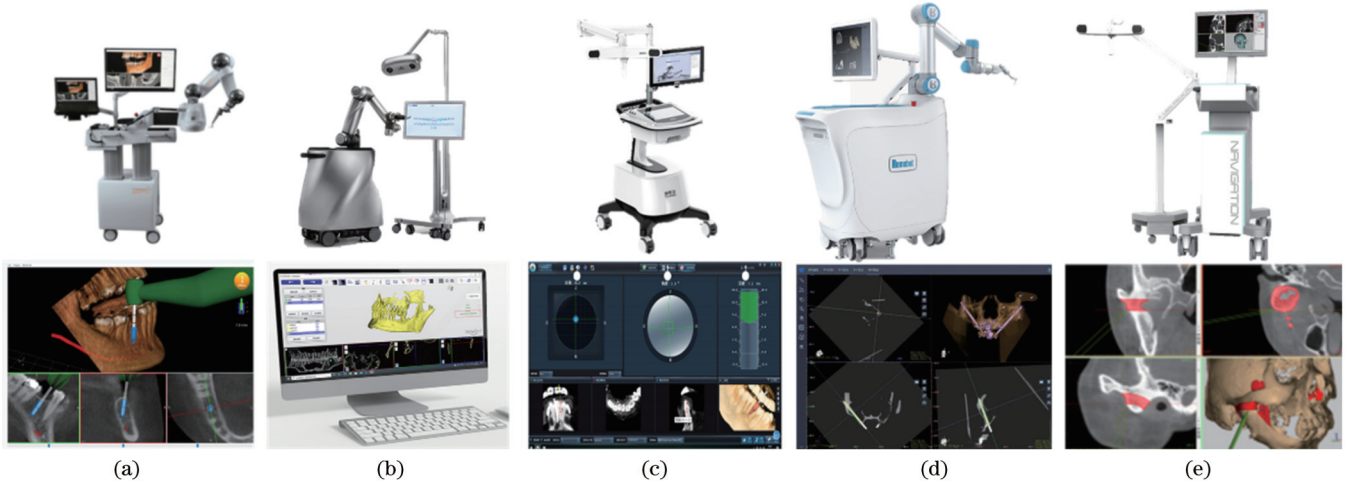


图 12 商用颅颌面外科手术导航系统。(a)Neocis Yomi 牙科手术机器人导航系统^[105];(b)“雅客智慧”口腔种植机器人系统^[106];(c)迪凯尔“易植美”口腔种植手术导航系统;(d)柏惠维康公司“瑞医博”口腔手术机器人系统;(e)优医基公司 AccuNavi-A 颌面修复手术导航系统

Fig. 12 Commercial craniomaxillofacial surgery navigation systems. (a) Yomi dental surgery robotic navigation system (Neocis©)^[105]; (b) dental implant robotic system (Yakerbot©)^[106]; (c) dental implant surgical navigation system (Dcarer©); (d) dental implant surgical navigation system (Remebot©); (e) AccuNavi-A maxillofacial prosthetic surgery navigation system (UEG©)

统、Novarad 公司研发的 VisAR 系统,以及中国台湾 Point Robotics Medtech 公司的 Point Kinguide 系统、北京天智航公司的天玑系列机器人、上海微创公司的鸿鹄骨科手术机器人等骨科手术导航与机器人设备均已

实现临床应用。此外,国内和华瑞博、罗森博特、骨圣元化、长木谷、威高、纳通、鑫君特、键嘉等企业也相继推出了骨科手术导航机器人,并实现产业转化落地。图 13 展示了部分商业化骨科手术导航系统。

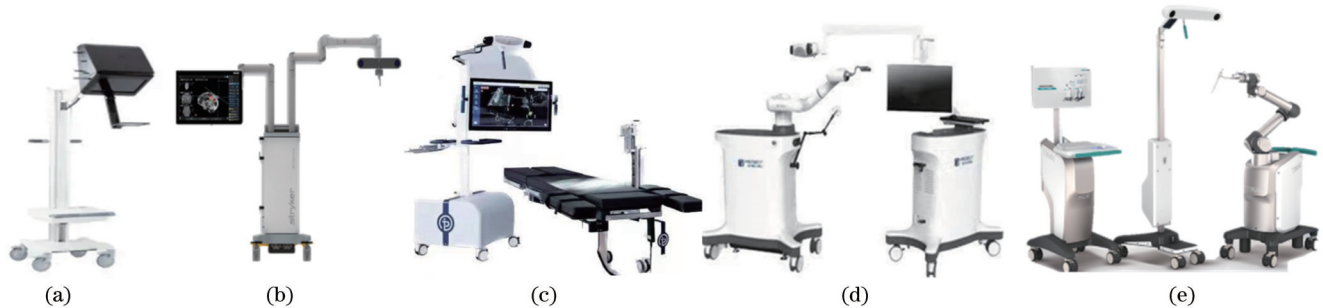


图 13 商用骨科手术导航系统。(a)Surgalign 公司的 Holo Portal 系统;(b)Stryker 公司的 Q 导航系统;(c)Point Robotics Medtech 公司的 Point Kinguide 系统;(d)上海微创医疗器械公司的鸿鹄骨科手术机器人系统;(e)北京天智航公司的天玑骨科手术机器人系统

Fig. 13 Commercial orthopedic surgery navigation systems. (a) Holo Portal system (Surgalign©); (b) Q navigation system (Stryker©); (c) Point Kinguide system (Point Robotics Medtech©); (d) SkyWalker™ system (Microport©); (e) TianJi™ system (TINAVI©)

4.4 经皮穿刺导航

在临床经皮穿刺过程中,医生需反复对比术前高分辨率 CT/MRI 图像与术中超声影像的纹理结构信息,凭经验识别超声影像中病灶及管道结构,并判断穿刺针与病灶的空间位置关系。上述过程耗时长且严重依赖医生经验,医生手术操作负担较重。多模态图像引导经皮穿刺导航旨在将术前影像与术中超声配准,提升术者对于手术区域的三维感知,增强规划路径显示,大幅提升手术效果。Hong 等^[112]提出了一种专用于腹部经皮穿刺治疗的超声引导穿刺导航系统,与单一影像引导相比,该系统的术前规划时间更短且穿刺更准确。Krucker 等^[113]研发了基于电磁跟

踪的超声引导穿刺导航系统,通过实时跟踪体表电磁传感器位置信息,实现因患者位姿变化导致的配准误差的补偿。此外,科研人员将机器人技术引入经皮穿刺导航系统以解决穿刺手术中医生手部抖动和用力不均导致的穿刺精度低、稳定性差的问题。目前商用穿刺介入导航系统已完成初代产品研发并逐步应用于临床。其中,奥地利 Interventional Systems、印度 Perfint Healthcare、美国 XACT Robotics 和新加坡 Biobot 等医疗设备公司构建了多款经皮穿刺导航系统。国内精励医疗和真健康医疗科技有限公司均推出了经皮穿刺导航系统,并应用于临床。图 14 展示了部分商业化经皮穿刺手术导航系统。

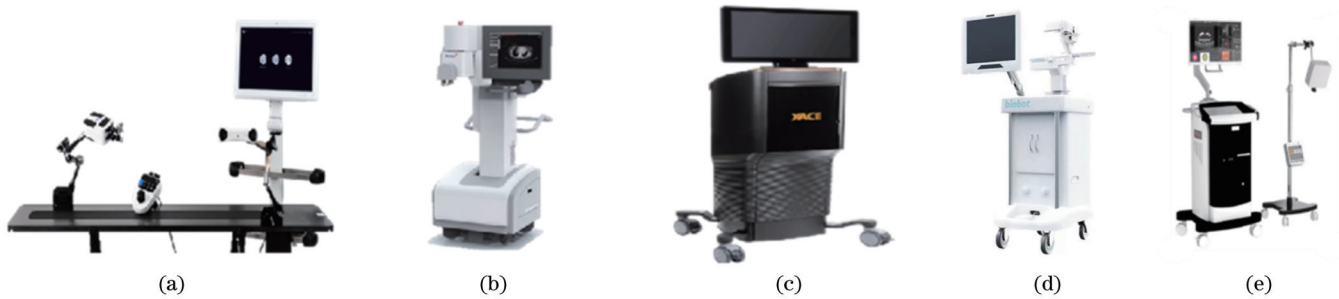


图 14 商用经皮穿刺导航系统。(a) Interventional Systems 公司的 Micromate™; (b) Perfint Healthcare 公司的 MAXIO; (c) XACT Robotics 公司的 XACT; (d) Biobot 公司的 iSR'obot; (e) 精励医疗公司的穿刺导航系统

Fig. 14 Commercial percutaneous puncture navigation systems. (a) Micromate™ (Interventional Systems©); (b) MAXIO (Perfint Healthcare©); (c) XACT (XACT Robotics©); (d) iSR'obot (Biobot©); (e) puncture navigation system (Jingmai©)

4.5 血管介入导航

血管介入是心血管疾病最常使用的治疗手段之一。手术过程中,医生首先将穿刺针从手腕或大腿根处的动脉血管插入人体体内,之后将导管套在导丝上送进血管,后沿着动脉血管上行至心脏,寻找病灶。手术过程中,导丝对血管壁误触,导管刺激和支架植入位置不当,可能会引起血管壁损伤、血管闭塞、局部夹层和胸痛等并发症。血管介入导航系统可在术前利用多模态图像分析技术,获取血管走形、血管病灶分级、介入位置等手术目标信息,辅助医生制定最佳手术方案,在手术过程中,利用术前术中的多模态图像配准融合技术,可弥补术中二维图像无法提供组织立体空间信息的缺陷,帮助医生实时精准定位导丝/导管和病灶三维空间位置,减小导丝导管误触风险,提升支架植入和球囊扩张等操作位置的准确性,从而有效降低手术风险^[114-115]。在商业化应用方面,以色列海法医院研制了心血管介入手术机器人(RNS)并首次开展临床实验,美国 Corindus Vascular Robotics 公司研制的 CorPath GRX 系统成为首个获得美国食品药品监督管理局(FDA)和欧盟认证(CE)批准的经皮冠状动脉介入医疗设备,法国 Robocath 公司开发的血管介入手术机器人 R-One™与微创合作,已完成国内的注册临床试验。另外,国内唯迈医疗、奥朋医疗、爱博医疗、睿心医疗、易度医疗等企业研发的血管介入手术机器人均已完成动物实验或进入注册临床阶段。图 15 展示了部分商业化血管介入导航系统。

5 未来发展方向分析与展望

综上所述,多模态图像引导的手术导航通过融合 X 光、内镜、超声、荧光等多模态影像的优势,综合运用多模态图像分割建模、手术方案决策、手术空间位姿标定与跟踪、多模影像配准融合以及光学/电磁手术器械跟踪定位等技术,在三维空间提供病灶的结构和功能信息,提升医生对血管、神经等重要组织结构的视觉感知和空间辨识力,大幅提升手术的安全性,缩短手术时间并提高手术效率。多模态图像引导手术导航在神经

外科、颅颌面矫正、骨科、穿刺及血管介入治疗等微创手术中已得到广泛应用,可辅助医生实现创伤小、疗效好、恢复快的精准诊疗,降低手术的并发症,成为当前国际精准诊疗的重大前沿热点。

多模态图像引导手术导航技术发展至今,其成熟程度以及应用稳定性高度依赖图像分割建模、术中定位系统、成像设备、配准精度、目标区域运动形变补偿、导航精度量化评估等方面,随着手术导航机器人应用场景的不断扩大和手术难度的增加,其仍面临诸多技术与应用挑战。主要包括:

1) 组织柔性形变

肝脏、肾脏等柔性器官受呼吸、心跳、手术器械碰撞、牵拉等影响发生形变位移,导致术中器官形状位置与术前影像存在差异,尽管经过空间位姿标定后术中与术前器官间刚性变换得到了补偿,但是形变位移产生的非刚性变换使得术前影像上规划的治疗方案、分割的组织模型无法与术中实际情况匹配,最终导致手术操作与术前规划偏差大。目前的手术导航主要通过运动跟踪方法,将静态三维术前影像与术中动态二维影像间的非刚性形变补偿,实现多模态影像融合导航。运动跟踪分为外部跟踪及内部跟踪方法,通过运动建模或非线性相关分析,从被跟踪的外部或内部多个标志点的呼吸运动规律推断出器官的整体运动趋势。但是由于只关注稀疏的标志点,而不是整个器官的形变位移,具有较少参数的模型无法准确完整地捕捉器官复杂运动,会因为模型中的微小错误导致目标位置的显著偏差。因此,通过术前三维影像动态重建技术,构建器官的个性化运动形变模型,依据术中实时影像预测不同状态下器官任意点的移动量,在精准量化的基础上实现非刚性形变补偿,是未来需要攻克的关键技术之一。

2) 动态三维可视化

手术导航效果的呈现,包括多模态影像间的动态融合显示、虚拟模型叠加影像的三维增强显示等。多源信息的融合、叠加本质是为了内容互补、扩展术野,但是术中场景复杂,如何将丰富的多源信息让医生直

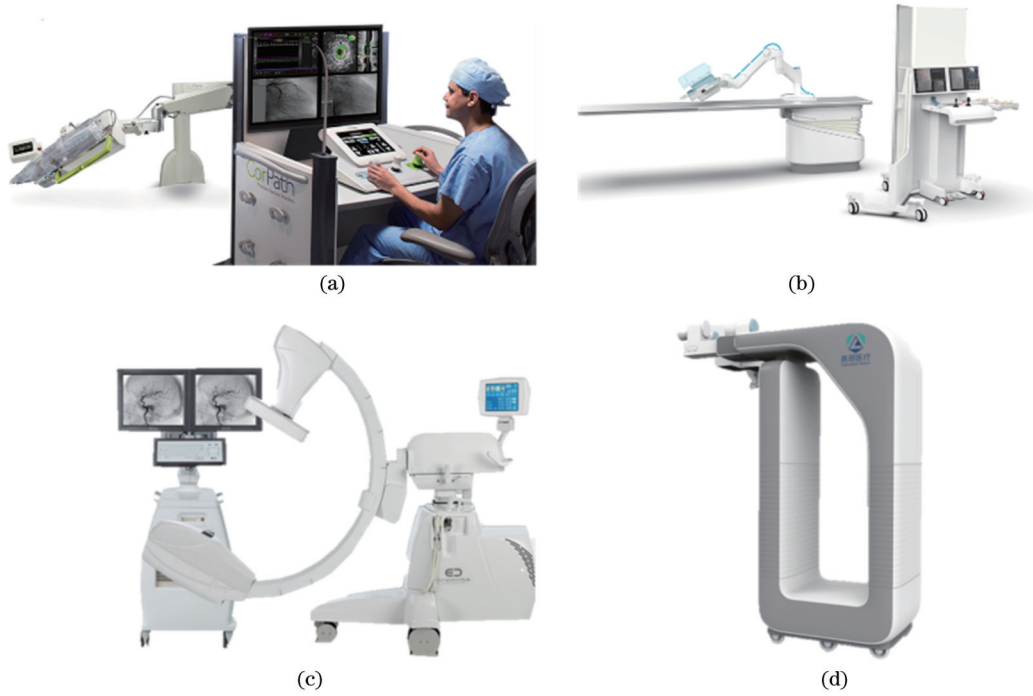


图 15 商用血管介入导航系统。(a)Corindus Vascular Robotics 公司 CorPath GRX 手术机器人系统;(b)Robocath 公司的 R-One™ 血管介入手术机器人系统;(c)唯迈医疗的 Alien 血管介入手术机器人系统;(d)奥朋医疗的 Allvas 血管介入手术机器人系统

Fig. 15 Commercial vascular intervention navigation systems. (a) CorPath GRX surgical robot system (Corindus Vascular Robotics©); (b) R-One™ vascular interventional surgical robot system (Robocath©); (c) Alien vascular interventional surgical robot system (WeMed©); (d) Allvas vascular interventional surgical robot system (Operaterobot©)

观可视、接受并快速理解,是未来研究的重点。现有的手术导航系统通过二维屏幕显示三维融合叠加的渲染结果,平面显示效果缺乏空间纵深感,医生无法直观立体地看到术中患者体内实时的治疗场景,导航过程中仍然需要医生自行构想病灶与周围组织结构及手术器械间的三维相对关系。术中操作的混合现实可视化引导可以为医生带来直观的立体视觉协助。借助头戴式混合现实眼镜等设备对患者体内手术全过程进行可视化,将多模态融合影像、组织模型及规划方案直接渲染到医生视野空间内患者身体的对应区域上,为医生提供精准立体可视化手术引导是未来需要攻克的关键技术之一。

3) 手术动态决策

现有手术导航系统存在术中动态决策能力不足的问题。具体表现为,术中实时数据更新有限,无法及时反映手术过程中病变组织结构和位置的变化,导致术前手术规划方案失效及手术实施依据丢失。在遇到意外情况或复杂病变时,医生可能需要根据术中实际情况做出规划调整,术中动态决策功能缺失将严重降低手术导航系统的临床适用性。为此,探索更加有效的术中实时成像和传感技术,并基于实时数据更新术前规划方案以反映解剖结构的动态变化,进而提高系统的动态决策能力将是手术导航系统未来需要攻克的关键技术之一。

4) 执行机构安全调控

手术导航机器人作为一种高精度、高效率的手术

辅助工具,其末端执行机构是执行手术规划方案进而完成精准手术的关键部件。然而,由于手术导航机器人需要在患者体内进行操作,因此对于末端执行机构的精度、可靠性和安全性要求非常高。首先,在手术导航机器人在进行手术辅助操作时,需要通过末端执行机构进行把持、切割、刨削等有创操作,如果末端执行机构出现故障或操作不当,可能会对患者造成极大的伤害。其次,末端执行机构的智能化和自适应性是未来发展的方向,但是在实现智能化和自适应性方面,还需要克服诸如算法不准确、数据不充分等问题。如何使末端执行机构具备更强的智能化和自适应性能力是未来需要攻克的关键技术之一。

因此,实现导航设备的高精度、智能化、轻量化、低成本是手术导航未来发展的主要方向。对于手术过程中软组织存在柔性形变影响导航精度的问题,进一步攻破配准技术、形变补偿、运动补偿、软组织张力感知的技术难点,开发适用于柔性环境的高精度非刚性配准算法、术中导航信息的动态三维可视化、手术方案的动态决策及执行结构的安全调控方案,实现术中导航动态感知及优化,是进一步提升当前手术导航系统实时性、便携性、精准性与实现智能化的关键。未来,随着多模态手术导航系统在临床中的应用场景不断扩大,手术导航技术的精准性与临床应用价值必然日趋凸显,将会为更多术者提供可靠的引导与辅助,为不同级别医疗机构提供同质化医疗服务提供有力保障,提

供规范化诊疗,对提高我国微创手术诊疗水平和普惠人民健康具有重要意义。

参 考 文 献

- [1] Su H, Kwok K W, Cleary K, et al. State of the art and future opportunities in MRI-guided robot-assisted surgery and interventions[J]. *Proceedings of the IEEE*, 2022, 110(7): 968-992.
- [2] Forner A, Reig M, Bruix J. Hepatocellular carcinoma[J]. *Lancet*, 2018, 391(10127): 1301-1314.
- [3] Wright J. Surgery: the eyes of the operation[J]. *Nature*, 2013, 502(7473): S88-S89.
- [4] Kang T W, Rhim H. Recent advances in tumor ablation for hepatocellular carcinoma[J]. *Liver Cancer*, 2015, 4(3): 176-187.
- [5] Zhu M L, Chen Z, Yuan Y X. DSI-net: deep synergistic interaction network for joint classification and segmentation with endoscope images[J]. *IEEE Transactions on Medical Imaging*, 2021, 40(12): 3315-3325.
- [6] Ji X, Feng M, Treb K, et al. Development of an integrated C-arm interventional imaging system with a strip photon counting detector and a flat panel detector[J]. *IEEE Transactions on Medical Imaging*, 2021, 40(12): 3674-3685.
- [7] Hu J, Luo Z W, Wang X, et al. End-to-end multimodal image registration via reinforcement learning[J]. *Medical Image Analysis*, 2021, 68: 101878.
- [8] Kochanski R B, Lombardi J M, Laratta J L, et al. Image-guided navigation and robotics in spine surgery[J]. *Neurosurgery*, 2019, 84(6): 1179-1189.
- [9] 陶陈凝, 郑臻荣. 面向手术导航的增强现实计算光谱成像系统[J]. *激光与光电子学进展*, 2022, 59(20): 2011014.
Tao C N, Zheng Z R. Augmented reality computational spectral imaging for surgical guidance[J]. *Laser & Optoelectronics Progress*, 2022, 59(20): 2011014.
- [10] Zhu J J, Li H, Ai D N, et al. Iterative closest graph matching for non-rigid 3D/2D coronary arteries registration[J]. *Computer Methods and Programs in Biomedicine*, 2021, 199: 105901.
- [11] Ambrosini P, Smal I, Ruijters D, et al. A hidden Markov model for 3D catheter tip tracking with 2D X-ray catheterization sequence and 3D rotational angiography[J]. *IEEE Transactions on Medical Imaging*, 2017, 36(3): 757-768.
- [12] Matl S, Brosig R, Baust M, et al. Vascular image registration techniques: a living review[J]. *Medical Image Analysis*, 2017, 35: 1-17.
- [13] Rodríguez E, Kypson A P, Moten S C, et al. Robotic mitral surgery at East Carolina University: a 6 year experience[J]. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2006, 2(3): 211-215.
- [14] Ishizawa T, Fukushima N, Shibahara J, et al. Real-time identification of liver cancers by using indocyanine green fluorescent imaging[J]. *Cancer*, 2009, 115(11): 2491-2504.
- [15] Miyata A, Ishizawa T, Tani K, et al. Reappraisal of a dye-staining technique for anatomic hepatectomy by the concomitant use of indocyanine green fluorescence imaging[J]. *Journal of the American College of Surgeons*, 2015, 221(2): e27-e36.
- [16] Chen D F, Ren J, Wang Y, et al. Intraoperative monitoring of blood perfusion in port wine stains by laser Doppler imaging during vascular targeted photodynamic therapy: a preliminary study[J]. *Photodiagnosis and Photodynamic Therapy*, 2016, 14: 142-151.
- [17] Wang Y, Xu Y X, Guo X H, et al. Enhanced antimicrobial activity through the combination of antimicrobial photodynamic therapy and low-frequency ultrasonic irradiation[J]. *Advanced Drug Delivery Reviews*, 2022, 183: 114168.
- [18] Herregodts S, Verhaeghe M, De Coninck B, et al. An improved method for assessing the technical accuracy of optical tracking systems for orthopaedic surgical navigation[J]. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2021, 17(4): e2285.
- [19] Burström G, Nachabe R, Homan R, et al. Frameless patient tracking with adhesive optical skin markers for augmented reality surgical navigation in spine surgery[J]. *Spine*, 2020, 45(22): 1598-1604.
- [20] Saeidi H, Ge J, Kam M, et al. Supervised autonomous electrosurgery via biocompatible near-infrared tissue tracking techniques[J]. *IEEE Transactions on Medical Robotics and Bionics*, 2019, 1(4): 228-236.
- [21] Sheinker A, Moldwin M B. Adaptive interference cancellation using a pair of magnetometers[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2016, 52(1): 307-318.
- [22] Marmulla R, Lüth T, Mühling J, et al. Markerless laser registration in image-guided oral and maxillofacial surgery[J]. *Journal of Oral and Maxillofacial Surgery*, 2004, 62(7): 845-851.
- [23] Ledderose G J, Stelter K, Leunig A, et al. Surface laser registration in ENT-surgery: accuracy in the paranasal sinuses-a cadaveric study[J]. *Rhinology*, 2007, 45(4): 281-285.
- [24] Marmulla R, Eggers G, Mühling J. Laser surface registration for lateral skull base surgery[J]. *Min-Minimally Invasive Neurosurgery*, 2005, 48(3): 181-185.
- [25] Malham G M, Munday N R. Comparison of novel machine vision spinal image guidance system with existing 3D fluoroscopy-based navigation system: a randomized prospective study[J]. *The Spine Journal*, 2022, 22(4): 561-569.
- [26] Li W J, Fan J F, Li S W, et al. Calibrating 3D scanner in the coordinate system of optical tracker for image-to-patient registration[J]. *Frontiers in Neurorobotics*, 2021, 15: 636772.
- [27] Chen L, Zhang F F, Zhan W, et al. Research on the accuracy of three-dimensional localization and navigation in robot-assisted spine surgery[J]. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2020, 16(2): e2071.
- [28] Swank M L, Alkire M, Conditt M, et al. Technology and cost-effectiveness in knee arthroplasty: computer navigation and robotics[J]. *American Journal of Orthopedics*, 2009, 38(2 Suppl): 32-36.
- [29] Zhou G, Chen X Q, Niu B L, et al. Intraoperative localization of small pulmonary nodules to assist surgical resection: a novel approach using a surgical navigation puncture robot system[J]. *Thoracic Cancer*, 2020, 11(1): 72-81.
- [30] Bonatti J, Vetrovec G, Riga C L, et al. Robotic technology in cardiovascular medicine[J]. *Nature Reviews Cardiology*, 2014, 11(5): 266-275.
- [31] Jeon S, Hoshier A K, Kim K, et al. A magnetically controlled soft microrobot steering a guidewire in a three-dimensional phantom vascular network[J]. *Soft Robotics*, 2019, 6(1): 54-68.
- [32] Kim Y, Parada G A, Liu S D, et al. Ferromagnetic soft continuum robots[J]. *Science Robotics*, 2019, 4(33): eaax7329.
- [33] Jeong S, Choi H, Go G, et al. Feasibility study on magnetically steerable guidewire device for percutaneous coronary intervention [J]. *International Journal of Control, Automation and Systems*, 2017, 15(1): 473-479.
- [34] Zhang S X, Yin M, Lai Z Y, et al. Design and characteristics of 3D magnetically steerable guidewire system for minimally invasive surgery[J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 4040-4046.
- [35] Kyriakides Y. Accuracy assessment of a novel optical image guided system for trans-nasal sinus and skull base surgeries[J]. *International Bulletin of Otorhinolaryngology*, 2020, 16(2): 41.
- [36] Trope M, Shamir R R, Joskowicz L, et al. The role of automatic computer-aided surgical trajectory planning in improving the expected safety of stereotactic neurosurgery[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2015, 10(7): 1127-1140.
- [37] Cong W J, Yang J A, Ai D N, et al. Quantitative analysis of

- deformable model based 3-D reconstruction of coronary artery from multiple angiograms[J]. IEEE Transactions on Biomedical Engineering, 2015, 62(8): 2079-2090.
- [38] Seitel A, Engel M, Sommer C M, et al. Computer-assisted trajectory planning for percutaneous needle insertions[J]. Medical Physics, 2011, 38(6): 3246-3259.
- [39] Fu T Y, Yang J, Li Q, et al. Groupwise registration with global-local graph shrinkage in atlas construction[J]. Medical Image Analysis, 2020, 64: 101711.
- [40] Song S, Yang J, Ai D N, et al. Patch-based adaptive background subtraction for vascular enhancement in X-ray cineangiograms[J]. IEEE Journal of Biomedical and Health Informatics, 2019, 23(6): 2563-2575.
- [41] Azzopardi G, Strisciuglio N, Vento M, et al. Trainable COSFIRE filters for vessel delineation with application to retinal images[J]. Medical Image Analysis, 2015, 19(1): 46-57.
- [42] Li Y L, Zhou S J, Wu J H, et al. A novel method of vessel segmentation for X-ray coronary angiography images[C]//2012 Fourth International Conference on Computational and Information Sciences, August 17-19, 2012, Chongqing, China. New York: IEEE Press, 2012: 468-471.
- [43] Hernandez-Vela A, Gatta C, Escalera S, et al. Accurate coronary centerline extraction, caliber estimation, and catheter detection in angiographies[J]. IEEE Transactions on Information Technology in Biomedicine, 2012, 16(6): 1332-1340.
- [44] Liu T, Tai Y H, Zhao C M, et al. Augmented reality in neurosurgical navigation: a survey[J]. The International Journal of Medical Robotics and Computer Assisted Surgery, 2020, 16(6): 1-20.
- [45] Minaee S, Minaei M, Abdolrashidi A. Deep-emotion: facial expression recognition using attentional convolutional network [J]. Sensors, 2021, 21(9): 3046.
- [46] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. arXiv preprint, 2014, arXiv:1412.7062.
- [47] Wang J C, Shu Y C, Lin C Y, et al. Application of deep learning algorithms in automatic sonographic localization and segmentation of the Median nerve: a systematic review and meta-analysis[J]. Artificial Intelligence in Medicine, 2023, 137: 102496.
- [48] Alqazzaz S, Sun X, Yang X, et al. Automated brain tumor segmentation on multi-modal MR image using SegNet[J]. Computer Vision and Image Understanding, 2019, 5: 209-219.
- [49] Christ P F, Elshaer M E A, Ettlinger F, et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields[M]//Ourselin S, Joskowicz L, Sabuncu M R, et al. Medical image computing and computer-assisted intervention-MICCAI 2016. Lecture notes in computer science. Cham: Springer, 2016, 9901: 415-423.
- [50] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [51] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [52] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [53] Isensee F, Jaeger P F, Kohl S A A, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature Methods, 2021, 18(2): 203-211.
- [54] Zhou Z W, Siddiquee M M R, Tajbakhsh N, et al. UNet++: a nested U-net architecture for medical image segmentation[M]//Stoyanov D, Taylor Z, Carneiro G, et al. Deep learning in medical image analysis and multimodal learning for clinical decision support. Lecture notes in computer science. Cham: Springer, 2018, 11045: 3-11.
- [55] Alami M, Alghamdi M. The role of generative adversarial network in medical image analysis: an In-depth survey[J]. ACM Computing Surveys, 2022, 55(5): 96.
- [56] Xue Y, Xu T, Zhang H, et al. SegAN: adversarial network with multi-scale L_1 loss for medical image segmentation[J]. Neuroinformatics, 2018, 16(3): 383-392.
- [57] Dai W, Dong N Q, Wang Z Y, et al. SCAN: structure correcting adversarial network for organ segmentation in chest X-rays[M]//Stoyanov D, Taylor Z, Carneiro G, et al. Deep learning in medical image analysis and multimodal learning for clinical decision support. Lecture notes in computer science. Cham: Springer, 2018, 11045: 263-273.
- [58] Dai Z Y, Yang R Q, Hang F, et al. Neurosurgical craniotomy localization using interactive 3D lesion mapping for image-guided neurosurgery[J]. IEEE Access, 2019, 7: 10606-10616.
- [59] Hu W R, Jiang H Y, Wang M. Flexible needle puncture path planning for liver tumors based on deep reinforcement learning [J]. Physics in Medicine & Biology, 2022, 67(19): 195008.
- [60] Luo M, Jiang H Y, Shi T Y. Multi-stage puncture path planning algorithm of ablation needles for percutaneous radiofrequency ablation of liver tumors[J]. Computers in Biology and Medicine, 2022, 145: 105506.
- [61] Shatrov J, Parker D. Computer and robotic-assisted total knee arthroplasty: a review of outcomes[J]. Journal of Experimental Orthopaedics, 2020, 7(1): 70.
- [62] Hafez M A, Chelule K L, Seedhom B B, et al. Computer-assisted total knee arthroplasty using patient-specific templating [J]. Clinical Orthopaedics & Related Research, 2006, 444: 184-192.
- [63] Xiao D Q, Wang L, Deng H, et al. Estimating reference bony shape model for personalized surgical reconstruction of posttraumatic facial defects[M]//Shen D G, Liu T M, Peters T M, et al. Medical image computing and computer assisted intervention-MICCAI 2019. Lecture notes in computer science. Cham: Springer, 2019, 11768: 327-335.
- [64] Xiao D Q, Deng H, Lian C F, et al. Unsupervised learning of reference bony shapes for orthognathic surgical planning with a surface deformation network[J]. Medical Physics, 2021, 48(12): 7735-7746.
- [65] Xiao D Q, Lian C F, Deng H, et al. Estimating reference bony shape models for orthognathic surgical planning using 3D point-cloud deep learning[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(8): 2958-2966.
- [66] Liu S Y, Fan J F, Song D P, et al. Joint estimation of depth and motion from a monocular endoscopy image sequence using a multi-loss rebalancing network[J]. Biomedical Optics Express, 2022, 13(5): 2707.
- [67] Yu T, Zheng S A, Zhang X W, et al. A novel computer navigation method for accurate percutaneous sacroiliac screw implantation[J]. Medicine, 2019, 98(7): e14548.
- [68] Gupta S, Gupta P, Verma V S. Study on anatomical and functional medical image registration methods[J]. Neurocomputing, 2021, 452: 534-548.
- [69] Fan J F, Yang J, Zhao Y T, et al. Convex hull aided registration method (CHARM) [J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 23(9): 2042-2055.
- [70] Sengupta D, Gupta P, Biswas A. A survey on mutual information based medical image registration algorithms[J]. Neurocomputing, 2022, 486: 174-188.
- [71] Chu Y K, Yang J, Ma S D, et al. Registration and fusion quantification of augmented reality based nasal endoscopic

- surgery[J]. *Medical Image Analysis*, 2017, 42: 241-256.
- [72] Giesel F L, Mehndiratta A, Locklin J, et al. Image fusion using CT, MRI and PET for treatment planning, navigation and follow up in percutaneous RFA[J]. *Experimental Oncology*, 2009, 31(2): 106-114.
- [73] Sundar H, Shen D G, Biros G, et al. Robust computation of mutual information using spatially adaptive meshes[M]//Ayache N, Ourselin S, Maeder A. *Medical image computing and computer-assisted intervention-MICCAI 2007. Lecture notes in computer science*. Heidelberg: Springer, 2007, 4791: 950-958.
- [74] Fu T Y, Fan J F, Liu D K, et al. Divergence-free fitting-based incompressible deformation quantification of liver[J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(3): 720-736.
- [75] Sentker T, Madesta F, Werner R. GDL-FIRE^{4D}: deep learning-based fast 4D CT image registration[M]//Frangi A F, Schnabel J A, Davatzikos C, et al. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11070: 765-773.
- [76] Uzunova H, Wilms M, Handels H, et al. Training CNNs for image registration from few samples with model-based data augmentation[M]//Descoteaux M, Maier-Hein L, Franz A, et al. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2017. Lecture notes in computer science*. Cham: Springer, 2017, 10433: 223-231.
- [77] Cheng X, Zhang L, Zheng Y F. Deep similarity learning for multimodal medical images[J]. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2018, 6(3): 248-252.
- [78] Du B, Liao J D, Turkbey B, et al. Multi-task learning for registering images with large deformation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(5): 1624-1633.
- [79] Fan J F, Cao X H, Yap P T, et al. BIRNet: brain image registration using dual-supervised fully convolutional networks[J]. *Medical Image Analysis*, 2019, 54: 193-206.
- [80] Blendowski M, Bouteldja N, Heinrich M P. Multimodal 3D medical image registration guided by shape encoder-decoder networks[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2020, 15(2): 269-276.
- [81] Sedghi A, Luo J, Mehrtash A, et al. Semi-supervised image registration using deep learning[J]. *Proceedings of SPIE*, 2019, 10951: 109511G.
- [82] Yoo I, Hildebrand D G C, Tobin W F, et al. ssEMnet: serial-section electron microscopy image registration using a spatial transformer network with learned features[M]//Cardoso M J, Arbel T, Carneiro G, et al. *Deep learning in medical image analysis, international workshop on multimodal learning for clinical decision support. Lecture notes in computer science*. Cham: Springer, 2017, 10553: 249-257.
- [83] 林立昊, 易见兵, 曹锋, 等. 多尺度并行全卷积神经网络的肺计算机断层扫描图像非刚性配准算法[J]. *激光与光电子学进展*, 2022, 59(16): 1617004.
- Lin L H, Yi J B, Cao F, et al. Non-rigid registration algorithm of lung computed tomography images based on multi-scale parallel full convolution neural network[J]. *Laser & Optoelectronics Progress*, 2022, 59(16): 1617004.
- [84] Zhu J J, Fan J F, Guo S A, et al. Heuristic tree searching for pose-independent 3D/2D rigid registration of vessel structures[J]. *Physics in Medicine & Biology*, 2020, 65(5): 055010.
- [85] Jeroen V H, Emmanuel A, Zheng G Y, et al. Deep learning-based 2D/3D registration of an atlas to biplanar X-ray images[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2022, 17(7): 1333-1342.
- [86] Miao S, Wang Z J, Liao R. A CNN regression approach for real-time 2D/3D registration[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1352-1363.
- [87] Vijayan R, Sheth N, Mekki L, et al. 3D - 2D image registration in the presence of soft-tissue deformation in image-guided transbronchial interventions[J]. *Physics in Medicine & Biology*, 2023, 68(1): 015010.
- [88] Varnavas A, Carrell T, Penney G. Fully automated 2D-3D registration and verification[J]. *Medical Image Analysis*, 2015, 26(1): 108-119.
- [89] Flach B, Brehm M, Sawall S, et al. Deformable 3D - 2D registration for CT and its application to low dose tomographic fluoroscopy[J]. *Physics in Medicine and Biology*, 2014, 59(24): 7865-7887.
- [90] 李文举, 孔德卿, 曹国刚, 等. 基于训练-推理解耦架构的 2D-3D 医学图像配准[J]. *激光与光电子学进展*, 2022, 59(16): 1610015.
- Li W J, Kong D Q, Cao G G, et al. 2D-3D medical image registration based on training-push understanding coupling architecture[J]. *Laser & Optoelectronics Progress*, 2022, 59(16): 1610015.
- [91] Wohlhart P, Lepetit V. Learning descriptors for object recognition and 3D pose estimation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 3109-3118.
- [92] Heinrich M P, Jenkinson M, Bhushan M, et al. MIND: modality independent neighbourhood descriptor for multi-modal deformable registration[J]. *Medical Image Analysis*, 2012, 16(7): 1423-1435.
- [93] Heinrich M P, Jenkinson M, Papież B W, et al. Towards realtime multimodal fusion for image-guided interventions using self-similarities[M]//International conference on medical image computing and computer-assisted intervention-MICCAI 2013. *Lecture notes in computer science*. Heidelberg: Springer, 2013, 8149: 187-194.
- [94] Wang Y F, Fu T Y, Wu C, et al. Multimodal registration of ultrasound and MR images using weighted self-similarity structure vector[J]. *Computers in Biology and Medicine*, 2023, 155: 106661.
- [95] Wei W, Haishan X, Alpers J, et al. A deep learning approach for 2D ultrasound and 3D CT/MR image registration in liver tumor ablation[J]. *Computer Methods and Programs in Biomedicine*, 2021, 206: 106117.
- [96] Markova V, Ronchetti M, Wein W, et al. Global multi-modal 2D/3D registration via local descriptors learning[M]//Wang L W, Dou Q, Fletcher P T, et al. *Medical image computing and computer-assisted intervention-MICCAI 2022. Lecture notes in computer science*. Cham: Springer, 2022, 13436: 269-279.
- [97] Dey D, Slomka P J, Gobbi D G, et al. Mixed reality merging of endoscopic images and 3-D surfaces[M]//Delp S L, DiGoia A M, Jaramaz B. *Medical image computing and computer-assisted intervention-MICCAI 2000. Lecture notes in computer science*. Heidelberg: Springer, 2000, 1935: 796-803.
- [98] Li W J, Fan J F, Li S W, et al. Homography-based robust pose compensation and fusion imaging for augmented reality based endoscopic navigation system[J]. *Computers in Biology and Medicine*, 2021, 138: 104864.[LinkOut]
- [99] Suciadi L P, William Y, Jorizal P, et al. Comparing lung CT in COVID-19 pneumonia and acute heart failure: an imaging conundrum[J]. *Cureus*, 2021, 13(5): e15120.
- [100] Behrens A, Guski M, Stehle T, et al. A non-linear multi-scale blending algorithm for fluorescence bladder images[J]. *Computer Science-Research and Development*, 2011, 26(1): 125-134.
- [101] Tan W, Tiwari P, Pandey H M, et al. Multimodal medical image fusion algorithm in the era of big data[J]. *Neural Computing and Applications*, 2020: 1-21.
- [102] Higgins W E, Helferty J P, Lu K K, et al. 3D CT-video fusion for image-guided bronchoscopy[J]. *Computerized Medical Imaging and Graphics*, 2008, 32(3): 159-173.
- [103] Chu Y K, Li X, Yang X L, et al. Perception enhancement using

- importance-driven hybrid rendering for augmented reality based endoscopic surgical navigation[J]. *Biomedical Optics Express*, 2018, 9(11): 5205-5226.
- [104] Tarsitano A, Ricotta F, Baldino G, et al. Navigation-guided resection of maxillary tumours: the accuracy of computer-assisted surgery in terms of control of resection margins—a feasibility study[J]. *Journal of Cranio-Maxillofacial Surgery*, 2017, 45(12): 2109-2114.
- [105] Bolding S L, Reebye U N. Accuracy of haptic robotic guidance of dental implant surgery for completely edentulous arches[J]. *The Journal of Prosthetic Dentistry*, 2022, 128(4): 639-647.
- [106] 白石柱, 任楠, 冯志宏, 等. 自主式口腔种植机器人手术系统动物体内种植精度的研究[J]. *中华口腔医学杂志*, 2021, 56(2): 170-174.
Bai S Z, Ren N, Feng Z H, et al. Animal experiment on the accuracy of the Autonomous Dental Implant Robotic System[J]. *Chinese Journal of Stomatology*, 2021, 56(2): 170-174.
- [107] Mathew K K, Marchand K B, Tarazi J M, et al. Computer-assisted navigation in total knee arthroplasty[J]. *Surgical Technology International*, 2020, 36: 323-330.
- [108] Jones C W, Jerabek S A. Current role of computer navigation in total knee arthroplasty[J]. *The Journal of Arthroplasty*, 2018, 33(7): 1989-1993.
- [109] Shimokawa N, Takami T. Surgical safety of cervical pedicle screw placement with computer navigation system[J]. *Neurosurgical Review*, 2017, 40(2): 251-258.
- [110] D'Souza M, Gendreau J, Feng A, et al. Robotic-assisted spine surgery: history, efficacy, cost, and future trends[J]. *Robotic Surgery (Auckland)*, 2019, 6: 9-23.
- [111] Balboni J M, Siddique K, Nomoto E K, et al. Novel use of robotics and navigation for anterior lumbar total disc replacement surgery[J]. *North American Spine Society Journal*, 2022, 9: 100097.
- [112] Hong J, Nakashima H, Konishi K, et al. Interventional navigation for abdominal therapy based on simultaneous use of MRI and ultrasound[J]. *Medical and Biological Engineering and Computing*, 2006, 44(12): 1127-1134.
- [113] Krücker J, Xu S, Glossop N, et al. Electromagnetic tracking for thermal ablation and biopsy guidance: clinical evaluation of spatial accuracy[J]. *Journal of Vascular and Interventional Radiology*, 2007, 18(9): 1141-1150.
- [114] Zhou Y J, Xie X L, Zhou X H, et al. Pyramid attention recurrent networks for real-time guidewire segmentation and tracking in intraoperative X-ray fluoroscopy[J]. *Computerized Medical Imaging and Graphics*, 2020, 83: 101734.
- [115] Ma H, Smal I, Daemen J, et al. Dynamic coronary roadmapping via catheter tip tracking in X-ray fluoroscopy with deep learning based Bayesian filtering[J]. *Medical Image Analysis*, 2020, 61: 101634.

Developments of Multimodal Image-Guided Surgical Navigation

Yang Jian^{1,2}, Wang Yuanyuan^{1,2}, Ai Danni^{1,2}, Song Hong³, Fan Jingfan^{1,2}, Fu Tianyu⁴,
Xiao Deqiang^{1,2}, Shao Long³, Gu Ying^{5**}, Wang Yongtian^{1,2*}

¹*School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China;*

²*Laboratory of Beijing Engineering Research Center of Mixed Reality and Advanced Display, Beijing 100081, China;*

³*School of Computer Science, Beijing Institute of Technology, Beijing 100081, China;*

⁴*School of Medical Technology, Beijing Institute of Technology, Beijing 100081, China;*

⁵*Department of Laser Medicine, The First Medical Center of Chinese PLA General Hospital, Beijing 100853, China*

Abstract

Significance Surgical navigation combines organ segmentation modeling, surgical planning, pose calibration and tracking, multimodal image registration, and fusion display technologies to enable surgeons to precisely locate lesions and surgical tools and to observe internal tissue through the tissue surface, which can significantly improve surgical safety and time efficiency. Conventional surgery usually uses two-dimensional (2D) unimodal images such as ultrasound, endoscopy, or X-ray images to guide the surgical process. However, unimodal images lack three-dimensional (3D) information and depend heavily on the surgeon's experience. In contrast, multimodal image-guided surgical navigation provides real-time instrument positions, as well as structural and functional information of the lesion in 3D space, helping the surgeons to effectively protect important tissue, vessels, and organs around the lesion, avoiding unnecessary damage, and reducing the probability of surgical complications, which has become an important tool for a variety of clinical surgical procedures.

Progress The surgical navigation system mainly consists of imaging devices, tracking and positioning core devices, end-effectors, surgical tools, and other hardware, and it combines modern imaging technology, stereotactic technology, computer technology, and artificial intelligence technology to enable patients to obtain safe, precise, and minimally invasive surgical treatment. The surgical navigation system involves core theories and methods in various aspects such as multimodal image segmentation and tissue modeling, surgical planning, pose calibration and tracking, multimodal image registration, and image fusion. The hardware components of multimodal image-guided surgical navigation mainly include intraoperative imaging devices, such as X-ray, ultrasound, and endoscopic imaging systems, tracking and positioning core

devices, such as optical and electromagnetic lasers, structured light positioning systems, and navigation actuation components, such as robotic arms and guidewire. The key technologies for multimodal image-guided surgical navigation include multimodal image segmentation and tissue modeling, surgical protocol decision making, surgical spatial calibration and tracking, multimodal image registration, and multi-source information fusion display. The segmentation and modeling technologies based on preoperative multimodal medical images can depict the spatial structure and position information of target tissue and organs, providing an important data base for preoperative surgical planning and intraoperative real-time guidance. The surgical plan decision is used to guide intraoperative surgical operations, and the surgical plan can be formulated by the relationship of the 3D model positions of tissue, organs, and lesions. Preoperative planning for different surgical procedures has a large variability and can be divided into two categories: surgical path planning and surgical scheme planning. The surgical navigation system is based on core tracking and positioning devices to track the real-time position of intraoperative surgical instruments and obtain the relative position relationship of the preoperative reconstructed model, intraoperative patients, and surgical instruments. Multimodal image registration aims to seek to coordinate transformation among multimodal medical images to make these images aligned and unified in the spatial coordinate system, which helps to obtain complementary tissue structure or functional information from different modalities. The fusion and display of multi-source information aim to integrate different images, tissue models, surgical protocols, tracking postures, and other information on the same coordinate system for 2D or 3D display, which overcomes the limitations of a single source in the information presentation and contributes to improving the precision of clinical diagnosis and treatment. At present, multimodal image-guided surgical navigation has become a powerful tool for precise treatment in clinical departments such as neurosurgery, craniomaxillofacial, orthopedics, percutaneous puncture, and vascular intervention, and it has important application prospects.

Conclusions and Prospects Multimodal image-guided surgical navigation provides structural and functional information of the lesion in 3D space by fusing multimodal images such as X-ray, endoscopy, ultrasound, and fluorescence, as well as integrating technologies such as multimodal image segmentation and tissue modeling, surgical planning and decision making, pose calibration and tracking, and multimodal image registration and fusion to improve the surgeon's visual perception and spatial recognition of important tissue such as blood vessels and nerves. This technique significantly improves the safety of surgery, shortens surgery time, and increases surgery efficiency. Multimodal image-guided surgical navigation has been widely used in minimally invasive surgeries such as neurosurgery, craniomaxillofacial orthopedics, orthopedics, puncture, and vascular interventions, which can assist surgeons to achieve precise treatment with less trauma, higher efficacy, and faster recovery and reduce complications of surgery, which is a major frontier hotspot in the international precision treatment. Achieving higher precision, higher intelligence, lighter weight, and lower cost of navigation devices are the main development directions of surgical navigation. How to further break the difficulties of image registration, deformation compensation, motion compensation, and soft tissue tension perception techniques, how to develop a high-precision non-rigid registration algorithm for flexible tissue deformation, and how to achieve dynamic visualization of intraoperative navigation information and decision making of surgical plans are essential to further improve the real-time, portability, accuracy, and intelligence of current surgical navigation systems. In the future, as the application scenarios of multimodal surgical navigation systems in the clinic continue to expand, the clinical application value of surgical navigation technology will become prominent, which will provide reliable guidance and assistance to more surgeons, and it is of great significance in improving the level of minimally invasive surgical treatment in China.

Key words surgical navigation; surgical robot; multimodal medical image; medical image processing; location and tracking; registration and fusion