



## 基于计算机视觉的人体内腔三维重建技术综述

吴海滨<sup>1</sup>,徐若彤<sup>1</sup>,王爱丽<sup>1</sup>,于晓洋<sup>1</sup>,岩堀 祐之<sup>2</sup>,赵蓝飞<sup>1</sup>,刘 赫<sup>1</sup>

(1.哈尔滨理工大学 黑龙江省激光光谱技术及应用重点实验室,哈尔滨 150080;

2.中部大学 计算机科学学院,日本 爱知 487-8501)

**摘 要:** 医生通过内窥镜观察的人体内腔显示为二维图像,不能立体地展现内腔环境中病灶、血管及邻近组织的关系,而内腔三维重建及可视化技术能够清晰、全面地展现病灶及其他组织的三维形态,更好地辅助医生进行精准的手术判断。将人体内腔环境中的三维重建技术分为主动式测量方法与被动式测量方法,分类综述基于结构光、飞行时间、双目立体视觉、单目视觉的内腔三维重建技术及发展现状。针对同时定位与地图构建的内腔三维重建法,分析对比内腔环境下的特征点检测与匹配的发展、方法及特点,并对人体内腔三维重建的难点和未来发展趋势进行展望。

**关键词:** 内腔三维重建;结构光;飞行时间;双目立体视觉;单目视觉;特征点检测;特征点匹配

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 吴海滨,徐若彤,王爱丽,等.基于计算机视觉的人体内腔三维重建技术综述[J].计算机工程,2021,47(10):1-15.

**英文引用格式:** WU H B, XU R T, WANG A L, et al. Overview of 3D human lumen reconstruction technology based on computer vision[J]. Computer Engineering, 2021, 47(10): 1-15.

## Overview of 3D Human Lumen Reconstruction Technology Based on Computer Vision

WU Haibin<sup>1</sup>, XU Ruotong<sup>1</sup>, WANG Aili<sup>1</sup>, YU Xiaoyang<sup>1</sup>, IWAHORI Yuji<sup>2</sup>, ZHAO Lanfei<sup>1</sup>, LIU He<sup>1</sup>

(1.Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application, Harbin University of Science and Technology, Harbin 150080, China; 2.Department of Computer Science, Chubu University, Aichi 487-8501, Japan)

**[Abstract]** The human lumen images observed by the doctor through the endoscope is shown in a two-dimensional style, and fail to present the three-dimensional relationships between the lesions, blood vessels and adjacent tissues in the human lumen. The three-dimensional human lumen reconstruction technology and visualization methods can clearly and comprehensively display the three-dimensional images of lesions and tissues, assisting doctors in precise operations. In this paper, the three-dimensional reconstruction technologies for human lumen are divided into the active measurement methods and passive measurement methods. On this basis, this paper summarizes the reconstruction technologies based on structured light, Time Of Flight (TOF), binocular stereo vision, and monocular vision in order, and discusses their development. Then focusing on the reconstruction technology using Simultaneous Localization and Mapping (SLAM), the paper analyzes and compares the development, methods and characteristics of the feature point detection and matching in human lumen. Finally, the difficulties and future development trends of the three-dimensional human lumen reconstruction technology are prospected.

**[Key words]** 3D lumen reconstruction; structured light; Time Of Flight (TOF); binocular stereo vision; monocular vision; feature point detection; feature point matching

**DOI:** 10.19678/j.issn.1000-3428.0061710

### 0 概述

基于视觉的三维重建技术是计算机视觉的一个重

要领域,该技术通过使用相关仪器来获取场景物体的数据图像,并对获取的数据信息进行分析处理,将真实场景刻画成符合计算机逻辑表达的数学模型,重建出

**基金项目:** 国家自然科学基金(61671190, 61801149); 日本学术振兴会科学基金(20K11873); 黑龙江省普通高校专项基金(LGYC2018JC050)。

**作者简介:** 吴海滨(1977—),男,教授、博士生导师,主研方向为机器视觉;徐若彤,硕士研究生;王爱丽(通信作者),副教授、博士;于晓洋,岩堀 祐之,教授、博士生导师;赵蓝飞,讲师、博士;刘 赫,副教授、博士。

**收稿日期:** 2021-05-20 **修回日期:** 2021-07-02 **E-mail:** aili925@hrbust.edu.cn

真实环境中物体的三维信息。三维重建技术目前主要应用于人工智能、自动驾驶、虚拟现实、安防监控、运动目标监测、行为分析和重点人群监护等领域,同时因其具有速度快、显示清晰、实时性好等优点,可以对如文物保护、游戏开发、建筑设计、临床医学等研究起到辅助的作用。

微创手术是利用现代医疗器械设备穿过人体表面的微小创口在人体内施行多人间接眼-手协作的手术新技术<sup>[1]</sup>。与传统外科手术或早期的微创手术相比,现代微创手术具有操作精准、出血量较少、患者术后恢复的速度较快等优势,日益受到患者的欢迎,并广泛应用于在内腔手术中。但是,外科医生通过内窥镜视频流的2D可视显示找到目标并执行复杂的手术,偶尔容易迷失方向、手眼失调,难以通过将腹腔镜视野与术前图像进行经验匹配来确定病灶部位,易引起术中失误。

近年来,微创手术逐渐与计算机技术相融合,如外科医生利用手术经验与图像处理技术相结合,使用内窥镜系统对病灶区域进行立体定位,打破了传统手术的局限性<sup>[2]</sup>。面对微创手术特殊复杂的环境以及临床微创手术中实时性的要求,同时定位与地图构建(Simultaneous Localization and Mapping, SLAM)技术逐渐应用于内窥镜视觉领域<sup>[3-5]</sup>,由数字化成像设备采集到的人体内腔信息通过计算机建立三维内腔模型,并结合医生视角、内腔三维模型、患者体表三维形貌生成内腔图像,能够更准确、真实地反映内腔结构及空间关系。在不影响医生及手术环境的前提下,满足普通微创手术中医生对内腔辅助观察的需求。该技术在一定程度上能够降低心跳、呼吸、手术因素对非刚性内腔表面的振动影响,进而提高特征信息选择和配准的准确度。再进一步,医务人员可以利用内腔先验知识,从生物力学角度建立内腔器官的运动与形变模型,减少术中操作失误,提高手术的准确性与科学性。

本文根据近年来的国内外研究现状,对内腔三维重建技术中的常用方法进行分类,重点研究内腔环境中基于计算机视觉领域的三维成像方法的优缺点和实际应用中的局限性等问题,分析不同方法的重建效果与应用场景,探究未来内腔环境中的三维重建技术的发展方向。最后,对单目视觉SLAM技术进行重点阐述。

## 1 内腔三维重建

三维重建技术是对三维物体或者场景图像描述的逆过程,以及利用数学公式根据二维影像恢复出物体表面三维信息的一种计算机技术。基于视觉的三维重建技术是利用相关设备获取单视图或者多视图图像的图像信息,结合图像处理技术对其进行处理、操作和分析,最后综合视觉计算等技术获取物体的三维信息<sup>[6]</sup>。该技术由于具有成本低廉、真实感强、全自动或半自动建模速度较快等优势,因此成为一个极具潜力的热门领域,在诸多方面有着很重要的应用。三维重建技术

在医学影像领域中,主要通过三维测量和实时重建将内腔环境呈现给医护人员,辅助医护人员术中判断,目前主要应用在虚拟手术、术前规划、牙科治疗、整形及假肢外科、病灶诊断等医学方面。

内腔三维测量技术可分为非接触式测量与接触式测量。接触式测量的典型方法为探针法,即使用探针接触物体表面进行重建,但会对所接触的物体表面造成一定的破坏或带来组织变形等危险情况,无法应用到医学微创手术中。鉴于人体内腔三维建模技术需要针对狭小、特征缺乏的内腔空间快速三维重建,有较多限制,目前主要应用的是非接触式测量,可分为主动式测量与被动式测量两大类,如图1所示。

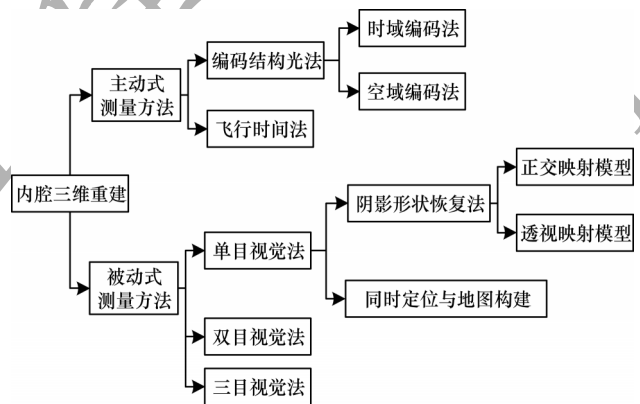


图1 内腔三维重建技术分类

Fig.1 Classification of 3D lumen reconstruction techniques

20世纪以来,国内外研究者逐步根据光学三维测量技术的硬件配置简单、测量精度高、点密度高、速度快、成本低等优点,设计了主动测量方法对人体内腔进行三维重建。主动式测量方法是将光源投射到内腔器官表面,然后通过内窥镜采集器官表面图像并获得扫描数据,根据接收返回的光波来获取物体的三维信息,从而重建内腔三维模型,主要包括结构光法、飞行时间(Time Of Flight, TOF)法等。

被动式测量可以根据系统使用的摄像头数量分为单目、双目、三目等。在待测物体周围的光照条件下,单目系统只利用一个摄像头获取待测物体的表面信息,以特征点为匹配基元进行两帧图像间的匹配,为三维重建的过程做铺垫。由于受人体内腔复杂环境的限制,基于单目视觉法的三维重建技术成为主流研究方法。

## 2 主动式测量方法

主动式三维测量方法主要包括编码结构光法和飞行时间法。前者是将特定的光线进行投影,摄像机在待测物体表面探测到相同形状光线的三维图像,进而通过光线的二维畸变信息得出人体内腔表面的三维数据;后者通过待测量物体在固定介质中飞行一定距离所消耗的时间,从而进一步获得人体内腔表面信息。

## 2.1 基于编码结构光法的内腔三维重建

基于编码结构光法的三维重建运用光学三角法的原理是将一幅或多幅编码图案投影到人体内腔中,由内窥镜采集内腔中投影得到的编码图像,结合三角法的数学原理将编码图像与编码图案对应点联系起来,得到人体内腔的三维信息,如图2所示。大量学者聚焦编码方法以提高内腔三维重建的测量精度、速度和准确度。

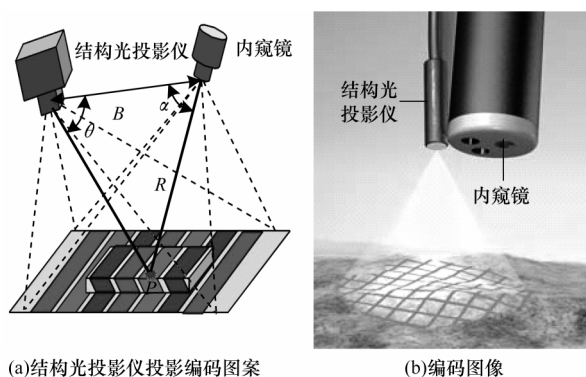


图2 结构光原理示意图

Fig.2 Schematic diagram of structured light principle

根据所使用编码图案类型和解码算法的不同,可以将编码结构光法分为时域编码法和空域编码法两类。时域编码法包括时间二进制码、时间 $n$ 元码、格雷码、时间相移法、格雷相移法等;空间编码法有De Bruijn编码、非正规编码、 $M$ 阵列编码等,这些方法的核心思想是保证局部编码在全局图像中的唯一性。

### 1) 时域编码法

时域编码法将多个不同的编码图案按照时间顺序投射到人体器官表面,同时将得到的编码图像序列组合起来进行解码。1994年,HANEISHI等<sup>[7]</sup>提出使用带有激光束投影的内窥镜来测量胃肠道表面的轮廓,通过计算激光头、摄像头和胃肠道内的反射点之间三角关系进而得到胃肠道的深度信息,从而实现胃肠道表面的轮廓重建<sup>[8]</sup>。该方法验证了结构光法对内腔进行三维重建的可行性,并对测量精度进行了评价,在投影没有失真的前提下,肠道表面测量误差小于0.16 mm。但是被测量物体在连续移动的过程中,光学三角法的测量速度比激光束的扫描速度慢,因而无法应用于临床医学中。

为提高空间分辨率的精度,在时域编码法的基础上引入了相位编码方式中的时间相移法,通常以条纹形式投影到内腔表面,经由器官调制后变形,最后通过相机采集获取内腔表面相位信息。1998年,ARMBRUSTER等<sup>[9]</sup>结合时间相移法,设计了对人体内腔进行三维重建的内窥镜系统。但是由于当时医疗内窥镜体积较大,因此无法在人体内腔里灵活移动。时间相移法中相位的求解在理论上不受背景光

强、内腔表面的反射率和纹理色彩等信息的干扰,因此相比传统的时域编码法,具有更高的空间分辨率和相位测量精度。在手术这种动态场景中,由于其多帧测量特性,时间相移法3D重建的精度容易受到投影仪、相机非线性和内腔环境变化的影响,因此该方法在内腔三维重建手术中应用有限。

### 2) 空域编码法

基于测量实时性和精度的需要,空域编码法是将向人体器官仅投射一幅特定的编码图案,再将编码图案与编码方式对照进行解码,从而重建出人体内腔的三维信息。2006年,HAYASHIBE等<sup>[10]</sup>利用激光扫描内窥镜系统实现了腹腔镜中手术区域的实时3D可视化,较好地重建了肝脏三维表面形貌,图3给出了条纹编码法的实现框图。但是仅通过器官表面形状信息估计其变形后的内部结构会产生较大误差,导致重建结果不准确,故需要结合其他多模态传感器的信息进行三维重建。

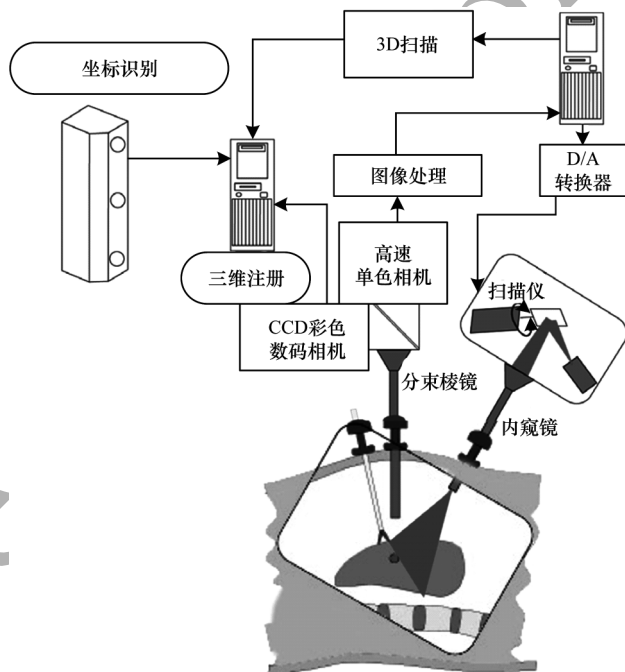


图3 条纹编码法

Fig.3 Fringe coding method

空域编码法是基于散斑投影的三维重建技术,散斑图案的不同会影响投射散斑的稳定性与三维重建的点云密度及空间分辨率。非正规编码是最古老的一种空域编码方法,该方法以节省时间为目的,使用暴力搜索各区域的编码信息从而生成编码图案<sup>[11]</sup>,因此该类方法生成的编码图案无法保证唯一性,且缺乏鲁棒性。2009年,斯特拉斯堡大学的ALBITARL等<sup>[12-13]</sup>基于单色 $M$ 阵列编码图案的建模方法,提出腹腔重建的构想。该方法采用几何图形代替传统的颜色信息进行编码,选择空心圆和实心圆简化图像分割的过程,同时利用条纹携带的方向

信息来确定观察的方向。2012年,其结合带有激光源的投影仪设备,又提出一种基于完备子映射(Perfect Submap, PSM)和亚完备子映射(Subperfect Submap, SPSM)的方法,加快了解码过程并实现了腹腔的三维重建。

2011年,CLANCY<sup>[14]</sup>等提出一种基于光谱识别的编码方法对肾脏进行表面重建。该方法利用超连续谱激光器在肾脏表面投影127个随机颜色的斑点,再根据这些投影点的波长对其进行分割和三维标定,重建出肾脏的表面轮廓。但是在标定时如果有投影点被遮挡,会导致肾脏质心的计算有误,最终重建结果出现偏差。

2014年,REITER等<sup>[15]</sup>使用德布鲁因序列编码法对腹腔进行三维重建,如图4所示。该方法通过LED蓝光投影仪对腹腔进行投影,蓝光经过二向色分束器传送至图案相机中进行解码,最后通过成像相机进行腹腔三维重建。该方法优点是与传统的白光光照条件相比,在只接收蓝光的图案相机内进行解码,实现过程变得更加简单。

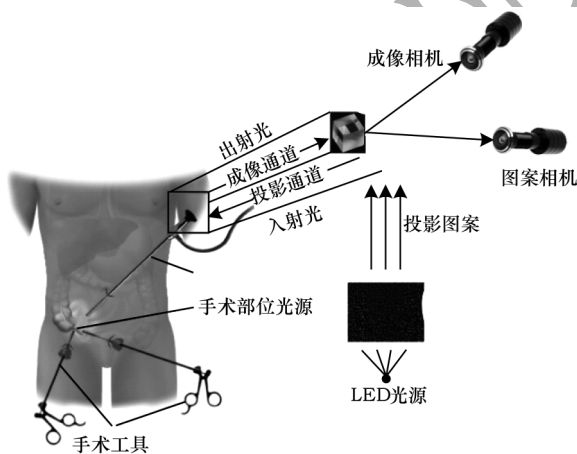


图4 德布鲁因序列法

Fig.4 DeBruine sequence method

2018年,LIU<sup>[16]</sup>等针对湿润、均匀、无纹理的口腔环境,结合激光编码投影帮助外科医生获取更多的特征点来实现口腔地图重建。但是受到口腔视场狭窄、图像配准错误等限制,构建地图具有挑战性,同时呼吸或舌动与手术工具相互作用引起的非刚性问题也是目前研究的难点之一。

结构光投影仪是结构光内腔三维重建技术的关键设备,在非接触测量的特性上又增加了全视场、高精度等优点。在数字投影设备出现之前,特殊的编码图案是通过在内窥镜镜头上覆盖透明掩膜经由光源照射实现的,导致编码图案种类单一。随着数字投影技术的发展,HAMADOU等<sup>[17]</sup>研发出一种新型的内窥镜,具有激光点阵投影、图像采集和白光照明3个功能,并且能在膀胱三维重建过程中实现小位移(每秒几毫米),采集速率达到25 frame/s,因此重建

点的密度很高,实现了膀胱表面的重建。不足之处是膀胱表面特征提取的难度较高,当投影仪投射的激光点不足8个时,膀胱纹理容易失真。2019年,YANG等<sup>[18]</sup>研发出一种柔性内镜导管投影仪对胃肠道进行投影,辅助外科医生计算组织器官的深度信息,从而完成胃肠道三维重建。这种投影仪具有白光照明模式和激光投影模式,可替代传统的照明系统。其直径为3.2 mm,可以灵活地在胃肠道内移动,摆脱了传统内窥镜导管直径过长的困扰。

结构光法大多基于散斑相关算法将编码图像与编码图案进行处理以获取场景的3D深度信息,因此需要大量的运算。为缩短运算的时间,提高内腔三维重建的效率,2020年,FURUKAWA等<sup>[19]</sup>提出了基于图卷积网络的自动校准方法,成功地从200帧图像序列中重建出猪胃3D表面。该方法能够快速地将编码图像与编码图案对应点进行匹配与自动校准,鲁棒性较强。

综上所述,结构光法因其具有硬件配置简单、测量精度高、点密度高、速度快、成本低等优点成为最具潜力的非接触式三维重建技术之一。并且,可同时输出内腔表面纹理图的优点,适合在光照条件不足或缺乏纹理的内腔场景下使用。

## 2.2 基于飞行时间法的内腔三维重建

飞行时间(TOF)法通过连续发射光脉冲到内腔器官表面,然后接收由内腔表面反射回去的光脉冲,根据其飞行时间来计算深度信息,如图5所示<sup>[20]</sup>。

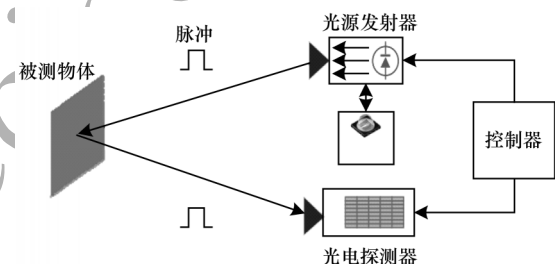


图5 TOF原理示意图

Fig.5 Schematic diagram of TOF principle

飞行时间法是采用相机测量由激光或LED提供的人造光信号的往返时间,来解析相机和拍摄对象之间距离的技术<sup>[21]</sup>。TOF相机通过发散面光源,在一定距离内光信息不会出现大量的衰减且检测速度较快,从而应用在三维重建中。2002年,TATAR等<sup>[22]</sup>利用TOF法中的距离和飞行时间的线性相关性,提出一种新型定位系统,可以为外科医生提供手术器械在病人体内的准确位置。但是TOF相机中存在半导体材料,测量精度易受温度变化的影响。TOF法通过图像传感器直接获取内腔的稠密深度信息与强度信息,在算法层面较为简单,其测量的数据精度与飞行时间的测量精度直接相关。2009年,PENNE等<sup>[23-24]</sup>针对病人三维呼吸

密度信息,提出使用TOF传感器重建病人胸腹的3D表面模型方案,如图6所示。该方案通过估计患者的三

维呼吸运动,结合非接触呼吸信号算法对患者胸腹进行拟合得到其三维模型。

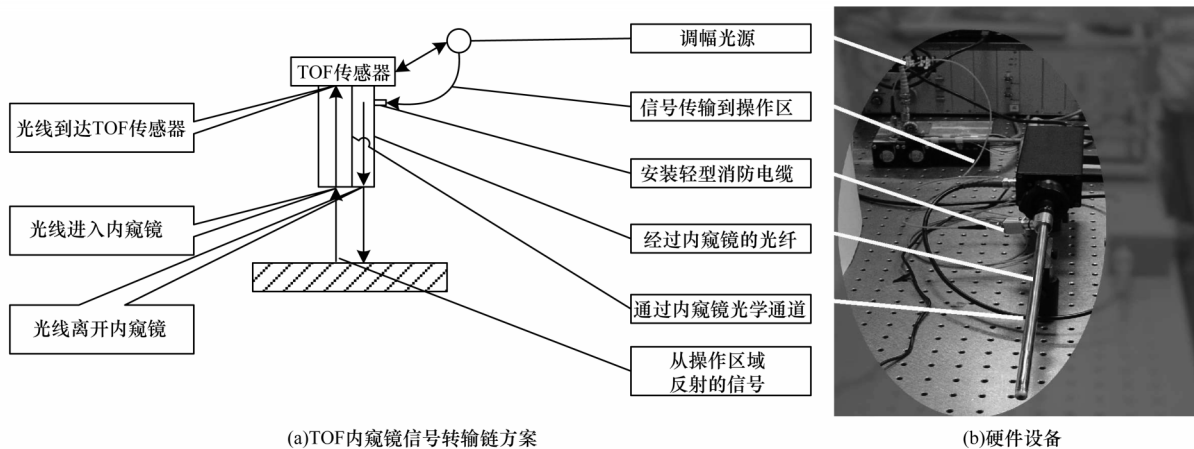


图6 TOF内窥镜信号传输链方案及硬件设置

Fig.6 Scheme of TOF endoscope signal transmission chain and hardware setup

TOF相机只能获取灰度和深度信息,为了使其获取更高品质的3D成像效果,2013年,KILGUS等<sup>[25]</sup>提出一种结合RGB颜色信息的飞行时间法对人体内腔进行三维重建。该方法根据颜色信息或光脉冲的飞行(往返)时间得到准确的距离信息,进而选取其作为重建指标。同时,TOF相机结构小巧便

于灵活移动,且功耗比较低,并且可以实时获取内腔的三维数据信息。2014年—2018年,ZVOLSKY等<sup>[26]</sup>、ZORRAQUINO等<sup>[27]</sup>和BRANDT等<sup>[28]</sup>研发一种用于胰腺检查的多模式三维重建技术,如图7所示,在正电子发射断层扫描的基础上结合飞行时间法实现了人体的胰腺重建。

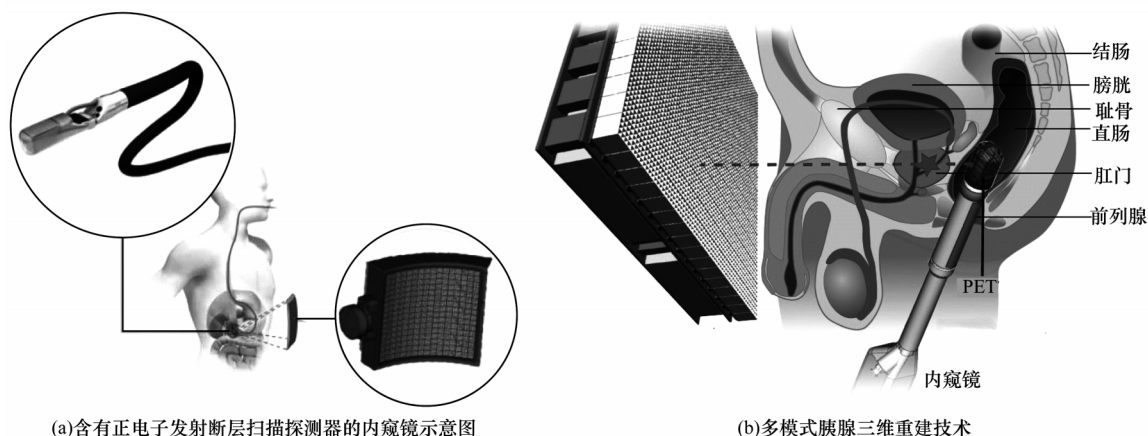


图7 胰腺重建示意图

Fig.7 Schematic diagram of pancreas reconstruction

虽然TOF相机操作简单,处理速度快,在远距离测量环境中具有较高的精度,但由于硬件条件的制约,目前市场上TOF相机易受高低温、真空等内腔环境因素的影响,三维重建的稳定性较低。同时,TOF相机读取图像不清晰会导致还原内腔深度图像精度较差。由于TOF相机应用成本较高,因此在内腔环境下利用飞行时间法进行三维重建的应用较少。

### 3 被动式测量方法

被动测量方法是在外界光源对被测物体的照射下,通过内窥镜等图像采集设备获取图像,再将所得图像进行立体匹配的三维建模方式。人体内腔环境

下的测量方法可以分为单目视觉法、双目立体视觉法、三目立体视觉法。

#### 3.1 基于双目立体视觉法的内腔三维重建

双目立体视觉系统模仿人类双眼视觉的原理,通过放置2台相机从不同的视角对目标物体进行观测,捕获同一个位置下的左右两侧图像,运用三角测量原理获得图像特征的深度等信息进行三维模型的建立,如图8所示。与其他三维重建技术相比,双目立体视觉法不需要过于复杂的硬件条件,通过相对便捷的操作方法即可实现三维重建,在医学图像处理、机器人视觉、物理科学和物体参数提取等领域中占有重要的地位,是目前三维重建技术中的热点之一。

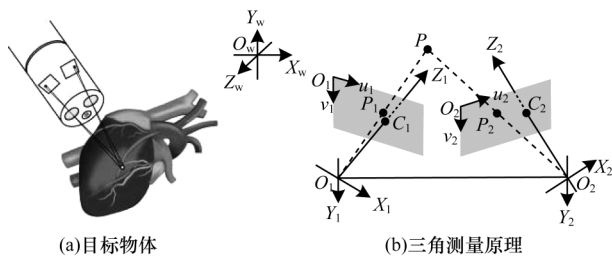


图8 双目立体视觉法原理示意图

Fig.8 Schematic diagram of binocular stereo vision method

双目视觉法利用2台摄像机观测同一内腔表面位置,在观测图像对应点的位置偏差的基础上,利用三角测量等数学原理估测内腔表面的深度信息。运用双目视觉法进行内腔三维重建主要包括内腔图像获取、内窥镜系统的摄像机标定、内腔图像的特征提取与匹配、立体匹配、内腔表面三维建模等过程。

在实现双目视觉法的整个过程中,立体匹配算法起到了关键作用,在内窥镜的应用场景中,通过在2个或多个观察点之间寻找对应关系,进而获得不同

点之间的视差,实现内腔三维信息投影,得到内腔三维模型。2004年,LAU等<sup>[29]</sup>针对传统局部匹配算法费时且精度不高的不足,开发一种表面跟踪算法,通过调整局部曲面结构的数值信息,从而直接推断出心脏的三维结构,该算法具有强稳定性,且计算效率较高。同年,STOYANOV等<sup>[30-31]</sup>针对软组织高光变形部分提出一种3D深度恢复方法,将图像校正与约束视差配准相结合,得到了腹腔的深度信息,同时保证了立体腹腔镜运动稳定性。特征匹配算法具有抗干扰性强、计算量小、速度快的优点。2010年—2013年,STOYANOV等<sup>[32-34]</sup>基于特征匹配通过结构传播对心脏表面稀疏重建。该方法首先组建特征点对的集合,然后获得点对之间的视差数据,最后估计每一帧的立体视差和连续帧之间的运动信息对软组织表面镜面反射和手术器具遮挡的影响,具有较好的鲁棒性。但是该方法得到稀疏的视差场,需要进一步通过复杂的插值计算才能得到较为稠密的心脏表面重建场景。

基于特征匹配的双目立体视觉法如图9所示。

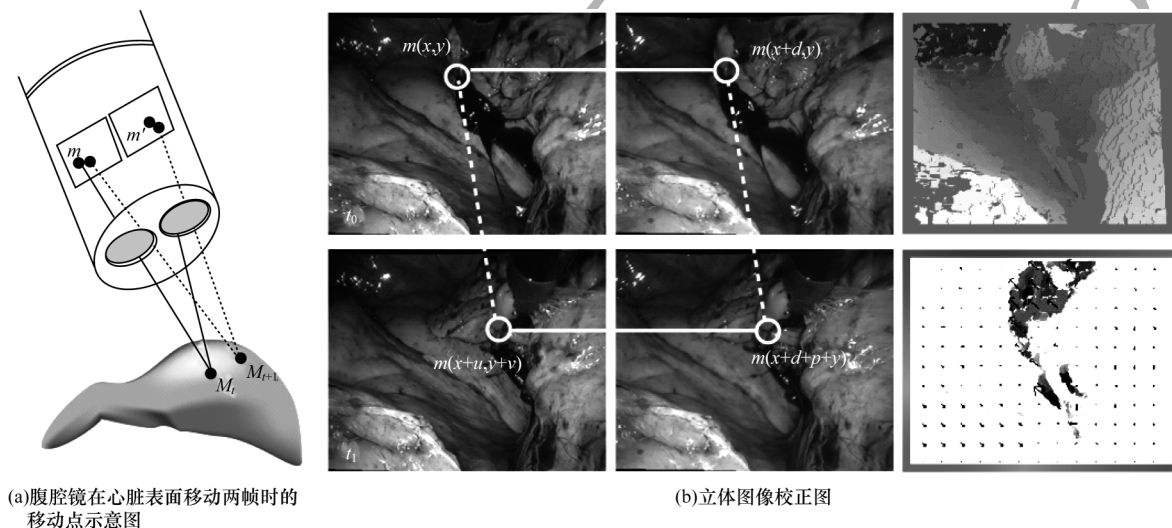


图9 基于特征匹配的双目立体视觉法

Fig.9 Binocular stereo vision method based on feature matching

2014年,TOTZ等<sup>[35]</sup>在稀疏场景重建的基础上采用迭代金字塔法进行序贯局部匹配传播,进而提出准稠密立体匹配的重建方法。与其他顺序处理图像金字塔方法相反,该方法从金字塔顶端开始垂直遍历至底端,并左右进行特征匹配以提高图像分辨率。2016年,PENZA等<sup>[36]</sup>提出基于滑窗和Census变换特征的软组织稠密三维重建方法,利用超像素方法对视差图进行细化。该方法能够有效应用于腹腔微创手术增强可视化,同时简化了术前模型与术中部位之间的配准过程。

基于特征匹配算法由于缺少致密的特征点,无法重建出稠密的内腔场景。基于特征区域的匹配算法通过在局部区域内提取目标进而获得致密的深度场景,但是在匹配过程中目标搜索策略速度低、耗时较长。2018年,BILEL等<sup>[37]</sup>提出一种基于小波分析的双目视觉法,将小波分析应用到了内腔图像分析中,在立体匹配环节大幅减小了计算量,提高了图像

的传播速度,改善了图像的分辨率。该方法通过小波分析处理左右视图的小波子带,根据双目视觉法的原理对立体内窥镜图像的局部特征进行增强处理,为内腔重建出更致密的三维图像打下了基础。但是该方法会增加镜面反射分量,同时在立体内窥镜图像受到较大遮挡的情况下,可能会影响重建效率。

在图像特性不连续的区域中常出现图像周围特性有阶跃变化的边缘特征,特征匹配解决了抗噪与非连续区域处的误匹配问题。2018年,WANG等<sup>[38]</sup>提出一种基于高斯加权变换的特征匹配算法实现了内腔三维重建。该算法针对边缘和不连续区域,利用图像熵和区域增长提取了遮挡导致的失配像素,与经典的立体匹配算法相比,立体匹配的平均相对误差缩小至8.48%,长度测量的平均相对误差减少至3.22%,可以有效地利用内窥镜对人体内腔进行测量

与重建。同时,其描述子具有良好的旋转不变性,且对光变化和噪声的鲁棒性较好。2019年,王博闻等<sup>[39]</sup>提出基于半全局块匹配的双目视觉法,实现了人体内腔稠密重建。但是该方法从立体内窥镜视频中提取出深度信息并进行地图稠密重建的计算时间

较长,不能满足手术指导的要求。2020年,WANG等<sup>[40]</sup>提出一种基于深度卷积网络的双目视觉算法对胃肠道进行重建。该算法利用尺度不变损失函数提高深度映射信息的精度和重建速度。图10为尺度不变损失函数的模型。

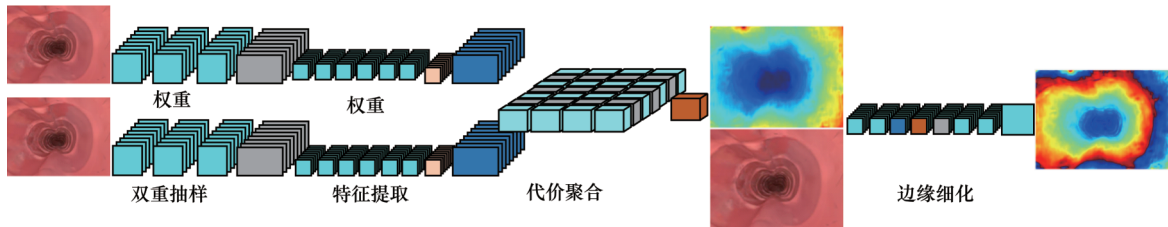


图10 尺度不变损失函数模型

Fig.10 The scale-invariant error function model

双目立体视觉技术具有抗干扰性能强、对环境要求不高、深度信息计算简单的优点,从而在人体内腔三维重建中得到应用。但由于双目立体视觉技术需要2个摄像头,并且对内窥镜直径的大小有较高要求,同时在内腔环境中可能出现遮挡器官或者视角受限等问题。视觉匹配和视差求取一直是双目立体视觉的研究热点,在匹配的过程中得到的特征点较为稀疏且特征点匹配耗时长。所以,目前主要是在工业内窥镜中得到广泛应用,而在医用手术内窥镜中的应用有限。

### 3.2 基于三目视觉法的内腔三维重建

在大视场测量环境中,单目立体视觉系统与双目立体视觉系统由于检测仪器视场角有限、获取信息单一等原因不适用于大尺寸物体的标定。基于上述原因,三目立体视觉技术逐渐发展起来。三目视觉法相比于双目视觉法增加1台摄像机,以中间相机为基础获得更可靠的深度信息。同时,单个像素点的视差可以进行多次计算减小误差,在一定程度上优于双目视觉方法。

根据摄像机的位置关系,三目视觉可分为直角三角形结构和共线结构2种。由于三目视觉法设备结构更加复杂,导致移动不方便且成本较高,因此实际应用情况甚少。在医学领域中,三目视觉法需要在单目视觉法的基础上增加2台摄像机,不利于医疗器械的小型化,会加大腔镜手术难度及图像采集过程的复杂度,因此,该方法不适用于人体内腔的三维重建。

### 3.3 基于单目视觉法的内腔三维重建

单目视觉法是利用一个相机拍摄目标物体图像进行三维重建技术。基于单目视觉法的三维重建方法包括阴影形状恢复法、光度立体视觉法、纹理法、轮廓法、调焦法等<sup>[6]</sup>。

在内腔环境下进行单目视觉的三维建模主要有2种实现方式:利用内腔图像本身的信息,通过特定的算法得到内腔的三维特征信息;通过对内窥镜系统的摄像机进行参数标定,进而得到被测点的深度信息。因为单目视觉法具有设备结构简单、使用方便和数据信息便于处理等优点,所以国内外的研究大多是运用单目视觉算法进行内腔三维重建。常用的方法是阴影形状恢复(Shape From Shading, SFS)法和同时定位和映射法。

#### 3.3.1 基于阴影形状恢复法的内腔三维重建

阴影形状恢复法作为一种被动式测量方法,利用单幅图像的明暗灰度变化来恢复物体表面各点的

相对高度或法向量<sup>[41]</sup>。早在1970年,HORN<sup>[42]</sup>提出该方法并对月球表面进行三维重建。SFS技术经过50多年的发展,许多新的算法不断涌现,因为其原理简单、操作简便,可以快速提取待测物体表面的三维信息,所以广泛应用于医学图像分析、物体变形矫正、生物识别技术、地貌测量等领域<sup>[43]</sup>。

在医学图像分析与重建中,外科医生使用内窥镜对人体内腔进行拍摄,根据图像灰度信息的明暗变化对内腔器官进行三维重建。SFS算法的便捷性在于只用1台摄像机,目前已经有较多相关研究将SFS算法应用于内腔三维重建手术中。传统的SFS方法大多为正交映射模型,假设光源为无限远处的点光源时,物体表面为朗伯体表面。这些假设条件虽然降低了该方法的复杂程度,但实际的物体表面并非理想的漫反射表面,导致重建误差较大。为了克服上述缺陷,OKATANI等<sup>[44]</sup>于1996年利用透视映射模型替换传统的正交映射模型,引入物体表面等距轮廓的概念,首次采用SFS算法对人体胃壁图像进行三维重建。但是由于人体内腔的特殊性,光线在器官表面上进行反射,无法解决胃部轮廓拓扑模糊等问题。

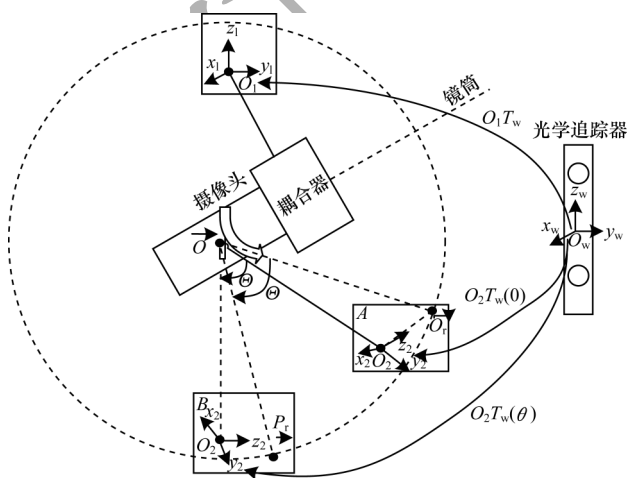
2004年,ARIEL等<sup>[45]</sup>提出一种基于快速行进法的透视映射SFS算法。与传统正交映射SFS算法相比,该算法可以对阑尾和盲肠等腔体进行重建,且对十二指肠、直肠重建的鲁棒性和准确性更好。2006年,LI等<sup>[46]</sup>提出一种基于三角形模型的SFS算法。该方法通过增强约束减小胃部表面边缘的旋转失真,并采用亮度误差因子来减弱透视辐照方程的非线性对重建误差的影响,最后由三角形曲面单元拼接成胃部表面。不足之处在于:重建结果的准确性相对较低,并且当内腔器官表面没有显著特征时,容易导致图像匹配算法失败。2009年,LAMATA等<sup>[47]</sup>基于朗伯体模型对器官上互反射、漫反射和高光部分根据其特征颜色的不同进行区分,通过调整反照率和光强度对肝脏进行三维重建后,使用迭代最近点算法并与术前模型配准,在肝切除手术得以应用。

由于器官表面反射特性存在多样性,且存在多种反射特性区域,因此外科医生在没有术前数据的前提下无法精确地恢复出深度信息。2012年,CIUTI、VISENTINI-SCARZANELLA等<sup>[48-49]</sup>提出一套完整的SFS校准方法,并假设光源位于接近器官表面而远离

光学中心,通过对器官表面具有镜面高光的部分三角剖分得到空间三维坐标。该方法在无任何术前数据的情况下,内窥镜装置根据计算的轨迹进行3D测量并最终实现胶囊自动导航。但是磁悬浮胶囊在运动过程中达不到理想状态,校准精度还需进一步提高。同年,TOKGOZOGLU等<sup>[50]</sup>提出一种基于彩色投影的SFS法,该方法可最大程度地减少由不同表面特性引起的强度变化,具有鲁棒性,但缺点在于易受浮雕模糊性的影响。2016年,LEI等<sup>[51]</sup>提出基于光度校准法的透视映射SFS法重建器官表面,该方法结合光流法将灰度梯度场的相对变化变为绝对变化,提高了器官表面重建的稳定性。

目前多数SFS算法是在获取的二维图像未失真的前提下进行内腔三维重建。但由于内窥镜系统具有宽视场的特性,获取的二维图像存在径向畸变(Radial Distortion, RD),容易造成重建效果误差较大。2015年,NUNO等<sup>[52]</sup>针对径向畸变和图像边缘分辨率降低等问题,提出一种基于近光源透视映射的阴影形状恢复(Perspective Shape From Shading, PSFS)算法。该算法通过建立RD模型,并对图像边缘分辨率降低的问题进行补偿,完成膝骨三维重建。

为提高三维重建的精度,WU等<sup>[53]</sup>提出了结合SFS算法和运动形状恢复法的内腔三维重建方法,如图11所示。该方法结合迭代最近点算法减少多幅人造脊椎图像中坐标系转换时的误差,提高了匹配率并且恢复出骨骼边界线。优点是显著加快算法的运行速度,缺点是由于骨骼表面过度平滑,重建的部分结果仍然精度较差。2018年,TURAN等<sup>[54]</sup>将上述方法用于胃肠表面重建,然而胃肠表面并非平滑区域,凹凸不平的表面会使梯度向量变化率较高,得到的灰度值也会低于真实值,导致重建误差较大。



2015年,LIN等<sup>[63]</sup>提出了基于SLAM恢复腹腔手术三维场景表面结构的方法,针对内腔图像的纹理特性绿色通道选取和反光区域的处理等方面做出了改进,并研究了新类型的图像特征即血管特征中的分支点。在检测到血管特征点后,又针对分支段进行联合检测和匹配分支段,以此匹配图像中的血管特征。最后在每帧图像中恢复出三维血管,通过血管的匹配将来自不同视角的三维血管整合在一起,得到一个全局的三维血管网络。

外科医生通过内窥镜观察到的内腔场景是散乱的,只有依赖临床经验才能判断病变组织,若要进行治疗时,也只能凭借经验去判断之前的位置,降低了手术的准确性和安全性。2016年,杨滨豪<sup>[64]</sup>提出了基于单目SLAM的内窥镜定位和胃肠特征地图的构建。该方法从效率与匹配正确率的角度选择了ORB算法进行特征点检测,并结合局部位姿优化算法与几何距离最小的三角测量法,通过对关键帧的重新选定以及对特征点的筛选来处理大量的数据冗余情况。但由于该环境为内窥镜运动轨迹非闭合的肠道,而局部趋于直线,因此与大部分内腔环境的闭合情况不同。

基于特征点法的SLAM虽然数据量有所减少,降低了计算复杂度,但相比于原始图,因其信息量的

减少也代表着错过一些有用的信息。为了充分利用图像信息,使关键帧跟踪和内腔重建地图更为精确,2019年,MAHMOUD等<sup>[65]</sup>提出基于单目ORB-SLAM的腹腔稠密三维重建。该方法先利用稀疏ORB-SLAM的检测匹配流程估计关键帧相机位姿,其中关键帧的选取依据视差准则来确定,再运用零均值归一化互相关(Zero mean Normalized Cross Correlation,ZNCC)和梯度Huber范数正则化器相结合的变分方法并行计算关键帧之间的稠密匹配。该方法采用单目视频输入,不需要任何基准点,也不需要外部跟踪器,且在猪腹腔视频序列上得到了验证和评估,说明其对严重的光照变化和不同场景纹理的鲁棒性强。该系统的主要局限性为:对软组织表面的纹理特征描述不具有代表性,存在重建后纹理失真的情况。同年,XIE等<sup>[66]</sup>结合在胃肠道中内窥镜的测量数据,在位姿优化和空间点定位方面引入了局部姿势优化算法和最小几何距离的三角测量算法。图13所示为基于最小几何距离的三角测量法的单目视觉SLAM算法框图。但是该方法存在低帧频下跟踪丢失的现象,需要改善当前系统中的帧速率,并通过使用压缩感知的单次测量以高帧速率重建序列图像,进而恢复出胃肠道3D形状与细致的表面纹理。

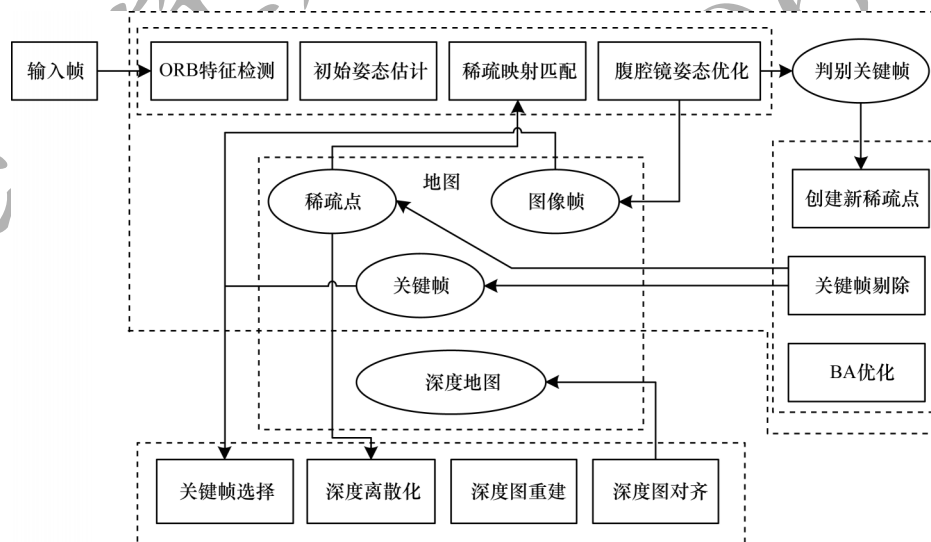


图 13 基于最小几何距离三角测量法的单目视觉 SLAM 算法

**Fig.13 Monocular visual SLAM algorithm based on triangular measurement algorithm with the minimum geometric distance**

PICCINELLI 等<sup>[67]</sup>通过结合单目 SLAM 算法获得的稀疏点云与患者的术前模型配准,对患者的内腔进行三维重建,但是无法解决软组织变形等问题。2021 年,LAMARCA 等<sup>[68]</sup>提出 Def SLAM 算法。该算法可以在变形场景中实时运行,并将计算拆分为变形跟踪线程和变形映射线程 2 个并行线程。变形跟踪线程用于估计摄像机姿态和场景的变形,变形映射线程应用于内窥镜的位姿估计,以便更好地适应内腔变形场景,从而生成人体内腔的精确 3D 模型,但是容易受照明不均匀的影响造成视觉纹理差,

且不适用于非等距变形的内腔重建。

近年来,单目内窥镜在内窥镜诊断方面的应用已经有了大量的研究和长足的发展。在微创手术中,人体组织会有形变流血的现象,并且往往不具备强边缘特征,从而表现出有高光和镜面反射的特点。面对这种复杂的微创手术环境,单目SLAM方法具有较高的鲁棒性,并且能实时处理软组织序列图像。因此,单目SLAM技术在临床微创手术中的应用及推广具有十分重要的现实意义。表1给出了人体内腔三维重建方法性能比较与分析。

表1 人体内腔三维重建方法的性能分析比较

Table 1 The performance analysis and comparison of 3D reconstruction technologies in human lumen

方法	优点	缺点	重建效果	应用情况
时域编码法	易于实现,空间分辨率高	无法对动态场景进行测量	受投影仪空间分辨率的限制,匹配精度不高	需投影多幅图案,只适用于静态场景测量
编码结构光法	测量速度快,实现简单且精度较高	在解码阶段,空间邻近点信息的丢失会导致重建误差的产生	需要采用预先设计好的一幅编码图案进行编码,才能得到较好的重建效果	适用于动态场景三维信息的获取
飞行时间法	可以快速响应	系统误差与随机误差对测量结果的影响很大,需要进行后期数据处理	读取的图像并不清晰,获取的深度图像无法还原内腔的细节特征	易受环境因素影响,复杂环境下系统稳定性不高,分辨率低
双目视觉法	具备第三维深度信息,在三维模型重建上更为便捷	镜头的畸变,特征点稀疏,特征点匹配耗时长	在基线距离不大的情况下,重建效果较好	在工业内窥镜中得到少量应用,在医用手术内窥镜中的应用不广泛
阴影恢复形状法	较为精细地恢复目标表面的细节纹理信息	数学计算较多,只能得到到相对曲面形状	易受光源影响,鲁棒性较差	对光源要求严格,难以应用较为复杂环境下的重建
单目视觉法	同时定位与地图构建	误差较小,重建精度较高,应用场景广泛	依赖于相对丰富的点云数据	对特征点提取依赖较高,适合对人体内腔进行三维重建

图像的特定模型可以是物体的边缘、轮廓以及各种连接点,它们被统称为图像特征。特征信息在物体识别、运动结构分析、纹理分类,特别是图像特征提取与匹配中有着广泛的应用<sup>[69]</sup>。基于特征点的内腔SLAM利用关键帧图像信息解算3D点云信息,每当系统获取到一帧图像时,首先提取内腔图像中的特征点,将每个特征点的描述子与前一帧图像中的特征点进行特征匹配,建立起特征点之间的一一对应关系,然后利用匹配的特征点在不同图像中的投影结果最小化重投影误差来估计相机的运动。该类方法的关键在于特征点检测、匹配和运动估计的处理过程,如图14所示。

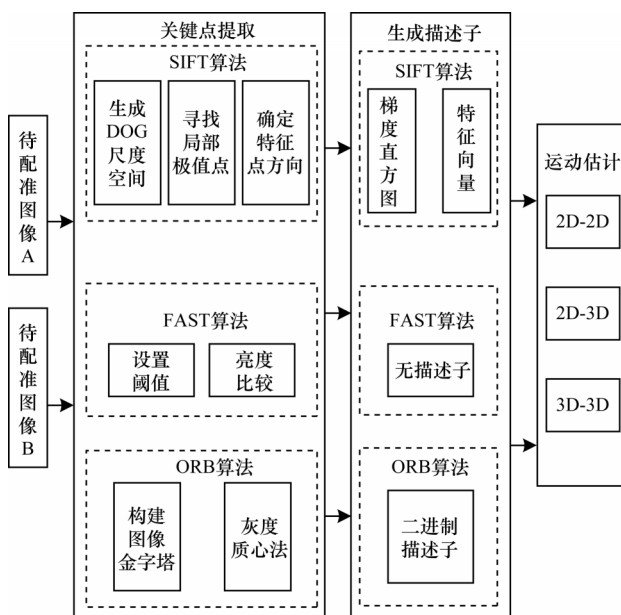


图14 特征点处理流程

Fig.14 The procedure of feature point processing

特征点检测与特征匹配方法如下:

#### 1) 特征点检测

特征点检测是图像处理中的基础步骤,首先将图像的每个像素进行检验,看其是否符合特征点要求,算法一般只检查图像的特征区域,这一步骤有效减少了算法的数据处理规模,提高了图像处理的效率。根据特征的数据类型可分为基于向量的特征检测和基于二值的特征检测2类:

(1) 基于向量的特征检测方法是通过像素点强度的变化得到特征点,利用特征点周围的一小块像素区域计算出固定长度的描述子<sup>[70]</sup>。自20世纪80年代起,MORAVEC<sup>[71]</sup>提出一种角点算法,并将其定义为具有低“自相关性”的点,随后HARRIS等<sup>[72]</sup>为了解决文献[71]算法中强度值各向同性和响应噪声等问题,提出了用于检测图像特征的Harris角点,这种特征具有旋转不变性,但尺度变化较大。

随着时间的推移,不断有研究者对角点检测算法提出新的思路。2001年,POSTON<sup>[73]</sup>提出了一种基于累加点到弦长距离(Chord-to-Point Distance Accumulation, CPDA)的角点检测方法。该方法通过计算曲率来判定特征点,不涉及求导等复杂运算,运算速度快,且不需要考虑高斯平滑的尺度值设定,所以提取的特征点中除目标点外的冗余点较少,但不适用于人体内腔环境的鲁棒性稍差的角点。在CPDA角点检测方法的基础上,赵晓芳等<sup>[74]</sup>根据图像的拓扑结构设计了自适应矩形探测器,通过提取出视网膜血管的边缘线条并利用矩形窗检测法剔除对于血管形态来讲鲁棒性低的角点,留下视网膜纹理中的分支点。在检测成功的情况下,还需将特征点分成角点、端点、分支点等不同类别进行类统计。但矩形窗设定复杂,计算成本高,且人体内腔血管相

较视网膜血管更为复杂,检测环境中的不定因素更多,例如照明引起的反光等,需设计特定的算法来适应环境。

2004年,LOWE<sup>[75]</sup>完善总结了尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)算法,结合高斯金字塔尺度空间检测以及特征点梯度的概念,保证特征点的检测不会被图像尺度的变化以及图像的旋转所影响。SIFT算法的提取目标是那些像素值与像素值梯度方面有突变的点,例如斑点和角点等。该算法中描述子的匹配重复率高,但由于SIFT算法在匹配环节中需要统计的描述子数量为单个特征点的128维向量,所以算法效率较低。考虑在内腔背景下斑点与角点的鲁棒性较差,内腔出现出血点就会导致特征点的增加或消去,匹配效果较差。

为了进一步减少特征检测算法的计算量,2008年,BAY等<sup>[76]</sup>以SIFT算法为基础提出了SURF(Speeded Up Robust Features)特征点检测算法,其对特征点主方向的分配不是通过梯度计算,而是利用对图像的Haar小波响应运算,提高计算图像梯度的效率,其运算效率是SIFT算法的3~7倍左右,但对特征点的区别性能有明显的下降。

2017年,卢鑫<sup>[77]</sup>提出了邻域模块检测算法,用于检测视网膜血管图像中的分叉点和交叉点。该方法在提取出血管形态后使用Sobel算子进行增强处理,再采用二维最大熵进行阈值分割。优点是采用形态学中的腐蚀法对血管进行窄化,仅保留了血管宽度中的一个像素,优化了Harris算法并通过检测特征点邻域内的像素值区分出特征点的类型。缺点是需要进行二次模板的设定,通过二次检测来区分出特征点中的分支点,算法复杂度高,导致效率低。

(2)基于二值的特征检测方法,与基于向量方法最大的不同在于其描述子为二值序列,由0和1组成一定长度。2011年,CALONDER等<sup>[78]</sup>提出了二值特征计算方法BRIEF(Binary Robust Independent Elementary Feature),摒弃了利用区域灰度直方图描述特征点的传统方法,在计算速度上有巨大的提升,极大地降低了特征匹配的时间。但由于BRIEF描述子不具有方向性,当内窥镜在内腔中大角度旋转时会对匹配有很大的影响。RUBLEE等<sup>[79]</sup>提出了一种结合FAST(Features from Accelerated Segment Test)角点检测和BRIEF描述子的ORB算法,改进了BRIEF描述子对图像噪声敏感和不具备旋转不变性的缺点,在速度方面也有了较大提升,适合在内腔环境下进行实时三维重建。

## 2)特征匹配

特征匹配是基于特征的三维重建流程中最重要的一步,根据提取待测图像及目标图像的部分特征或是主要特征构成特征空间,然后通过对比进行匹配。2011年,华南理工大学赵晓芳<sup>[80]</sup>提出基于结构特征的视网膜血管性态识别技术。该方法基于节点

之间的相互位置固定不变以及三角形的稳定性,对于1个节点找出距离它最近的3个不共线特征点,形成三角形结构,进行基于特殊结构的匹配。该方法鲁棒性强,可以在图像的尺度变化与旋转变化中达到很好的匹配效果。

2015年,LIN等<sup>[81-82]</sup>提出了基于血管线条的线匹配算法,在进行匹配操作时,基于两关键帧的左、右图像块像素的相似度,首先对特征点做块匹配,然后通过极线几何的线对应,避免进行遍历整个图像像素点的缓慢匹配。在此基础上,通过血管段上的特征点,将点匹配延伸至血管的线匹配。

以上2种方法由于内腔图像特征的局部特性原因产生误匹配点对,因此需要利用随机样本一致性(Random Sample Consensus, RANSAC)方法对匹配结果进一步优化,筛选出多余的匹配点对并剔除误匹配点对。该方法根据观测的匹配数据构建一个数学模型,然后将匹配结果反复迭代以更新参数,找出异常的匹配点对。2017年,SONG等<sup>[83]</sup>在内腔三维重建中获取了密集的表面关键特征点,用SURF算法建立描述子进行匹配,而后经过刚体位姿估计得到平移参数和旋转参数,利用RANSAC过滤异常值,优化匹配结果。

特征匹配也可以根据2幅内腔图像之间的相似性测度,选取两者完全相符或部分相符的血管等信息进行配准。目前使用较多的检索方法主要有最小距离度量、相关函数度量与概率度量3种。2019年,MAHMOUD等<sup>[65]</sup>提出ZNCC和梯度Huber范数正则化器相结合的变分方法,并行计算关键帧之间的稠密匹配。除常用的块匹配外,还有适用于不同应用场景的基于特殊结构的特征匹配,例如基于线条的匹配与基于各种稳定几何结构的匹配。针对当人体内腔中特征点匹配时,在考虑块匹配的同时可以用血管线条作为辅助。

## 4 存在的问题与研究方向

医学图像三维重建技术是目前图像处理领域中流砥柱,在医学研究方面具有创新意义,在医学教育方面也发挥着指导作用。目前的人体内腔三维重建方法仍然存在一定问题,下文对其未来的发展方向和问题优化作出讨论。

### 4.1 三维重建算法的误差问题

人体内腔三维重建技术鲁棒性相对较差,在内腔中存在的光线不足、高光、液体较多等问题,都会影响重建效果。还有在匹配时,匹配的误差也对导致重建结果不够准确。与其他用于捕获3D图像的激光扫描方法相比,TOF相机速度更快,每秒可提供多达160次图像获取,具有较高的实时性。同时,TOF相机受光线的影响较小,不需要立体视觉中的复杂算法即可从输出信号中提取距离信息,且算法不受物体上图案的干扰。

双目视觉系统运用外部光线对被测物体进行照射,当光源产生明暗变化时,对被测物体进行特征提取与匹配的效果将受到影响。同时,在运用双目视觉系统进行三维成像时,所需算法的运算量将占用较多运算资源,由于该方法提取距离信息的算法较为复杂,因此实现实时三维立体成像比较困难。单目内窥镜具有灵活性好、结构简单的特点,而且单目在同一时刻只需要处理单幅软组织图像,实时性更好。尤其是SLAM算法针对狭小、特征缺乏的内腔空间,可以快速、准确地获取更真实的三维重建效果,比如ORB-SLAM系统得到的稀疏点云地图在内腔三维重建系统中有很好的应用。

#### 4.2 三维重建算法对病灶的特征增强与精准定位

在一些医学检测手段的成像过程中,患者的生理运动或自主性运动会使扫描数据发生相位偏移,致使重建图像含有运动伪影,从而降低成像的质量,而且严重的伪影会影响医生对病灶的精准定位。因此,人体内腔的三维重建算法对于病灶的定位与增强现实仍有待加强,为主刀医师提供多方位视角,进行局部病灶与邻近结构关系的辨识,方便医生选择最佳手术路径,更好地为病人服务。如在解剖性肺段切除术中通过3D重建准确定位肺动脉、静脉、支气管分支、肺边缘达到精准切除靶病灶,符合外科手术原则,减少了误伤血管概率及术中出血量,缩短手术时间,提高术后生活质量。

#### 4.3 内窥镜三维重建技术与CT三维影像重建信息的融合

运用CT技术进行临床诊断能够较快定位患者病灶进行及时有效的治疗,因而将常规CT与三维重建技术相结合,有利于更加准确地定位病灶位置,揭示病灶特征与组织之间的关系。进一步,将内窥镜三维重建技术与CT三维影像重建信息融合,在图像处理平台进行优化处理,使诊断的准确率得到有效提升,对疾病的发现与诊断具有重要意义,对术前完整性评估具有良好价值。

#### 4.4 优化三维重建算法运算量

运算量大是人体内腔三维重建技术面临的主要难题,过大的运算量会占用较长的时间,导致三维重建的效率低下。例如,双目视觉方法由于受光照、图像特征等因素的影响,在潮湿、昏暗的内腔环境下对特征匹配的要求更加严格。为了优化后续的位姿估计,需要剔除误匹配点对,造成算法计算量剧增,对实时应用场景提出较大挑战。由于曝光不均、噪声等因素通常会引起病灶区数据漏检、边界模糊等问题,进行3D重建前通常会通过精确超分辨率重建进行预处理,也需要大量运算进行迭代计算等。

人体内腔三维重建技术的运算量过大对重建的实时性造成很大影响。因此,合理利用计算机技术,采用GPU计算、网络云计算等多种方式节省时间,提高运行速度及效率是解决该问题的关键。

## 5 结束语

人体内腔三维重建技术目前已经得到广泛的应用,并随着科学技术的进步不断地更新迭代,其方法对临床医学产生了深远影响。医学图像的三维重新在临床诊断中发挥着越来越重要的作用,也在一定程度上加速了临床医学的发展进程。本文分类介绍了人体内腔环境下的三维重建方法并给出相应的评析,阐述当前三维重建算法存在的问题并对未来的研究工作进行了展望,但是这些方法要应用到实际中还需要更进一步的研究和考察。因此,在人体内腔三维重建领域中,根据优化算法对病灶精确定位、结合多传感器融合图像信息进行内腔表面精准三维重建,将三维重建方法与已有技术相结合等研究方向,都需要进行更加深入细致的研究。

## 参考文献

- [1] ROSEN M, PONSKY J. Minimally invasive surgery[J]. Endoscopy, 2001, 33(4): 358-364.
- [2] 费保蔚. 计算机辅助外科手术的方法和进展[J]. 国际生物医学工程杂志, 2015, 20(4): 624-631.
- [3] FEI B W. The method and development of computer-assisted surgery[J]. International Journal of Biomedical Engineering, 2015, 20(4): 624-631. (in Chinese)
- [4] DISSANAYZKE G, DURRANT-WHYTE H, BAILEY T. A computationally efficient solution to the Simultaneous Localisation and Mapping (SLAM) building problem[C]// Proceedings of IEEE International Conference on Robotics & Automation, Washington D. C., USA: IEEE Press, 2000: 1009-1014.
- [5] DURRANT-WHYTE H, MAJUMDER S, THRUN S, et al. A Bayesian algorithm for simultaneous localisation and map building[J]. Robotics Research, 2003, 264(3): 49-60.
- [6] DISSANAYZKE G, NEWMAN P, CLARY S, et al. A solution to the simultaneous localisation and mapping building problem[J]. IEEE Transactions of Robotics and Automation, 2013, 17(3): 229-241.
- [7] 佟帅, 徐晓刚, 易成涛, 等. 基于视觉的三维重建技术综述[J]. 计算机应用研究, 2011, 28(7): 2411-2417.
- [8] TONG S, XU X G, YI C T, et al. Overview on vision-based 3D reconstruction[J]. Application Research of Computers, 2011, 28(7): 2411-2417. (in Chinese)
- [9] HANEISHI H, OGURA T, MIYAKE Y. Profilometry of a gastrointestinal surface by an endoscope with laser beam projection[J]. Optics Letters, 1994, 19(9): 601-603.
- [10] 郑太雄, 黄帅, 李永福, 等. 基于视觉的三维重建关键技术综述[J]. 自动化学报, 2020, 46(4): 631-652.
- [11] ZHENG T X, HUANG S, LI Y F, et al. Key techniques for vision based 3D reconstruction: a review [J]. Acta Automatica Sinica, 2020, 46(4): 631-652. (in Chinese)
- [12] ARMBRUSTER K, SCHEFFLER M. Messendes 3D-endoskop[EB/OL]. [2021-04-01]. <https://philpapers.org/rec/ARMMD>.
- [13] HAYASHIBE M, SUZUKI N, NAKAMURA Y. Laser-scan endoscope system for intraoperative geometry acquisition

- and surgical robot safety management[J]. *Medical Image Analysis*, 2006, 10(4): 509-519.
- [11] SALVI J, PAGÈS J, BATLLE J. Pattern codification strategies in structured light systems [J]. *Pattern Recognition*, 2004, 37(4): 827-849.
  - [12] ALBITAR C, GRAEBLING P, DOIGNON C. Fast 3D vision with robust structured light coding[C]//*Proceedings of IEEE SPIE'09*. Washington D. C., USA: IEEE Press, 2009: 125-136.
  - [13] MAURICE X, ALBITAR C, DOIGNON C, et al. A structured light-based laparoscope with real-time organs' surface reconstruction for minimally invasive surgery[C]//*Proceedings of 2012 IEEE International Conference on Engineering in Medicine and Biology Society*. Washington D. C., USA: IEEE Press, 2012: 5769-5772.
  - [14] CLANCY N T, STOYANOV D, MAIER-HEIN L, et al. Spectrally encoded fiber-based structured lighting probe for intraoperative 3D imaging[J]. *Biomedical Optics Express*, 2011, 2(11): 3119-3128.
  - [15] REITER A, SIGARAS A, FOELER D, et al. Surgical structured light for 3D minimally invasive surgical imaging [C]//*Proceedings of IEEE/RSJ International Conference on Intelligent Robots & Systems*. Washington D. C., USA: IEEE Press, 2014: 1282-1287.
  - [16] QIU L, REN H L. Endoscope navigation and 3D reconstruction of oral cavity by visual SLAM with mitigated data scarcity [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2018: 1354-1476.
  - [17] HAMADON A B, DAUL C, SOUSSEN C, et al. A novel 3D surface construction approach: application to three-dimensional endoscopic data [C]//*Proceedings of IEEE International Conference on Image Processing*. Washington D. C., USA: IEEE Press, 2010: 235-246.
  - [18] YANG S P, SEO Y H, HWANG K, et al. Switchable dot projection module for 3D flexible endoscopes [C]//*Proceedings of 2019 IEEE International Conference on Optical MEMS and Nanophotonics*. Washington D. C., USA: IEEE Press, 2019: 346-358.
  - [19] FURUKAWA R, OKA S, KOTACHI T, et al. Fully auto-calibrated active-stereo-based 3D endoscopic system using correspondence estimation with graph convolutional network [C]//*Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference*. Washington D. C., USA: IEEE Press, 2020: 4357-4360.
  - [20] 李占利,周康,牟琦,等. TOF相机实时高精度深度误差补偿方法[J]. *红外与激光工程*, 2019, 48(12): 263-272.  
LI Z L, ZHOU K, MU Q, et al. TOF camera real-time high precision depth error compensation method[J]. *Infrared and Laser Engineering*, 2019, 48(12): 263-272. (in Chinese)
  - [21] 魏加立,曲慧东,王永宪,等. 基于飞行时间法的3D相机研究综述[J]. *红外技术*, 2021, 43(1): 60-67.  
WEI J L, QU H D, WANG Y X, et al. Research review of 3D cameras based on time-of-flight method[J]. *Infrared Technology*, 2021, 43(1): 60-67. (in Chinese)
  - [22] TATAR F, MOLLINGER J R, DULK R, et al. Ultrasonic sensor system for measuring position and orientation of laparoscopic instruments in minimal invasive surgery [C]//*Proceedings of the 2nd Annual International IEEE-EMB Special Topic Conference on Micro Technologies in Medicine & Biology*. Washington D. C., USA: IEEE Press, 2002: 113-125.
  - [23] PENNE J, SCHALLER C, HORNEGGER J, et al. Robust real-time 3D respiratory motion detection using time-of-flight cameras [J]. *International Journal of Computer Assisted Radiology and Surgery*, 2008, 3(5): 427-431.
  - [24] PENNE J, SCHRAUDER T, SCHNEIDER A, et al. Time-of-flight 3D endoscopy [C]//*Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*. London, UK: [s. n.], 2009: 467-474.
  - [25] HAASE S, WASZA J, KILGUS T, et al. Laparoscopic instrument localization using a 3D time-of-flight/RGB endoscope [C]//*Proceedings of 2013 IEEE Workshop on Applications of Computer Vision*. Washington D. C., USA: IEEE Press, 2013: 478-486.
  - [26] ZVOLSKY M, CSERKASZKY A, CUCCIATI G, et al. Monte-carlo simulations for an endoscopic time-of-flight PET detector [C]//*Proceedings of 2014 IEEE Nuclear Science and Medical Imaging Conference*. Washington D. C., USA: IEEE Press, 2014: 253-268.
  - [27] ZORRAQUINO C, BUGALHO R, ROLO M, et al. Asymmetric data acquisition system for an endoscopic PET-US detector [J]. *IEEE Transactions on Nuclear Science*, 2016, 63(1): 213-221.
  - [28] BRANDT O, MUNWES Y. Commissioning and first image reconstruction with a new time-of-flight PET prototype [C]//*Proceedings of 2018 IEEE Conference on Nuclear Science and Medical Imaging*. Washington D. C., USA: IEEE Press, 2018: 134-146.
  - [29] LAU W W, RAMEY N, CORSO J J, et al. Stereo-based endoscopic tracking of cardiac surface deformation [C]//*Proceedings of International Conference on Medical Image Computing & Computer-Assisted Intervention-Miccai*. Saint-malo, France: [s. n.], 2004: 357-368.
  - [30] STOYANOV D, DARZI A, YANG G Z. Dense 3D depth recovery for soft tissue deformation during robotically assisted laparoscopic surgery [C]//*Proceedings of International Conference on Medical Image Computing & Computer-Assisted Intervention-Miccai*. Saint-malo, France: [s. n.], 2004: 587-596.
  - [31] STOYANOV D, MYLONAS G P, DELIGIANNI F, et al. Soft-tissue motion tracking and structure estimation for robotic assisted MIS procedures [J]. *Medical Image Computing and Computer-Assisted Intervention*, 2005, 8(2): 139-146.
  - [32] STOYANOV D, SCARZANELLA M V, PRATT P, et al. Real-time stereo reconstruction in robotically assisted minimally invasive surgery [C]//*Proceedings of IEEE International Conference on Medical Image Computing & Computer-assisted Intervention*. Washington D. C., USA: IEEE Press, 2010: 387-396.
  - [33] STOYANOV D. Stereoscopic scene flow for robotic assisted minimally invasive surgery [C]//*Proceedings of IEEE International Conference on Medical Image Computing & Computer-*

- assisted Intervention. Washington D. C. , USA; IEEE Press, 2012; 578-587.
- [34] CHANG P L, STOYANOV D, DAVISON A J, et al. Real-time dense stereo reconstruction using convex optimisation with a cost-volume for image-guided robotic surgery[C]// Proceedings of IEEE International Conference on Medical Image Computing & Computer-Assisted Intervention. Washington D. C. , USA; IEEE Press, 2013; 588-598.
- [35] TOTZ J, THOMPSON S, STOYANOV D, et al. Fast semi-dense surface reconstruction from stereoscopic video in laparoscopic surgery [C]// Proceedings of IEEE International Conference on Information Processing in Computer-Assisted Interventions. Washington D. C. , USA; IEEE Press, 2014; 763-777.
- [36] PENZA V, ORTIZ J, MATTOS L S, et al. Dense soft tissue 3D reconstruction refined with super-pixel segmentation for robotic abdominal surgery [J]. International Journal of Computer Assisted Radiology and Surgery, 2016, 11(2): 197-206.
- [37] SDIRI B, KAANICHE M, CHEIKH F A, et al. Efficient enhancement of stereo endoscopic images based on joint wavelet decomposition and binocular combination [J]. IEEE Transactions on Medical Imaging, 2019, 38(1): 33-45.
- [38] WANG D, LIU H, CHENG X. A miniature binocular endoscope with local feature matching and stereo matching for 3D measurement and 3D reconstruction[J]. Sensors, 2018, 18(7): 2243.
- [39] 王博闻. 基于ORB-SLAM2的医学影像三维重建[D]. 太原: 中北大学, 2019.  
WANG B W. 3D reconstruction of medical images based on ORB-SLAM2[D]. Taiyuan: North University of China, 2019. (in Chinese)
- [40] WANG Z. Deep convolutional network for stereo depth mapping in binocular endoscopy[J]. IEEE Access, 2020, 7: 12421-12432.
- [41] 王国琿, 韩九强, 张新曼, 等. 一种新的基于从明暗恢复形状的月球表面三维形状恢复算法[J]. 宇航学报, 2009, 30(6): 2265-2269.  
WANG G H, HAN J Q, ZHANG X M, et al. A new three-dimensional reconstruction algorithm of the lunar surface based on shape from shading method [J]. Journal of Astronautics, 2009, 30(6): 2265-2269. (in Chinese)
- [42] HORN B K P. Height and gradient from shading [J]. International Journal of Computer Vision, 1990, 5(1): 37-75.
- [43] 王国琿, 王建, 孙帅. SFS三维重构技术的发展现状与趋势研究[J]. 光学仪器, 2011, 33(4): 90-94.  
WANG G H, WANG J, SUN S. Research on development status and trend of SFS-based 3D reconstruction technique [J]. Optical Instruments, 2011, 33(4): 90-94. (in Chinese)
- [44] OKATANI T, DEGUCHI K. Shape reconstruction from an endoscope image by Shape from shading technique for a point light source at the projection[J]. Computer Vision and Image Understanding, 1997, 66(2): 19-131.
- [45] TANKUS A, SOCHEN N, YESHURUN Y. Reconstruction of medical images by perspective shape-from-shading[C]// Proceedings of the 17th International Conference on Pattern Recognition. Washington D. C. , USA; IEEE Press, 2004; 158-269.
- [46] LI J, REN A M, ZHANG J N. 3D reconstruction by perspective shape from shading using linearized triangular element surface model [C]// Proceedings of 2006 International Conference on Mechatronics and Automation. Harbin, China; [s. n. ], 2006; 378-386.
- [47] LAMATA P, MORVAN T, REIMERS M, et al. Addressing shading-based laparoscopic registration[C]// Proceedings of World Congress on Medical Physics and Biomedical Engineering. Munich, Germany; [s. n. ], 2009; 235-246.
- [48] CIUTI G, VISENTINI-SCARZANELLA M, DORE A, et al. Intra-operative monocular 3D reconstruction for image-guided navigation in active locomotion capsule endoscopy [C]// Proceedings of IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechanics. Washington D. C. , USA; IEEE Press, 2012; 768-774.
- [49] VISENTINI-SCARZANELLA M, STOYANOV D, YANG G Z. Metric depth recovery from monocular images using Shape-from-shading and specularities[C]// Proceedings of IEEE International Conference on Image Processing. Washington D. C. , USA; IEEE Press, 2013; 457-468.
- [50] TOKGOZOGLU H N, MEISNER E M, KAZHDAN M, et al. Color-based hybrid reconstruction for endoscopy [C]// Proceedings of Computer Vision & Pattern Recognition Workshops. Washington D. C. , USA; IEEE Press, 2016; 92-101.
- [51] LEI L, LI J H, LIU M Q, et al. Shape from shading and optical flow used for 3-dimensional reconstruction of endoscope image[J]. Acta Oto-Laryngologica, 2016, 23: 1190-1192.
- [52] GONCALVES N, ROXO D, BARRETO J, et al. Perspective shape from shading for wide-FOV near-lighting endoscopes[J]. Neurocomputing, 2015, 150: 136-146.
- [53] WU C Y, NARASIMHAN S G, JARAMAZ B. A multi-image shape-from-shading framework for near-lighting perspective endoscopes [J]. International Journal of Computer Vision, 2010, 86(2/3): 211-228.
- [54] TURAN M, PILAVCI Y Y, GANIYUSUFOGLU I, et al. Sparse-then-dense alignment-based 3D map reconstruction method for endoscopic capsule robots[J]. Machine Vision and Applications, 2018, 29(2): 345-359.
- [55] SMITH R, SELF M, CHEESEMANN P. A stochastic map for uncertain spatial relationships[C]// Proceedings of the 4th IEEE International Symposium on Robotics Research. Washington D. C. , USA; IEEE Press, 1988, 29: 467-474.
- [56] 瑚琦, 蔡文龙, 卢定凡, 等. 单目视觉惯性的同时定位与地图构建综述[J]. 软件导刊, 2020, 19(7): 275-280.  
HU Q, CAI W L, LU D F, et al. Survey on monocular visual inertial SLAM algorithms[J]. Software Guide, 2020, 19(7): 275-280. (in Chinese)
- [57] MOUNTNEY P, STOYANOV D, DAVISON A, et al. Simultaneous stereoscope localization and soft-tissue mapping for minimal invasive surgery [C]// Proceedings of IEEE International Conference on Medical Image Computing & Computer-Assisted Intervention. Washington D. C. , USA; IEEE Press, 2006; 347-356.

- [58] DAVISON A J, REID I D, MOLTON N D, et al. MonoSLAM: real-time single camera SLAM [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 1052-1067.
- [59] GRASA O G, CIVERA J, MONTIEL J M M. EKF monocular SLAM with relocalization for laparoscopic sequences [C]//Proceedings of IEEE International Conference on Robotics and Automation. Washington D. C., USA: IEEE Press, 2011: 4816-4821.
- [60] GRASA Ó G, BERNAL E, CASADO S, et al. Visual SLAM for handheld monocular endoscope [J]. IEEE Transactions on Medical Imaging, 2014, 33(1): 135-146.
- [61] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORBSLAM: a versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [62] 揭云飞,王峰,智凯旋,等. 视觉SLAM系统分析[J]. 电脑知识与技术, 2018, 14(19): 227-229.  
JIE Y F, WANG F, ZHI K X, et al. An system analysis of visual SLAM [J]. Computer Knowledge and Technology, 2018, 14(19): 227-229. (in Chinese)
- [63] LIN B, SUN Y, SANCHEZ J E, et al. Efficient vessel feature detection for endoscopic image analysis [J]. IEEE Transactions on Bio-Medical Engineering, 2015, 62(4): 1141-1150.
- [64] 杨滨豪. 基于单目SLAM的内窥镜定位和胃肠特征地图构建[D]. 南昌:南昌大学, 2016.  
YANG B H. Gastrointestinal endoscope localization and feature map reconstruction based on monocular SLAM [D]. Nanchang: Nanchang University, 2016. (in Chinese)
- [65] MAHMOUD N, COLLINS T, HOSTETTLER A, et al. Live tracking and dense reconstruction for handheld monocular endoscopy [J]. IEEE Transactions on Medical Imaging, 2019, 38(1): 79-89.
- [66] XIE C, YAO T, WANG J, et al. Endoscope localization and gastrointestinal feature map construction based on monocular SLAM technology [J]. Journal of Infection and Public Health, 2020, 13(9): 1314-1321.
- [67] PICCINELLI N, ROBERTI A, TAGLIABUE E, et al. Rigid 3D registration of pre-operative information for semi-autonomous surgery [C]//Proceedings of 2020 IEEE International Symposium on Medical Robotics. Washington D. C., USA: IEEE Press, 2020: 323-336.
- [68] LAMARRCA J, PARASHAR S, BARTOLI A, et al. DefSLAM: tracking and mapping of deforming scenes from monocular sequences [J]. IEEE Transactions on Robotics, 2020, 58(1): 1-13.
- [69] 林育斌. 基于计算机视觉的三维重建技术研究[D]. 天津: 天津大学, 2010.  
LIN Y B. Research of 3D reconstruction based on computer vision [D]. Tianjin: Tianjin University, 2010. (in Chinese)
- [70] 向龙海. 基于特征的软组织3D视觉重建与跟踪[D]. 成都: 电子科技大学, 2019.  
XIANG L H. Feature based 3D visual reconstruction and tracking of soft tissue [D]. Chengdu: University of Electronic Science and Technology of China, 2019. (in Chinese)
- [71] MORAVEC H P. Rover visual obstacle avoidance [C]//Proceedings of IEEE IJCAI'81. Washington D. C., USA: IEEE Press, 1981: 785-790.
- [72] HARRIS C, STEPHENS M. A combined corner and edge detector [EB/OL]. [2021-04-01]. <http://citeseerx.ist.psu.edu/>.
- [73] HAN J H, POSTON T T. Chord-to-point distance accumulation and planar curvature: a new approach to discrete curvature [J]. Pattern Recognition Letters, 2001, 22(10): 1133-1144.
- [74] 赵晓芳, 林土胜. 视网膜血管图像特征点自动提取和分类 [J]. 计算机工程与应用, 2011, 47(8): 14-17.  
ZHAO X F, LIN T S. Automatic detection and classification of feature points from retinal fundus images [J]. Computer Engineering and Applications, 2011, 47(8): 14-17. (in Chinese)
- [75] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [76] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359.
- [77] 卢鑫. 视网膜血管识别技术与算法实现 [D]. 成都: 电子科技大学, 2017.  
LU X. Research and implementation on retinal vessel recognition technology [D]. Chengdu: University of Electronic Science and Technology of China, 2017. (in Chinese)
- [78] CALONDER M, LEPETIT V, ÖZUYSAL M, et al. BRIEF: computing a local binary descriptor very fast [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(7): 1281-1298.
- [79] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF [C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2011: 6-13.
- [80] 赵晓芳. 视网膜血管图像处理的若干关键问题研究 [D]. 广州: 华南理工大学, 2011.  
ZHAO X F. Research on some key problems in processing of retinal vessel image [D]. Guangzhou: South China University of Technology, 2011. (in Chinese)
- [81] LIN B, YU S, SANCHEZ J, et al. Vesselness based feature extraction for endoscopic image analysis [C]//Proceedings of the 11th IEEE International Symposium on Biomedical Imaging. Washington D. C., USA: IEEE Press, 2014: 581-597.
- [82] LIN B X. Visual SLAM and surface reconstruction for abdominal minimally invasive surgery [D]. Tampa, USA: University of South Florida, 2015.
- [83] ZHANG T, WU K Z, SONG J W, et al. Convergence and consistency analysis for a 3D invariant-EKF SLAM [J]. IEEE Robotics & Automation Letters, 2017, 2(2): 1158-1169.