

基于结构光和 CT 的背部点云配准算法

沈春梅, 刘凡*, 朱佳乐

河海大学计算机与信息学院, 江苏 南京 210098

摘要 针对现有深度学习算法在进行医学图像配准时出现的模型对噪声敏感、精度较低等问题, 提出一种基于结构光和 CT 图像的点云配准算法, 旨在提高配准精度、算法鲁棒性的同时, 避免图像采集过程中 X 射线带来的辐射。首先, 向术中患者体表投射编码后的结构光, 并使用主成分分析方法获取主轴矫正术后患者的体表点云; 然后, 对术前患者进行 CT 扫描并进行三维重建及采样获取术前患者的体表点云。最后, 构建一种基于特征重用和注意力机制的动态图卷积网络模型结合图像迭代配准算法进行人体背部点云配准。整个过程较好地融合了两种模态的信息, 且具有无辐射、精度较高、用时较短等优点。

关键词 多模态点云配准; 注意力机制; 多尺度特征融合; 主轴矫正

中图分类号 TP391.4 **文献标志码** A

DOI: 10.3788/LOP231456

Back Point Cloud Registration Algorithm Based on Structured Light and CT

Shen Chunmei, Liu Fan*, Zhu Jiale

School of Computer and Information, Hohai University, Nanjing 210098, Jiangsu, China

Abstract. A point cloud registration algorithm based on structured light and CT images is proposed to address model sensitivity to noise and low accuracy issues in medical image registration using existing deep learning algorithms. The primary aim is to improve registration accuracy and algorithm robustness whilst avoiding radiation caused by X-rays during image acquisition. First, project encoded structured light onto the body surface of an intraoperative patient, and principal component analysis is used to obtain the point cloud for the body surface after spindle correction. Second, CT scans are performed on preoperative patients and three-dimensional reconstruction and sampling are conducted to obtain their body surface point clouds. Finally, a dynamic graph convolutional network model based on feature reuse and attention mechanisms is constructed, combined with an image iterative registration algorithm for point cloud registration on the back of the human body. The entire process effectively integrates the information from the two modes, with the advantages of no radiation, high accuracy, and short time consumption.

Key words multi-modal point cloud registration; attention mechanism; multi-scale feature fusion; spindle correction

1 引言

点云配准算法的典型代表是基于迭代最近点的配准法^[1], 其通过逐点遍历的方式进行计算, 但计算量大、时间长、对噪声敏感且在点云空间位置相差过大、角度旋转过大的情况下有着极高的配准失败率。因此, 有学者进行了一系列的改进^[2-4], 但基于迭代最近点改进的点云配准算法大都依赖于良好的初始位姿, 当存在噪声及异常值时, 效果较差。

近年来, 随着深度学习的快速发展, 点云配准在三维重建、逆向工程、医疗影像^[5-8]等领域得到广泛应用。为了充分利用不同成像技术^[9-10]的优势, 临床上通常需要采用多模态医学图像配准方法, 以实现不同模态数据之间相互补充和增强从而提高配准的精度。2D 与 3D 医学图像配准^[11-13]是医学图像配准中常见的任务之一, 例如: 李文举等^[14]针对 2D-3D 医学图像配准效率低等问题, 提出一种融合注意力机制的 ResNet^[15], 实现了 X-ray 与 CT 图像的多模态医学

收稿日期: 2023-06-05; 修回日期: 2023-07-06; 录用日期: 2023-08-08; 网络首发日期: 2023-08-21

基金项目: 国家自然科学基金(62372155)、装备预研教育部联合基金(8091B032157)

通信作者: *fanliu@hhu.edu.cn

图像配准;此外,李文举等^[16]针对现有配准算法精度较低、耗时较多等问题,提出一种训练-推理解耦架构,实现了 X-ray 和 CT 图像的配准。然而,X-ray 图像只提供目标的二维投影信息,且增加了医生和患者被辐射的风险。多模态的三维医学图像配准^[17]可以提供更全面、准确的病情信息,有助于病灶分析、手术规划和治疗监测等,核磁共振(MRI)图像可提供丰富的软组织对比度,CT 图像提供较好的骨骼结构显示,两者的配准可以获得更全面的解剖信息。Blendowski 等^[18]提出一种弱监督多模态图像配准方法,实现了 CT 和 MRI 图像的多模态对齐;Cheng 等^[19]通过二元分类器及自动编码器进行头部 CT 与 MRI 图像块之间的配准,提高了算法的准确率和鲁棒性。MRI 图像获取成本较高,而结构光点云获取方式具有非接触、扫描范围大、无辐射、成本较低等优点,通过结构光与 CT 图像的配准,可以将二者的三维信息对齐,实现更精确的空间匹配,但目前基于结构光进行医学图像配准的研究较少。雷超等^[20]为结构光相机点云和 CT 图像配准找到了一种最优的算法组合,提出一种快速、廉价且精准的算法,但没有研究算法的鲁棒性。

综上所述,本文提出一种基于结构光及 CT 的背部点云配准算法。通过基于主成分分析的点云预处理模块,有效解决术前 CT 图像点云与术中结构光点云之间旋转平移差异过大的问题;使用基于特征重用和注意力机制的编解码模块对齐点云特征,在保证配准时间的前提下,进一步提高两种模态点云的配准精度

及鲁棒性。

2 结构光与 CT 点云获取

2.1 结构光点云获取

结构光系统的标定主要用来获取摄像机和投影仪的内参、外参及镜头的畸变系数,标定的精度决定着采集到的患者点云精度。该部分主要包括摄像机、投影仪标定。投影仪是主动发出光源的设备,与摄像机被动接收光线正好相反,投影仪无法像摄像机那样拍摄得到像素坐标值。因此,可以将投影仪视作逆相机,作用是投射经过调制的结构光图案到被测物体上。投影仪将这些结构光图案投影到被测物体上后,摄像机会捕捉到物体表面反射的光,形成带有结构的图像。通过分析这些结构图案的变化,可以推断出物体的形状和表面细节,从而生成点云数据。对于摄像机的标定,采集 20 个不同角度的棋盘格标定板图像,检测特征点获取像素坐标值,求解内外参初值并估计畸变系数,通过最大似然估计优化参数,并计算重投影误差,若小于 0.2 pixel 则输出摄像机参数。使用张正友标定法进行摄像机的标定。首先,将棋盘格标定模板看作世界坐标系中的 $Z=0$ 平面,通过拍摄多角度的棋盘格图像计算角点约束信息获得像素坐标值;然后,通过单应性矩阵求解内外参初始值、最小二乘法估计畸变系数、最大似然估计优化参数;最后,计算重投影误差,替换误差较大的棋盘格图像,迭代上述过程,直至重投影误差小于设置的阈值。摄像机标定算法流程图如图 1 所示。

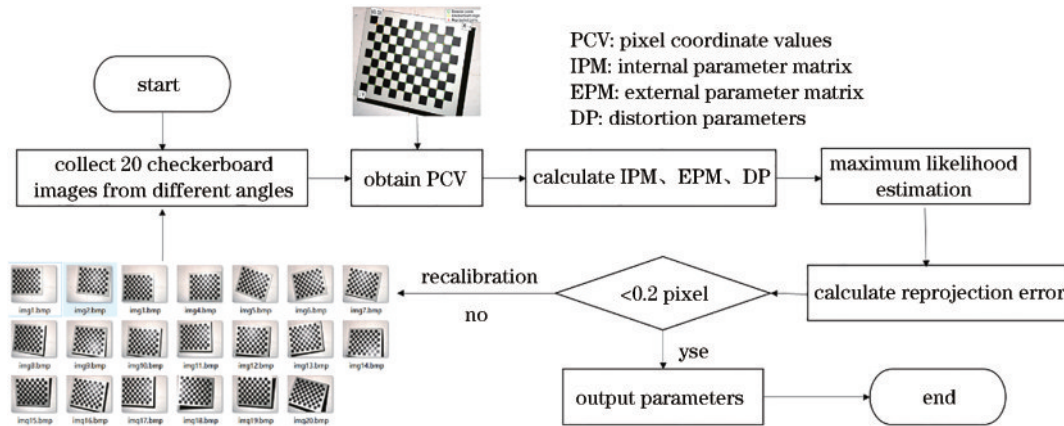


图 1 摄像机标定流程图

Fig. 1 Flow chart of the camera calibration

对于投影仪的标定,通过采集纵横互补格雷码图像,解码得到解码值,使用局部单应性矩阵求解亚像素值,计算重投影误差,若小于 0.2 pixel,则输出投影仪参数。标定流程如图 2 所示。

在投影仪标定过程中,由于解码后的像素精度只有像素级别,为了提高解码的精度,使用局部单应性矩阵对标定过程中投影仪像素的求取进行改进,从而得到更精确的亚像素级别的像素坐标值。具体做法

是通过计算摄像机下角点周围 $5 \text{ pixel} \times 7 \text{ pixel}$ 区域内的像素值,使得投影仪下的像素值与使用局部单应性矩阵变换后的摄像机角点矩形区域的像素值差值最小化,从而找到这个局部单应性矩阵 \hat{H} ,如式(1)所示:

$$\hat{H} = \underset{H}{\operatorname{argmin}} \sum_{\forall p} \|q - Hp\|, \quad (1)$$

式中: $H \in \mathbb{R}^{3 \times 3}$; p 代表摄像机矩形区域内像素点集; q

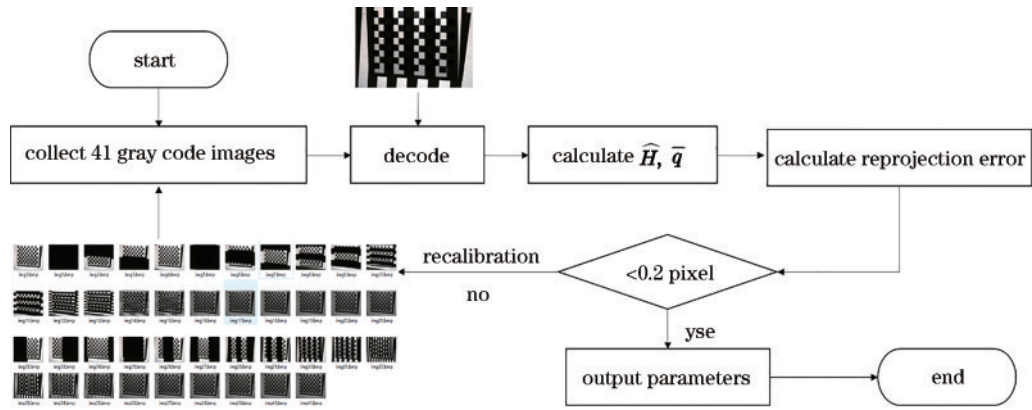


图 2 投影仪标定流程图

Fig. 2 Flow chart of the Projector calibration

代表解码后的对应投影仪像素点集。

$$\bar{q} = \hat{H} \cdot \bar{p} \quad (2)$$

对于位于区域中间位置的目标角点 \bar{p} , 通过局部单应性矩阵 \hat{H} , 得到最终投影仪角点像素坐标 \bar{q} :

最终完成整个系统的标定, 得到的标定结果如表 1 所示。

表 1 标定结果

Table 1 Calibration results

Calibration parameter	Video camera	Projector
Internal parameter matrix	$\begin{bmatrix} 2218.104 & 0 & 640.718 \\ 0 & 2218.241 & 460.704 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1112.843 & 0 & 523.886 \\ 0 & 1113.644 & 216.825 \\ 0 & 0 & 1 \end{bmatrix}$
Distortion coefficient	$(-0.01625, 0.35658, -0.00107, 0.00001, -2.76695)$	$(0.06391, -0.16930, -0.00063, -0.00039, 0.02259)$
Relative rotation matrix	$\begin{bmatrix} 0.96787 & 0.04739 & -0.24696 \\ -0.04239 & 0.99877 & 0.02553 \\ 0.24787 & -0.01424 & 0.96869 \end{bmatrix}$	
Relative translation matrix	$[221.190 \quad 172.222 \quad -87.051]$	
Error of reprojection / pixel	0.12269	0.14582

为了进一步获取被测物体的点云, 首先向被测物体投射时序横纵正反互补的格雷码图像, 归一化解码得到每个像素点的解码值, 然后使用三角测量原理计算出世界坐标系下被测物体表面点的三维坐标, 最后得到被测物体表面的三维点云。以人体模型和脊椎模型作为被测物体, 其采集过程及获取的点云结果如图 3 所示。

在点云采集的过程中, 由于摄像机和投影仪存在一些标定误差, 格雷码投影序列在编码过程中存在不可测点, 即使采用了正反互补格雷码的改进编码方式, 依旧会受到环境光照、物体反光等干扰因素的影响, 导致采集到的点云产生噪声数据, 如图 4 中圆圈框选出来的点, 这部分点远离主体点云, 因此也被称作离群点。这部分噪声离群点在模型提取局部特征时会产生干扰, 从而降低配准精度, 所以需要选用合适的方式去除这些噪声离群点。使用统计滤波作为去除离群噪声的方式, 设置临近点集数量为 50, 距离方差阈值

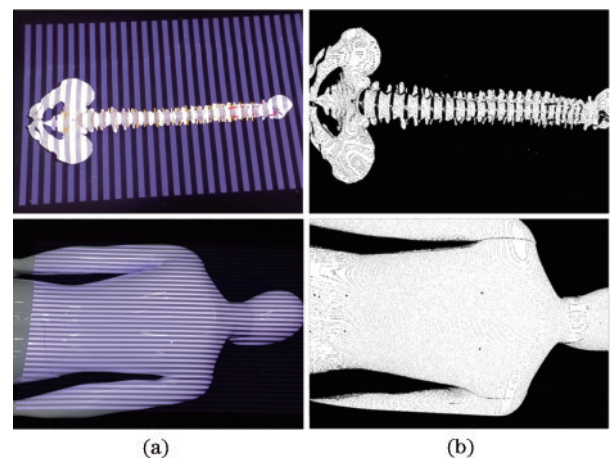


图 3 点云采集。(a) 脊椎模型和人体模型; (b) 脊椎模型和人体模型点云

Fig. 3 Collection of point cloud. (a) Spine model and human body model; (b) vertebral and anthropomorphic point clouds

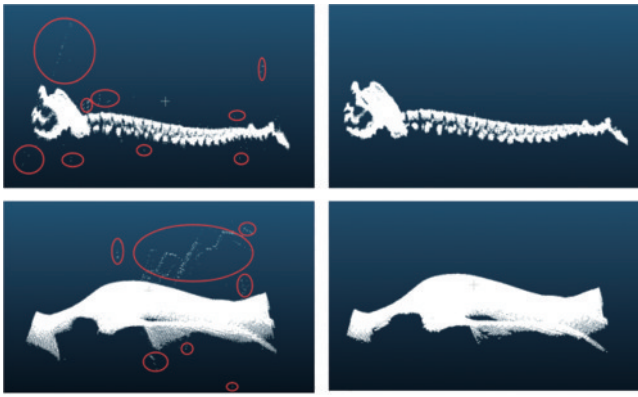


图 4 去噪前后的对比

Fig. 4 Comparison before and after denoising

为 1, 从而将离群噪点准确去除, 去除噪点前后对比图如图 4 所示。

2.2 CT 点云获取

CT 成像的结果只有主视图即横截面图, 对于医生来说难以判断患者疾病病灶的情况, 所以一般都会重建出三视图包括冠状面、矢状面以及参考面图。通过 SIEMENS 的 EMOTION 16 层螺旋 CT 扫描机采集层厚为 5 mm、螺距为 1.05、分辨率为 512×512 的胸腹部平扫 CT 图像序列, 并使用 Mimics 软件建立其三视图, 如图 5 所示。

从 CT 三视图可以看到较为全面的人体信息, 但

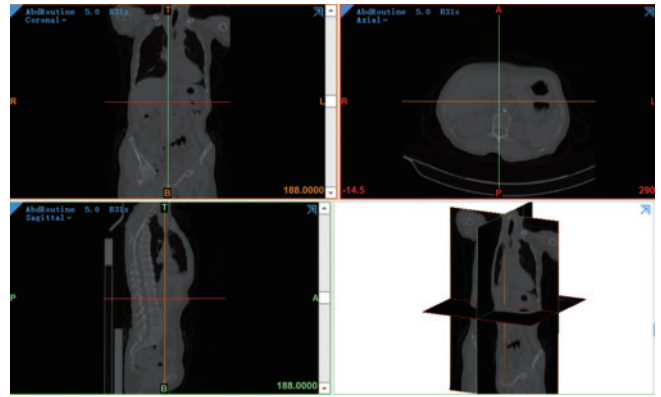


图 5 人体 CT 三视图与参考面图

Fig. 5 Three views and reference view of human CT

对于病灶的详细情况却难以观察到, 此时需要对 CT 进行三维重建, 从而获得人体部位的解剖结构信息以更好地进行手术的规划和模拟, 提高手术的成功率。使用 Mimics 软件对人体的体表以及骨骼进行三维重建。通过设置 CT 值(单位为 Hu) 大于 -200 Hu、小于 50 Hu, 将人体表面及骨骼的三维模型进行重建和分离, 并去除内脏等的干扰, 得到如图 6 所示的三维模型图。将术前的 CT 图像空间与术中患者的结构光图像进行配准, 对重建后的 CT 体表的三维模型进行采样。使用 point cloud library (PCL) 采样 CT 人体体表模型的点云, 如图 7 所示。

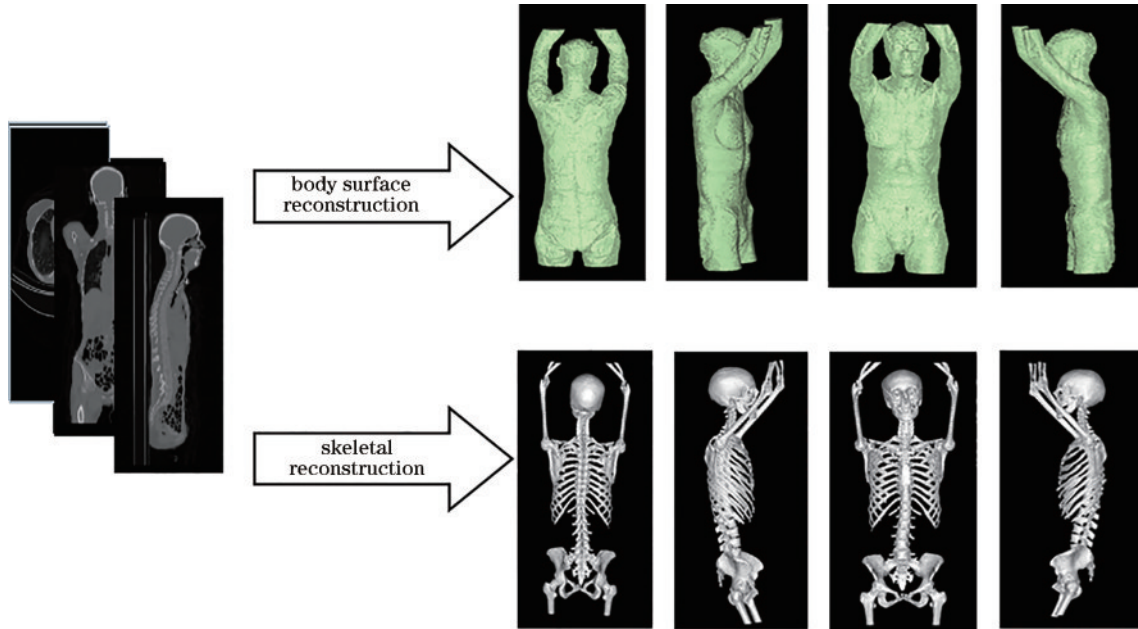


图 6 CT 体表和骨骼重建三维模型图

Fig. 6 3D model of CT body surface and bone reconstruction

3 点云配准

得到两种模态的人体体表点云后, 应用相同的点云分割算法来识别和提取人体背部区域。首先通过结构光系统获取患者背部在实际空间中的点云 $P =$

$\{p_1, \dots, p_i, \dots, p_N\} \in \mathbb{R}^{N \times 3}$, 通过 CT 等医学影像获取患者在图像空间中的背部点云 $Q = \{q_1, \dots, q_i, \dots, q_N\} \in \mathbb{R}^{N \times 3}$, 其中, p_i, q_i 分别表示这两个点云中的点坐标, 点的数量为 N , 3 表示该点的三维坐标。将这两种模态的配准看作刚体配准问题, 即找到一个最佳的刚性变化

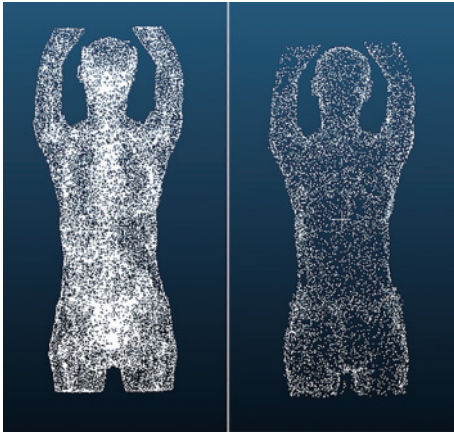


图 7 采样前后的 CT 人体体表模型点云

Fig. 7 Point clouds of CT human body surface model before and after sampling

矩阵 \mathbf{G} (旋转矩阵 $\mathbf{R} \in SO(3)$, 平移向量 $\mathbf{t} \in \mathbb{R}^3$), 使得变换前后点云对应点距离最小, 如式(3)所示。

$$\operatorname{argmin}_{\mathbf{G}} \left\| \sum_{k=1}^K \mathbf{p}_k - (\mathbf{R}\mathbf{q}_k + \mathbf{t}) \right\|_2, \quad (3)$$

式中: \mathbf{p}_k 与 \mathbf{q}_k 表示待配准的两个点云的对应点; K 表示对应点的数量。

3.1 基于主轴矫正的主成分分析模块

由于术前和术中的点云由不同的设备获取, 是两种不同模态的数据, 如果直接进行配准, 在两点云初始位姿差异较大的情况下, 效果比较差。因此, 采用主成分分析方法 (PCA) 进行点云初始位姿的调整, 解决主轴反向的问题。该模块的工作流程如下:

1) 初始化结构光点云 \mathbf{S} 和 CT 点云 \mathbf{Q} 作为输入, 点的数量为 N , 计算两个点云的质心 $\bar{\mathbf{s}}$ 与 $\bar{\mathbf{q}}$, 如式(4)、(5)所示:

$$\bar{\mathbf{s}} = \frac{1}{N} \sum_{i=1}^N \mathbf{s}_i, \quad (4)$$

$$\bar{\mathbf{q}} = \frac{1}{N} \sum_{i=1}^N \mathbf{q}_i. \quad (5)$$

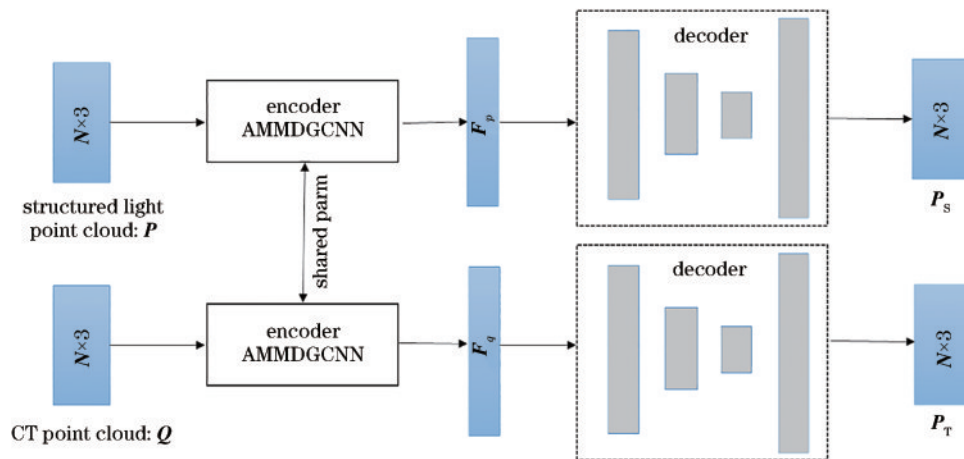


图 8 编解码模块

Fig. 8 Encoder-decoder module

2) 计算点云 \mathbf{S} 和 \mathbf{Q} 的协方差矩阵 \mathbf{C}_S 和 \mathbf{C}_Q , 如式(6)、(7)所示:

$$\mathbf{C}_S = \frac{1}{N} \sum_{i=1}^N (\mathbf{s}_i - \bar{\mathbf{s}})(\mathbf{s}_i - \bar{\mathbf{s}})^T, \quad (6)$$

$$\mathbf{C}_Q = \frac{1}{N} \sum_{i=1}^N (\mathbf{q}_i - \bar{\mathbf{q}})(\mathbf{q}_i - \bar{\mathbf{q}})^T. \quad (7)$$

3) 对协方差矩阵 \mathbf{C}_S 和 \mathbf{C}_Q 进行奇异值分解得到协方差矩阵对应的特征向量和特征值, 并从大到小排序特征值, 取前 K 特征值的特征向量作为表示当前数据的主成分向量, 组成矩阵 \mathbf{M}_S 和 \mathbf{M}_Q , 由此计算出两片点云之间的旋转矩阵 \mathbf{R} 和平移向量 \mathbf{t} , 如式(8)、(9)所示:

$$\mathbf{R} = \mathbf{M}_S \mathbf{M}_Q^T, \quad (8)$$

$$\mathbf{t} = \bar{\mathbf{q}} - \mathbf{R} \cdot \bar{\mathbf{s}}. \quad (9)$$

4) 通过 PCA 得到变换矩阵, 但是由于协方差矩阵的特征向量有正负两个取值, 因此使用矫正矩阵进行 PCA 主轴方向的矫正。通过计算变换后的结构光点云与 CT 点云在 xoy 、 xoz 、 zoy 坐标系下的投影重叠度, 找到投影重叠度最大的变换方式所对应的矫正矩阵 \mathbf{W} , 如式(10)所示。

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \end{bmatrix}, \quad (10)$$

式中: w_1, w_2, w_3 取值只有 -1 或 1 , 1 表示与主轴方向同向, -1 表示与主轴方向反向。

通过这种方式解决了使用 PCA 变换时出现的主轴反向对齐问题, 从而正确调整两片点云的初始位置, 得到矫正后的结构光点云 \mathbf{P} 和预估的变换矩阵 \mathbf{G}_{est} , 完成结构光点云和 CT 点云的初始配准。

3.2 点云配准

3.2.1 编解码模块

编解码模块如图 8 所示, 使用并行的编码器 multi-modal dynamic graph CNN network with attention mechanism and multi-scale feature fusion (AMMDGCNN) 对结构光点云与 CT 点云分别进行特征提取, 两个分支

共享权重不但减少了模型整体的参数量,而且可以促进结构光点云和 CT 点云两种不同模态信息的融合,从而提高模型的配准效果。每个编码器分支的输入维度是 $N \times 3$,假设经过 PCA 模块后的结构光点云和 CT 点云分别为 \mathbf{P} 和 \mathbf{Q} ,经过 AMMDGCNN 编码器后得到特征 \mathbf{F}_p 和 \mathbf{F}_q ,其维度为 D 。为了更好保证提取的特征向量能够有效表征几何结构且对于旋转和平移敏感,在每个编码器分支后面连接 1 个由 Leaky ReLU 激活的 4 个全连接层组成的解码器网络,解码器输入和输出维度分别为 $(D, D), (D, D/2), (D, D/4), (D/4, N \times 3)$,最终解码器模块从特征中恢复点云 \mathbf{P}_s 和 \mathbf{P}_T 。并将恢复后的结构光点云 \mathbf{P}_s 和 CT 点云 \mathbf{P}_T 与初始输入的点云之间的倒角距离作为损失函数训练编码器,这种方式有助于训练编码器模块在意识到旋转差异的情况下学习并生成有意义的特征,可以更好地训练编码器对旋转平移敏感的特征的提取能力,促进两个模态数据之间的融合和交互,以提高配准的准确性和鲁棒性。

所提 AMMDGCNN 编码器如图 9 所示,在 DGCNN^[21]的基础上加入多尺度特征融合及注意力机制。将多尺度特征融合引入 DGCNN 中,即将不同层动态图卷积的特征按通道维度进行拼接,不同尺度的特征可以捕捉目标的多个层次特征,如边缘、角点等,在增加一定

模型参数的情况下,提供更全面的目标特征描述。使用注意力机制对边缘卷积模块进行改进,基于索引的并行注意力机制包括 3 部分:第 1 部分是基于索引的空间注意力机制,将进行边缘卷积后的特征图 \mathbf{F} 作为输入,此时特征图 \mathbf{F} 的维度为 $f \times N \times k, f$ 表示通道的维度, N 表示点云中的 N 个点, K 表示 N 个点通过 KNN 构成邻域图的邻接点。沿着通道维度分别进行平均池化和最大池化得到维度为 $1 \times N \times K$ 的两个特征图 \mathbf{A} 和 \mathbf{B} ,最大池化时记录下最大值的索引位置。然后将特征图 \mathbf{A} 和特征图 \mathbf{B} 进行拼接并进行两次 3×3 的卷积,形成 $2 \times N \times K$ 的特征图,使得经过平均池化和最大池化索引操作得到的特征信息交互,再按通道维度将特征图 \mathbf{F}_1 拆分并还原特征图 \mathbf{B}_1 最大池化索引时的位置信息。第 2 部分是通道注意力机制,先进行全局平均池化再进行两次 3×3 卷积。第 3 部分是多级特征融合,将基于索引的空间注意力机制得到的特征图 \mathbf{A}_1 和 \mathbf{B}_1 点乘,再使用特征图 \mathbf{D} 进行加权及 Sigmoid 函数,最终形成 1 个与原始特征图尺寸相同的权重特征图 \mathbf{P} 。基于索引的并行注意力机制在空间注意力部分使用最大池化操作记录最大值的索引位置,一方面保留了特征图的底层信息,另一方面能有效抑制数据中的干扰,索引空间注意力与通道注意力相结合提高了边缘卷积模块提取特征的能力。

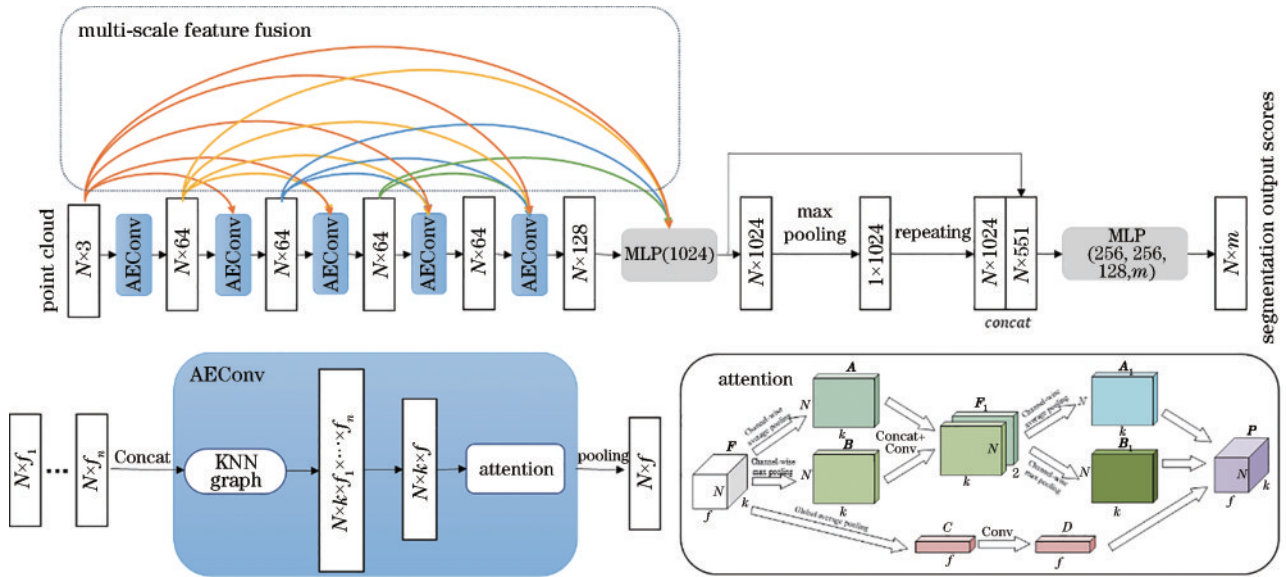


图 9 AMMDGCNN 结构

Fig. 9 AMMDGCNN structure

3.2.2 点云配准

模块配准的具体流程如下:假设输入的结构光点云 $\mathbf{P}_s \in \mathbb{R}^{N_1 \times 3}$ 和 CT 点云 $\mathbf{P}_T \in \mathbb{R}^{N_2 \times 3}$,其中, N_1 和 N_2 代表点云中点的数量。配准的目的是找到刚体变换参数 $\mathbf{G} \in SE(3)$ (旋转矩阵 $\mathbf{R} \in SO(3)$, 平移向量 $\mathbf{t} \in \mathbb{R}^3$) 使得结构光点云和 CT 点云的特征差异最小化,获得最佳配准,如式(11)所示:

$$\operatorname{argmin}_{\mathbf{R} \in SO(3), \mathbf{t} \in \mathbb{R}^3} \|\phi(\mathbf{P}_T) - \phi(\mathbf{R}\mathbf{P}_s + \mathbf{t})\|, \quad (11)$$

式中: ϕ 表示 AMMDGCNN 编码器,用于点云特征提取, $\phi: \mathbb{R}^{N \times 3} \rightarrow \mathbb{R}^K, K$ 表示提取特征的维数,本研究使用的维度为 1024。

二维图像中的图像迭代配准算法 LK^[22]需要多次迭代获取图像的最佳配准参数,在三维点云中这种方式显然不适用,因为在点云中迭代计算雅可比矩阵会

产生极高的计算成本。因此,使用逆合成(IC)算法将模板点云逆变换到源点云,此时只需计算一次雅可比矩阵,大大降低了计算成本,如式(12)所示:

$$\operatorname{argmin}_{\mathbf{G}^{-1}} \left\| \phi(\mathbf{P}_S) - \phi(\mathbf{G}^{-1} \mathbf{P}_T) \right\|_2^2, \quad (12)$$

那么雅可比矩阵可以表示为

$$\mathbf{J} = \frac{\partial \phi[\mathbf{G}^{-1}(\boldsymbol{\xi}) \cdot \mathbf{P}_T]}{\partial \boldsymbol{\xi}^T} \in \mathbb{R}^{K \times 6}, \quad (13)$$

式中: $\mathbf{G}^{-1} = \exp\left(-\sum_i \xi_i T_i\right)$, $\boldsymbol{\xi}$ 表示对应的旋转平移参数,是一个6维向量, T 表示指数映射生成器。

在具体求解雅可比矩阵的时候,通过给定一个无穷小扰动,使用数值的有限差分来近似雅可比矩阵的每一列。得到雅可比矩阵后,求出矩阵对应的旋转平移参数 $\boldsymbol{\xi}$, 其中, \mathbf{J}^+ 是雅可比矩阵 \mathbf{J} 的 Moore-Penrose 逆矩阵:

$$\boldsymbol{\xi} = \mathbf{J}^+ [\phi(\mathbf{P}_S) - \phi(\mathbf{P}_T)]. \quad (14)$$

通过配准算法中的迭代计算,得到新的旋转平移参数 $\boldsymbol{\xi}$, 从而得到变换矩阵的变化 $\Delta \mathbf{G}$, 对源点云进行更新。通过不断更新变化,当达到迭代的次数或者变化量小于阈值的时候,得到最终的预测变换矩阵 \mathbf{G}_{est} , 如式(15)所示:

$$\mathbf{G}_{\text{est}} = \Delta \mathbf{G}_n \cdots \Delta \mathbf{G}_i \cdots \Delta \mathbf{G}_0. \quad (15)$$

4 实验结果与分析

4.1 实验设置

系统硬件环境与软件安装包如表2所示。采集到的结构光和CT背部点云数据较少,以其中880个点云作为训练集,220个点云作为测试集。本实验的数据使用公开的 ModelNet40^[23] 数据集作为预训练数据集,在 ModelNet40 数据集训练的基础上,使用采集的背部点云数据进行迁移学习。配准网络训练过程中对编解码模块采用 Adam 优化器,学习率为 0.001,训练的 epoch 为 200,最大迭代次数是 10, batch size 为 8,对输入点云进行 2048 个点的采样。

4.2 评价指标

对于点云配准效果的评价,除了通过主观的视觉效果外,还可以通过对比真实变换矩阵和预估变换矩阵之间的均方误差 (E_{MSE})、均方根误差 (E_{RMSE}) 以及绝

表2 系统硬件环境与软件安装包

Table 2 System hardware environment and software installation package

Hardware and software name	Configuration
Video camera	MER-125-30UC 1292×964
Projector	DLP LightCrafter 4500
Host hardware configuration	Intel(R) Core i7-6700HQ, 24 GB
Operating system	Windows10
Programming language	Python3.8
Integrated development environment	Pycharm 2020.2.3 Pro, Qt Designer
Dependency package	Open3d, PyQt, PyTorch, OpenCV, PCL
Database	MySQL5.7

对平均误差 (E_{MAE}) 来进行客观评价。

对于旋转变换,用 \mathbf{R}_g 表示真实的旋转矩阵、 \mathbf{R}_e 表示估计的旋转矩阵,其均方误差如(16)所示,均方根误差如式(17)所示,绝对平均误差如式(18)所示:

$$E_{\text{MSE}}(\mathbf{R}) = \frac{1}{M} \sum_{i=1}^m (\mathbf{R}_g - \mathbf{R}_e)^2, \quad (16)$$

$$E_{\text{RMSE}}(\mathbf{R}) = \sqrt{E_{\text{MSE}}(\mathbf{R})}, \quad (17)$$

$$E_{\text{MAE}}(\mathbf{R}) = \frac{1}{M} \sum_{i=1}^m |\mathbf{R}_g - \mathbf{R}_e|. \quad (18)$$

同理,对于平移矩阵,用 \mathbf{T}_g 表示真实的旋转矩阵、 \mathbf{T}_e 表示估计的旋转矩阵,其误差也可以通过均方误差、均方根误差、绝对平均误差表示,如式(19)~(21)所示:

$$E_{\text{MSE}}(\mathbf{T}) = \frac{1}{M} \sum_{i=1}^m (\mathbf{T}_g - \mathbf{T}_e)^2, \quad (19)$$

$$E_{\text{RMSE}}(\mathbf{T}) = \sqrt{E_{\text{MSE}}(\mathbf{T})}, \quad (20)$$

$$E_{\text{MAE}}(\mathbf{T}) = \frac{1}{M} \sum_{i=1}^m |\mathbf{T}_g - \mathbf{T}_e|. \quad (21)$$

4.3 实验结果

4.3.1 无噪声点云配准实验

这部分实验主要测试无噪声点云在给定随机旋转角度 $[0^\circ, 45^\circ]$ 及平移距离 $[0, 0.8]$ 时的配准误差。配准误差越低代表模型的配准性能越好,从表3可以看出,在没有噪声的情况下,所提点云配准算法的旋转误差和平移误差最小,证明了其在无噪声点云配准中的有效性。

表3 无噪声点云在随机旋转条件下的配准误差

Table 3 Experimental results of point cloud registration without noise

Method	Rotation			Translation		
	$E_{\text{MSE}}(\mathbf{R})$	$E_{\text{RMSE}}(\mathbf{R})$	$E_{\text{MAE}}(\mathbf{R})$	$E_{\text{MSE}}(\mathbf{T})$	$E_{\text{RMSE}}(\mathbf{T})$	$E_{\text{MAE}}(\mathbf{T})$
ICP ^[1]	775.8915	27.8548	5.5384	0.0188	0.1369	0.0299
GICP ^[3]	249.3943	15.7922	1.6934	0.0074	0.0860	0.0082
PointNetLK ^[24]	6.9356	2.6335	0.2178	0.0002	0.0141	0.0009
Proposed method	0.1147	0.3386	0.1629	1.57×10^{-6}	0.0013	0.0001

4.3.2 有噪声点云配准实验

结构光和 CT 采集点云会出现噪声,且两者点云密度不一致,因此给点云配准带来了困难。为了验证所提算法对于噪声以及不同密度点云配准的有效

性,为点云中的每一个点添加均值为 0、方差为 0.01 的高斯噪声。同时给定随机的旋转角度 $[0^\circ, 45^\circ]$ 及平移距离 $[0, 0.8]$ 。在测试集上的测试结果如表 4 所示。

表 4 有噪声点云在随机旋转条件下的配准误差

Table 4 Experimental results of point cloud registration with noise

Method	Rotation			Translation		
	$E_{MSE}(R)$	$E_{RMSE}(R)$	$E_{MAE}(R)$	$E_{MSE}(T)$	$E_{RMSE}(T)$	$E_{MAE}(T)$
ICP ^[1]	809.3575	28.4492	6.9102	0.0279	0.1673	0.0489
GICP ^[3]	436.7152	20.8977	3.5435	0.0101	0.1006	0.0171
PointNetLK ^[24]	232.4861	15.2475	6.1024	0.0065	0.0812	0.0452
Proposed method	4.5131	2.1244	1.8336	0.0004	0.0206	0.0165

从表 3、表 4 可以看出,GICP 作为 ICP 的改进算法,相比较 ICP 降低了误差,不过由于其本质还是对点集中对应点进行变换矩阵的求解,这在有噪声的点云配准中效果会变差,因此 GICP 的配准误差虽然小于 ICP,但仍然大于所提算法。在有噪声的情况下,PointNetLK 得到的特征是不准确的,而所提算法与其余 3 种算法相比有着显著的优势,体现了其对噪声的鲁棒性。

4.3.3 人体模型背部点云配准实验

这部分实验主要展示同一个人体模型背部的结构光扫描点云与 CT 扫描点云的配准精度和效果。

不同算法在人体模型背部点云配准的可视化结果如图 10 所示,红色的是结构光扫描的点云,蓝色的是 CT 提取的点云。从图 10 可以看出,ICP 算法和 PointNetLK 算法无法处理这种大角度旋转的情况,配准误差较大,GICP 算法在配准的过程中出现了反向配准的情况,而所提算法通过主轴矫正改进 PCA 配准,使用主轴矫正的方式,避免了反向配准的情况。所提基于注意力机制及多尺度特征融合的编码器提取的特征更加准确有效,因此取得了几种算法中最好的配准效果。

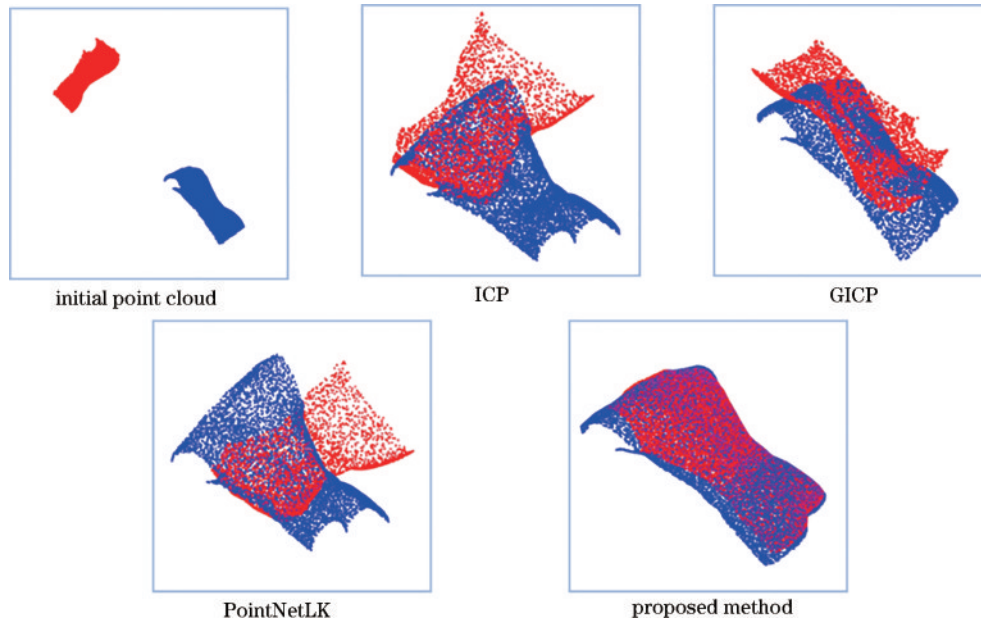


图 10 人体背部点云配准结果图

Fig. 10 Point cloud registration results of human back

为了更好地评估算法的有效性,对比每种算法的配准精度以及配准所需要的时间。这部分的配准精度评价指标为两个空间中的对应点之间的距离平均值,即均方根距离误差(RMS),单位为 mm,这个值越小代表配准的误差越小,精度越高。同时计算每一个算法配准的时间,记作 t ,单位为 s。人体模型背部点

云配准结果如表 5 所示。可以看出,除了所提算法,其余 4 种算法的误差都非常大,所提算法的均方根距离误差只有 1.45 mm。同时对比时间消耗可知,所提算法虽然不是用时最短的算法,但是也可以在 2 s 左右完成配准的任务,且达到较高的精度,因此具有一定的实用性。

表 5 人体背部点云配准结果表

Table 5 Registration results of point cloud on human back

Method	ICP ^[1]	GICP ^[3]	PointNetLK ^[24]	Proposed method
RMS/mm	368.83	401.33	287.74	1.45
t/s	6.414	0.471	2.045	2.062

5 结 论

针对医学图像配准精度较低及易受噪声干扰等问题,提出一种基于 AMMDGCNN 与 PCA 相结合的人体背部点云配准算法。首先,使用改进的 PCA 配准模块调整结构光点云的初始位置,解决主轴反向的问题。然后,设计 AMMDGCNN 作为特征提取器,引入解码器模块,使编码器提取到对旋转和平移更加敏感的特征,有利于两种模态的对齐。最后,利用 LK 算法进行点云配准。实验结果表明,所提基于结构光和 CT 的人体背部点云配准算法在无噪声及有噪声的情况下都取得了比其他算法更好的效果,实现了对患者术前和术中空间的匹配。在下一步工作中,将研究更好的配准算法应用于其他模态医学图像的配准。

致谢 感谢江苏高校“青蓝工程”资助,感谢河海大学朱永泰同志提供的指导和帮助。

参 考 文 献

- [1] Besl P J, McKay N D. A method for registration of 3-D shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239-256.
- [2] Granger S, Pennec X. Multi-scale EM-ICP: a fast and robust approach for surface registration[M]//Heyden A, Sparr G, Nielsen M, et al. Computer vision-ECCV 2002. Lecture notes in computer science. Heidelberg: Springer, 2002, 2353: 418-432.
- [3] Segal A, Haehnel D, Thrun S. Generalized-ICP[C]//Robotics: Science and Systems V, June 28-July 1, 2009, Seattle, USA. Robotics: Science and Systems Foundation, 2009.
- [4] Yang J L, Li H D, Campbell D, et al. Go-ICP: a globally optimal solution to 3D ICP point-set registration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(11): 2241-2254.
- [5] Li J, Deng Z Q, Shen N Y, et al. A fully automatic surgical registration method for percutaneous abdominal puncture surgical navigation[J]. Computers in Biology and Medicine, 2021, 136: 104663.
- [6] Shi W, Wang L P. High-accuracy digital volume correlation-based point cloud registration for 3D reconstruction[J]. Proceedings of SPIE, 2022, 12319: 1231907.
- [7] Park S, Kim S, Seo H. Study on representative parameters of reverse engineering for maintenance of ballasted tracks[J]. Applied Sciences, 2022, 12(12): 5973.
- [8] 杨泽鑫,叶勤,王旭飞,等.面向城市场景异源多时相点云的自动配准[J].中国激光,2023,50(10):1010004. Yang Z X, Ye Q, Wang X F, et al. Automated registration of cross-source and multi-temporal point clouds in urban areas[J]. Chinese Journal of Lasers, 2023, 50(10): 1010004.
- [9] 李奇明,任杰骥,裴晓寒,等.基于多模态数据协同的弱纹理表面高精度点云配准算法[J].光学学报,2022,42(8):0810001. Li Q M, Ren J J, Pei X H, et al. High-accuracy point cloud matching algorithm for weak-texture surface based on multi-modal data cooperation[J]. Acta Optica Sinica, 2022, 42(8): 0810001.
- [10] 刘磊,李元祥,倪润生,等.基于卷积与图神经网络的合成孔径雷达与可见光图像配准[J].光学学报,2022,42(24):2410002. Liu L, Li Y X, Ni R S, et al. Synthetic aperture radar and optical images registration based on convolutional and graph neural networks[J]. Acta Optica Sinica, 2022, 42(24): 2410002.
- [11] Liu S, Yang B, Wang Y, et al. 2D/3D multimode medical image registration based on normalized cross-correlation[J]. Applied Sciences, 2022, 12(6): 2828.
- [12] 陈向前,郭小青,周钢,等.基于深度学习的2D/3D医学图像配准研究[J].中国生物医学工程学报,2020,39(4):394-403. Chen X Q, Guo X Q, Zhou G, et al. 2D/3D medical image registration using convolutional neural network[J]. Chinese Journal of Biomedical Engineering, 2020, 39(4): 394-403.
- [13] Shabnam S, Diana P, Jennie M S, et al. An efficient hybrid method for 3D to 2D medical image registration [J]. International Journal of Computer Assisted Radiology and Surgery, 2022, 17(7): 1313-1320.
- [14] 李文举,孔德卿,曹国刚,等.基于注意力残差网络的跨模态医学图像配准[J].计算机仿真,2022,39(11):224-229. Li W J, Kong D Q, Cao G G, et al. Cross-modal medical image registration based on attention residual network[J]. Computer Simulation, 2022, 39(11): 224-229.
- [15] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [16] 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-inference decoupling architecture[J]. Laser & Optoelectronics Progress, 2022, 59(16): 1610015.
- [17] 程天宇,顾菊平,华亮,等.基于几何代数SURF的三维医学图像配准研究[J].仪器仪表学报,2022,43(9):227-237. Cheng T Y, Gu J P, Hua L, et al. Research on 3D medical image registration based on geometric algebra

- SURF[J]. Chinese Journal of Scientific Instrument, 2022, 43(9): 227-237.
- [18] 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.
- [19] 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.
- [20] 雷超, 贾于, 王宏, 等. 结构光相机点云图与 CT 图像配准算法对比研究[J]. 医疗卫生装备, 2022, 43(3): 8-13, 28. Lei C, Jia Y, Wang H, et al. Comparative study on registration algorithm between point cloud image and CT image of structured light camera[J]. Chinese Medical Equipment Journal, 2022, 43(3): 8-13, 28.
- [21] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): 146.
- [22] Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision[C]//Proceedings of the 7th international joint conference on Artificial intelligence-Volume 2, August 24-28, 1981, Vancouver, BC, Canada. New York: ACM Press, 1981: 674-679.
- [23] Wu Z R, Song S R, Khosla A, et al. 3D ShapeNets: a deep representation for volumetric shapes[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1912-1920.
- [24] Aoki Y, Goforth H, Srivatsan R A, et al. PointNetLK: robust & efficient point cloud registration using PointNet [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 7156-7165.