

光学学报

基于 Swin-Transformer 迭代展开的有限角 CT 图像重建用于 PTCT 成像

袁伟^{1,2}, 席雅睿^{1,2}, 谭川东^{1,2}, 刘川江^{1,2}, 朱国荣^{1,2}, 刘丰林^{1,2*}¹重庆大学 ICT 研究中心光电技术及系统教育部重点实验室, 重庆 400044;²重庆大学工业 CT 无损检测教育部工程研究中心, 重庆 400044

摘要 针对相对平行直线扫描 CT (PTCT) 图像重建存在的有限角伪影问题, 提出一种学习局部和非局部正则项的深度迭代展开方法。该方法将具有固定迭代次数的梯度下降算法迭代展开到神经网络, 利用具有坐标注意力 (CA) 机制的卷积模块和 Swin-Transformer 模块作为迭代模块交替级联部署, 构成端到端的深度重建网络。卷积模块学习局部正则化, 其中 CA 用于减少图像过平滑; Swin-Transformer 模块学习非局部正则化, 提高网络对图像细节的恢复能力; 在相邻模块间, 使用迭代连接 (IC) 增强模型提取深层特征的能力, 提高每次迭代的效率。通过消融实验验证了网络各部分的有效性, 并在两种类型的数据集上进行实验, 结果证明了本文方法的效果。实验结果表明, 本文方法在抑制 PTCT 重建图像有限角伪影的同时, 能较好地保留重建图像细节, 提高重建图像质量。

关键词 X 射线光学; 计算机断层成像; 相对平行直线扫描; 图像重建; 有限角; 深度学习

中图分类号 TP391 文献标志码 A

DOI: 10.3788/AOS231823

1 引言

计算机断层成像 (CT) 是一种基于 X 射线透射, 依据多角度投影重建物体内部结构的成像技术, 因其无损、直观等特点广泛应用于医疗诊断和工业无损检测等领域^[1-3]。传统医用 CT 采取圆周扫描方式, 采用滑环部件, 成像系统体积大、结构复杂。2014 年, Liu 等^[4]提出相对平行直线扫描 CT (PTCT), 通过平板探测器 (FPD) 和射线源相对检测对象做平行直线运动的方式获取投影数据。由于单段 PTCT 扫描可获得一定等效投影角度的投影, 通过多次 PTCT 扫描可获得不小于 180° 的投影数据用于图像重建。该成像方式系统结构简单, 在工业上具有较好的应用前景。

实际检测中存在难以实现多次 PTCT 扫描成像的情况, 如高压电缆在役检测。高压电缆通常固定于地下隧道, 检测现场空间狭窄, 受检测环境和检测系统结构的限制, 只进行单段 PTCT 扫描成像^[5]。单段 PTCT 只能获得有限角度的等效投影数据, 因而其重建问题属于有限角 CT 重建, 本文重点研究 PTCT 有限角图像重建问题。

针对 PTCT 的图像重建, 伍伟文等^[6]推导了 PTCT 的滤波反投影 (PTCT-FBP) 算法; 李雷等^[7]提

出求导-希尔伯特变换-反投影 (PTCT-DHB) 算法, 将 PTCT-FBP 中的滤波函数分解为求导和希尔伯特变换, 以提升抗噪性能。上述重建算法对数据完备性具有较高要求, 对于无法通过多段直线扫描采集完备投影数据的检测场景, 该算法得到的重建图像边缘处会出现严重的伪影。

基于深度学习的有限角 CT 图像重建已取得显著效果, 深度学习类方法主要包括图像域后处理、双域恢复及基于模型的数据驱动重建方法^[8]。图像域后处理^[9-10]直接将低质量图像作为网络输入, 该方法可以有效去除伪影, 但是未考虑数据一致性, 难以准确恢复图像细节信息。双域恢复^[11-12]将投影域插值和图像域后处理相结合进行图像重建, 然而由于投影域插值误差的存在, 双域处理的最终图像可能会受二次伪影的影响。基于模型的数据驱动方法将传统迭代与深度学习相结合, 这样可使深度网络在理论层面更具可解释性的同时提高网络的稳定性^[13-14]。Chen 等^[15]提出基于专家评估的重建网络 (LEARN), 采用卷积神经网络 (CNN) 学习梯度下降算法中的正则项, 该算法在稀疏角重建中取得一定效果; Xiang 等^[16]提出 FISTA-Net, 结合近端映射将快速迭代收缩阈值算法展开到神经网络, 并在 CT 成像等逆问题中证明了该方法的有效性; Hu 等^[17]提出一种基于深度迭代

收稿日期: 2023-11-22; 修回日期: 2023-12-29; 录用日期: 2024-01-16; 网络首发日期: 2024-02-20

基金项目: 国家自然科学基金 (62171067)、重庆市自然科学基金 (CSTB2022NSCQ-MSX1311)

通信作者: *liufl@cqu.edu.cn

优化的残差学习(DIOR)方法,结合迭代优化和基于残差域的深度神经网络,提高了网络的收敛性和泛化能力。

然而,在基于模型的数据驱动方法中,以CNN为主要结构的深度神经网络往往关注局部邻域信息而忽略非局部特征^[18-19]。基于Transformer结构的网络由于具备对远程信息建模的能力,逐渐在图像复原等任务中表现出优异性能,例如SwinIR^[20]、Uformer^[21]、Restormer^[22]等典型网络。另一方面,迭代算法的研究表明,非局部特征可以改善细节^[23],这对有限角CT重建具有重要意义。

受到LEARN在稀疏CT重建中良好表现的启发,本文针对PTCT有限角重建提出一种可学习的局部和非局部正则化网络模型。该模型将迭代重建策略展开到神经网络,将具有坐标注意力(CA)机制^[24]的卷积模块和Swin-Transformer(SWIN-TRANS)^[25]模块分别作为局部和非局部正则化网络,取代LEARN中简单的CNN正则项,以避免重建图像可能造成的过平滑。局部正则项用于抑制伪影并通过CA增强图像特征,非局部正则项用于提升模型的细节恢复能力,并通过引入迭代连接(IC)提取深层特征以进一步恢复图像细节。结合网络模块构成,将该模型记为STICA-Net。针对不同数据的对比实验和消融实验验证本文方法的可行性和有效性。

2 几何模型

PTCT几何模型如图1所示。如图1(a)所示,PTCT三维几何模型^[7]中射线源在不同位置发出射线,穿透检测对象,并由另一侧的探测器采集投影数据。以检测对象中心 O 为原点,建立坐标系 $O-xyz$, x 轴与射线源的运动轨迹平行,以射线源运动方向为正方向, y 轴与探测器平面垂直,以探测器方向为正方向。以过射线源轨迹并垂直于探测器的平面为研究对象,建立图1(b)所示的PTCT二维几何模型,其中 S_0 和 S_D 分别为射线源轨迹与中心 O 、与探测器轨迹的距离。

为说明PTCT图像重建属于有限角问题,结合投影数据在频域的分布进行分析。设 l 为穿过图像 $f(x, y)$ 的任意一条射线,与 x 轴交于 A 点, OA 长度为 s , θ 为 x 轴正方向与 l 垂线的夹角, $\theta \in \left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$ 。则PTCT的二维投影数据 $p(s, \theta)$ 可表示为

$$p(s, \theta) = \frac{1}{|\cos \theta|} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x + y \tan \theta - s) dx dy, \quad (1)$$

式中: $\delta(\cdot)$ 为狄拉克函数, $x + y \tan \theta - s$ 为射线的投影线。

沿探测器方向对式(1)进行傅里叶变换,得到

$$p(\omega, \theta) = \frac{1}{|\cos \theta|} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \exp[-2\pi i \omega (x + y \tan \theta)] dx dy = |\sec \theta| F(\omega, \omega \tan \theta), \quad (2)$$

式中: ω 为频率; $F(\omega, \omega \tan \theta)$ 为投影数据在频域中的分布。 $F(\omega, \omega \tan \theta)$ 的函数表达式为

$$F(\omega, \omega \tan \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \exp[-2\pi i \omega (x + y \tan \theta)] dx dy. \quad (3)$$

由式(3)可得到PTCT二维投影数据在频域的分

布图,如图1(c)所示。数据分布整体为一个矩形,两个闭合的等腰三角形部分为数据缺失区域。在实际扫描情形中,探测器和射线源的平移轨迹为有限长度的线段, $\theta \neq \frac{\pi}{2}$,即投影数据在傅里叶空间中取不到 $\omega = 0$ 直线上除原点外的任何点。随着平移轨迹长度的增加,数据缺失区域在傅里叶空间中的变化表现为两个等腰三角形的底边缩短、高度增加。因此,PTCT扫描

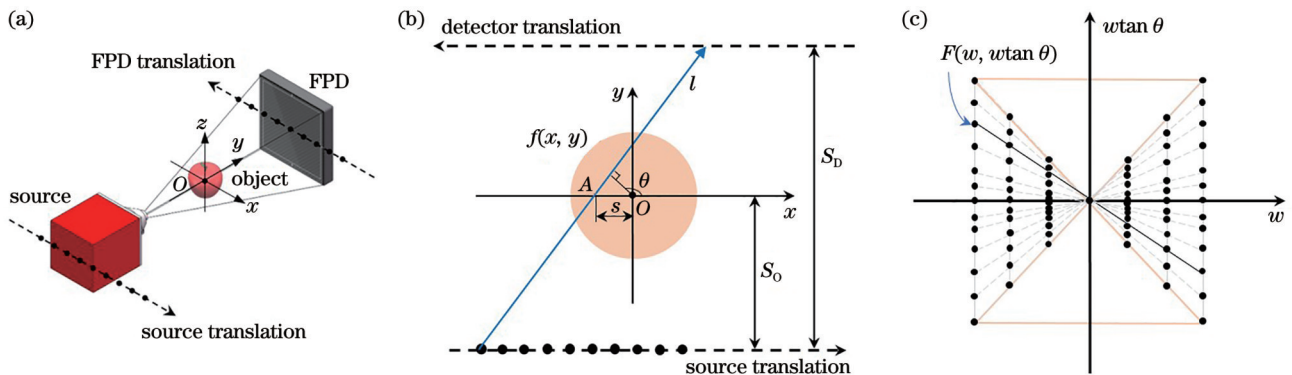


图1 PTCT几何模型示意图。(a)PTCT三维几何模型;(b)PTCT二维几何模型;(c)PTCT投影数据分布

Fig. 1 Schematic diagram of PTCT geometric model. (a) PTCT three-dimensional geometric model; (b) PTCT two-dimensional geometric model; (c) PTCT projection data distribution

方式下的图像重建属于有限角问题。

设 $\theta_{\min} \in \left[0, \frac{\pi}{2}\right]$ 表示射线源轨迹最右端位置与探测器轨迹最左端位置中心射线的垂线与 x 轴正方向的夹角, $\theta_{\max} \in \left[\frac{\pi}{2}, \pi\right]$ 表示射线源最左端位置与探测器轨迹最右端位置中心射线的垂线与 x 轴正方向的夹角。则可定义 PTCT 有限角等效的扫描角度 β :

$$\beta = \theta_{\max} - \theta_{\min} \quad (4)$$

3 本文方法

3.1 网络结构

LEARN 将具有固定迭代次数的梯度下降算法展开到神经网络, 将 CNN 作为迭代模块取代梯度下降算法中的正则项, 网络模型^[15]如图 2 所示。其中 m 、 n 表

示特征图的长与宽, c 表示通道数, $iter\ k$ 表示第 k 个迭代模块。每个 CNN 迭代模块包含的三个分支分别为用于保证数据一致性的分支、用于残差学习的短连接, 以及正则项网络分支。梯度下降算法在迭代模块中的图像更新过程可以显式地表示为

$$\mathbf{u}_k = \mathbf{u}_{k-1} - \left[\lambda_{k-1} \mathbf{A}^T (\mathbf{A} \mathbf{u}_{k-1} - \mathbf{p}) + f_{\text{network}}(\mathbf{u}_{k-1}, \Phi_{k-1}) \right], \quad (5)$$

式中: \mathbf{u}_k 和 \mathbf{u}_{k-1} 分别为第 k 和第 $k-1$ 个迭代模块的输出; $\lambda_{k-1} \mathbf{A}^T (\mathbf{A} \mathbf{u}_{k-1} - \mathbf{p})$ 为数据一致性项; \mathbf{A} 为由扫描几何和体积确定的系统矩阵, 也称为前向投影算子; \mathbf{A}^T 为反投影算子, 上标 T 为转置符号; \mathbf{p} 为测量的投影正弦图数据; λ 为权重系数, 在模型中设置为可学习的参数; $f_{\text{network}}(\mathbf{u}_{k-1}, \Phi_{k-1})$ 为迭代过程中可学习的正则项网络; Φ_{k-1} 为网络参数。在图(2)中, $f_{\text{network}}(\mathbf{u}_{k-1}, \Phi_{k-1})$ 即为 CNN 迭代模块中的 3 个卷积层。

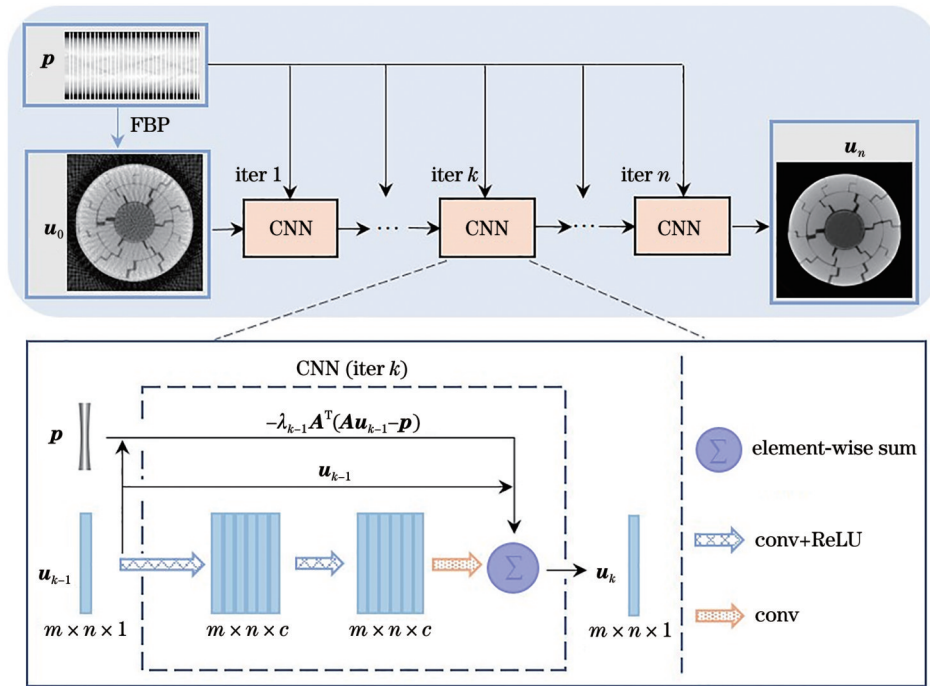


图 2 LEARN 结构示意图

Fig. 2 Schematic diagram of LEARN structure

考虑到有限角重建的 CT 图像存在大量信息丢失的问题, 本文基于 LEARN 的模型框架, 提出使用可学习的局部和非局部正则化网络取代 LEARN 中 3 层简单的 CNN 正则项, 设计了如图 3 所示的 STICA-Net 模型。STICA-Net 将用于非局部正则化的 Swin-Transformer 模块和具有 CA 机制的 CNN 局部正则化模块, 作为迭代模块交替部署。初始输入为 \mathbf{u}_0 和扫描角度为 β 时的有限角投影数据 \mathbf{p} , 其中 \mathbf{u}_0 由投影数据 \mathbf{p} 通过 FBP 算法重建获得。

3.2 迭代模块

以 CNN 为主要结构的局部正则化模块可以取得良好的伪影抑制效果, 但容易造成细节丢失。为此, 引

入 Swin-Transformer 学习非局部正则化。在如图 3 所示的非局部正则化迭代模块中, \mathbf{u}_{k-2} 首先通过一个卷积层进行补丁嵌入, 再经过两个使用默认参数设置的连续 Swin-Transformer 模块。Swin-Transformer 由非重叠窗口的多头自注意力 (WMSA) 及移位窗口的多头自注意力 (SWMSA) 模块组成, 如图 4 所示。WMSA 可以学习小尺寸窗口区域不同像素间的远程信息相关性, SWMSA 通过移位窗口改善不同窗口之间的交互。每个 Swin-Transformer 模块依次由层归一化 (LN)、多头自注意力机制、残差连接和多层感知机 (MLP) 构成。最后将 Swin-Transformer 的输出经过一个卷积层得到单通道的特征图, 用于与另外两个分

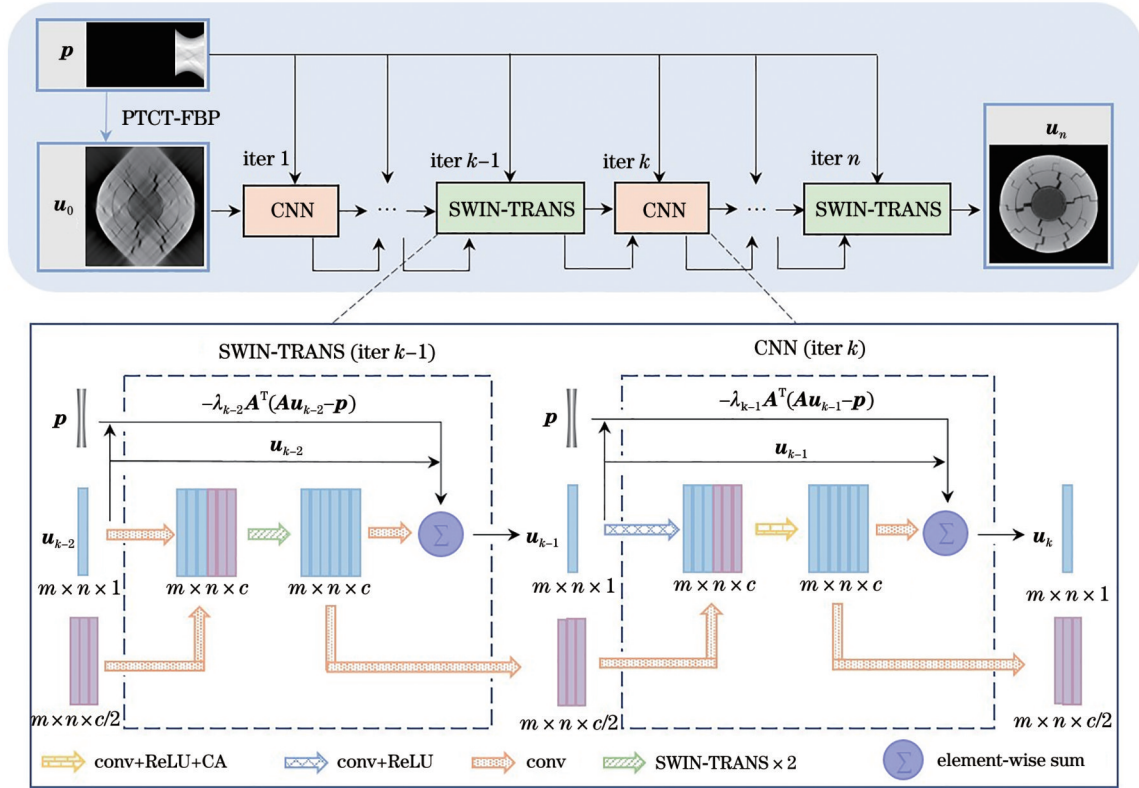


图 3 STICA-Net 结构示意图

Fig. 3 Schematic diagram of STICA-Net structure

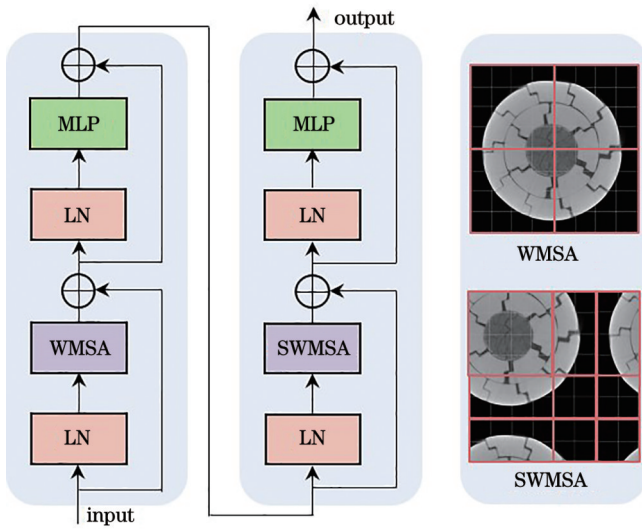


图 4 Swin-Transformer 结构示意图

Fig. 4 Schematic diagram of Swin-Transformer structure

支进行逐元素求和。因此,非局部正则化迭代模块通过获得特征图中任意像素之间的联系,更好地恢复图像细节。

局部正则化模块中,仅以简单的三个卷积层作为迭代模块难以提供有效的正则化约束,重建图像往往存在过平滑的问题。为此,在卷积模块中加入 CA 机制,用于特征增强。在如图 3 所示的基于 CNN 的局部正则化中, u_{k-1} 首先经过一个卷积层与线性整流

(ReLU) 激活函数进行特征提取,再经过一次卷积与 ReLU 激活操作后,使用 CA 强化特征表达,最后通过一个卷积层输出单通道的特征图,用于与另外两个分支进行逐元素求和。本文使用的 CA 结构示意图如图 5 所示,其中 concat 表示将特征图沿通道方向串联, sigmoid 用于将输入映射到 0、1 之间,同时移除了原始 CA 在卷积和 ReLU 激活函数之间的批量归一化(BN)层,这样有利于网络收敛。CA 模块沿通道方向将输入特征图分解为两个一维特征编码,通过在水平和垂直方向聚合特征,最终得到两个方向上的注意力权重,两组权重互补地应用于输入特征图,使得 CA 能够更准确地定位感兴趣对象的确切位置,因此对重要特征的表达得到强化,单独使用卷积层造成的图像过平滑现象得到缓解。

为建立不同迭代模块间特征图的联系,在网络中加入迭代连接(IC),如图 3 中迭代模块的下分支所示。将迭代模块内第一次卷积结果的特征图通道数减半,将其与上一次迭代连接传递的特征图在通道维度串联。通过对第二次卷积结果进行额外卷积,得到通道数为 $\frac{c}{2}$ 的特征图,即传送至下一次迭代的信息。通过增加 IC 结构可以扩展网络深度,提取更深层次的细节特征。

STICA-Net 局部正则化模块中,除 CA 的卷积使用默认 1×1 的卷积核外,其余卷积的卷积核均为 5×5 ,

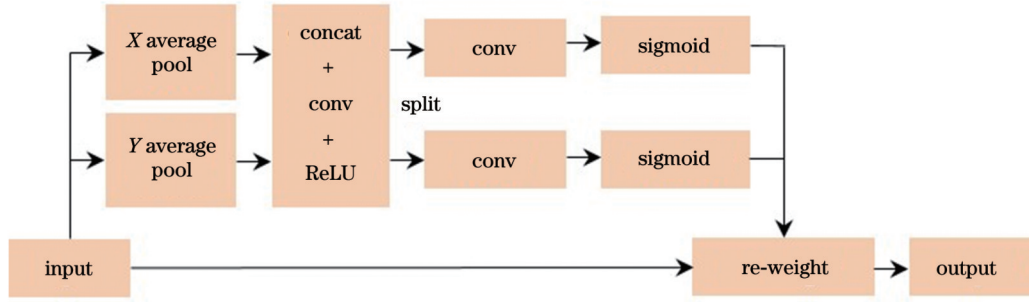


图 5 CA 机制结构示意图

Fig. 5 Schematic diagram of CA mechanism structure

与 LEARN 保持一致。非局部正则化模块中的卷积主要用于补丁嵌入以及输出单通道的特征图,选择稍小的 3×3 卷积核,以减少模型参数量,节约计算资源。迭代连接结构中的卷积,与之同理,选择 3×3 的卷积核。

3.3 损失函数

本文使用均方误差(MSE)和感知损失构建损失函数,若单独使用 MSE 作为损失函数,可能导致重建的图像模糊,使图像缺乏高频信息。MSE 为重建图像和标签图像的均方误差,用于衡量两幅图像中每个像素点之间的误差,计算公式为

$$L_u = \frac{1}{N} \sum_{i=1}^N \|\hat{\mathbf{u}}_i - \mathbf{u}_i\|_2^2, \quad (6)$$

式中: N 为训练样本数量; $\hat{\mathbf{u}}_i$ 为迭代网络重建的图像; \mathbf{u}_i 为标签图像。

感知损失将标签图像和重建图像经过相同的特征提取网络并对提取的特征进行比较,使得更深层次网

络的图像特征接近。本文使用经预训练的 VGG16 网络^[26]作为特征提取器,VGG16 包括 13 个卷积层,将 2、4、7 和 10 个卷积层输出的特征用于定义感知损失:

$$L_{\text{VGG}} = \frac{1}{N} \sum_{i=1}^N \sum_{j=2}^J \|\phi_j(\hat{\mathbf{u}}_i) - \phi_j(\mathbf{u}_i)\|_2^2, \quad j \in \{2, 4, 7, 10\}, \quad (7)$$

式中: ϕ_j 为第 j 个卷积层的特征提取器; J 为卷积层数。最终的损失函数为

$$L = L_u + \alpha L_{\text{VGG}}, \quad (8)$$

式中: α 为权重系数,设置为 0.001^[27],用于平衡两种损失。

4 实 验

4.1 实验设置

本文使用公开的 2DeteCT 数据集^[28]和实际扫描获取的碳纤维复合芯导线(ACCC)数据集综合评估所提出的深度迭代网络。实验参数设置如表 1 所示。

表 1 实验参数设置

Table 1 Experimental parameter settings

Item	2DeteCT		ACCC
Scanning method	RCT	PTCT	PTCT
Scan angle / ($^\circ$)	90	$90(\beta)$	$90(\beta)$
Tube voltage / kV	—	—	100
Tube current / μA	—	—	83
S_0 / mm	100	100	80
S_D / mm	240	240	259
Detector pixel size / mm	0.085	0.085	0.085
Array length / pixel	256	256	1536
Reconstruction matrix size / (pixel \times pixel)	256×256	256×256	256×256
Number of sampling points	150	150	150
Sampling model	—	Equi-spatial	Equi-spatial

2DeteCT 数据集用于仿真,以验证本文方法的可行性。该数据集将熔岩石、咖啡豆、核桃等样本混合物作为检测对象,样本对象在样本间和样本内的形状和密度都具有较高的自然可变性。数据集包括 5000 张

切片,随机选取 600 张切片作为训练集,100 张切片作为验证集,200 张切片作为测试集。将数据集中的图像作为标签图像(真值),利用标签图像模拟扇形束投影,生成圆周 CT(RCT)有限角或 PTCT 等效扫描角

度 β 为 90° 的投影数据。

ACCC 数据集用于实际数据实验, 以进一步评估本文方法的有效性。图 6 所示为基于采集的 ACCC 数据搭建的 PTCT 实验系统, 该系统由射线源、平板探测器、直线滑台和转台组成。单段 PTCT 扫描过程中, 检测对象固定于转台上, 射线源和平板探测器在直线滑台上做相对反向运动。进行多段扫描时, 将转台带动检测对象旋转一定角度, 再进行下一段直线扫描。通过 4 段扫描获取足够的投影数据并进行重建, 将重建结果作为标签图像, 伪影图像由单段 PTCT 扫描获取。ACCC 数据集共包括 3600 张切片, 随机选取的训练集、测试集、验证集数量与 2DeteCT 数据集的相同。

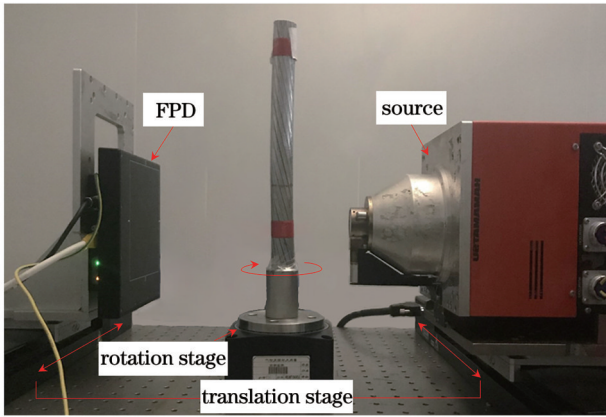


图 6 PTCT 实验系统

Fig. 6 Experimental system of PTCT

所提出的网络模型基于 pytorch 框架编写, 迭代过程中 PTCT 的前向投影与反投影算子通过 ASTRA Toolbox^[23] 中的 CUDA 内核实现。实验硬件平台配置

为 CPU: AMD Ryzen Threadripper 3960X, 3.79 GHz, GPU: NVIDIA GeForce RTX 4090×2。模型采用 AdamW 算法优化, 学习率初始值为 0.0001, 并逐渐减小至 0。网络迭代模块数 $n = 20$, 通道数 $c = 96$ 。

使用 FBP、SIRT、SwinIR、FISTA-Net、LEARN 作为实验的对比方法。其中 SIRT 为迭代重建算法, 设置其迭代次数为 1000 次, 以保证算法足够收敛。SwinIR 为具有 Swin-Transformer 结构的图像域后处理网络, FISTA-Net 和 LEARN 为基于深度迭代展开的代表性方法, 分别采用原文默认的 7 个和 50 个迭代模块。深度学习方法中 epochs 数量均为 200。使用均方根误差 (RMSE)、峰值信噪比 (PSNR) 和结构相似性指数 (SSIM) 三个量化指标来综合评价重建图像质量。一般情况下, RMSE 越小, PSNR 和 SSIM 越大, 则图像质量越好。

4.2 仿真数据实验结果

4.2.1 RCT 对比实验

为证明所提出的 STICA-Net 模型对于有限角 CT 重建具有优势, 首先通过 RCT 有限角对比实验进行验证。图 7 所示为不同方法针对测试图像的重建结果示例, 图像右下角为感兴趣区域放大图, 从箭头处可以观察到不同重建图像具有明显差异的特征。由于大范围投影数据缺失, FBP 重建图像中存在严重的有限角伪影, SIRT 虽然能够抑制部分伪影, 但对图像细节信息的恢复能力依然较差。相比之下, 深度学习方法显著提高了图像质量。整体看来, SwinIR 能够还原大部分特征, 但对具体细节结构的恢复还有些欠缺。FISTA-Net 和 LEARN 均使用卷积层进行重建约束, 图像存在不同程度的过平滑现象, 相比之下, LEARN 能够更好地恢复细节。

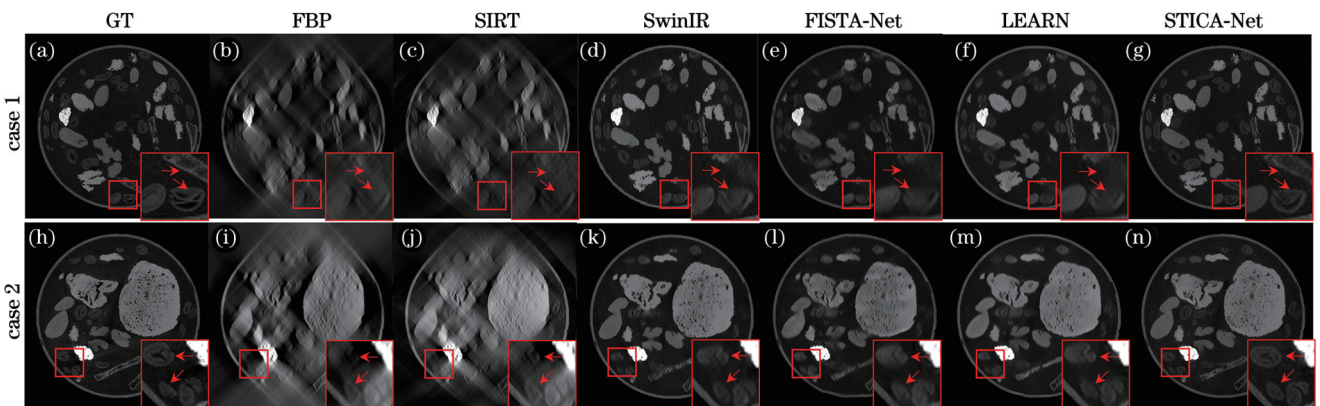


图 7 2DeteCT 数据集下不同方法的 RCT 重建结果。(a)~(g) 示例一重建结果; (h)~(n) 示例二重建结果

Fig. 7 RCT reconstruction results of 2DeteCT dataset obtained by different methods. (a)~(g) Reconstruction results of case 1; (h)~(n) reconstruction results of case 2

STICA-Net 同时考虑图像非局部信息和像素局部邻域信息, 能够恢复更多的纹理细节。实验结果表明, 本文方法能够有效地用于 RCT 有限角重建, 相比对比方法取得最佳结果。

4.2.2 PTCT 对比实验

PTCT 图像重建属于有限角问题, 为验证 STICA-Net 对于 PTCT 有限角重建同样具有良好的重建效果, 使用同样的 2DeteCT 数据集进行对比

实验。图 8 为测试图像的重建结果示例，STICA-Net 方法在 PTCT 有限角重建中依然具

有较好的表现，重建图像在视觉上与真值图像最为接近。

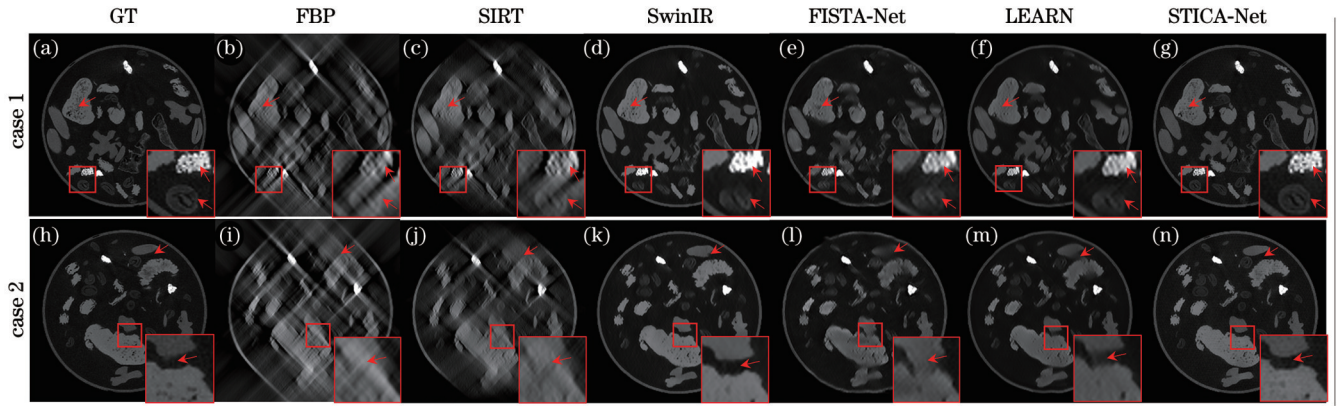


图 8 2DeteCT 数据集下不同方法的 PTCT 重建结果。(a)~(g) 示例一重建结果；(h)~(n) 示例二重建结果

Fig. 8 PTCT reconstruction results of 2DeteCT dataset obtained by different methods. (a)~(g) Reconstruction results of case 1; (h)~(n) reconstruction results of case 2

表 2 为图 8 中示例图像的客观评价指标对比。从表中可看出，深度学习的方法相较于传统方法，所有指标均得到明显提升。在两个示例图像中，SwinIR 的 SSIM 高于 LEARN，而 STICA-Net 在三个评价指标中均取得最佳数值。STICA-Net 相比 SwinIR，两

个示例图像的 SSIM 分别提高了 0.0049 和 0.0071。相比于具有次优 PSNR 和 RMSE 的 LEARN 网络，STICA-Net 针对两个示例图像的 PSNR 分别提高了 2.1668 dB 和 2.5951 dB，RMSE 分别减小了 0.0050 和 0.0049。

表 2 2DeteCT 数据集下不同方法的重建图像指标

Table 2 Reconstruction image indexes of different methods on 2DeteCT dataset

Method	PSNR of case 1 /dB	SSIM of case 1	RMSE of case 1	PSNR of case 2 /dB	SSIM of case 2	RMSE of case 2
FBP	22.0108	0.3628	0.0793	20.4884	0.3350	0.0945
SIRT	26.0630	0.6650	0.0499	25.2571	0.6447	0.0546
SwinIR	33.1356	0.9354	0.0220	32.9674	0.9355	0.0225
FISTA-Net	32.4324	0.8980	0.0239	32.4536	0.8982	0.0238
LEARN	34.1125	0.9223	0.0197	34.4578	0.9240	0.0189
STICA-Net	36.2793	0.9403	0.0147	37.0529	0.9426	0.0140

PTCT 对比实验的主观图像评价与定量评价指标结果表明，本文方法可用于解决 PTCT 有限角问题，并实现高质量的图像重建。对比 PTCT 与 RCT 有限角重建图像的伪影分布可以发现，二者均在数据缺失区域出现严重伪影，经过 STICA-Net 网络重建后，获得了类似的实验结果。

4.2.3 消融实验

为验证模型中不同网络结构在 PTCT 有限角重建中的有效性，基于 PTCT 有限角为 90° 的重建图像进行消融实验。2DeteCT 数据相比 ACCC 数据具有更加丰富的图像特征，增加不同网络模块对重建结果的影响更容易在重建结果中体现，因此本节基于 2DeteCT 数据集进行分析。STICA-Net 具有 20 个迭代模块，本文以具有 20 个迭代模块的 LEARN 作为基准网络，损失函数为 L_v ，用 L20 表示，采取渐进式验证策略。增加感知损失 L_{VGG} 的 L20 网络用 L20L 表示，

ST-Net 表示在 L20L 中增加了 Swin-Transformer 非局部正则化模块。同理，STIC-Net 表示在前者基础上增加了迭代连接的 IC 结构，STICA-Net 表示增加了 CA 机制的本文方法。

消融实验结果示例如图 9 所示，通过观察不同的重建图像框选区域放大图以及箭头处的特征差异，可以清楚地看到，随着网络结构的增加，图像质量也在不断提升。其中，仅使用 20 个迭代模块的 L20 在图像特征形状和细节结构方面均出现严重失真，得益于 VGG16 作为特征提取器，增加感知损失 L_{VGG} 后的 L20L 可以将重建图像与真值图像靠近，但是残余的伪影仍然影响观察。ST-Net 中增加 Swin-Transformer 模块，使得图像质量得到明显提升，仅局部存在少量伪影和轻微模糊。而 STIC-Net 由于增加了 IC 结构，提高了每次迭代的效率，能够提取到更深层次的图像特征，原始图像中的有限角伪影得到抑制。STICA-Net

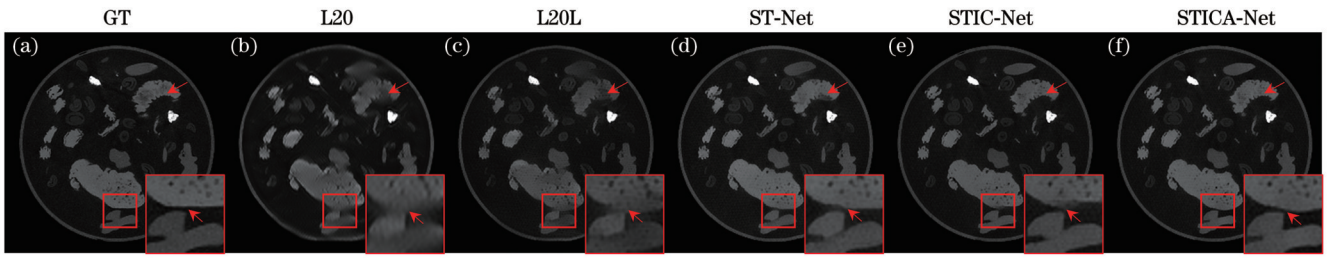


图 9 不同消融网络在 2DeteCT 数据集上的消融实验重建结果

Fig. 9 Reconstruction results of ablation experiment of different ablation networks on 2DeteCT dataset

通过增加 CA 模块进行特征增强,使得图像质量得到了进一步改善,在伪影抑制和细节恢复两方面均实现了较好的结果。

图 10 为测试集中所有样本在不同消融网络中指标数值分布的箱线图,四分位距(IQR)即箱体高度,用于

衡量图中数据点的离散程度。整体来看,逐个增加网络结构后各指标数值得到明显改善。STICA-Net 相比基准网络 L20, RMSE 的中位数降低了 31.55%, SSIM、PSNR 的中位数分别提高了 6.64% 和 10.97%。IQR 有所减小,表明本文方法的离散程度更小,稳定性更高。

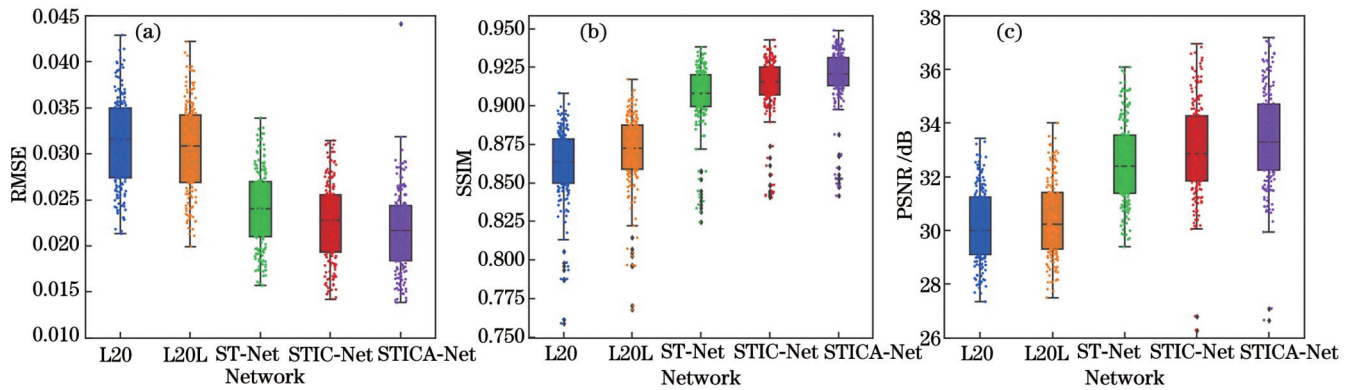


图 10 消融网络测试集样本指标数值分布的箱线图。测试集样本在不同消融网络中的(a)RMSE、(b)SSIM 和(c)PSNR 数值分布
Fig. 10 Box diagrams of numerical distribution of sample indexes of ablation network for test set. Numerical distributions of (a) RMSE, (b) SSIM, and (c) PSNR of test set samples in different ablation networks

4.3 实际数据实验结果

PTCT 扫描方式在在役检测中具有显著优势,为使研究工作更加贴合实际,开展了基于 ACCC 数据集的实验。如图 11 所示,每张图像的右下角为框选区域的放大图,箭头处表示 ACCC 内部芯棒出现的裂纹缺

陷。整体看来,传统方法在大角度数据缺失的情况下,重建图像仍然存在大量伪影。FISTA-Net 和 LEARN 的重建图像伪影得到了基本抑制, FISTA-Net 虽然呈现出相比 LEARN 更好的重建结果,但细节特征仍然存在模糊。本文方法在图像细节复原和伪影抑制方面

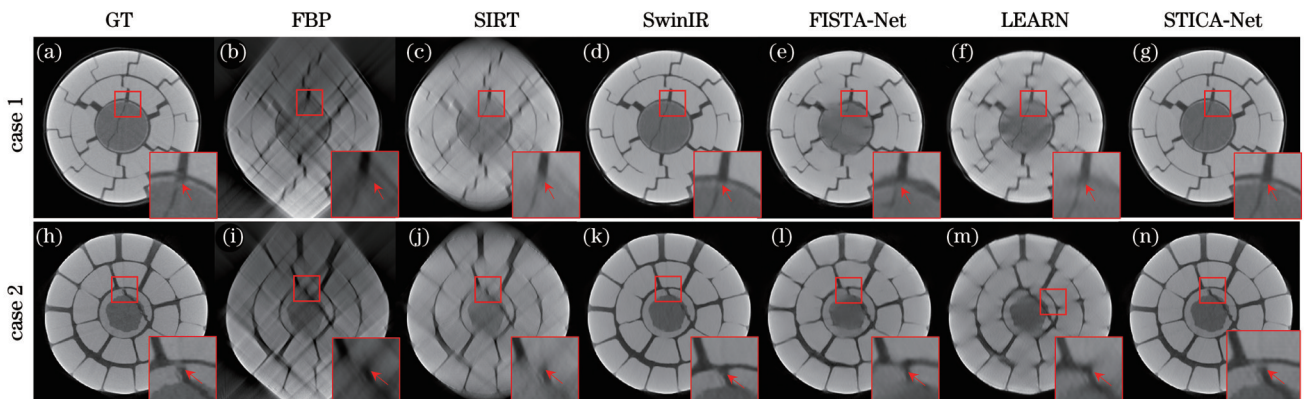


图 11 ACCC 数据集下不同方法的 PTCT 重建结果。(a)~(g) 示例一重建结果;(h)~(n) 示例二重建结果

Fig. 11 PTCT reconstruction results of ACCC dataset obtained by different methods. (a)~(g) Reconstruction results of case 1; (h)~(n) reconstruction results of case 2

均呈现出较好的效果,可以更清楚地看到箭头处的图像特征。从箭头处可以观察到,SwinIR虽然能够较为清晰地恢复特征信息,但细节处的重建结果并不如STICA-Net精确,这是因为基于后处理的深度重建方法严重依赖训练数据的质量,且没有考虑数据一致性。相比之下,STICA-Net的重建结果与真值图像更为接近。

表3为上述两个ACCC示例的评价指标对比,STICA-Net相比SwinIR,两个示例图像的PSNR分别提高了4.72%和5.53%,SSIM分别提高了2.88%和1.59%,RMSE分别降低了15.94%和19.32%。综合来看,在伪影消除和细节恢复方面,本文方法优于其他方法。

表3 ACCC数据集下不同方法的重建图像指标
Table 3 Reconstruction image indexes of different methods on ACCC dataset

Network	PSNR of case 1 /dB	SSIM of case 1	RMSE of case 1	PSNR of case 2 /dB	SSIM of case 2	RMSE of case 2
FBP	12.8890	0.4429	0.2268	13.7951	0.4646	0.2043
SIRT	20.7439	0.7661	0.0918	20.3331	0.7498	0.0962
SwinIR	31.9921	0.9292	0.0251	33.6882	0.9381	0.0207
FISTA-Net	29.2278	0.8625	0.0346	30.9193	0.8929	0.0285
LEARN	26.4515	0.8396	0.0476	24.4574	0.8379	0.0599
STICA-Net	33.5012	0.9560	0.0211	35.5502	0.9530	0.0167

4.4 讨论

基于深度学习的方法模型参数量、浮点运算次数以及测试集每个样本推理所需的平均时间如表4所示。可以看到,后处理网络SwinIR不涉及前向投影和反投影,在比FISTA-Net参数量多的情况下仍然获得

了最少的推理时间。基于模型的迭代展开方法随着参数量、浮点运算次数的增加,推理时间也更长。而STICA-Net虽然比LEARN具有更少的迭代模块,但Swin-Transformer结构的加入,使得在提高网络性能的同时,参数量、浮点运算次数也相应地增加。

表4 不同模型的参数量、浮点运算次数和推理时间

Table 4 Number of parameters, number of floating-point operations, and reasoning time of different models

Model	SwinIR	LEARN	FISTA-Net	STICA-Net
Number of parameters / 10^6	1.22	3.01	0.52	5.38
Number of floating-point operations / 10^9	81.22	196.61	51.00	352.90
Time /ms	34.00	235.26	54.81	314.28

另一方面,观察发现,在实验平台有限显存的范围内,STICA-Net随着迭代模块数量的增加,图像质量也逐渐改善。然而,具有20个迭代模块的STICA-Net仅能将batch size设置为2,综合考虑计算效率、图像质量及内存消耗,本文未使用更多的模块进行实验。

5 结论

针对PTCT图像重建的难点,通过建立PTCT几何模型,从理论层面分析了PTCT图像重建属于有限角问题,并在此基础上提出了STICA-Net模型。该模型在用于局部正则化的卷积模块中引入CA机制,减少图像平滑;通过设计非局部正则化的Swin-Transformer模块提升模型的细节恢复能力;增加迭代连接用于提取深层次特征。消融实验验证了模型中每一部分结构对重建图像的增益效果。相比于对比算法,图像质量得到明显提升,定量评价指标在不同数据类型中均取得最佳数值。综合结果表明,本文方法在

PTCT有限角伪影抑制和细节恢复方面优于对比算法,能够实现高质量的图像重建,有利于推动PTCT的在役检测应用。

本文方法的不足之处在于,虽然消融实验证明Swin-Transformer结构的增加给图像结果带来明显提升,但该结构需要更多的内存存储权重及中间特征,这限制了本文使用更高分辨率的图像进行实验。后续研究将对网络模块进行进一步完善,以使网络更加轻量化。

参考文献

- [1] Withers P J, Bouman C, Carmignato S, et al. X-ray computed tomography[J]. Nature Reviews Methods Primers, 2021, 1: 18.
- [2] 戈文杰,余海军,陈杰,等.基于求导-希尔伯特变换-反投影的源直线扫描计算机断层成像解析重建[J].光学学报,2022,42(11):1134025.
Ge W J, Yu H J, Chen J, et al. Analytical reconstruction for source translation scanning computed tomography based on derivative-Hilbert transform-back projection[J]. Acta Optica Sinica, 2022, 42(11): 1134025.
- [3] 周鹏伍,席雅睿,朱国荣,等.复合扫描计算机断层成像研究

- [J]. 仪器仪表学报, 2022, 43(12): 178-187.
- Zhou P W, Xi Y R, Zhu G R, et al. Research on the hybrid scanning computed laminography[J]. Chinese Journal of Scientific Instrument, 2022, 43(12): 178-187.
- [4] Liu F L, Yu H Y, Cong W X, et al. Top-level design and pilot analysis of low-end CT scanners based on linear scanning for developing countries[J]. Journal of X-Ray Science and Technology, 2014, 22(5): 673-686.
- [5] 廖明娟, 李雷, 段晓礁, 等. 源平移扫描局部CT成像及其检测在役高压电缆[J]. 光学学报, 2022, 42(16): 1611002.
- Liao M J, Li L, Duan X J, et al. Local source-translation CT imaging for testing in-service high-voltage cables[J]. Acta Optica Sinica, 2022, 42(16): 1611002.
- [6] 伍伟文, 全超, 刘丰林. 相对平行直线扫描CT滤波反投影图像重建[J]. 光学学报, 2016, 36(9): 0911009.
- Wu W W, Quan C, Liu F L. Filtered back-projection image reconstruction algorithm for opposite parallel linear CT scanning[J]. Acta Optica Sinica, 2016, 36(9): 0911009.
- [7] 李雷, 谭川东, 廖明娟, 等. 基于Radon逆变换的相对平行直线扫描CT解析重建[J]. 光学学报, 2021, 41(6): 0611003.
- Li L, Tan C D, Liao M J, et al. Analytic reconstruction for parallel translational computed tomography based on Radon inverse transform[J]. Acta Optica Sinica, 2021, 41(6): 0611003.
- [8] 邸江磊, 林俊成, 钟丽云, 等. 基于深度学习的稀疏或有限角度CT重建方法研究综述[J]. 激光与光电子学进展, 2023, 60(8): 0811002.
- Di J L, Lin J C, Zhong L Y, et al. Review of sparse-view or limited-angle CT reconstruction based on deep learning[J]. Laser & Optoelectronics Progress, 2023, 60(8): 0811002.
- [9] Wang J X, Liang J, Cheng J Y, et al. Deep learning based image reconstruction algorithm for limited-angle translational computed tomography[J]. PLoS One, 2020, 15(1): e0226963.
- [10] Fu T Y, Qiu S, Wang Y, et al. Limited-angle tomography reconstruction with a U-Conv-Swin-Net[EB/OL]. [2023-11-09]. <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/epdf/10.1002/xrs.3367>.
- [11] Ghani M U, Karl W C. Integrating data and image domain deep learning for limited angle tomography using consensus equilibrium[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), October 27-28, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 3922-3932.
- [12] Zhang Q Y, Hu Z L, Jiang C H, et al. Artifact removal using a hybrid-domain convolutional neural network for limited-angle computed tomography imaging[J]. Physics in Medicine & Biology, 2020, 65(15): 155010.
- [13] Huang Y X, Preuhs A, Lauritsch G, et al. Data consistent artifact reduction for limited angle tomography with deep learning prior[M]//Knoll F, Maier A, Rueckert D, et al. Machine learning for medical image reconstruction. Lecture notes in computer science. Cham: Springer, 2019, 11905: 101-112.
- [14] 耿梦凡, 张虎, 李哲, 等. 基于迭代优化展开的Cherenkov激发的荧光扫描成像重建算法[J]. 中国激光, 2023, 50(15): 1507106.
- Geng M F, Zhang H, Li Z, et al. A reconstruction algorithm for cherenkov-excited luminescence scanning imaging based on unrolled iterative optimization[J]. Chinese Journal of Lasers, 2023, 50(15): 1507106.
- [15] Chen H, Zhang Y, Chen Y J, et al. LEARN: learned experts' assessment-based reconstruction network for sparse-data CT[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1333-1347.
- [16] Xiang J X, Dong Y G, Yang Y J. FISTA-net: learning a fast iterative shrinkage thresholding network for inverse problems in imaging[J]. IEEE Transactions on Medical Imaging, 2021, 40(5): 1329-1339.
- [17] Hu D L, Zhang Y K, Liu J, et al. DIOR: deep iterative optimization-based residual-learning for limited-angle CT reconstruction[J]. IEEE Transactions on Medical Imaging, 2022, 41(7): 1778-1790.
- [18] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7794-7803.
- [19] Xia W J, Yang Z Y, Lu Z X, et al. RegFormer: a local-nonlocal regularization-based model for sparse-view CT reconstruction[J]. IEEE Transactions on Radiation and Plasma Medical Sciences, 2024, 8(2): 184-194.
- [20] Liang J Y, Cao J Z, Sun G L, et al. SwinIR: image restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), October 11-17, 2021, Montreal, BC, Canada. New York: IEEE Press, 2021: 1833-1844.
- [21] Wang Z D, Cun X D, Bao J M, et al. Uformer: a general U-shaped transformer for image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 17662-17672.
- [22] Zamir S W, Arora A, Khan S, et al. Restormer: efficient transformer for high-resolution image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 5718-5729.
- [23] van Aarle W, Palenstijn W J, Cant J, et al. Fast and flexible X-ray tomography using the ASTRA toolbox[J]. Optics Express, 2016, 24(22): 25129-25147.
- [24] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [25] Liu Z, Lin Y T, Cao Y, et al. Swin Transformer: hierarchical Vision Transformer using Shifted Windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), June 20-25, 2021, Montreal, QC, Canada. New York: IEEE Press, 2021: 9992-10002.
- [26] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9906: 694-711.
- [27] Ge Y S, Su T, Zhu J T, et al. ADAPTIVE-NET: deep computed tomography reconstruction network with analytical domain transformation knowledge[J]. Quantitative Imaging in Medicine and Surgery, 2020, 10(2): 415-427.
- [28] Kiss M B, Coban S B, Batenburg K J, et al. 2DeteCT-A large 2D expandable, trainable, experimental Computed Tomography dataset for machine learning[J]. Scientific Data, 2023, 10: 576.

Limited-Angle CT Image Reconstruction Based on Swin-Transformer Iterative Unfolding for PTCT Imaging

Yuan Wei^{1,2}, Xi Yarui^{1,2}, Tan Chuandong^{1,2}, Liu Chuanjiang^{1,2}, Zhu Guorong^{1,2}, Liu Fenglin^{1,2*}

¹*ICT Research Center, Key Laboratory of Optoelectronic Technology & Systems, Ministry of Education, Chongqing University, Chongqing 400044, China;*

²*Industrial CT Non-Destructive Testing Engineering Research Center, Ministry of Education, Chongqing University, Chongqing 400044, China*

Abstract

Objective Computed tomography (CT) is an imaging technique that employs X-ray transmission and multi-angle projection to reconstruct the internal structure of an object. Meanwhile, it is commonly adopted in medical diagnosis and industrial non-destructive testing due to its non-invasive and intuitive characteristics. Parallel translational computed tomography (PTCT) acquires projection data by moving a flat panel detector (FPD) and a radiation source in parallel linear motion relative to the detection object. This method has promising applications in industrial inspection. Due to the limitations of the inspection environment and the structure of the inspection system, there are scenarios where it is difficult to realize multi-segment PTCT scanning and imaging, and only single-segment PTCT scanning and imaging can be performed. Since the single-segment PTCT can only obtain the equivalent projection data at a limited angle, its reconstruction problem belongs to limited-angle CT reconstruction. Images reconstructed by traditional algorithms will suffer from serious artifacts. Deep learning-based limited-angle CT image reconstruction has yielded remarkable results, among which model-based data-driven methods have caught much attention. However, such deep networks with CNNs as the main structure tend to focus on the local neighborhood information of the image and ignore the non-local features. Additionally, research on iterative algorithms shows that non-local features can improve detail preservation, which is important for limited-angle CT reconstruction.

Methods To address the limited-angle artifact in PTCT image reconstruction, we propose a deep iterative unfolding method (STICA-Net, Fig. 3) that learns local and non-local regular terms. The method unfolds a gradient descent algorithm with a fixed number of iterations to a neural network and utilizes convolutional modules with the coordinate attention (CA) mechanism and Swin-Transformer modules deployed as iterative modules in alternating cascades to form an end-to-end deep reconstruction network. The convolution module learns local regularization, in which CA is leveraged to reduce image smoothing. The Swin-Transformer module learns non-local regularization to improve the network's ability to restore image details. Among neighboring modules, iterative connection (IC) is adopted to enhance the model's ability to extract deeper features and improve the efficiency of each iteration. The employed experimental comparison methods are FBP, SIRT, SwinIR, FISTA-Net, and LEARN. The quality of the reconstructed image is comprehensively evaluated by utilizing three sets of quantitative indicators of root mean square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM). Meanwhile, comparison experiments are conducted on both simulated and real datasets to verify the feasibility of the proposed method. Additionally, we perform ablation experiments to confirm the effectiveness of each component of the network.

Results and Discussions We present the results of a contrast experiment of 90° limited-angle rotational scanning CT using the simulation data 2DeteCT dataset. The results demonstrate the effectiveness of the STICA-Net method for limited-angle reconstruction (Fig. 7). It is noted that PTCT image reconstruction is a limited-angle problem. To verify STICA-Net's effectiveness in PTCT limited-angle reconstruction, we employ the same dataset to generate projection data with an equivalent scanning angle of 90° via PTCT scanning, and then compare different methods. The results of both subjective image evaluation (Fig. 8) and quantitative evaluation index (Table 2) show that STICA-Net can solve the limited-angle problem of PTCT and achieve high-quality image reconstruction. By building the PTCT experimental platform (Fig. 6), the actual dataset of carbon fiber composite core wire (ACCC) is obtained. The two example results (Fig. 11) of the ACCC dataset indicate that the reconstructed images of the traditional method still contain a significant number of artifacts in the absence of large-angle data. However, the artifacts in the reconstructed images of FISTA-Net and LEARN have been significantly reduced. Although FISTA-Net produces better reconstruction results than LEARN, the details are still somewhat blurred. Compared with the suboptimal SwinIR, the PSNR of STICA-Net increases by 4.72% and 5.53%, the SSIM rises by 2.88% and 1.59%, and the RMSE decreases by 15.94% and 19.32% respectively.

Meanwhile, ablation experiments verify the effectiveness of different network structures in PTCT limited-angle reconstruction. Figure 10 demonstrates clear improvement in the numerical values of each index as network structures are added incrementally.

Conclusions To deal with the difficulty of PTCT image reconstruction, we theoretically conclude that PTCT image reconstruction is a limited-angle problem by building a PTCT geometric model, and then propose the STICA-Net model. Ablation experiments confirm the effectiveness of each model component in improving the reconstructed image. Compared to the contrast algorithm, the proposed method significantly improves image quality and yields the best quantitative evaluation indicators across different data types. Additionally, comprehensive results demonstrate that the proposed method outperforms the contrast algorithm in terms of PTCT limited-angle artifact suppression and detail recovery, and high-quality image reconstruction can be achieved. This is beneficial for promoting the in-service detection application of PTCT. However, the method's limitation is that although the ablation experiments demonstrate that the inclusion of the Swin-Transformer structure enhances image results, more memory is needed to store weights and intermediate features, which restricts the utilization of higher-resolution images in our study. In the future, the network module will be further improved to make the network more lightweight.

Key words X-ray optics; computed tomography; relatively parallel translational scanning; image reconstruction; limited angle; deep learning