面向激光光条图像修复的循环相似度映射网络

冀振燕[1],韩梦豪[1],宋晓军[1],冯其波[2]
(1.北京交通大学软件学院,北京 100044;2.北京交通大学理学院,北京 100044)

摘 要: 基于线结构光的视觉检测技术广泛应用于工业检测,现场动态采集的激光光条图像通常含有光斑和局部断裂,影响光条中心提取精度进而影响测量精度,因此,需要构建图像修复模型去除光斑修复断裂.现有图像修复模型在RGB数据集上的修复效果显著,但不适用于激光光条灰度图像的修复.因此本文提出面向激光光条图像修复的循环相似度映射网络(Recurrent Similarity Mapping Network,RSM-Net),以循环网络为主体,采用软编码Pconv(Partial convolution)层取代部分原始Pconv层,强化特征学习能力;设计非对称相似度模块,降低图像背景特征对修复的负面影响;设计含有多尺度结构相似性(Multi-Scale Structural SIMilarity,MS-SSIM)损失项的混合损失函数,精确地引导光条结构信息的还原,实现高精度的激光光条图像修复.实验验证RSM-Net在小光斑区域、大光斑区域和断裂区域的修复精度均优于所对比的主流图像修复模型.

关键词: 深度学习;图像修复;循环神经网络;相似度映射
中图分类号: TP183 **文献标识码:** A **文章编号:** 0372-2112(2022)05-1234-09
电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20211538

Recurrent Similarity Mapping Network Oriented to Laser Stripe Image Inpainting

JI Zhen-yan[1],HAN Meng-hao[1],SONG Xiao-jun[1],FENG Qi-bo[2]
(1. *Department of Software*, *Beijing Jiaotong University*, *Beijing* 100044, *China*;
2. *Department of Science*, *Beijing Jiaotong University*, *Beijing* 100044, *China*)

Abstract: The vison inspection technology based on line structured light is broadly used in industrial inspection. The images dynamically collected onsite usually contain stripe adhesions and local fractures, which influences the center line extraction accuracy and further measurement accuracy. Thus, it is necessary to construct an image inpainting model to remove the stripe adhesions and fix the local fractures. The existing image inpainting models can achieve high accuracy on RGB datasets, but cannot adapt to laser stripe grayscale images. Therefore, a laser stripe image inpainting network RSM-Net(Recurrent Similarity Mapping Network) is proposed. The kernel of RSM-Net is a recurrent neural network. RSM-Net replaces some original Pconv(Partial convolution) layers with soft-coding Pconv layers to strengthen the feature learning ability. The asymmetric similarity module is designed to decrease the negative impact of image background features on the restoration. A mixed loss function containing the multi-scale structural similarity(MS-SSIM) loss term is designed to precisely guide the restoration of stripe structural information and realize the high-precision laser stripe image inpainting. The experiments demonstrate that the proposed RSM-Net outperforms the compared state-of-the-art image inpainting models in inpainting accuracy for small and large stripe adhesions, and local fractures.

Key words: deep learning;image inpainting;recurrent neural network;similarity mapping

1 引言

结构光视觉检测具有高精度、非接触、结构简单、范围广、实时性强等特点,被广泛应用于轨道交通、机器人视觉引导、地形检测、物体型面检测等工业检测领域,并成为当下的研究热点.

由于结构光视觉检测设备多为室外现场在线监测,容易受到环境干扰、被测物体表面散射和外界杂散光的影响,所采集激光光条图像含有大量区域粘连(光

收稿日期:2021-11-17;修回日期:2022-03-13;责任编辑:朱梅玉
基金项目:国家自然科学基金面上项目(No.52175493);国家自然科学基金重点项目(No.51935002)

斑区域)和局部光条断裂,影响后续中心线提取精度并进而影响测量精度. 因此,构建图像修复模型对激光光条图像进行高精度地去噪和修复,对于工业检测具有重要意义.

图像修复模型主要利用已知的图像信息去修复图像中缺失(待修复)区域的图像信息. 图像修复模型的发展经历了2个阶段:传统算法和深度学习模型. 传统算法可以分为基于扩散的图像修复[1,2]和基于样本的图像修复[3,4]. 基于深度学习的图像修复模型以生成对抗网络(Generative Adversarial Network, GAN)模型[5~7]为主流,文献[8]提出部分卷积网络(Partial convolution, Pconv),通过渐进式修复实现了对不规则孔洞的较高精度修复. 文献[9]提出重建-生成路径的网络提供多样化的修复方案,生成路径在基本的数据分布基础上可构建不同特色的修复图像. 文献[10]和文献[11]改进了文献[8]提出的部分卷积层,文献[10]提出基于门控卷积的修复网络,为特征图添加软编码,增添了网络学习的灵活性和泛化性;文献[11]提出可学习的注意图(Learnable Attention Map, LAM),以取优的软掩码模式将掩码区域取值较高的子区域设为新的掩码区域. 图像修复还可借助图像的边缘信息在图像中填充细节纹理[12~14].

现有的图像修复模型对于人脸、风景等RGB图像修复效果显著,可以满足视觉自然、语义合理等要求,但对于激光光条灰度图像的修复效果不佳,因为颜色相似度信息缺失. 此外,激光光条图像的修复除了对于视觉和语义有要求外,对光条边缘像素点的位置准确度也有要求,边缘像素点位置的准确度直接影响后续光条中心线提取精度,并最终影响被测物体的几何参数解耦精度. 尽管灰度激光光条图像的修复对于基于多线激光视觉的工业检测十分重要,但目前该领域的研究以传统算法为主,精度不高、适应性不强,鲜有采用深度学习方法的研究工作发表,本文的研究为解决基于多线激光视觉的工业检测领域所存在的问题提供了一个新的思路.

针对上述激光光条图像的修复问题,本文提出以循环网络为主体的相似度映射图像修复模型RSM-Net (Recurrent Similarity Mapping Network). 该模型对循环网络进行了如下改进:首先引入软编码Pconv层替换原始Pconv层,强化特征学习能力,有效提升精度;其次提出非对称相似度度量,兼顾特征向量的角度差异和源特征向量响应值,降低图像背景特征对修复的负面影响,更有效地将光条特征信息应用于未知区域的修复优化;设计含有多尺度结构相似性(Multi-Scale Structural SIMilarity, MS-SSIM)损失项的混合损失函数,契合激光光条图像强结构、强纹理的特征,以精确引导光条结构信息的还原,实现高精度的光条图像修复. 本文还构建了多线激光光条图像数据集,实验表明所提出模型在小光斑、大光斑、断裂区域的修复精度和效果均优于所对比的主流修复模型.

2　基于循环神经网络的相似度映射网络

针对多线激光光条图像修复问题,本文提出如图1所示的RSM-Net模型,模型由2个模块组成:相似度映射修复模块和特征融合模块. 相似度映射修复模块的主体是循环神经网络,负责图像特征的修复和修复优化;本模型融合各层循环的输出特征作为后续输入特征以提高模型精度,特征融合模块负责各层循环输出特征的融合以及融合之后的特征处理. 模型损失函数结合绝对差异损失项($L_{hole}+L_{valid}$)、风格损失项(L_{style})、感知损失项($L_{perceptual}$)、全变分损失项(L_{tv})和多尺度结构相似度损失项(L_{ms}),各损失项的作用详见第2.2节. 模型的输入是mask和训练图像的结合,具体结合方法是将2个图像对应的张量进行point-wise乘法运算.

2.1　相似度映射修复模块

相似度映射修复模块负责图像纹理特征的修复和修复后的优化,其输入阶段是连续的2个软编码Pconv层,第一层在修复图像的同时进行降采样操作,使得网络整体在较低的分辨率环境下执行,减少模型的计算复杂度. 输入图像分辨率为384×384,经过两个连续的软编码Pconv层后特征图的分辨率降为192×192.

模块的主体是循环神经网络,通过带自反馈的神经元构成循环,每层循环采用相同结构和参数变量,循环使得网络的输出不仅和当前的输入相关,还和上层循环的输出相关. 循环神经网络的优势:共享参数,因此可以用较少的参数实现纵深较大的网络结构;有利于内嵌的非对称相似度模块在不同层循环上的上下文信息联动,充分发挥循环神经网络的记忆功能. 本模块的循环次数设置为6,输入特征图分辨率为192×192.

考虑到显存负载的限制,每层循环初始是2个连续的原始Pconv层,原始Pconv层负责图像的渐进式修复,Pconv层后是一个类U-Net[15]的编码-解码子网络,编码层是步长为2的卷积层对特征图进行降采样,每个降采样操作将特征图的分辨率尺寸降为原来的1/2;解码层是步长为2的转置卷积对特征图进行升采样,每个升采样操作将特征图的分辨率尺寸升到原来的2倍. 编码-解码子网络总共包含3个编码层和3个解码层. 为了进一步增加循环网络的纵深,在编码层和解码层中间加入5层普通卷积,这5层卷积不改变特征图的分辨率. 非对称相似度模块内嵌在第一和第二个解码层之间,负责特征图修复特征的优化. 此外,编码层的压缩特征通过跳跃连接前向传递给对应分辨率的解码层,指导

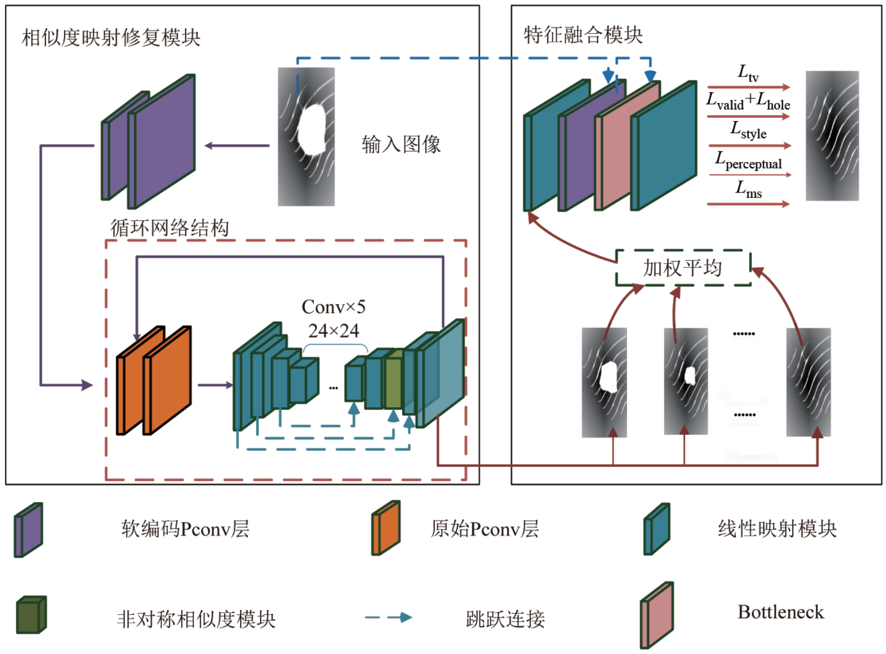


图1 RSM-Net框架

解码层进行低分辨率抽象特征的高分辨率具象化.

2.1.1 软编码Pconv层

RSM-Net模型中的软编码Pconv层是基于NVIDIA团队提出的Partial convolution(Pconv)[8]操作优化的到的.

(1) Pconv层

Pconv[8]曾采用迭代更新的mask策略实现渐进式的图像修复,mask在前向传播过程中逐渐缩小,可利用的图像信息越来越丰富,有利于修复精度的提升.

由于原始Pconv层的掩码取值采用硬编码策略(非0即1),缺少对不同区域mask的区分度;并且特征更新的缩放因子随M^c的值单调递减,但是不同mask区域对应的特征区域对图像待修复区域的影响不是简单的单调递减,因此Pconv层的缩放因子策略不能恰当地反映真实的特征影响分布. 因此,构建软编码Pconv层对原始Pconv层进行改进.

(2) 软编码Pconv层

软编码Pconv层对原始Pconv层进行了2项改进:学习型mask更新策略和非对称高斯缩放因子.

(a) 学习型mask更新策略

原始Pconv层的$k_{1/n}$滤波器矩阵的值均为$1/n$,且固定不变,因此不具有学习性. 软编码Pconv层对滤波器进行改进,将改进后的滤波器定义为K_m,采用深度学习中具有学习性的卷积核滤波器,即软编码滤波器,改进后$M^c=M\odot K_m$. 改进后的滤波器更加适合不规则形状的mask. mask更新公式如下:

$$g_M\left(M^c\right)=(\mathrm{ReLU}(M^c))^a \tag{1}$$

其中,ReLU是激活函数,该激活函数保证更新后的待修复区域mask值置零,与之前相同;在已修复区域上(ReLU正值区间)的mask值具有自适应性. α是比例超参数,用于增强ReLU正值区间的掩码值,当$\alpha=0$时,$g_M(M^c)$会退化为原始Pconv层的$f_M(M^c)$.

(b) 非对称高斯缩放因子

如前文所述,原始Pconv层缩放因子采用单调变化的$f_A(M^c)$,该缩放因子随着卷积窗口内已修复特征(对应mask值为1)的增多缩小,这种考虑过于简单和片面.

特征的缩放因子影响输出特征的响应值,应考虑下述2种情况:首先,当卷积窗口滑动的位置远离待修复区域时,当前卷积窗口对应的M^c值通常是最大的(因为区域M内的值均非0);而当卷积窗口与待修复区域有交集或在待修复区域之中时,随着卷积窗口从待修复区域内部向边界区域移动时(即M^c逐渐变大时),窗口内特征对待修复区域的影响应该是逐渐增大的;在图像修复领域,对修复结果影响最大的特征通常是在待修复区域边界附近,因此远离待修复区域的卷积窗口(M^c数值较大)对应的响应值应设置为较小的值. 综上所述,本文采用非对称高斯函数表示特征缩放因子,表示如下:

$$g_A\left(M^c\right)=\begin{cases} a\exp\left(-\gamma_l\left(M^c-\mu\right)^2\right), & M^c<\mu \\ 1+(a-1)\exp\left(-\gamma_r\left(M^c-\mu\right)^2\right), & \text{其他} \end{cases} \tag{2}$$

其中,a、γ_l、γ_r分别是式(2)的超参数,μ是非对称高斯函

数左右分支的分界点. 该函数形状如图2所示. 图中竖线是μ所在位置,代表待修复区域边缘区域. 由图可见,μ的右半区间代表远离待修复区域的特征区域,其缩放因子要小于μ对应的值;同时μ左半区间代表待修复区域内部的特征区域,该区域缩放因子也小于μ对应的值. 远离待修复区域的特征区域有完整的图像信息,而待修复区域内的图像信息是相对混乱的,因此μ右半区间的缩放因子水平要高于左半区间.

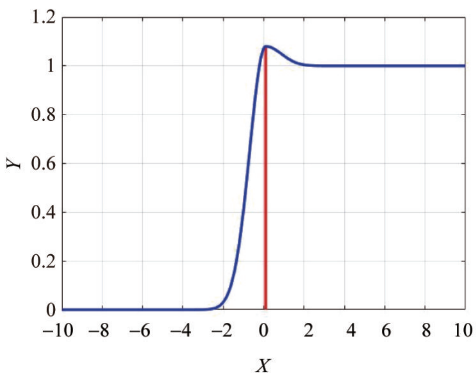


图2　非对称高斯函数示意图

与Pconv层相比,本文引入的软编码Pconv层具有如下优势:采用卷积滤波器提升学习能力,使得模型参数更新更灵活;特征权值更新采用非对称高斯函数,更全面地量化不同特征区域对待修复区域的影响程度.

2.1.2　非对称相似度模块

为进一步优化图像修复效果,本模块依据最近邻相似度搜索思想设计了非对称相似度模块,该模块内嵌在解码层中间. 其优化的作用范围是当前循环第二次原始Pconv层对应mask和上层循环第二次原始Pconv层对应mask的差异区域部分. 非对称相似度模块的作用是对每层循环起始Pconv层的修复结果进行特征优化,以提高修复精度. 非对称相似度模块利用全局的特征相似性进行优化,解决了传统卷积的局部感受野问题.

常用的相似度度量有余弦距离、欧式距离等,当前RGB图像修复领域更多应用余弦距离. 然而,由于激光光条图像是灰度图像,输入的R、G、B分量相同,与RGB图像相比,其特征缺少色度信息,因此在前向传播中容易生成更多的冗余特征图,冗余特征图之间特征相似,用余弦距离会导致两特征向量之间的相似度多接近1,区分度较低;此外,在激光光条图像中,响应值更大的特征往往更加清晰,对待优化区域的影响较大,因此设置更大的相似度映射;响应值较小的值可能对应的是背景信息,应设置较小的相似度映射. 为了使修复特征更加平滑,防止特征的局部落差过大(修复视觉效果上边界感明显),目标特征点(x,y)和源特征点(x',y')之间的相似度通过计算目标特征点和以源特征点(x',y')为中心的$k\times k$区域内的所有源特征点的相似度平均值得到,计算如下:

$$\mathrm{avg\ sim}^i_{(x,y),(x',y')}=\frac{\sum\limits_{p,q\in\{-k,k\}}\mathrm{sim}^i_{(x,y),(x+p,y+p)}}{k\times k}\tag{3}$$

其中,$k\times k$是以源特征点(x',y')为中心的区域大小. 为了得到目标特征点(x,y)的完整相似度映射,式(3)计算得到的值会通过softmax进行处理得到$\mathrm{score}^i_{(x,y),(x',y')}$,这样能够保证$(x,y)$和所有源特征点的相似度映射之和为1,从而计算最后的目标特征值. 此外,本模块的循环网络结构设计不仅能够以较低的训练参数代价获得高纵深的模型,还能实现非对称相似度模块的上下文联动. 具体而言,第i层循环非对称相似度模块的相似度映射结果不仅包含当前循环的计算结果,还综合了第i−1层前向传递的相似度映射,表示如下:

$$\mathrm{score}^i_{(x,y),(x',y')}=\lambda\mathrm{score}'_{(x,y),(x',y')}+(1-\lambda)\mathrm{score}^{i-1}_{(x,y),(x',y')}\tag{4}$$

其中,$\mathrm{score}^i_{(x,y),(x',y')}$是当前循环softmax处理得到的相似度映射,$\mathrm{score}^{i-1}_{(x,y),(x',y')}$是上一层最终计算得到的相似度映射. 为了防止上一层未修复区域的混乱特征误导当前循环特征的优化,λ的最终值基于上一层的mask;如果当前为第一层循环($i=1$),则λ取值为1.

计算完第i层最终的相似度映射,目标特征点的优化特征值由相似度映射和源特征点进行加权求和得到,计算如下:

$$f^i_{x,y}=\sum_{x'\in\{1,W\},y'\in\{1,H\}}\mathrm{score}^i_{(x,y),(x',y')}\cdot f^i_{x',y'}\tag{5}$$

其中,$f^i_{x,y}$是第i层循环目标特征点的优化值,$f^i_{x',y'}$是第i层源特征点(x',y')的特征值. 优化后的特征图和对应mask特征图的点乘结果就是该循环层最终的输出特征图.

2.2　特征融合模块

相似度映射修复模块包含多个循环层,每层输出一个修复后的特征图. 考虑到每层循环对最终修复结果都有贡献,如果以最后一层的输出特征图作为最终输出前向传播,有可能导致网络纵深过大,产生梯度消失问题. 因此特征融合模块负责将多个循环层的输出特征图进行特征融合. 考虑到不同层的输出特征图对应的mask区域大小是不同的,直接point-wise求和取平均会导致图像特征不连续,图像区域出现明显的断层问题.

针对特征直接求和取平均的断层问题,本模块采用更加细粒度的特征融合机制,单独计算特征图中各位置的特征点;考虑到循环结构中的特征修复和特征修复优化的计算都引入了掩码值,在计算各位置的融

合特征时也要结合每个循环的掩码值，最终的特征融合公式如下：

$$\bar{f}_{x,y,z}=\frac{\sum_{i=1}^{N}f_{x,y,z}^{i}}{\sum_{i=1}^{N}m_{x,y,z}^{i}} \tag{6}$$

其中，(x,y,z)是特征点在特征图上的位置，$\bar{f}_{x,y,z}$是位于(x,y,z)位置的融合特征值，$f_{x,y,z}^{i}$是第i层最终输出的位于(x,y,z)的特征值，$m_{x,y,z}^{i}$是第i层循环对应的mask特征图上位于(x,y,z)的掩码值．由式(6)可知，当某一层(x,y,z)的特征点处于未修复状态，则对应的特征值不计入最终的融合特征．

特征融合之后，为了还原到网络的输入分辨率，首先对融合的特征进行上采样，具体采用转置卷积；上采样后图像分辨率增大，特征可能出现损失，因此采用跳跃连接将原始图像的信息与上采样特征图的信息融合，利用原始图像的已知区域特征指导上采样特征图对应区域的特征细化；细化后的特征图先经过一层软编码Pconv层进行处理，再通过一个Bottleneck层融合通道特征实现一定程度地模型减参．软编码Pconv层和Bottleneck层的输出通道维连接后通过一个卷积层输出最终的修复图像．

除了网络结构，本模块的另一创新点是网络的损失函数设计．设计了结合绝对差异损失项、风格损失项、感知损失项、全变分损失项和多尺度结构相似度(Multi-Scale Structural Similarity，MS-SSIM)损失项的损失函数．损失函数表示如下：

$$\begin{aligned}L_{\text{total}}=&\lambda_{\text{hole}}L_{\text{hole}}+\lambda_{\text{valid}}L_{\text{valid}}+\lambda_{\text{perceptual}}L_{\text{perceptual}}\\&+\lambda_{\text{style}}L_{\text{style}}+\lambda_{\text{tv}}L_{\text{tv}}+\lambda_{\text{ms}}L_{\text{ms}}\end{aligned} \tag{7}$$

绝对差异损失项包括2部分：已知区域差异L_{valid}和mask区域差异L_{hole}. L_{valid}使得不需要修复的已知区域的特征差异收敛，有利于mask区域的修复；L_{hole}使得mask区域的特征差异收敛，使修复效果尽可能向ground-truth靠近．感知损失$L_{\text{perceptual}}$能够保证修复图像在风格上与ground-truth图像相同，从另一个角度保证修复结果与源图像的特征一致性，在图像修复领域中被广泛应用．风格损失项L_{style}与感知损失项功能相似，都是保证修复前后图像的特征风格一致性，但是具体思路不同，在图像修复中通常联合使用以达到更好的修复效果．全变分损失项L_{tv}用于维持图像修复效果的平滑性，防止图像局部区域尤其是修复边界区域出现局部特征落差过大的问题．MS-SSIM损失项L_{ms}能够保证激光光条的结构特征被更加准确地修复还原，并且考虑了不同图像分辨率的问题．

MS-SSIM损失项从亮度、对比度和结构3个方面约束修复图像和ground-truth图像之间的差异，包含M个分辨率，其由原始分辨率做2倍下采样，迭代$M-1$次得到．MS-SSIM损失项表达如下：

$$L_{\text{ms}}=1-\left[l_M(x,y)\right]^{\alpha_M}\sum_{i=1}^{M}\left[c_i(x,y)\right]^{\beta_i}\left[s_i(x,y)\right]^{\gamma_i} \tag{8}$$

其中，$l_M(x,y)$是第M个分辨率的亮度分量，$c_i(x,y)$和$s_i(x,y)$分别是第i个分辨率的对比度分量和结构分量，α_M是亮度分量对应的比重参数，β_i和γ_i分别是对比度分量和结构分量的比重参数．

MS-SSIM可最大程度地保留图像边缘和纹理结构等高频信息，适合激光光条图像的修复．

综上所述，RSM-Net具有以下优势：(1)采用非对称相似度模块取代余弦相似度模块．在角度之外考虑了源特征向量的激活值，激活值较低表示特征不明显或是背景信息；(2)采用软编码Pconv层替代传统的Pconv层，增加了渐进式修复过程中的可学习性，同时用更契合实际的特征缩放函数描述不同特征向量对待优化区域特征的影响；(3)引入MS-SSIM损失项，加强对图像结构特征的还原，以更精确地修复激光光条的边缘信息．

3 激光光条图像修复实验

为了验证RSM-Net在激光光条图像上的性能优势，设计实验对比了RSM-Net模型和多个主流图像修复模型的修复精度．

3.1 实验数据集

目前缺少高质量的激光光条图像数据集，为了验证模型有效性，构建了多线激光光条图像数据集，数据来源于企业采用轮对智能在线监测设备在列车真实运行环境下所采集的图像．

采集到的轮对激光光条图像尺寸为1280×1028，考虑到完整图像分辨率过大，导致实际训练效率低、显存负载压力过大，难以实现，且光斑区域和断裂区域都是局部问题，其修复更多利用周边局部区域的信息，因此本文的训练数据集是对采集到的原始轮对光条图像裁剪而成．考虑到训练模型的泛化性和通用性，本文对所采集图像中的区域进行随机裁剪，裁剪尺寸是384×384.

随机裁剪的区域背景信息所占比例不能过高，过高会导致裁剪的图像中大部分都是黑色背景，不利于模型提取光条特征．因此本实验采用的裁剪方案是一种基于比例阈值的窗口扫描法，即高光区域(值为255的像素点区域)像素点个数/总像素点个数(384×384)的比例不低于0.0001. 光条训练数据集包含16478张图像，每张图像分辨率384×384，训练数据集样例如图3所示．

轮对光条数据集还包括光斑区域测试集和断裂区域测试集，样例如图4所示．光斑区域测试集包含254

图3　训练集示意图

张图像,断裂区域测试集包含70张图像.测试掩码图区域为人工标注获得,ground-truth图像是由摄像机在理想黑暗环境下采集.

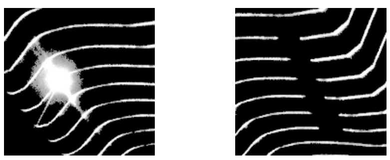


图4　测试集示意图

3.2　基准方法

实验对比了4个主流图像修复模型:经典模型Pconv-Net[8](Partial convolution Network)、对partial convolution进行了优化的LBAM[11](Learnable Bidirectional Attention Maps)、发表于2019年CVPR会议的代表模型PIC-Net[9](Pluralistic Image Completion Network)、高精度模型RFR-Net[16](Recurrent Feature Reasoning Network).

3.3　评价指标

本实验所采用的图像修复精度指标为:平均绝对误差MAE、均方误差MSE、峰值信噪比PSNR和结构相似性SSIM(Structural Similarity).MSE相较于MAE对异常点更为敏感,如果有个别修复后的像素值与ground-truth值差异较大,会引起MSE更大的偏差.可视误差PSNR客观性强,但与人类视觉系统的主观评价有出入,原因在于人的视觉系统对图像的主观感受不仅受到图像像素差异的影响,而且人眼对亮度的灵敏度比色度高.SSIM可以同时提取亮度、对比度、结构三种图像特征,更加符合人类的主观感受.

3.4　实验环境和模型超参数

实验环境硬件配置为:Intel i9-10900X处理器、NVIDIA GeForce RTX 3080显卡.软件配置为:CUDA10.1+Cudnn7.6.5并行计算框架、Python3.7编程语言、Pytorch1.7深度学习框架.RSM-Net模型和中间模型采用的超参数相同,参数值设置见表1.

表1　模型超参数

参数	释义	取值
batch_size	训练时批量输入的每批图像数	2
iteration	模型训练的迭代次数	3×10^5
optim	模型训练的优化器	Adam
lr	模型优化器的学习率	2×10^{-4}
λ_{valid}	有效区域绝对损失项的权重	1
λ_{hole}	待修复区域绝对损失项的权重	6
λ_{style}	风格损失项的权重	120
$\lambda_{perceptual}$	感知损失项的权重	0.05
λ_{tv}	全变分损失项的权重	0.1
λ_{ms}	MS-SSIM损失项的权重	0.1

3.5　实验结果及分析

本实验分阶段提出3种模型:Variant1、Variant2、RSM-Net.如表2所示,Variant1和Variant2的主要区别在于相似度模块采用的相似度度量不同,前者计算特征向量之间的余弦相似度,后者计算优化的非对称相似度.Variant2和RSM-Net的区别在于Pconv层,前者采用传统Pconv-Net模型的编码层,后者采用优化的软编码Pconv层.

表2　Variant1,Variant,RSM-Net模型结构对比

模型	相似度映射修复模块	特征融合模块
Variant1	循环网络+余弦相似度模块+原始Pconv层	特征融合+L_{total}损失函数
Variant2	循环网络+非对称相似度模块+原始Pconv层	特征融合+L_{total}损失函数
RSM-Net	循环网络+非对称相似度模块+软编码Pconv层	特征融合+L_{total}损失函数

3.5.1　光斑区域实验对比

设计实验对比了Variant1、Variant2、RSM-Net模型和4个主流图像修复模型在多线激光光条图像数据集上的光斑区域修复精度.实验结果如表3所示,最佳结果加粗显示.Variant1在所有精度指标上都超越了所对比的主流模型.Variant2对Variant1进行了优化,将余弦相似度模块改进为非对称相似度模块,明显提升了精度.RSM-Net在4个精度指标上全面超越Variant2,其对Variant2的优化是将传统Pconv层改进为软编码Pconv层,软编码Pconv层增加了渐进式修复的可学习性,同时用非对称高斯函数描述不同图像区域特征对待修复区域特征优化的影响分布.

图5、图6给出了RSM-Net模型和4个baseline模型对2个具体的光斑区域实例进行修复的实验结果对比.图5展示小光斑区域的修复效果对比,图6展示大光斑区域的修复效果对比.

如图5所示,RSM-Net模型在小光斑区域可以实现

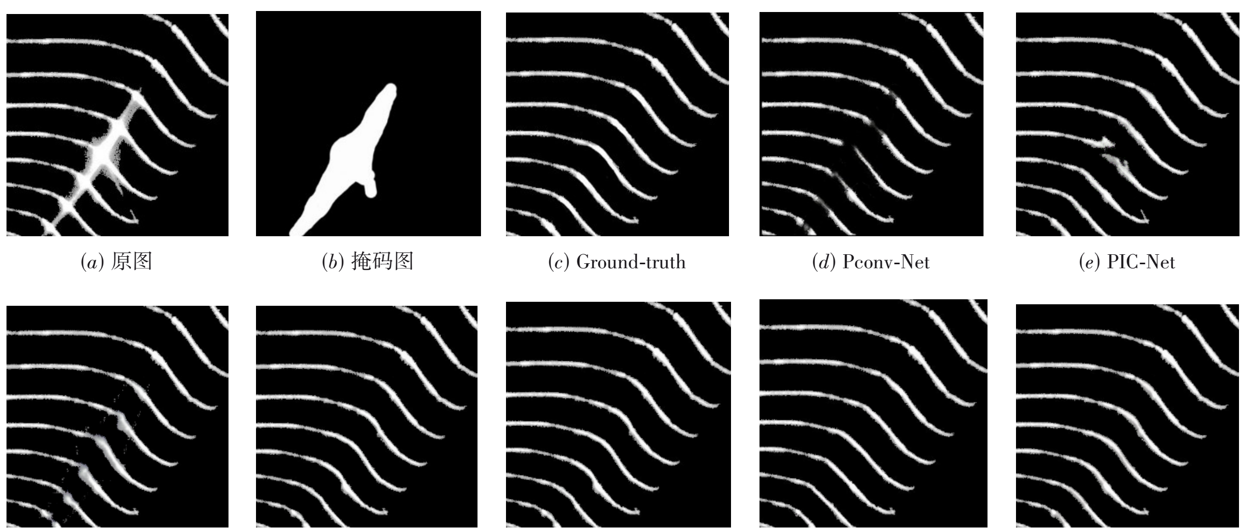
表3　模型精度对比（光斑区域）

模型	指标			
	值越低,精度越高		值越高,精度越高	
	MAE	MSE	PSNR	SSIM
Pconv-Net[8]	3.0331	493.5388	21.1976	0.959149
PIC-Net[9]	2.9202	426.6285	21.8303	0.932834
LBAM[11]	2.9605	507.9912	21.0722	0.957249
RFR-Net[16]	2.8189	481.4003	21.3057	0.962518
Variant1	2.4387	388.7194	22.2344	0.964595
Variant2	2.0234	277.3013	23.7013	0.968528
RSM-Net	**1.7837**	**217.6479**	**24.7533**	**0.972458**

精确的修复；Variant1，Variant2 和 RFR-Net 在该区域也能实现较精确的修复；而 LBAM 的修复效果欠佳，虽然能修复光条，但是修复的光条存在宽度不均衡、边缘不平滑的问题；PIC-Net 和 Pconv-Net 则无法完整地修复光条，修复结果中仍然存在局部断裂，Pconv-Net 断裂情况尤为明显.

如图6所示，RSM-Net模型在大光斑区域仍然可以实现精确的修复，而后6种模型的修复结果都存在不同程度的局部断裂，其中 LBAM 和 PIC-Net 存在光条左右部分连接不匹配问题.

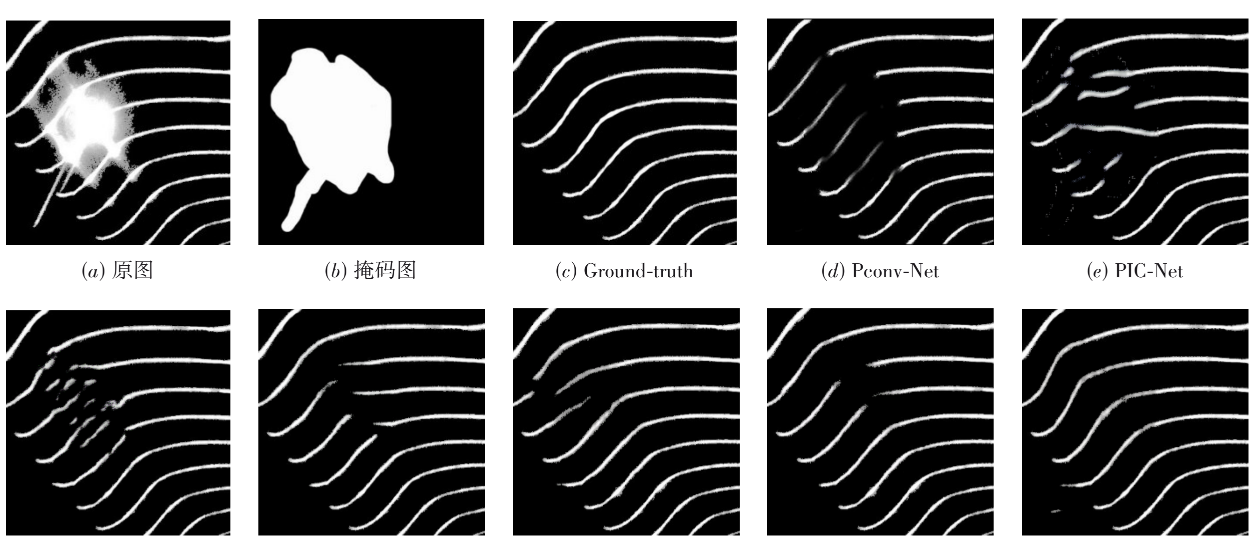
综上所述，本文提出的 RSM-Net 模型无论在小光斑区域还是大光斑区域上的修复效果都显著优于当前主流的图像修复模型.



(*a*) 原图　(*b*) 掩码图　(*c*) Ground-truth　(*d*) Pconv-Net　(*e*) PIC-Net

(*f*) LBAM　(*g*) RFR-Net　(*h*) Variant1　(*i*) Variant2　(*j*) RSM-Net

图5　修复样例(小光斑区域)



(*a*) 原图　(*b*) 掩码图　(*c*) Ground-truth　(*d*) Pconv-Net　(*e*) PIC-Net

(*f*) LBAM　(*g*) RFR-Net　(*h*) Variant1　(*i*) Variant2　(*j*) RSM-Net

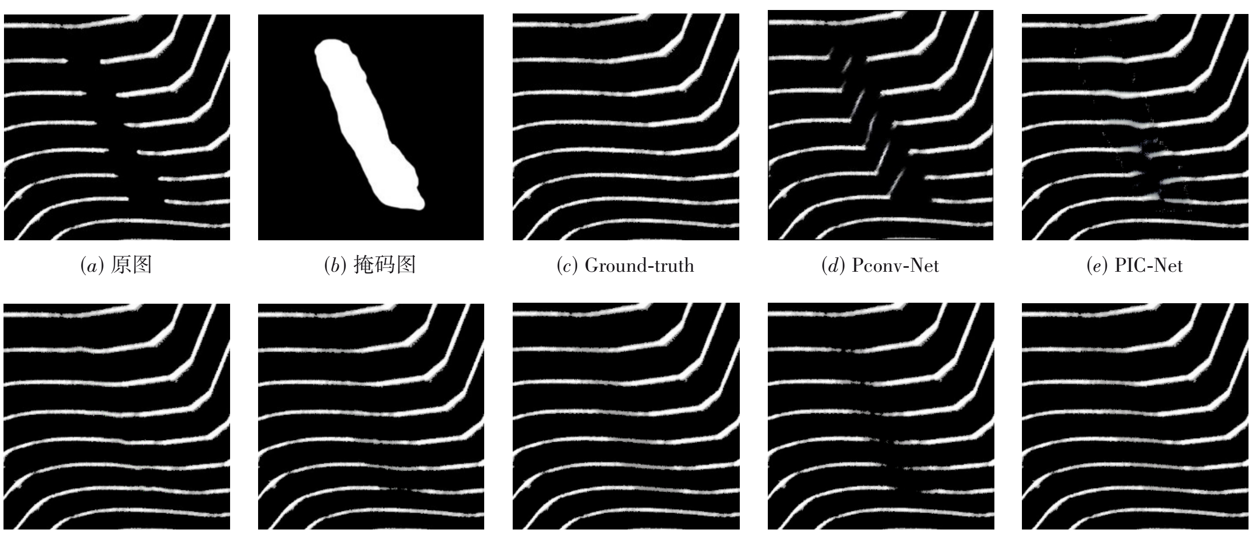
图6　修复样例(大光斑区域)

3.5.2 断裂区域实验对比

本文设计实验对比了 Variant1、Variant2、RSM-Net 模型和4个主流图像修复模型在轮对多线激光光条图像断裂区域的修复精度,实验结果如表4和图7所示.表4中最佳实验结果加粗显示.与光斑区域相比,光条断裂区域的待修复特征简单,在前3个精度指标上,LBAM 模型表现更佳,但是 RSM-Net 在最契合主观视觉判断的 SSIM 指标上结果最佳.整体精度对比,Variant2 和 RSM-Net 各有优势;图像修复效果对比图如图7所示,RSM-Net 在光条断裂区域的修复效果与 ground-truth 图像基本一致;Variant1 生成的光条不光滑;而 Variant2 和 RFR-Net 修复的光条部分存在局部断裂问题;LBAM 能完整修复光条,但是该模型修复的光条走向与 ground-truth 不完全贴合,存在局部突出问题;PIC-Net 和 Pconv-Net 都存在明显的修复断裂问题,并且 Pconv-Net 还存在光条左右部分连接不匹配问题.

表4　模型精度对比(断裂区域)

模型	指标			
	值越低,精度越高		值越高,精度越高	
	MAE	MSE	PSNR	SSIM
Pconv-Net[8]	4.7255	819.4150	18.9958	0.929092
PIC-Net[9]	2.6930	395.1818	22.1628	0.936993
LBAM[11]	**1.4592**	**157.6043**	**26.1551**	0.959429
RFR-Net[16]	1.7819	207.5207	24.9602	0.967641
Variant1	1.7822	193.2016	25.2707	0.969305
Variant2	1.5875	158.2233	26.1381	0.970926
RSM-Net	1.6486	158.2327	26.1378	**0.971448**



(*a*) 原图　(*b*) 掩码图　(*c*) Ground-truth　(*d*) Pconv-Net　(*e*) PIC-Net

(*f*) LBAM　(*g*) RFR-Net　(*h*) Variant1　(*i*) Variant2　(*j*) RSM-Net

图7　修复样例(断裂区域)

综上所述,本文提出的 RSM-Net 在光条局部断裂区域的修复效果显著优于当前主流的图像修复模型.

4 结论

针对激光光条灰度图像修复问题,本文提出基于深度学习的修复模型 RSM-Net,可高精度地修复激光光条图像的光斑和断裂,修复精度和效果均优于当前主流图像修复模型.考虑到激光检测多用于嵌入式检测设备,设备资源有限,对运行其上的模型有轻量化要求,未来将探索模型的轻量化,在满足精度要求的同时实现模型的轻量.

参考文献

[1] BERTALMIO M, SAPIRO G, CASELLES V, et al. Image inpainting[C]//Proceedings of the 27th Annual Conferenc on Computer graphics and interactive techniques, New York: ACM, 2000: 417-424

[2] 高成英, 徐仙儿, 罗燕媚, 等. 基于稀疏表示的物体图像修复[J]. 计算机学报, 2019, 42(9): 1953-1965.
GAO C Y, XU X E, LUO Y M, et al. Object image inpainting based on sparse representation[J]. Chinese Journal of Computers, 2019, 42(9): 1953-1965. (in Chinese)

[3] CRIMINISI A, PÉREZ P, TOYAMA K. Region filling and object removal by exemplar-based image inpainting [J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200-1212.

[4] CHENG W H, HSIEH C W, LIN S K, et al. Robust algorithm for exemplar-based image inpainting[C]//Proceedings of International Conference on Computer Graphics, Imaging and Visualization. Beijing: IEEE, 2005: 64-69.

[5] PATHAK D, KRÄHENBÜHL P, DONAHUE J, et al. Context encoders: feature learning by inpainting[C]//2016 IEEE conference on Computer Vision and Pattern Recognition. Las Vegas, NV: IEEE, 2016: 2536-2544.

[6] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics(ToG), 2017, 36(4): 1-14.

[7] SAGONG M C, SHIN Y G, KIM S W, et al. PEPSI: fast image inpainting with parallel decoding network[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach, CA: IEEE, 2019: 11352-11360.

[8] LIU G, REDA F A, SHIH K J, et al. Image inpainting for irregular holes using partial convolutions[C]//Proceedings of the European conference on computer vision (ECCV). Munich :Springer, 2018: 85-100.

[9] ZHENG C X, CHAM T J, CAI J F. Pluralistic Image Completion[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach, CA: IEEE, 2019:1438-1447.

[10] YU J H, LIN Z, YANG J M, et al. Free-form image inpainting with gated convolution[C]//2019 IEEE/CVF International Conference on Computer Vision(ICCV). Seoul: IEEE, 2019: 4470-4479.

[11] XIE C H, LIU S H, LI C, et al. Image inpainting with learnable bidirectional attention maps[C]//2019 IEEE/CVF International Conference on Computer Vision(ICCV). Seoul: IEEE, 2019: 8857-8866.

[12] XIONG W, YU J H, LIN Z, et al. Foreground-aware image inpainting[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach, CA: IEEE, 2019: 5833-5841.

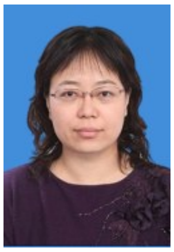
[13] NAZERI K, NG E, JOSEPH T, et al. Edgeconnect: structure guided image inpainting using edge prediction[C]//2019 IEEE/CVF International Conference on Computer Vision Workshops(ICCVW). Seoul: IEEE, 2019: 3265-3274.

[14] WANG M H, CHEN X M, CHEN W Q, et al. MSMC-net: Image inpainting using deep multi-scale and multi-connection networks[C]//2019 IEEE International Conference on Multimedia & Expo Workshops. Shanghai: IEEE, 2019: 683-686.

[15] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Munich: Springer, 2015: 234-241.

[16] LI J Y, WANG N, ZHANG L F, et al. Recurrent feature reasoning for image inpainting[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA: IEEE, 2020: 7757-7765.

作者简介



冀振燕 女,1972年4月出生,河南人.博士,现为北京交通大学软件学院副教授,博士生导师.主要研究领域为计算机视觉、推荐系统等.
E-mail: zhyji@bjtu.edu.cn

韩梦豪 男,1999年5月出生,山西人.现为北京交通大学软件学院硕士研究生.主要研究方向为计算机视觉、深度学习等.

宋晓军 男,1995年9月出生,山东人.研究生,毕业于北京交通大学软件学院.主要研究方向为计算机视觉、深度学习等.

冯其波 男,1962年5月出生,广东人.博士,现为北京交通大学理学院教授,博士生导师.研究方向为光学测量、铁路安全测量技术、仪器仪表、计算机视觉等.