

轻量级无监督深度学习水下图像拼接算法^{*}胡帅晨^{1,2} 张鹏^{1,2} 崔敏³(1. 中北大学仪器与电子学院 太原 030051; 2. 中北大学电子测量技术国家重点实验室 太原 030051;
3. 中北大学半导体与物理学院 太原 030051)

摘要: 传统拼接方法在复杂场景下表现不佳,且有监督方法标注数据困难,现有的无监督图像拼接方法模型参数大,拼接时间长。因此,提出了一种轻量级无监督深度学习的图像拼接框架,分为无监督图像变形网络和无监督图像融合网络两个阶段。在图像变形网络中,使用 MobileNetV2 作为网络主干,结合 ECA 模块得到图像的变形信息。图像融合模块使用 UNeXt 作为主干网络得到图像重叠区域的接缝线进行无缝拼接,通过添加 AG 模块以及改进 Tokenized MLP 模块来提高精度。此外,由于缺乏水下图像拼接的数据集,本文还构建了真实场景下的无监督水下图像拼接的数据集,并在该数据集以及公开数据集 UDIS-D 上比较了 SIFT+Ransac, ORB+Ransac, UDIS 算法和 UDIS++ 算法,实验结果表明,本文算法在保证拼接精确度的同时将模型的参数量下降了 74%,拼接速度提升了 46%。

关键词: 水下图像;图像拼接;MobileNetV2;UNeXt

中图分类号: TP391;TN0 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Lightweight unsupervised deep learning algorithm for
underwater image stitchingHu Shuaichen^{1,2} Zhang Peng^{1,2} Cui Min³(1. School of Instrumentation and Electronics, North Central University, Taiyuan 030051, China;
2. State Key Laboratory of Electronic Test Technology, North Central University, Taiyuan 030051, China;
3. School of Semiconductors and Physics, North Central University, Taiyuan 030051, China)

Abstract: Traditional stitching methods perform poorly in complex scenes, and supervised methods face challenges due to the difficulty of annotating data. Existing unsupervised image stitching methods suffer from large model parameters and long stitching times. Therefore, a lightweight unsupervised deep learning-based image stitching framework is proposed, which consists of two stages: an unsupervised image deformation network and an unsupervised image fusion network. In the image deformation network, MobileNetV2 is used as the backbone, combined with the ECA attention mechanism module to obtain image deformation information. The image fusion module employs UNeXt as the backbone network to generate seamless stitching by identifying the seam lines in the overlapping regions of the images. The accuracy is improved by incorporating the AG module and enhancing the tokenized MLP module. Additionally, due to the lack of datasets for underwater image stitching, a real-world unsupervised underwater image stitching dataset is constructed. Comparative experiments are conducted on this dataset and the publicly available UDIS-D dataset, evaluating SIFT+Ransac, ORB+Ransac, UDIS, and UDIS++ algorithms. The experimental results demonstrate that the proposed algorithm reduces the model parameters by 74% and improves stitching speed by 46% while maintaining stitching accuracy.

Keywords: underwater image; image stitching; MobileNetV2; UNeXt

0 引言

水下图像是水下自主机器人实现导航、定位和测绘等

功能的重要方式^[1],其质量和分辨率直接影响到后续的分析 and 应用效果。然而,由于水下环境的特殊性,如光线衰减、水质浑浊等因素,使水下图像的采集面临诸多挑战。此

外,单张图像所能覆盖的范围有限,难以满足需求。因此,图像拼接技术在水下图像处理中扮演着至关重要的角色。

图像拼接是指将具有重叠区域的图像合并成一张宽视角或全景图的技术。在水下图像处理领域,有效的图像拼接不仅可以扩大视野范围,还能提高图像的利用率和分析效率。传统图像拼接方法主要基于特征匹配和透视变换。特征匹配包括特征点检测、特征点描述符提取和描述符匹配算法^[2]。一些常见的特征点检测方法,如 ORB + Ransac^[3]、SIFT + Ransac^[4]等,往往受限于水下图像特有的低对比度和纹理不明显等问题,导致拼接精度不高,拼接速度慢,甚至出现错误匹配的情况^[5]。

近年来,基于深度学习的图像拼接算法逐渐成为研究热点^[6]。这些方法通过训练神经网络模型来学习图像的特征表示和拼接规则,相比传统算法在特征提取方面能够学习图像高级特征,具有更好的精度和鲁棒性^[7]。基于深度学习的图像拼接算法可以分为有监督算法和无监督算法。有监督的图像拼接算法依赖成对输入图像及其对应的标注拼接结果,流程包括数据准备、特征提取、监督训练和预测后处理。而无监督算法不需要标注,流程包括图像预处理、特征匹配、几何估计、自监督优化和拼接融合。Detone 等^[8]在 2016 年首次提出了一种有监督 DL 卷积神经网络来估计一对图像之间的相对单应性,输入两张图,它直接产生两幅图像之间的同形关系。Nie 等^[9]提出了一个大基线 DL 单应性模型,用于估计不同特征尺度下参考图像和目标图像之间的精确投影变换,但由于合成图像中缺乏真实的场景视差,其泛化能力受到限制。Nie 等^[10]提出的 VFIS-Net 是有监督的深度学习图像拼接框架,可以实现自由视角的图像拼接,增强了网络在真实数据集下灵活视图的泛化能力。但是该算法不能实现任意分辨率

图像的拼接,以上这些基于有监督学习的方法,由于真实的图像拼接数据集获取难度大,标记数据集时间成本高,只能在人工合成的数据集上完成训练,这导致算法在真实场景下表现不佳。随后,Nie 等^[11]在 VFIS-Net 的基础上提出的图像拼接算法 UDIS 是无监督图像拼接的开山之作,他们发现图像拼接后像素域上的不对齐在特征域上得到了削弱,因此提出了基于两阶段的图像拼接策略。Nie 等^[12]后续在单应性矩阵估计中引入样条运动,通过局部扭曲方式完成图像配准,使得网络在大视差场景下更具鲁棒性。但这些算法存在训练时间长、计算复杂度高且拼接速度较慢等问题。

针对传统拼接算法的不足、有监督学习数据标注困难以及现有的无监督深度学习网络模型复杂、训练时间长、拼接速度慢^[13],且对于复杂的水下环境拼接效果并不理想。因此,本文针对水下场景图像的拼接问题,提出了一种轻量级的基于无监督深度学习的水下图像拼接算法,无需手动标注训练数据。在图像变形阶段利用卷积神经网络自动提取图像的特征,并通过学习得到参考图像与目标图像之间的变换矩阵,结合薄板样条变换(thin plate splines, TPS)方法得到对齐后的图像。在图像融合部分将对齐后的图像作为输入,利用卷积神经网络输出合成掩码,这个掩码是一个与图像相同尺寸的矩阵,掩码中的每个值表示在该位置参考图像和目标图像的融合比例。根据掩码的值将对齐后的图像按比例融合,得到无缝的拼接图像。

1 无监督图像拼接网络结构设计

本文提出的轻量级无监督深度学习水下图像拼接网络由无监督图像变形网络和无监督图像融合网络两部分组成。框架如图 1 所示。

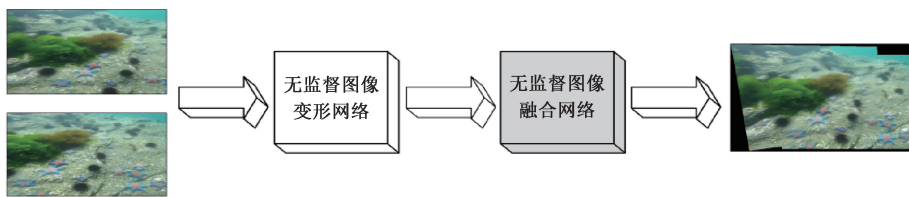


图 1 无监督图像拼接网络框架

Fig. 1 Unsupervised image Mosaic network framework

1.1 无监督图像变形网络

在无监督图像变形网络中,网络结构如图 2 所示,使用 MobileNetV2^[14]作为主干网络,MobileNetV2 是一种轻量级的卷积神经网络,通过倒置残差结构、深度可分离卷积和线性瓶颈层平衡了计算效率和特征表达能力。在传统的残差结构中,输入和输出通道数较多,而中间的卷积层通道数较少,会导致低维空间中的信息丢失。而在倒置残差模块中,输入和输出通道数较少,中间深度可分离卷积层(DWise)通道数较多,形成一个倒置的瓶颈结构。这样的设计可以在高维空间中进行更有效的特征提取,避

免低维信息过早压缩导致信息丢失。深度可分离卷积通过将传统卷积分解为深度卷积和逐点卷积分别处理空间维度和通道维度的信息,显著降低参数量和计算成本,同时保留特征提取能力。线性瓶颈层是指在残差块的降维阶段,移除非线性激活函数(ReLU),仅保留线性变换。这是因为降维后的低通道空间直接使用 ReLU 等非线性激活函数,可能导致信息丢失。在倒置残差中使用了跳跃连接,有助于缓解梯度消失的问题。

另外,本文在倒置残差模块的第一个卷积层之后插入通道注意力机制模块(efficient channel attention, ECA),

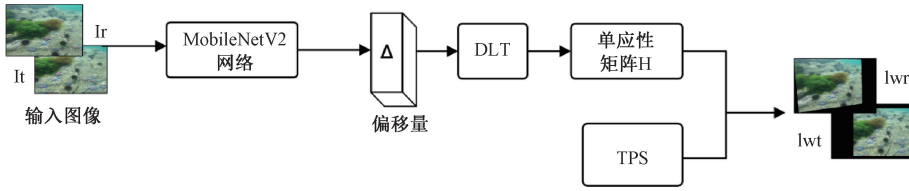


图 2 无监督图像变形网络框架

Fig. 2 Unsupervised image deformation network framework

如图 3 所示。ECA^[15] 模块是一种轻量且高效的通道注意力机制,工作原理是通过自适应地确定一维卷积的核大小 k ,以学习每个通道相对于其他通道的重要性,从而提升整体性能。同时,ECA 模块的计算复杂度较低,不会显著增加模型的计算开销,并通过注意力机制有效提升特征提取效率。ECA 模块结构如图 4 所示,先将输入的特征图进行全局平均池化操作,然后进行卷积核大小为 k 的 1 维卷积操作,经过 sigmoid 激活函数得到各个通道的权值,将权值与原始输入特征图对应元素相乘,得到最终的特征图。通过提取的特征点对,进行单应性矩阵(H)的估计,再结合 TPS 方法,能够在保证重叠区域对齐精准的同时,保持非重叠区域的形状。 H 用于描述了两个平面之间的几何变换关系,主要解决图像对的线性变换任务,如平移、旋转等问题。TPS 是一种常用于图像配准和变形的非刚性变换方法,它通过一组控制点,将一幅图像中的关键点映射到另一幅图像中的对应位置,解决图像对的非线性变换问题,如拉伸、弯曲、压缩等。通过这两个方法联合优化,提高了处理图像视差的能力。

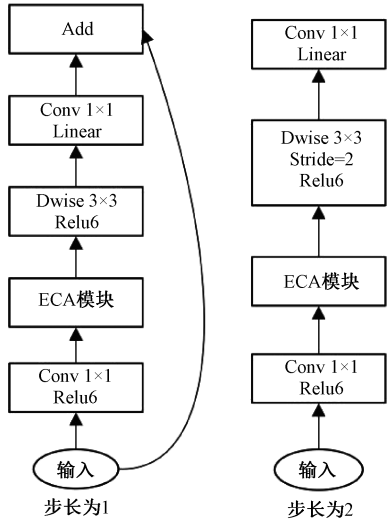


图 3 改进 MobileNetV2 网络结构

Fig. 3 Improved MobileNetV2 network architecture

在具体实现中,输入图像通过 MobileNetV2 一系列精心设计的层次结构高效地提取图像特征。首先,输入图像经过预处理后,被送入一个标准的 3×3 卷积层进行初步特征提取。随后,图像进入多个倒置残差块,每个块先通过 1×1 点卷积扩展通道数,然后通过 ECA 模块对高维特

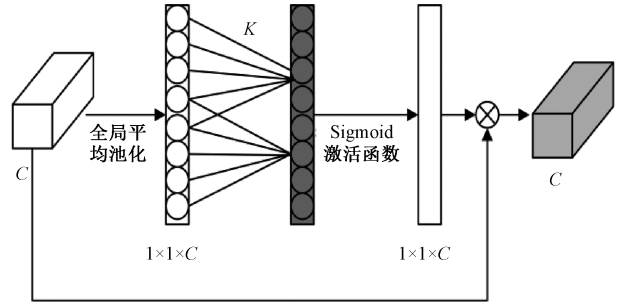


图 4 ECA 模块结构图

Fig. 4 ECA module structure diagram

征进行通道注意力校准,接着使用深度可分离卷积进一步提取特征,并通过 1×1 卷积降维,同时采用线性激活函数。在特征提取的最后阶段,使用 1×1 点卷积生成最终的特征图。在上下文相关层(CCL)匹配特征点得到偏移量 Δ ,CCL 是一种用于深度学习中的特征匹配技术,它通过计算特征图之间的相关性来捕捉图像间的语义关系,它能够在没有标签的情况下自动寻找不同图像之间的对应关系。然后使用直接线性变换(DLT)算法计算张量单应性矩阵 H ,DLT 公式如式(1)、(2)所示, H 结合 TPS 变换,实现图像之间的平滑过渡和变形,再结合偏移量 Δ 得到最终变换后的图像对。TPS 公式如式(3)、(4)所示。

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = H \cdot \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (1)$$

$(x_i, y_i, 1)^T$ 和 $(u_i, v_i, 1)^T$ 两幅图像对应的特征点对, H 是一个 3×3 的齐次矩阵,其形式为:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (2)$$

TPS 样条函数的形式为:

$$\mathcal{F}(x, y) = \alpha_1 + \alpha_x x + \alpha_y y + \sum_{i=1}^p w_i U(\| (x_i, y_i) - (x, y) \|) \quad (3)$$

其中,前 3 个系数 ($\alpha_1, \alpha_x, \alpha_y$) 表示能通过所有控制点 (x_i, y_i) 最逼近 x' (或 y') 的线性平面, p 是控制点个数, w_i 表示每个控制点的控制权重, $U(\| (x_i, y_i) - (x, y) \|)$ 表示核函数, $U(r) = r^2 \log(r)$ 。扭曲前的点坐标为 P , $g(x, y)$ 和 $h(x, y)$ 分别表示 x 和 y 方向的形变,则扭

曲之后对应点的坐标为 P_0 。

$$\begin{cases} x_0 = x - g(x, y) \\ y_0 = y - h(x, y) \end{cases} \quad (4)$$

1.2 无监督图像融合网络

本文的无监督图像融合网络框架如图 5 所示,使用改进 UNeXt 作为主干网络。UNeXt 是一种基于卷积多层感知器(MLP)的图像分割网络,编码器-解码器体系结构,解码器和编码器都包含卷积和 Tokenized MLP 两个阶段,每个编码器块减少特征分辨率 2 倍,每个解码器块增加特征分辨率 2 倍。在编码器中,前 3 个块是卷积,后两个是 Tokenized MLP 块,在解码器中,前两个块是 Tokenized MLP,后面 3 个是卷积块。编码器的卷积块使用带有池窗口 2×2 的最大池化层,而解码器的卷积块使用双线性插值层对特征图进行上采样。使用双线性插值而不是转置卷积,因为转置卷积基本上是可学习的上采样,会产生更多可学习的参数。Tokenized MLP 模块结合了卷积神经网络和 MLP 的优点,用于深层特征提取。Tokenized MLP 块通过标记化、轴向移位、Shifted MLP 处理以及残差连接

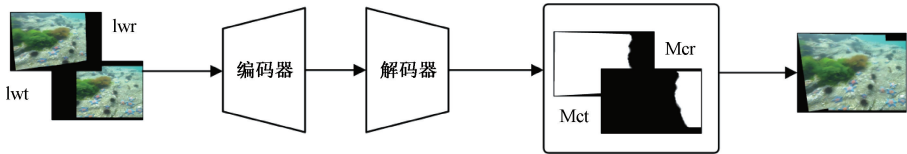


图 5 无监督图像融合网络框架

Fig. 5 Unsupervised image fusion network framework

UNeXt 网络通过减少卷积层数量,采用更大的特征尺寸进行下采样,并集成了 Tokenized MLP 模块以降低计算复杂度,同时维持分割精度。然而,由于其感受野相对受限,可能导致该网络更侧重于局部特征的捕捉,而对全局信息的理解不够充分。为了在保持模型轻量化的同时提升分割精度,对 UNeXt 网络实施了改进。第 1 个改进是 Tokenized MLP 模块中增加一条新路径,如图 6 所示,先在 height 方向平移特征,经过卷积操作后,再通过 width 方向平移的

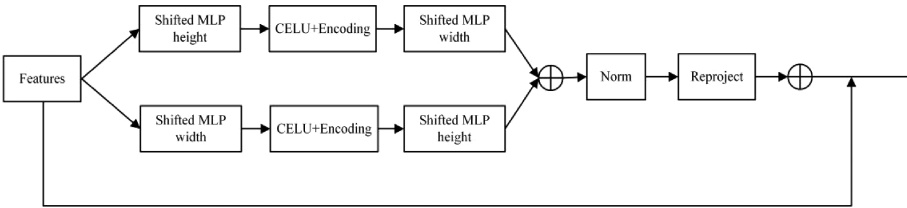


图 6 改进 Tokenized MLP 结构

Fig. 6 Improved Tokenized MLP structure

具体实现的过程,使用网络的编码器逐步减小输入图像的空间尺寸,同时增加特征图的深度分别从参考图像(I_{wrf})和目标图像(I_{wtf})中提取语义特征,并共享权重。这些特征通常包括边缘、纹理、形状等信息。使用跳跃连接,通过从 I_{wrf} 的特征中减去 I_{wtf} 的特征并将残差在每个分辨

等步骤,实现了对图像特征的有效提取和建模,同时保持了较高的计算效率和较低参数复杂度。具体工作原理是首先使用 3×3 卷积将输入特征映射到嵌入维度 E,这个 E 就是标记(token)的数量。接着,特征被送入两个 Shifted MLP 进行处理。Shifted MLP 是一种变体的 MLP 模型,它通过引入位移操作来增强模型的特征提取能力。与传统的 MLP 不同,Shifted MLP 在每个 MLP 块之间加入了位移操作,使得模型能够更好地捕获图像中的局部信息。第 1 个 Shifted MLP 沿宽度方向移动特征,这个 MLP 通过深度卷积层(DWConv)来传递特征,并使用高斯误差线性单元(GELU)激活函数完成激活操作。DWConv 的使用有助于编码位置信息,并且相比标准卷积使用的参数更少,提高了效率。GELU 激活函数的目的是通过引入随机性来提高神经网络的性能和泛化能力。第 2 个 Shifted MLP 则沿高度方向移动特征。移位操作使得同一通道下包含不同 token 的信息。最后使用一个残差连接,将原始 token 添加为残差,然后利用归一化层,将输出特征传递到下一个块。

Shifted MLP 处理。最终将原始路径与新路径的输出特征融合,以增强细节信息的提取能力。第 2 个改进是在前 3 个模块的编码器与解码器之间的跳跃连接中引入注意力门(attention gate, AG)模块,AG 模块是一种软注意力机制,通过动态权重分配引导网络聚焦于目标相关区域,同时抑制无关背景干扰。能够有效过滤编码阶段上采样过程中产生的冗余特征,减少噪声和伪影的干扰,从而提升计算效率和分割精度。改进的 UNeXt 网络结构如图 7 所示。

率下传递到解码器。跳跃连接将编码器层的输出与解码器层的输出直接连接,允许网络在不同层次之间传递特征信息,保留了更多的空间信息,从而提高了分割的精度和效率。解码器接收来自编码器的语义特征以及跳跃连接传递来的残差信息,通过上采样和卷积操作,将这些特征

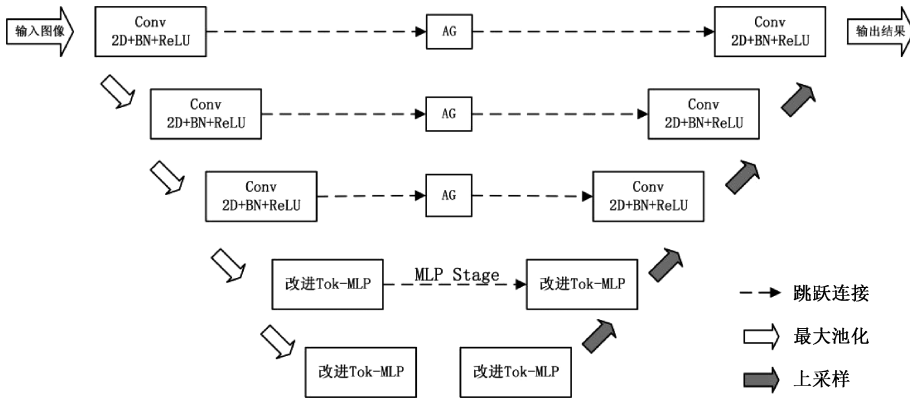


图 7 改进 UNeXt 网络结构

Fig. 7 Improve the UNeXt network structure

重新组合成拼接后的图像。本文设置最后一层的滤波器数量和激活函数分别为 1 和 sigmoid, 这样的设置使得输出是一个二值化的掩码矩阵, 其中 1 表示参考图像的区域, 0 表示目标图像的区域。以预测扭曲参考图像的掩码矩阵 M_{cr} 。扭曲目标图像的另一个掩码矩阵 M_{ct} 可以通过简单的后处理获得。最后将变形后的图片 (I_{wr} 和 I_{wt}) 与掩码矩阵 (M_{cr} 和 M_{ct}) 组合得到无缝拼接的图像 S , S 的计算公式如式(5)所示。

$$S = M_{cr} \times I_{wr} + M_{ct} \times I_{wt} \quad (5)$$

2 实验设计与结果对比分析

2.1 实验环境与数据集

本文的网络模型的训练在 Ubuntu22.4 操作系统中完成。CPU 为 Intel(R)Core(TM)i9-10980XE CPU@3.00 GHz, 系统内存 8 GB, GPU 为 NVIDIA 3090Ti。训练基础软件环境为 CUDA 12.4、python 3.8.5、pytorch 1.7.1、OpenCV 4.5.5、Tensorflow 0.8.2, 网络采用 adma 优化器, 在无监督变形网络中, 学习率设置为 0.000 1, 批次大小为 4, 训练 100 个 epoch。无监督融合网络使用变形网络的训练结果进行再一次的训练, 学习率设置为 0.000 1, 批次大小为 4, 训练 50 个 epoch。

2.2 数据集

本文自建的水下图像数据集中包括 3 000 对图像, 按照 8:1:1 的比例划分训练集、验证集、测试集。分别包含低亮度、非均匀光照、蓝绿色调等多种情况, 尽可能包含水下图像出现的所有结果。并且随机抽数据集中的图像对进行翻转、旋转等操作来提高数据的多样性和模型的泛化能力。

2.3 评价指标

本文分别使用结构相似性 (SSIM)、峰值信噪比 (PSNR)、拼接时间和模型参数量作为主要评价指标, SSIM 和 PSNR 的计算公式如式(6)~(8)所示, SSIM 值越高, 表示图像之间越相似。PSNR 的值越高, 表示处理后的图像与原始图像的相似度越高^[16]。

$$SSIM(x, y) = \frac{(2u_x u_y + c_1)(\sigma_{xy} + c_2)}{(u_x^2 + u_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

其中, σ_x 和 σ_y 分别代表图像 x 和 y 的标准差; σ_{xy} 表示 x 与 y 的协方差; u_x 和 u_y 分别代表图像 x 和 y 的均值; c_1 、 c_2 是常数。

$$MSE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n [I(i, j) - K(i, j)]^2 \quad (7)$$

$$PSNR = 20 \cdot \lg \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (8)$$

其中, MSE 表示图像间的均方误差, MAX_I 表示图像中所有像素点的最大值的二机制位数, 若图像是 8 位的, MAX_I 就为 255。

2.4 实验结果与分析

为了验证本文算法的性能, 在本文的数据集上进行了测试, 选取 SIFT+Ransac、ORB+Ransac 算法和 UDIS、UDIS++ 算法分别作为传统方法和深度学习方法的代表, 对 SSIM、PSNR、模型参数量和拼接时间做对比, 模型参数量的对比只针对深度学习的方法。拼接时间是 10 对分辨率为 $4\ 896 \times 6\ 528$ 的待拼接图像拼接成功的平均时间, 测试结果如表 1 所示。

表 1 本文数据集上的比较结果

Table 1 Comparison results on the dataset used in this study

方法	平均 PSNR	平均 SSIM	模型参 数量/M	平均拼接 时间/s
SIFT+Ransac	20.237	0.694	—	15.56
ORB+Ransac	19.563	0.607	—	7.89
UDIS	24.362	0.732	96.57	3.36
UDIS++	26.782	0.885	77.96	2.83
本文算法	28.561	0.913	20.14	1.52

从表中可以看出, 本文的算法的平均 PSNR、平均 SSIM 以及平均拼接时间 3 个指标均优于 SIFT+Ransac、ORB+Ransac 和 UDIS 算法, 虽然平均 PSNR 和平均

SSIM 略高于 UDIS++ 算法,但本文算法的模型参数量相对 UDIS++ 算法减少了 74%,并且拼接时间均减少了将近一半。实验结果表明,本文提出的图像拼接网络模型不仅轻量化,而且在保证拼接精度的同时也大幅提高了图像拼接的速度,证明了本文网络的有效性和可靠性。

为了进一步验证本文算法的适用性,将 UDIS、UDIS++ 和本文算法在公开数据集 UDIS-D 上进行测试,测试结果如表 2 所示。本文算法的平均 PSNR 和 SSIM 均高于 UDIS 算法,与 UDIS++ 算法指标接近,再次验证了本文算法的精度。

表 2 UDIS-D 数据集上的比较结果

方法	平均 PSNR	平均 SSIM
UDIS	21.17	0.648
UDIS++	25.43	0.838
本文算法	25.02	0.874

为了直观的对比,本文对部分测试结果进行展示,分

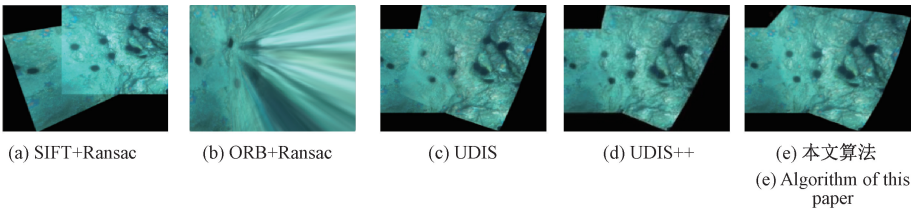


图 10 蓝色调图像拼接结果
Fig. 10 Blue-tone image stitching result

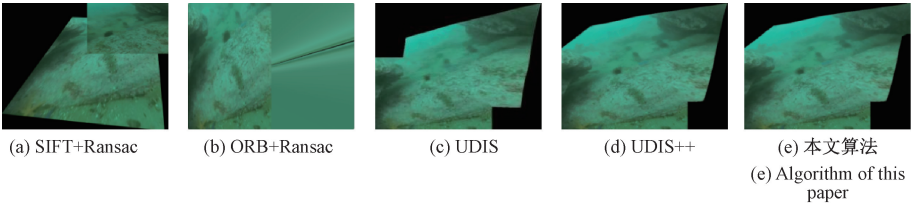


图 11 绿色调模糊图像拼结果
Fig. 11 Green-tone image stitching result

从上面的实验结果中可以看出,SIFT+Ransac 和 ORB+Ransac 这两个传统的拼接算法在上述两种情况下的拼接结果很差,UDIS 算法结果相对较好,但也能看到有较明显的接缝线,并且拼接后图像变得模糊,而本文算法和 UDIS++ 算法的拼接结果都比较清晰,没有出现模糊和接缝线的问题。

3 结 论

本研究提出了一种轻量级无监督深度学习方法,用于水下图像拼接,旨在解决传统水下图像拼接算法中存在的计算复杂度高、对光照变化敏感等问题。通过引入

别选取蓝色调模糊图像对和绿色调模糊图像对。拼接前的图像对如图 8、9 所示,对应的拼接结果如图 10、11 所示。

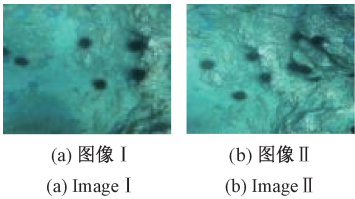


图 8 蓝色调待拼接图像对
Fig. 8 Blue tone to be stitched image pair

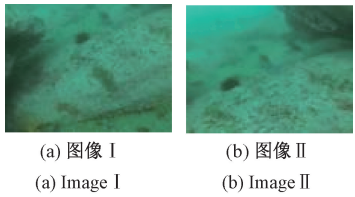


图 9 绿色调待拼接图像对
Fig. 9 Green tone to be stitched image pair

先进的深度学习架构和优化策略,能够在保证拼接质量的同时,显著降低计算资源的需求,为实时或近实时的水下图像处理提供了可能。与现有的先进方法相比,本文算法在公开数据集上均展现出了优异的性能,特别是在处理具有挑战性的水下环境(如低光照、浑浊水质)时,仍能保持较高的拼接精度和很快的拼接速度。本研究不仅为水下图像拼接领域带来了新的技术突破,也为利用深度学习技术解决海洋科学、水下考古、生态监测等领域的实际问题开辟了新途径。在接下来的工作中,将进一步挖掘算法的优化空间,追求更高的拼接效率与更优的质量。

参考文献

- [1] 刘贤三, 栾振东, 黄端. 基于水下机器人的海底视频图像拼接技术探讨[J]. 广西科学, 2023, 30(3): 543-550.
LIU X S, LUAN ZH D, HUANG D. Discussion on submarine video image Mosaic technology based on underwater robot[J]. Guangxi Science, 2023, 30(3): 543-550.
- [2] 李英浩, 牛少彰, 郝启贤, 等. 基于 Shi-Tomasi 和 RootSIFT 的多尺度曲率特征图像拼接算法[J]. 电子测量技术, 2024, 47(14): 42-48.
LI Y H, NIU SH ZH, HAO Q X, et al. Multi-scale curvature feature image stitching algorithm based on Shi-Tomasi and RootSIFT[J]. Electronic Measurement Technology, 2024, 47(14): 42-48.
- [3] 王朝虎, 卢洪义, 吴文勇, 等. 航空发动机叶型图像检测神经网络配准算法[J]. 电子测量技术, 2025, 48(8): 55-70.
WANG CH H, LU H Y, WU W Y, et al. Neural network registration algorithm for aeroengine blade profile image detection[J]. Electronic Measurement Technology, 2025, 48(8): 55-70.
- [4] 孙艳丽, 周伟, 张亚周, 等. 基于 SIFT 的航空侦查图像拼接算法[J]. 电子测量技术, 2020, 43(19): 112-116.
SUN Y L, ZHOU W, ZHANG Y ZH, et al. Aerial reconnaissance image stitching algorithm based on SIFT[J]. Electronic Measurement Technology, 2020, 43(19): 112-116.
- [5] 李晓明, 郝沙沙, 陈双慧. 结合先验知识的海底图像配准方法[J]. 计算机辅助设计与图形学学报, 2023, 35(11): 1743-1750.
LI X M, HAO SH SH, CHEN SH H. Submarine image registration method combined with prior knowledge[J]. Journal of Computer-Aided Design & Computer Graphics, 2023, 35(11): 1743-1750.
- [6] 杨利春, 田彬, 党建武. 基于深度学习的图像拼接算法研究综述[J]. 计算机应用研究, 2024, 41(7): 1930-1939.
YANG L CH, TIAN B, DANG J W. A review of image mosaic algorithms based on deep learning[J]. Application Research of Computers, 2024, 41(7): 1930-1939.
- [7] 赵帅, 张春堂, 樊春玲. 基于深度特征提取的深海序列图像拼接网络[J]. 电子测量技术, 2025, 48(3): 180-187.
ZHAO SH, ZHANG CH T, FAN CH L. Deep-sea sequence image stitching network based on depth feature extraction [J]. Electronic Measurement Technology, 2025, 48(3): 180-187.
- [8] DETONE D, MALISIEWICZ T, RABINOVICH A. Deep image homography estimation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 5668-5676.
- [9] NIE L, LIN C, LIAO K, et al. Deep rectangling for image stitching: A learning baseline[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5740-5748.
- [10] NIE L, LIN C, LIAO K, et al. A view-free image stitching network based on global homography[J]. Journal of Visual Communication and Image Representation, 2020, 73: 102950.
- [11] NIE L, LIN C, LIAO K, et al. Unsupervised deep image stitching: Reconstructing stitched features to images[J]. IEEE Transactions on Image Processing, 2021, 30: 6184-6197.
- [12] NIE L, LIN C, LIAO K, et al. Parallax-tolerant unsupervised deep image stitching[C]. 2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 7365-7374.
- [13] 梁镇锋, 夏海英, 谭玉枚, 等. 基于无监督深度学习的航拍图像拼接算法[J]. 北京航空航天大学学报, 2025, 51(7): 2437-2449.
LIANG ZH F, XIA H Y, TAN Y M, et al. Aerial image Mosaic algorithm based on unsupervised deep learning [J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(7): 2437-2449.
- [14] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 4510-4520.
- [15] 刘林, 廖子阳. 基于多通道图像的 ECA-CNN WiFi FTM 室内定位算法[J]. 仪器仪表学报, 2024, 45(10): 323-332.
LIU L, LIAO Z Y. ECA-CNN WiFi FTM indoor positioning algorithm based on multi-channel images[J]. Chinese Journal of Scientific Instrument, 2024, 45(10): 323-332.
- [16] 练洋奇, 陈新度, 吴磊, 等. 高分辨率瓷砖图像实时拼接算法[J]. 电子测量技术, 2023, 46(7): 144-150.
LIAN Y Q, CHEN X D, WU L, et al. Real-time stitching algorithm for high-resolution tile images[J]. Electronic Measurement Technology, 2023, 46(7): 144-150.

作者简介

胡帅晨(通信作者), 硕士研究生, 主要研究方向为数字图像处理。

E-mail: hushuaichen@126.com

张鹏, 副教授, 硕士生导师, 主要研究方向为智能感知、图像处理、自动化测试与控制。

E-mail: sxyczhangpeng@126.com

崔敏, 副教授, 硕士生导师, 主要研究方向为测试计量技术及仪器、自动化测试与控制、惯性感知与测控技术。

E-mail: cmcm_1980930@163.com