

改进基于多相机的无监督学习图像拼接算法

杨旭朝 雷志勇 王娇娇

(西安工业大学电子信息工程学院 西安 710021)

摘要:首先,针对大范围特征不显著地面图像的获取,由于传统单相机受到视场角和分辨率以及外界环境的影响,提出基于多相机对大范围地面的图像进行采集。其次针对大范围特征不显著地面图像使用传统方法拼接中存在特征提取能力差以及拼接产生的错位、伪影和结构变形问题,提出对无监督学习图像拼接的框架结构改进,提升图像拼接质量。最后借助图像拼接的评价指标对改进前后的拼接图像进行评价。实验结果表明,改进后方法不仅可以有效的解决传统方法基于特征不显著以及纹理相似度高的地面图像拼接中产生的伪影和失真,而且解决了无监督图像拼接过程中产生的结构的变化。改进后的拼接方法拼接后的质量优于传统方法的拼接,并且改进后的算法的迁移性很高,不止针对无显著特征的大范围地面,而且可以广泛用于不同场景下的大范围大基线图像拼接。

关键词:多相机;特征不显著;无监督学习;质量评价;大范围图像拼接

中图分类号: TCP391 **文献标识码:** A **国家标准学科分类代码:** 520.604

Improved unsupervised learning based on multiple cameras image stitching algorithm

Yang Xuzhao Lei Zhiyong Wang Jiaojiao

(Key Electronic Information Engineering College, Xi'an Technological University, Xi'an 710021, China)

Abstract: Firstly, for the acquisition of large-area feature-insignificant ground images, we propose the acquisition of large-area ground images based on multiple cameras because the traditional single camera is affected by the field of view and resolution as well as the external environment. Secondly, we propose to improve the framework structure of unsupervised learning image stitching to improve the image stitching quality. Finally, the stitched images before and after the improvement are evaluated with the help of the evaluation index of image stitching. The experimental results show that the improved method can not only effectively solve the artifacts and distortions generated in the traditional method based on the ground image stitching with insignificant features and high texture similarity, but also solve the changes of the structure generated in the unsupervised image stitching process. The quality of the improved stitching method in this paper is better than that of the traditional method of stitching, and the improved algorithm is highly migratory and can be widely used for stitching a large range of large baseline images in different scenes, not only for a large range of ground without significant features.

Keywords: multi-camera; feature insignificance; unsupervised learning; quality evaluation; large scale image stitching

0 引言

传统普通单相机拍摄图像时存在相机视场角和分辨率无法兼顾的问题,尽管全景相机能解决但是全景相机采集到的图像存在畸变和失真,同时成本较高不适合低成本场景^[1]。为了使用普通单相机拍摄宽视场,提出基于多相机图像拼接技术。通过对多相机采集到的图像进行拼接

完成大范围图像重构^[2]。其中,图像拼接技术是指将两幅或多幅含有重叠部分的图像,通过图像配准和图像融合,拼接成一幅包含各图像信息的高分辨率、宽视角图像的技术。

传统图像拼接技术最先应用在遥感技术领域^[3],其核心是图像配准。传统的图像拼接解决方案大多是基于特征的方法,其中特征检测是关键,直接地影响到拼接性能。

然后使用匹配的特征建立一个参数化的图像对齐模型。通过这个模型可以将目标图像扭曲,使其与参考图像对齐。最后,缝合后的图像可以通过给每个像素分配像素值来获得^[4]。如文献[5-8]提出在不同的领域和场景下进行图像的拼接。

传统的基于特征的解决方案通过如下两种主流方式来缓解。第1种方法是通过将目标图像与参考图像尽可能地对齐来消除伪影^[9-18]。该方法将一幅图像分割成不同的区域,并计算出这些方法将图像划分为不同的区域并计算每个不同区域的单应矩阵。通过对这些区域施加空间变化的扭曲进行对齐,结果显示该区域伪影显著减少。第2种方法是通过研究最佳接缝来隐藏接缝。扭曲的图像^[19-23]。然后,根据两个区域形成一个缝合的图像。尽管基于特征的解决方案可以显著减少大多数场景中的伪影。但是,它们仍然严重依赖特征检测,因此缝合性能在很大程度上依赖于特征检测,因此缝合性能在特征少或分辨率低的场景中,拼接性能会急剧下降甚至失败。

由于卷积神经网络(CNN)具有较强的特征提取能力,因此,最近基于深度学习方面的应用已经在各种领域取得较大的成果,如深度估计^[24]、光流估计领域,如深度估计^[25]、光流估计^[26]、单目视觉^[27]、扭曲矫正^[28]。越来越多的研究人员尝试将CNN应用于图像拼接。在文献[29-30]中,CNNs仅用于提取特征点,而在文献[31-33]中,CNN被提议用于拼接具有固定观看位置的图像。但是,这些方法不是一个完整的基于学习的框架或者只能用于缝合具有固定视角的图像,而不是任意视角的图像。然后,无视图的深度图像拼接方法^[34]被提出来以同时克服这两个问题。然而,所有的解决方案都是有监督的方法,而且由于无法获得真实场景中的缝合标签,所以获取真正的基于有监督方法的图像缝合数据集工作量较大,不符合实际场景中的图像拼接。

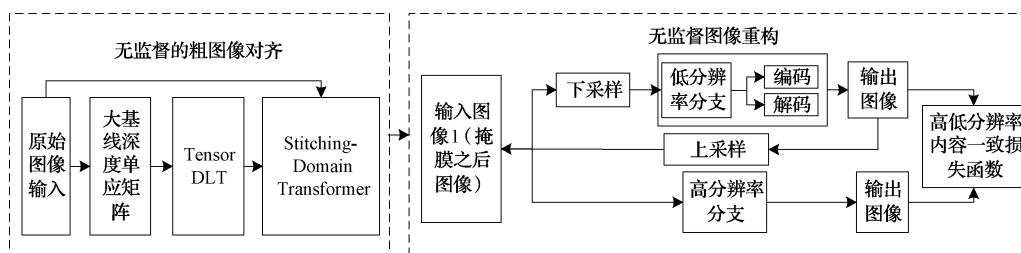


图2 无监督的图像对齐与重构

其中图像对齐模块是对两张输入的图像通过深度无监督单应矩阵进行对齐,并使用缝合域转换层对不同重叠率输入的图像输出不同分辨率的缝合图像。同时,图像重构模块分为两个分支即高分辨率分支和低分辨率分支。

2.1 图像对齐原理

给定两幅高分辨率的输入图像,首先使用深度单应网络以无监督的方式估计单应性。其次输入图像通过被扭曲

基于上述存在的各种问题,本文改进了基于无监督的深度图像拼接框架,包括无监督的图像粗对齐阶段改进和无监督的图像重建阶段改进,并提供了第1个用于无监督深度图像拼接的大型真实地面数据集,其中包括不同场景下的地面图像。同时,希望它能作为一个地面基准数据集,并促进其他相关的研究工作。经本文改进后的算法迁移性较高、鲁棒性较好,包括在真实场景中的单应估计方案和图像拼接方案。

1 图像的采集改进

传统大范围地面图像采集是基于单相机进行图像采集,通过移动相机的位置进行大范围图像的采集,该方法在进行图像采集时,容易受到外界环境的影响,导致待拼接图像质量差。因此,本文提出基于多相机对大范围特征不显著地面图像进行采集,通过该方式可有效减少外界干扰如天气、光照等因素的影响。为图像拼接提供优质图像。本文多相机的摆放形式包括两种,平行摆放和非平行位置摆放。多相机的摆放位置示意图如图1所示。

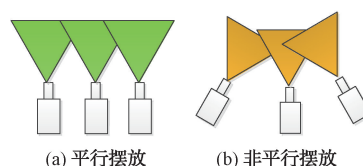


图1 多相机摆放位置示意图

其中,调整多相机之间的距离和夹角之间的关系,获取包含重叠区域的图像。

2 基于无监督学习的图像拼接

本文网络由两个部分组成包括图像对齐和图像重构,其中重构部分包括两个分支即低分辨率和高分辨率分支,其网络模型整体^[36]如图2所示。

进行对齐。其中深度单应矩阵的求解是图像对齐的关键。

单应估计是一种基本的图像对齐方法。通过提取和匹配稀疏特征点来实现,在弱光、低纹理、纹理相似、稀疏特征点图像中容易出错。同时,以前的深度单应方法要么使用合成图像进行有监督的学习,这两种方法都忽略了处理存在深度差异和移动对象在实际应用中的重要性,并且深度单应矩阵的求解的准确性也影响了图像重构的质量,

为了克服这些问题,本文使用一种新的改进后的架构设计,即无监督深度单应方法。改进后的网络结构解决了大

基线的图像对齐问题并提升了对齐的程度,改进后的深度单应网络主干结构如图3所示。

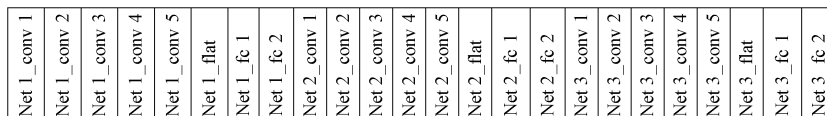


图3 改进后的深度单应矩阵循环网络

2.2 无监督重构网络的网络结构

鉴于单一同构图只能表示同一深度的空间变换的局限性,在图像对齐阶段,输入图像不能在真实世界的数据集完全对齐。为了突破单一同构图的问题,本文改进了文献[35]提出了从特征到像素重建缝合

后的图像。图4所示为所提出的无监督深度图像拼接模型框架的概况。重建网络可以由两个分支实现,低分辨率变形分支(图4上部)和高分辨率精化分支(图4下部),分别学习图像拼接的变形规则和提高分辨率。

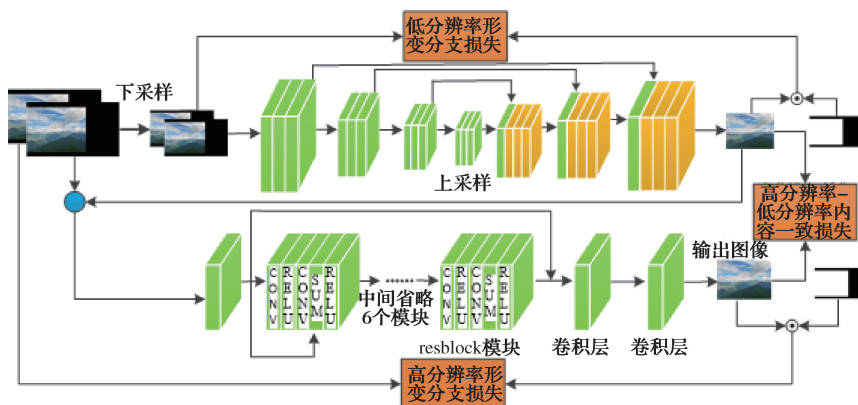


图4 无监图像拼接网络结构模型

2.3 重构网络结构改进

1) 低分辨率网络改进

仅在高分辨率下重建图像分支是不合理的,因为感受野会随着分辨率的提高而相对减少随着分辨率的提高而相对减少。为了确保网络的接收为了确保网络的感受野能够完全感知错位的区域(特别是在高分辨率和大视差的情况下视差),本文在原来低分辨率分支的网络基础上增加网络的深度,并引入残差网络解决增加网络深度可能带来的梯度消失和梯度爆炸的情况。如图5所示,扭曲的图像首先被下采样到一个低的分辨率,在本文的实现中定义为 256×256 。然后由5个卷积和5个反卷积层组成的编码-解码网络被用来重建缝合后的图像,模型示意图如图6所示。卷积层的过滤器数量被分别设置为64、64、64、64、128、128、256、256、256、256、512、512、256、256、256、128、128、64、64、64、64。此外,采用跳过连接的方式来连接低级和高级特征。在这个过程中,图像拼接的变形规则是通过内容掩码和接缝掩码来学习。

2) 高分辨率网络改进

在低分辨率的初始形变之后,本文改进了一个高分辨率的细化分支^[35],以提高分辨率并细化缝合后的图像。高分辨率分支的分辨率是指第一阶段的输出分辨率。实

际上,在本文数据集中,将分辨率大于 512×512 设置为高分辨率分支。该分支完全由卷积层组成,如图7所示,该分支可以处理任意分辨率的图片。改进后的网络由4个独立的卷积层和8个resblock块,其中除了最后一层的滤波器数被设置为3,其余每层的滤波器数都被设置为64。为了防止低层次的信息随着卷积网络的深入而逐渐被遗忘,改进后的网络在第1层与倒数第2、3层加入卷积层。此外,每个resblock都由conv、relu、conv、sum和relu组成。

3 实验及结果分析

3.1 实验环境及参数设置

本文的实验中数据集由两部分组成,包括自制地面数据集和无监督图像拼接数据集。其中数据集中在照片共11 638张,按照9:1的比例划分为训练集、验证集,其中训练集包含10 474张图片,验证集包含1 164张图片。本文选取6种客观评价指标对拼接后的图像进行评价,实验在Windows操作系统下进行,其中CPU为Intel(R)Core(TM)i7-10750H,内存为16 G, GPU为NVIDIA GeForce RTX1650,采用tensorflow作为深度学习框架。初始学习率为0.0001,批处理大小为30,优化算法采用随机梯度下降,衰减系数为0.4,最大迭代次数为100 000次。

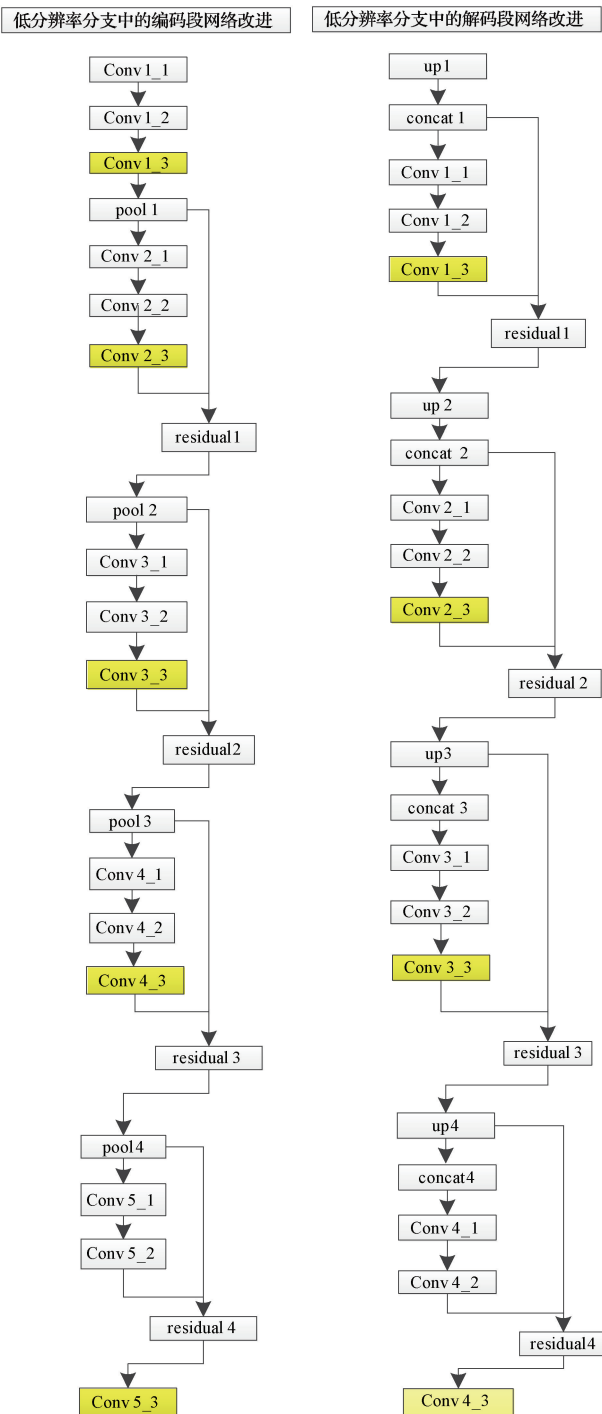


图5 低分辨率网络改进后示意图

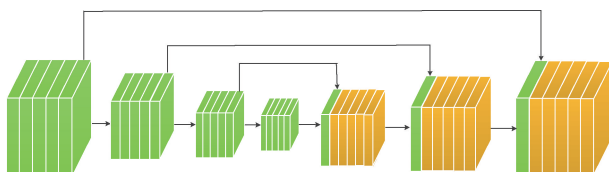


图6 低分辨率网络改进后模型

3.2 图像拼接评价指标

本文采用6种图像客观评价指标来衡量拼接图像的质量^[36],这些指标为拼接图像质量评价的常用指标,能够客观说明图像的拼接质量以及待拼接图像的重合率。包括均方误差(MSE)、峰值信噪比(PSNR)、均方根误差(RMSE)、熵值(Entropy)、信噪比(SNR)、结构相似度(SSIM)。

3.3 实验

为了验证本文改进后的算法与传统算法以及改进之前算法之间的差异,本文提出两种方案进行图像拼接。第1种方案,是选择纹理相似和特征不显著的地面图像作为待拼接图像,使用不同的算法进行图像拼接,并选择熵值作为评价指标进行对比。为了验证本文改进后拼接算法的迁移性和鲁棒性,使用第2种方案,即在不同场景下完成拼接,拼接完成之后使用图像拼接的6种客观评价指标进行评价。

1) 不同算法对比实验

实验通过改进模型实现了传统图像拼接的图像配准和融合,为检验训练后模型对图像的拼接效果,将其与传统经典算法和基于深度学习算法进行比较。本文选择SIFT+RANSAC经典算法和基于深度学习中的无监督(unsupervised)算法作为对比。

不同算法在不同场景下的图像拼接效果如图8所示,其中包含不同的场景如本文使用不同算法在相同拼接对象下进行实验,实验结果显示基于算法SIFT+RANSAC拼接后的图像均存在伪影,如基于SIFT+RANSAC算法拼接后红色框标记区域,这也是传统图像拼接存在的普遍问题。其次,基于无监督算法拼接后的图像虽然解决了传统图像拼接过程中的伪影,但是存在两种问题,第1种即特征提取能力较弱,图8(a3)、(a5)和(a6)中无监督图像拼接后图像,如红框标记所示。第2种问题即拼接后的图像重现结构变形和失真,图8(a4)、(a7)和(a8)中无监督图像拼接后图像(红框标记)。

经过上述实验基于SIFT+RANSAC、unsupervised算法拼接之后的图像质量较差。主要存在的问题在于传统方法拼接之后产生的伪影和无监督算法拼接之后图像中的物体产生扭曲和变形。

最后,经过本文改进后的无监督图像拼接算法拼接后的图像不仅解决了传统图像拼接后出现的伪影,而且解决了未改进之前无监督图像拼接出现的特征不显著和结构失真和变形的问题,图8(e)的拼接结果表明,大大提升的图像拼接的质量。同时,借助客观评价指标进行评价,本文选用Entropy评价指标对基于不同算法拼接后的图像进行评价并进行分析,如表1所示。

通过表1得出,基于无监督算法拼接后的图像大多数比SIFT+RANSAC算法拼接后的图像质量优。本文算法的拼接质量均优于Unsupervised算法和SIFT+

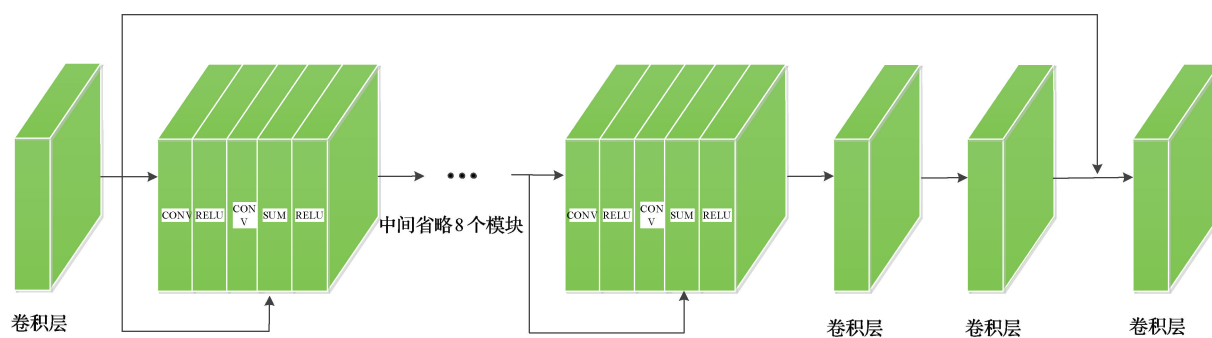
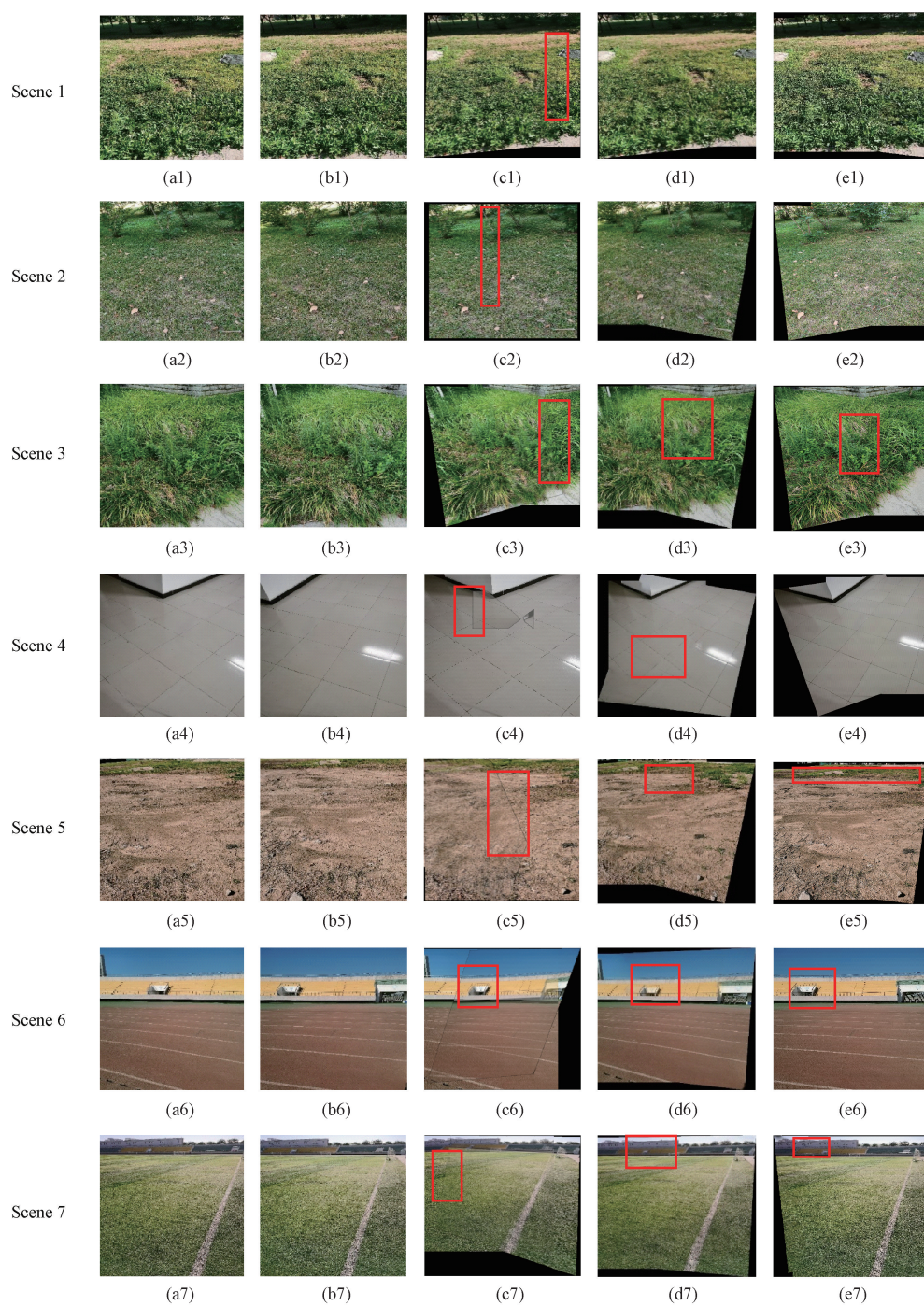


图7 高分辨率网络改进后模型



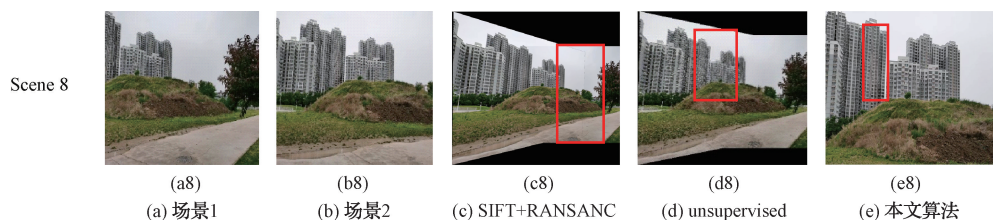


图8 不同算法在不同场景下的图像拼接效果

表1 不同算法不同场景拼接图像的 Entropy 指标

	SIFT+RANSANC	Unspervised	本文
Scene1	7.580 7	7.614 1	7.527 5
Scene2	7.299 7	7.336 6	7.432 9
Scene3	7.074 0	7.160 1	7.139 9
Scene4	6.659 5	6.104 1	6.855 2
Scene5	7.254 4	6.923 4	7.441 7
Scene6	7.154 2	7.325 6	7.490 5
Scene7	7.091 7	7.552 9	7.674 3
Scene8	6.713 6	7.185 0	7.612 6

RANSANC 算法拼接的质量。不同算法的图像拼接熵值如图9所示,由图9可知,本文算法的拼接质量优于 SIFT+RANSANC 算法和 Unspervised 算法的拼接质量。

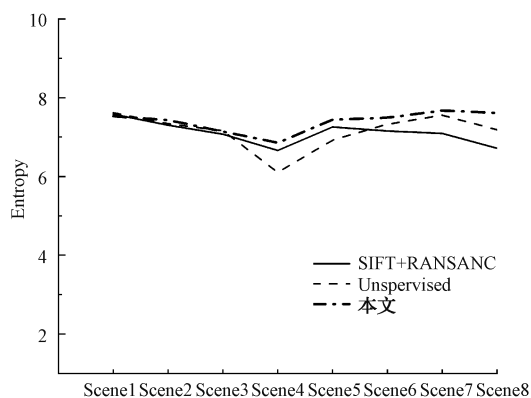


图9 不同算法的图像拼接熵值对比示意图

2) 基于本文算法在不同场景下的图像拼接

实验选取8种不同场景下的图像,如图10所示。使用图像拼接的评价指标对不同场景下拼接的图像进行评价,如表2所示。

综上所述,在特征不显著地面图像的拼接技术中,传统方法不适用于在不同视角下进行拼接,拼接后的图像大多数情况下会出现伪影。为解决该问题,本文采取基于深度学习的无监督方法代替传统方法,不仅实现了该条件下的无显著特征地面图像拼接,而且实现了在不同真实场景下的图像拼接。改进后的神经网络可通过训练学习图像间的特征与对应关系,实现图像配准并得到深

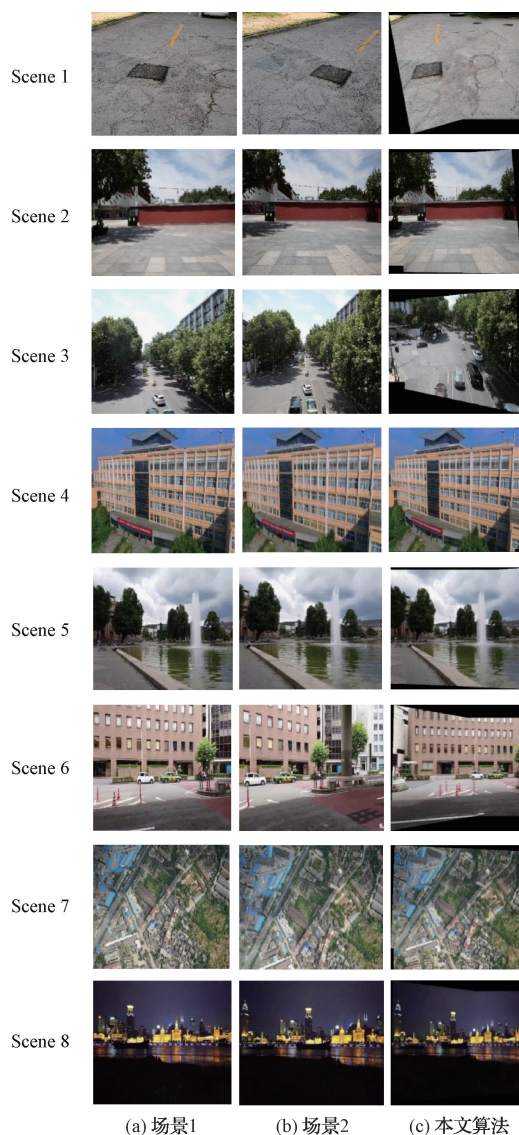


图10 基于改进后拼接算法不同场景下的实验

度单应性矩阵,完成图像的精对齐。同时,通过改进图像重构模块中的网络结构,最终得到高质量的拼接图像。表2中 SSIM 指标表明,本文算法适合不同重叠率下的图像拼接。通过对比实验的结果表明,改进后的无监督网络模型相比于传统方法的图像拼接具备更好的拼接能力。

表 2 改进后的无监督算法不同场景下融合图像评价指标均值

Experiment	MSE	PSNR	RMSE	Entropy	SNR	SSIM
1	2 446.279 1	14.245 7	2 446.3	6.354 7	8.505 4	0.196 6
2	3 483.969 0	12.710 1	3 484	7.477 9	8.110 4	0.576 4
3	4 581.587 1	11.520 6	4 581.6	7.543 1	6.785 2	0.523 2
4	2 933.770 5	13.456 5	2 933.8	7.651 0	7.524 7	0.240 0
5	3 879.901 8	12.242 6	3 879.9	7.889 4	7.406 3	0.482 3
6	5 282.973 9	10.902 0	5 283	7.254 2	6.124 4	0.320 1
7	3 559.933 4	12.616 4	3 559.9	7.464 8	6.134 0	0.255 6
8	1 904.009 8	15.334 1	1 904	7.275 7	3.352 4	0.702 6

4 结 论

本文针对大范围特征不显著的图像拼接存在的问题,诸如,伪影、结构失真等。通过使用多相机获取大范围图像并改进基于深度学习的无监督算法完成图像拼接。所改进算法主要针对拼接图像的两大模块,即图像对齐和图像重构。通过对原始网络中的参数和神经网络的层数进行调整并引入残差网络。其次,本文为了验证改进后的算法的鲁棒性,将不同场景下的图像和以及不同重叠率下的图像进行实验。实验结果表明,经本文改进后的算法不仅对特征不显著的图像之间的拼接效果远远优于传统算法的拼接,而且在普通场景下的实用性也优于现存的各种算法。因此,改进后的基于无监督的图像拼接算法对于图像拼接的质量和效率都在原来的基础上得到了大幅度的提升。

参 考 文 献

[1] 宫妍,位冲冲. 图像拼接关键技术研究综述[J]. 电脑知识与技术,2021,17(30):106-108.

[2] 姚万业,鲍珣,刘彤宇. 多相机图像拼接算法的改进[J]. 科学技术与工程,2019,19(28):227-232.

[3] 陈小青. 无人机遥感影像拼接技术的研究[D]. 贵阳:贵州大学,2021.

[4] 卢奇. 基于改进 SIFT 的图像拼接算法研究及其评价[D]. 上海:上海海洋大学,2019.

[5] 陈宁,刘志坚,苏雪平,等. 基于改进的 SIFT 算法的集成电路图像拼接[J]. 国外电子测量技术,2021,40(6):159-164.

[6] 李纪强,孔令华,花海燕. 基于图像拼接的直线轴承长度测量方法[J]. 电子测量技术,2022,45(8):148-154.

[7] 何赞泽,张帆,刘昊,等. 风机叶片无人机红外热图像拼接方法[J]. 电子测量与仪器学报,2022,36(7):44-53.

[8] 邓亮,滕奇志,何海波. 基于块匹配与多级采样的薄片孔隙图像拼接[J]. 电子测量技术,2022,45(15):106-114.

[9] GADDAM V R, RIEGLER M, EG R, et al. Tiling in

interactive panoramic video: Approaches and evaluation[J]. IEEE Transactions on Multimedia, 2016,18(9): 1819-1831.

[10] WANG L, YU W, LI B. Multi-scenes image stitching based on autonomous driving[C]. IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2020:694-698.

[11] LAI W, GALLO O, GU J, et al. Video stitching for linear camera arrays [C]. British Machine Vision Conference,2019.

[12] ANDERSON R, GALLUP D, BARRON J T, et al. Jump: Virtual reality video [J]. Transactions on Graphics, 2016, 35(6):198.

[13] KIM H, LIM H, RO Y. Deep virtual reality image quality assessment with human perception guider for omnidirectional image [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30: 917-928.

[14] GAO J, KIM S J, BROWN M S. Constructing image panoramas using dual-homography warping [C]. CVPR, 2011: 49-56.

[15] LIN W, LIU S, MATSUSHITA Y, et al. Smoothly varying affine stitching[C]. CVPR, 2011: 345-352.

[16] ZARAGOZA J, CHIN T, TRAN Q, et al. As-projective-as-possible image stitching with moving DLT[C]. IEEE Conference on Computer Vision and Pattern Recognition,2013: 2339-2346.

[17] CHANG C, SATO Y, CHUANG Y. Shape-preserving half-projective warps for image stitching[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition,2014: 3254-3261.

[18] CHANG C, CHUANG Y. A line-structure-preserving approach to image resizing[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012:1075-1082.

[19] CHEN Y, CHUANG Y. Natural image stitching with the global similarity prior [C]. European

- Conference on Computer Vision, 2016: 186-201.
- [20] LIN C, PANKANTI S, RAMAMURTHY K N, et al. Adaptive as-natural-as-possible image stitching[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015:1155-1163.
- [21] LI J, WANG Z, LAI S, et al. Parallax-Tolerant image stitching based on robust elastic warping[J]. IEEE Transactions on Multimedia, 2018, 20 (7): 1672-1687.
- [22] LEE K, SIM J. Warping residual based image stitching for large parallax [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 8195-8203.
- [23] LI J, DENG B, TANG R, et al. Local-adaptive image alignment based on triangular facet approximation [J]. IEEE Transactions on Image Processing, 2019, 29: 2356-2369.
- [24] ZHANG Z, XU C, YANG J, et al. Progressive hard-mining network for monocular depth estimation[J]. IEEE Transactions on Image Processing, 2018, 27(8): 3691-3702.
- [25] SUN D, YANG X, LIU M, et al. PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 8934-8943.
- [26] TIAN L, TU Z, ZHANG D, et al. Unsupervised learning of optical flow with cnn-based non-local filtering [J]. IEEE Transactions on Image Processing, 2020, 29: 8429-8442.
- [27] 林立雄, 郑佳春, 黄国辉, 等. 基于卷积神经网络与扩展卡尔曼滤波的单目视觉惯性里程计[J]. 仪器仪表学报, 2021, 42(10): 188-198.
- [28] HOANG V, TRAN D, NGUYEN G N, et al. Deep feature extraction for panoramic image stitching[C]. Asian Conference on Intelligent Information and Database Systems, 2020: 141-151.
- [29] SHI Z, LI H, CAO Q, et al. An image mosaic method based on convolutional neural network semantic features extraction [J]. Journal of Signal Processing Systems, 2020, 92(10): 435-444.
- [30] ZHAO Y, YE W, YU K, et al. Attentive deep stitching and quality assessment for 360° omnidirectional images[J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 14(1): 209-221.
- [31] LAI W, GALLO O, GU J, et al. Video stitching for linear camera arrays [C]. British Machine Vision Conference, 2019.
- [32] SHEN C, JI X, MIAO C. Real-time image stitching with convolutional neural networks [C]. 2019 IEEE International Conference on Real-time Computing and Robotics (RCAR), 2019: 192-197.
- [33] 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.
- [34] NIE L, LINC, LIAO K, et al. Learning edge-preserved image stitching from large-baseline deep homography [J]. Neurocomputing, 2022, 491: 533-543.
- [35] 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.
- [36] 陈国洋, 吴小俊, 徐天阳. 基于深度学习的无监督红外图像与可见光图像融合算法[J]. 激光与光电子学进展, 2022, 59(4): 151-160.

作者简介

杨旭朝, 硕士研究生, 主要研究方向为图像处理、深度学习等。

E-mail: 414792038@qq.com