

〈图像处理与仿真〉

## 稀疏深度特征红外图像拼接算法

胡俊伟, 王诗薇, 杨默远

(云南北方光电仪器有限公司, 云南 昆明 650114)

**摘要:** 为解决红外图像拼接过程中存在的红外特征少、特征匹配效果差等问题, 本文提出稀疏深度特征红外图像拼接 (Sparse Depth Feature infrared image Stitching, SDFS) 算法。该算法先基于卷积神经网络提取密集深度特征图, 然后从特征图中计算和描述稀疏特征点, 提高特征点的提取质量; 然后提出使用 K-近邻搜寻法完成稀疏特征点粗匹配, 再通过动态距离比策略精细化匹配结果, 提升匹配精度; 最后依据匹配结果计算单应矩阵进行图像投影变换, 并使用自适应因子加权融合完成图像无缝融合拼接。实验结果证明该算法鲁棒性高, 可有效适应不同场景的红外图像拼接, 拼接准确率和显示效果都高于常用的基于 SIFT、SURF 特征提取的拼接算法。

**关键词:** 红外图像拼接; 稀疏特征; 拼接优化; 特征匹配

中图分类号: TN911.73

文献标识码: A

文章编号: 1001-8891(2025)05-0584-07

## Sparse Depth Feature Infrared Image Stitching Algorithm

HU Junwei, WANG Shiwei, YANG Moyuan

(Yunnan North Optoelectronic Instrument Co., Ltd, Kunming 650114, China)

**Abstract:** To solve the problems of limited infrared features and poor feature matching in the process of infrared image stitching, this paper proposes an algorithm called sparse depth feature infrared image stitching (SDFS). The algorithm first extracts a dense depth feature map using a convolutional neural network, then calculates and describes sparse feature points from the feature map to enhance the quality of feature point extraction. Next, the K-nearest neighbor search method is used to perform coarse matching of sparse feature points, followed by the application of a dynamic distance ratio strategy to refine the matching results and improve matching accuracy. Finally, based on the matching results, a homography matrix is calculated for image projection transformation, and adaptive factor-weighted fusion is used to achieve seamless fusion and splicing of the image. Experimental results show that the algorithm exhibits high robustness and can effectively adapt to infrared image stitching in different scenes. The stitching accuracy and display effect outperform commonly used stitching algorithms based on SIFT or SURF feature extraction.

**Key words:** infrared image stitching, sparse features, seam optimization, feature matching

### 0 引言

红外成像技术通过捕捉物体发出的红外辐射, 转化为可视化图像, 能够在夜间或能见度较低的环境中进行有效成像。红外技术在安防监控、军事侦查、搜索救援、热故障检测、医学诊断等领域得到广泛的应用, 红外图像的价值也日益凸显<sup>[1]</sup>。受限于成像器件

的尺寸、阵列规模、像素间距等因素, 当前红外产品普遍面临着成像视场较小的问题, 为了解决这一问题, 迫切需要从技术层面进行创新和突破。综合考虑成本、风险、可扩展性、研发周期等因素, 相较于研制大视场成像器件, 采用图像拼接技术实现大范围观察, 优势更大, 是现阶段解决红外视场较小问题的主要技术。

收稿日期: 2024-04-29; 修订日期: 2024-09-23.

作者简介: 胡俊伟 (1997-), 男, 硕士, 工程师, 主要研究方向为计算机视觉与图像处理。E-mail: hujunwei\_zs@foxmail.com。

通信作者: 王诗薇 (1991-), 女, 博士, 高级工程师, 主要研究方向为计算机视觉。E-mail: wangshiwei\_124@163.com。

基金项目: 国防科技基础加强计划资助 (2021-JCJQ-JJ-1020)。

图像拼接技术是指将具有重叠区域的多幅图像配准对齐并得到各个图像之间的空间变换矩阵,利用所求矩阵将图像变换至同一坐标系下,并将重叠区域进行图像融合得到一幅广视角、高分辨率图像<sup>[2]</sup>。目前,可见光图像拼接技术已取得了一定的成果,并已广泛应用于市场,例如在民用领域的安全驾驶全景影像系统、360°虚拟现实(Virtual Reality, VR)显示等,通过准确的特征提取、图像配准和图像融合,得到高质量的全景图像,扩大用户的感受野<sup>[3]</sup>。然而,相对于可见光图像,红外图像的拼接面临着更多的挑战。红外图像通常具有较低的成像分辨率,并且图像边缘细节信息弱,主要以灰度信息呈现。这些特性导致在实际特征提取过程中,所能提取到的特征点少或基本不能匹配成功,从而影响拼接效果,甚至输出错误拼接结果。为了得到较好的拼接效果,通常需要在拼接前进行预先的手动校准,以确保图像间的准确对齐和特征点匹配。目前图像拼接技术主要是采用 SIFT (Scale Invariant Feature Transform)<sup>[4-6]</sup>, FAST (Features from Accelerated Segment Test)<sup>[7-8]</sup>, ORB (Oriented FAST and Rotated BRIEF)<sup>[9-10]</sup>, SURF (Speeded Up Robust Features)<sup>[11]</sup>等传统算法<sup>[12-13]</sup>提取特征点/线,进行匹配来完成图像拼接,这些算法可以在一定程度上应对视角变化和光照条件波动带来的影响,然而在红外场景中,这些特征提取算法的实际提取效果较差,鲁棒性低。

基于深度学习的卷积神经网络 (Convolutional Neural Networks, CNN) 在特征提取方面展现出了卓越的性能,通过多层次的架构可捕捉图像的深层次特征信息,其鲁棒性优于传统算法,且能在面对图像尺度、旋转、缩放等多重变化时,依然保持较高的识别和处理能力<sup>[14-15]</sup>。相关基于深度学习的图像拼接集中于可见光领域,包括深度学习估计单应矩阵<sup>[16-17]</sup>、无监督学习<sup>[18]</sup>和端到端输出拼接<sup>[19]</sup>等一系列算法,而关

于基于神经网络的红外图像拼接研究较少。为此,本文提出稀疏深度特征红外图像拼接算法 (Sparse Depth Feature infrared image Stitching, SDFS), 该算法基于 CNN 提取密集特征图,再将密集特征图中的特征点抽象为深度稀疏特征点,用于表示图像特征,进一步使用粗-细两步匹配策略保证匹配点对质量,计算单应矩阵图像完成拼接,最后构建自适应因子优化拼缝。通过与基于 SIFT、SURF 等的拼接算法进行仿真对比实验,证明 SDFS 算法的拼接准确性与鲁棒性优于目前常用算法。

## 1 SDFS 算法原理

### 1.1 总体流程

图像拼接的难点在于特征提取、特征匹配、图像拼缝处理。特征提取是拼接的基础,尤其在面对红外图像特征信息急剧减少的情况下,特征的提取是决定拼接效果的上限。匹配策略影响特征是否有效利用,直接影响到拼接结果的正确性,决定了拼接效果的下限。图像拼缝处理是提升图像拼接视觉效果的重要步骤,其目的是实现拼接区域的平滑过渡和无缝连接,以满足人眼对视觉效果的高要求。本文提出的 SDFS 算法着重对拼接的三大难点进行研究,在特征提取方面,使用卷积神经网络来提取深度密集特征图,参考 D2-Net 网络<sup>[20-21]</sup>将特征图提取为稀疏深度特征,通过更深层次的信息来保证算法的鲁棒性;在匹配策略方面采用了粗-细的策略,具体步骤为使用 K-近邻算法 (K-Nearest Neighbor classification, KNN)<sup>[22]</sup>完成对特征的初步筛选,再使用动态距离比进行精细匹配,通过粗-细策略保证特征匹配的可靠性;在拼缝处理部分,区别于传统固定因子加权融合,本文提出了自适应因子,优化拼缝显示效果。SDFS 算法总体流程如图 1 所示。

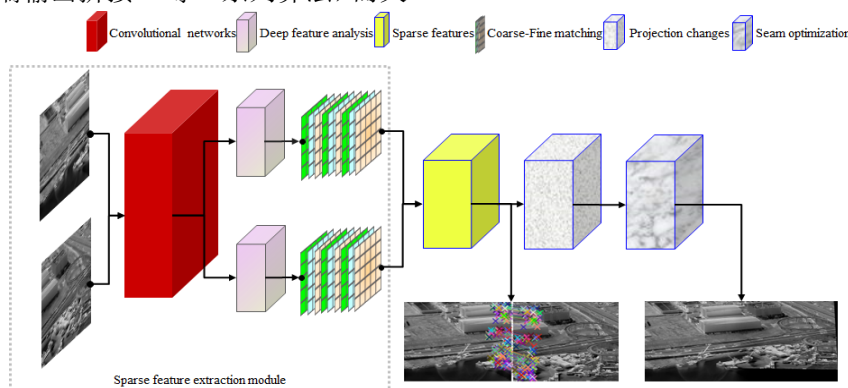


图1 SDFS 总体流程图

Fig.1 Overall flow chart of SDFS

SDFS 算法总体流程分为四部分：稀疏深度特征提取、粗-细匹配、图像投影变化和拼缝优化。其中，图像投影变化部分原理明确，主要过程是从匹配特征中计算单应性矩阵（Homography Matrix），通过单应性矩阵可将一幅图像中的点映射到另一幅图像的对应位置；在得到单应性矩阵后，可以将图像的4个角点 $((0, 0)$ 、 $(width, 0)$ 、 $(width, height)$ 、 $(0, height)$ )映射到变换后的图像中，这4个点代表了原始图像的边界，通过它们变换后的位置，可以预测整个图像在透视变换后的形状和尺寸，减少拼接结果右侧的无效区域。SDFS 算法未对投影变化过程进行优化，因此在下文不再赘述，本文着重对特征提取、匹配和拼缝优化三部分展开阐述。

## 1.2 稀疏深度特征提取模块

在卷积神经网络中，随着网络深度的增加，提取到的特征越抽象化，适应性更强。但网络深度增加也带来了一系列的挑战，首先模型的复杂度随之增加，这不仅增加了计算资源的消耗，还可能导致过拟合问题，影响模型的泛化能力；其次深度网络训练过程中常遇到梯度消失问题，致使模型训练无法收敛，且由于多次下采样导致特征定位信息基本丧失。在红外图像拼接领域，期望的特征是能应对全天候辐射变化带来图像灰度的影响，具有一定的尺度不变性和旋转特性，同时要求能精准定位图像相似区域，以此保证图像拼接的效果。因此，本文提出了稀疏深度特征提取模块，该模块通过一定深度的卷积网络提取密集深度特征图，再由特征解析部分获取稀疏特征，使特征具有高层次信息的同时保留位置信息。具体地，密集特征提取网络有10个卷积层（包含ReLU激活函数），

3个池化层，1个特征解析模块，总体结构如图2。考虑拼接需要像素级精准定位相似位置，卷积层与池化层没有进行高倍的下采样，因为高倍下采样容易导致重要信息的丢失，尤其是图像的边缘和细节的部分，不利于图像间的配准，因此在本文网络中只有4倍下采样。对于特征维度（通道）的设计，尽可能保证一定的通道数量，输出512维特征图，提供足够多的深度密集特征以支撑特征解析。特征提取网络模块的输入原本为单通道红外图像，综合考虑模型训练时数据不局限于红外图像，故将红外图像扩充为3通道，网络模型输入为 $512 \times 640 \times 3$ ，此外，模型卷积核大小都为 $3 \times 3$ ，卷积步长1，填充1，分别在第1、3、5、8个卷积层，将卷积输出通道分别设置为64、128、256、512，即卷积层只对特征深度信息加强，特征尺度分别由两个最大池化层控制，密集特征最终输出尺寸为 $128 \times 160 \times 512$ 。

从密集特征中提取稀疏特征，本文采用基于D2-Net的双最大筛选策略进行选取，充分使用特征的高维信息，并保证特征点在特征图中为关键信息。设密集特征网络输入图像 $I$ 尺寸为 $w \times h$ ，网络输出为 $F$ ， $F = I(I)$ ， $F \in R^{n \times h \times w}$ ， $n$ 表示特征维度，则将第 $k$ 维特征图和某个像素的特征向量式(1)和式(2)表示如下：

$$D^k = F_{:,k}, \quad D^k \in R^{h \times w} \quad (1)$$

$$d_{ij} = F_{ij,:}, \quad d_{ij} \in R^n \quad (2)$$

将稀疏深度特征 $P(i,j)$ 表示为式(3)、(4)：

$$P(i, j) = \max_{\text{area}} D_{ij}^m \quad (3)$$

$$m = \arg \max D_{ij}^k \quad (4)$$

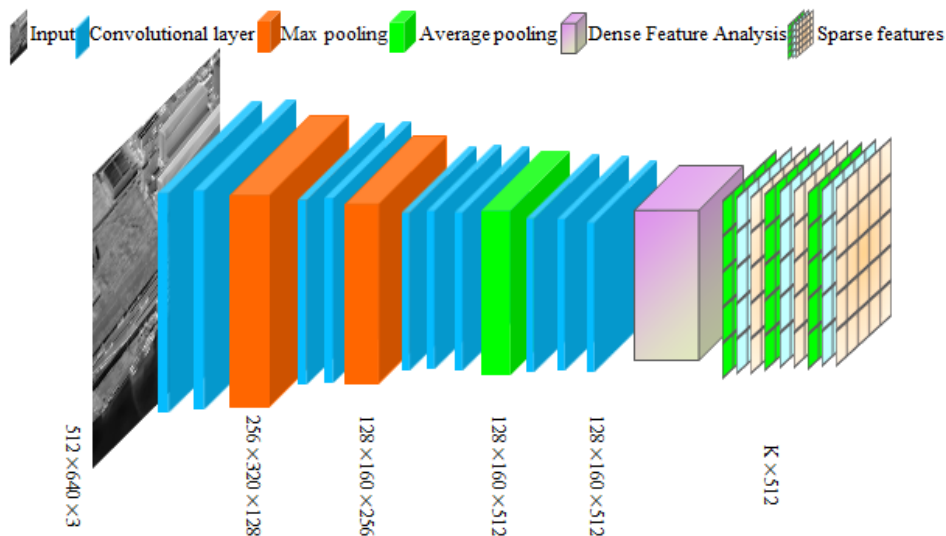


图2 稀疏深度特征提取模块结构图

Fig.2 Structure diagram of the sparse depth feature extraction module

即某个特征点既是该点特征向量在  $n$  维通道中的最大(设为第  $m$  维最大),也是第  $m$  维特征图中局部最大。局部限制在  $3 \times 3$ ,  $5 \times 5$  或更大区域,本文中使用  $3 \times 3$  的区域选择保证足够数量的特征,最后输出稀疏特征尺度为  $K \times 512$ ,其中  $K$  依据每幅图像输信息不同而变化。以图2中  $640 \times 512$  的图像输入为例,在密集特征解析前,密集特征图大小为  $128 \times 160 \times 512$ ,在第一个  $3 \times 3$  窗格,假设窗格中坐标(1,2)的值点是窗格中最大,同时是512维通道中相同坐标下的最大,则该点为解析点,以此类推,512维中每一维因为图像信息分布不同,每个通道解析出的稀疏点数量也不一致,在此用  $K$  表示每个维度解析出的稀疏特征点数量。本文考虑在训练时数据的多尺度变化已可保证特征一定尺度的适应性,故舍弃原网络中提取稀疏特征的三层尺度复杂结构,而是只保留一层来减少计算量。网络的损失函数沿用三元组距离排序损失函数,密集提取网络先使用 ImageNet 数据集<sup>[23]</sup>完成预训练,学习泛化特征,加速密集特征提取模块的卷积网络收敛,整个网络再采用 MegaDepth 数据集<sup>[24]</sup>进行训练,利用数据集基于深度信息的图像像素对应关系,使训练中有充分的稀疏点对应关系,实现提高网络密集特征提取以及稀疏特征点解析的准确性。

### 1.3 粗到细的特征匹配策略

在图像匹配、拼接、辨识等领域,一般使用随机采样一致性算法(Random Sample Consensus, RANSAC)<sup>[25]</sup>完成特征的匹配过程,该方法对传统方法提取的特征匹配效果较好,但对卷积网络提取的特征效果较差。主要原因有两方面,一是当 RANSAC 迭代次数受限时,得到的匹配结果往往不是最佳,且偶然存在很差的情况,致使匹配结果具有一定随机性;二是稀疏特征点会存在一定比例的不符合模拟模型的数据,这导致 RANSAC 匹配效果不佳。基于以上问题,本文提出粗-细的特征匹配策略,先使用 KNN 算法进行粗匹配,其对异常值和噪声有着较高的容忍度,且支持多分类,再使用动态距离比完成精细化匹配,进一步优化特征匹配结果,以支撑拼接的精度。

粗匹配过程中的 KNN 算法步骤为:将两拼接图像定义为训练样本(左图)和测试样本(右图),计算左图和右图稀疏特征点集中每个样本点的欧式距离,并将距离值排序,选取前  $K$  个最小距离的样本;依据  $K$  个样本的标签进行投票决策,获取分类结果<sup>[25]</sup>。在本文中取  $K=2$ ,即在匹配中搜寻的是最近邻与次近邻的匹配结果。细匹配过程中,通常认为最近邻匹配点的距离与次近邻匹配点的距离越小,则该点的匹配质量越高,准确性越可靠。一般以固定阈值因子来筛

选匹配对,阈值选取越大,选取的匹配对越多,可能包含的错误匹配对越多,导致计算单应矩阵时误差越大;阈值越小,剩余的匹配对数量越少,数据较少,拟合的单应矩阵误差也会较大,阈值选定后不能适应各种场景。故在细匹配过程中,引入平均距离和因子结合的筛选策略,设最近邻匹配对为  $m$ ,次近邻匹配对为  $n$ ,记  $m, n$  中第  $k$  对匹配结果的距离分别为  $\text{dis}_{k_m}$  和  $\text{dis}_{k_n}$ ,计算  $m$  的距离和  $n$  的距离差,再统计均值  $\text{disAvg}$ ,设定当  $m$  的距离小于  $n$  距离与  $\alpha$  倍均值距离之差时,认为匹配对  $m$  质量可靠,具体计算公式如式(5)和式(6):

$$\text{disAvg} = \frac{\sum_{k=0}^N (\text{dis}_{k_m} - \text{dis}_{k_n})}{N} \quad (5)$$

$$\text{dis}_{k_m} < \text{dis}_{k_n} - \alpha \cdot \text{disAvg} \quad (6)$$

式(5)中:  $N$  为匹配对的数量,本文中  $\alpha=1.8$ ,该策略将筛选时的阈值与当前两幅图像的信息引入,阈值依据不同图像粗匹配结果具有一定的可变性,以此适应不同场景特征匹配。此外,计算均值时也可采取距离排序后去除一定比例的较大距离值来进一步保证匹配对的质量。

### 1.4 图像拼接平滑优化

依据匹配对计算单应矩阵完成投影变化后,受到光照等的影响,如若直接拼接图像,在图像重叠区域会有非常明显的拼缝,影响视觉效果。传统方法使用加权因子(像素至拼缝距离/重叠区域宽度)进行拼缝处理,其是线性变化,实际优化效果不佳。本文提出自适应因子加权融合来进行拼缝平滑优化,充分考虑融合像素与拼接图像重叠区域的关系,实现拼缝处的自然过渡,该方法具体步骤如下:首先,计算左右两图拼接的重叠区域:先计算右图(0,0)、(0,height)两点经投影变化后横坐标的最小值,再由左图图像宽减去该值,得到重叠区域像素宽度  $d$ ;然后计算加权因子  $\beta$ ,如式(7):

$$\beta = \frac{1}{1 + e^{-k \cdot x}}, x \in \left(-\frac{d}{2}, \frac{d}{2}\right) \quad (7)$$

式(7)中:  $k=0.04$ ,引入指数分布,让融合过程更为平滑;最后图像加权融合,公式(8)如下:

$$I(i, j) = \begin{cases} I_1(i, j), & i \in (0, w-d) \\ (1-\beta) \cdot I_1(i, j) + \beta \cdot I_2(i, j), & i \in (w-d, w) \\ I_2(i, j), & i \in (w, w') \end{cases} \quad (8)$$

式中:  $w$  为原图的图像宽;  $w'$  为拼接后图像宽度;  $I_1$  和  $I_2$  分别表示输入的左图和已经投影变化的右图。

2 实验分析与结果

实验硬件平台为 Intel i7-12700 2.1GHz, 16GB 内存, 无 GPU, 软件平台为 Windows11 操作系统, 实验过程中对比的 SIFT、SURF 拼接算法皆为基于 OpenCV3 的库函数调用, SURF 算法的 hessian 阈值为 500, 其余参数为默认值, 软硬件平台一致。在进行拼接效果展示时, 以未进行拼缝优化的中间拼接输出结果进行对比和计算拼接准确率, 选取 50 组待拼接红外图像, 涉及室内、室外多地形, 水平、俯瞰多视角等场景参与实验。本文中使用正确率和错误率两个评价指标, 拼接错误率 Err 的定义为拼接错误图像数  $M$  (未能计算单应矩阵、图像投影变化错误) 和拼接错位像素误差大于 4 像素的图像数  $E$  加权和与总测试图像数  $N$  之比, 公式(9)如下:

$$Err = \frac{M + 0.5E}{N} \times 100\% \tag{9}$$

式中引入了拼接错位图像数, 以符合实际效果的反映, 一般 4 像素左右的错位通过加权拼缝优化算法可以直接消除, 一旦错位较大, 实际优化结果仍然存在一定的错位, 不符合人眼观察习惯。正确率 Acc 的定义为拼接正确图像数 (错位误差小于 4 像素) 与总测试图像数之比。

实验结果如表 1, 表中 SDFS\*算法表示不使用拼缝优化算法, 便于统计图像拼接时像素错位情况, 从

实验数据看, SIFT 和 SURF 算法的错误率均高于 10%, 而本文提出的 SDFS\*算法错误率低, 且准确率为 88%高于 SIFT 和 SURF 算法。

表 1 实验结果准确率对比  
Table 1 Comparison of the accuracy of the test results

Methods	SIFT	SURF	SDFS*
Acc	76	80	88
Err	17	14	8

实验效果如图 3 所示, 分别选取了(a)城市场景、(b)港口场景、(c)戈壁场景和(d)室内场景四种情况进行对比。对于城市场景, 全图信息丰富、纹理等明显, 三种算法拼接都未发生错位情况, SDFS 的拼缝优化过度自然; 对于港口场景, 图像相对(a)来说特征减少, 重叠区域上部分以单一建筑为主, 且较为简单, SIFT 和 SURF 拼接有 4 个像素左右的错位, SDFS\*未发生错位; 对于戈壁场景, 图像纹理信息少, 且地形具有一定相似性, SIFT 和 SURF 都不能有效计算单应矩阵, 故输出两图水平直拼效果图, 而 SDFS\*拼接效果优秀, 无错位; 室内场景细节信息少且图像存在一定的拖影, SIFT 和 SURF 拼接失败, SDFS\*虽然输出结果有一定的形变, 但其正确输出拼接结果, 优于 SIFT 和 SURF。

图 4 展示了 SIFT、SURF 和 SDFS\*算法在不同红外场景的特征匹配对数。在红外图像场景信息丰富

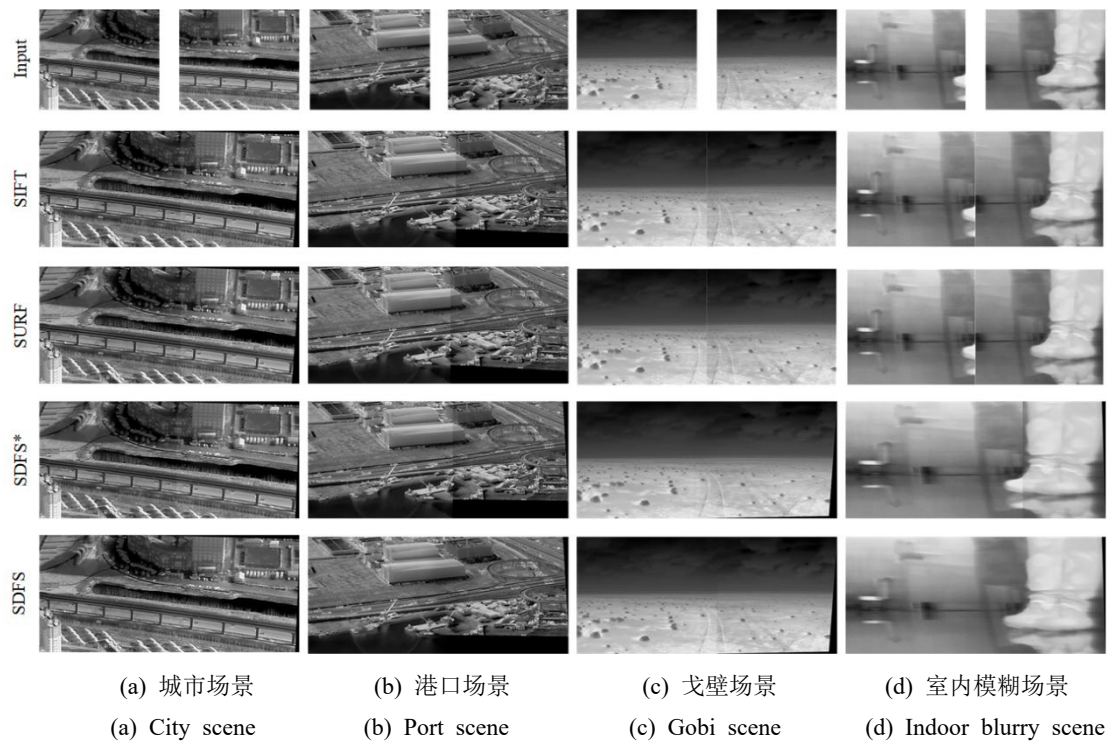


图 3 实验结果对比

Fig.3 Comparison of experimental results

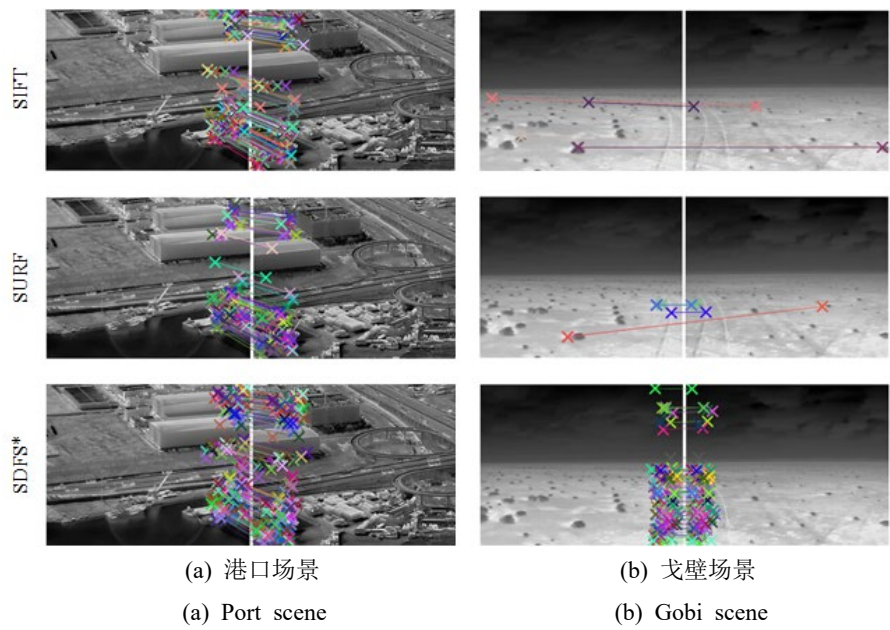


图 4 特征匹配对数量示意

Fig.4 Schematic diagram of the number of feature matching pairs

时，SIFT 和 SURF 基本能正确输出拼接结果，但存在拼接错位等情况，在场景简单、特征信息不足的条件 下，SIFT 和 SURF 基本不能提取足够数量的特征点， 或提取特征点质量不高，导致特征匹配对数量少或错 误匹配，无法计算单应矩阵，从而不能正确输出拼接 结果；SDFS 算法在场景信息丰富时拼接效果稍强于 SIFT 和 SURF，在场景信息有限的情况下可以保证高 质量的匹配对，拼接效果大幅度领先。

综合实验结果，SDFS 在细节丰富的俯视城市场 景(图 3(a))、拼缝纹理信息相对较少的港口(图 3(b))、 纹理相似的戈壁场景(图 3(c))和运动模糊的室内场 景(图 3(d))的拼接准确率都优于 SURF 算法。图 4 将图 3(b)和(c)两场景各对比算法提取匹配的特征点 展示，在纹理较少、且纹理高相似的场景下，SDFS 提 取的特征点匹配数量远多于 SURF，进一步说明 SDFS 不仅能有效应对细节丰富的红外图像，而且在纹理较 少、纹理相似的红外图像上有更强的适应性，且拼接 准确率高于对比方法。

表 2 为 SIFT、SURF 和 SDFS\*算法在相同硬件环 境下的运行时间对比，SDFS 算法的稀疏深度特征提 取模块在 CPU 端运行需要较长时间，导致程序运行 时间较长。

表 2 程序运行时间对比

Table 2 Comparison of program running times

Method	SIFT	SURF	SDFS*
Time	0.354	0.190	2.103

### 3 结论

红外图像拼接应用领域广泛，考虑到可见光拼接 方法在面对红外图像信息急剧下降的情况下极易失 效，本文提出了稀疏深度特征红外图像拼接算法，该 算法在特征提取方面引入卷积神经网络，在匹配策略 方面采用分粗-细两步策略保证匹配对质量，在拼缝优 化方面提出平滑优化方式，从三个方面提高红外图像 拼接的质量与可靠性。经实验证明，其效果优于现阶 段通用的算法。本文研究工作还存不足，一是由于图 像数据限制，SDFS 算法实验数据较少，未能充分验证 其适用性，未来需采集更多数据充分验证对比；二是 由于算法采用卷积网络且未在 GPU 上加速运行，其拼 接速度有进一步优化的空间，未来需进一步将深度特 征解析部分使用硬件加速计算，以此解决算法运行 时间过长的问 题，以推进算法的实际应用。

### 参考文献：

[1] 李相迪, 黄英, 张培晴, 等. 红外成像系统及其应用[J]. 激光与红外, 2014, 44(3): 229-234.

LI Xiangdi, HUANG Ying, ZHANG Peiqing, et al. Infrared imaging system and applications[J]. *Infrared Technology*, 2014, 44(3): 229-234.

[2] 王向军, 张嘉豪, 王越. 基于位姿数据校正的遥感红外图像相似性先 验拼接[J]. 红外技术, 2019, 41(12): 1091-1099.

WANG Xiangjun, ZHANG Jiahao, WANG Yue. Image stitching of remote sensing infrared images with similarity prior based on correction of pose data [J]. *Infrared Technology*, 2019, 41(12): 1091-1099.

- [3] 周登科, 郭星辰, 史凯特, 等. 风电场无人机巡检红外叶片图像拼接算法[J]. *红外技术*, 2023, **45**(11): 1161-1168.  
ZHOU Dengke, GUO Xingchen, SHI Kaite, et al. Infrared blade image stitching algorithm for unmanned aerial vehicle inspection in wind farms[J]. *Infrared Technology*, 2023, **45**(11): 1161-1168.
- [4] Lowe D G. Distinctive Image features from scale-invariant key points[J]. *International Journal of Computer Vision*, 2004, **60**(2): 91-110.
- [5] 卢泉, 杨振华, 黄粒峰. 改进最佳缝合线的红外图像拼接方法[J]. *红外技术*, 2022, **44**(6): 580-586.  
LU Quan, YANG Zhenhua, HUANG Lifeng. Improved infrared image stitching method for optimal suture [J]. *Infrared Technology*, 2022, **44**(6): 580-586.
- [6] 王昕平, 张森林, 刘妹琴, 等. 基于多尺度图像融合和 SIFT 特征的水下图像拼接研究[J]. *计算机应用与软件*, 2021, **38**(5): 213-217, 230.  
WANG Xinping, ZHANG Senlin, LIU Meiqin, et al. Research on underwater image stitching based on multi-scale image fusion and SIFT features[J]. *Computer Applications and Software*, 2021, **38**(5): 213-217, 230.
- [7] Rosten E, Porter R, Drummond T. Faster and better: a machine learning approach to corner detection[J]. *IEEE Trans Pattern Anal Mach Intell*, 2008, **32**(1): 105-119.
- [8] 邓秋菊, 王宁, 陆阳. 基于全局和局部特征的红外图像快速拼接研究[J]. *激光杂志*, 2023, **44**(10): 84-88.  
DENG Qiuju, WANG Ning, LU Yang. Research on fast infrared image stitching based on global and local features[J]. *Laser Journal*, 2023, **44**(10): 84-88.
- [9] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF[C]//*ICCV*, 2011, **11**(1): 2564-2571.
- [10] 彭畅, 李广泽, 张晓阳, 等. 基于改进的 ORB 算法的红外遥感图像拼接研究[J]. *控制工程*, 2020, **27**(8): 1332-1336.  
PENG Chang, LI Guangze, ZHANG Xiaoyang, et al. Research on infrared remote sensing image stitching based on improved ORB algorithm [J]. *Control Engineering*, 2020, **27**(8): 1332-1336.
- [11] Bay H, Tuytelaars T, Gool L V. SURF: speeded up robust features[C]//*European Conference on Computer Vision*, 2006: 404-417.
- [12] 车畅, 兰文宝. Canny 算法和中值滤波法的红外全景图像拼接[J]. *激光杂志*, 2020, **41**(5): 109-113.  
CHE Chang, LAN Wenbao. Canny algorithm and median filtering method for infrared panoramic image stitching [J]. *Laser Journal*, 2020, **41**(5): 109-113.
- [13] 蒋敏, 蒋品群, 宋树祥, 等. 基于点线配准和超像素分割的图像拼接算法[J]. *计算机应用与软件*, 2023, **40**(10): 250-254, 267.  
JIANG Min, JIANG Pinqun, SONG Shuxiang, et al. Image stitching algorithm based on point line registration and superpixel segmentation[J]. *Computer Applications and Software*, 2023, **40**(10): 250-254, 267.
- [14] 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): 1-11.  
YANG Lichun, TIAN Bin, DANG Jianwu. A review of image stitching algorithms based on deep learning [J]. *Computer Application Research*, 2024:1-11.
- [15] YAN Ni, MEI Yupeng, XU Ling, et al. Deep learning on image stitching with multi-viewpoint images: a survey[J]. *Neural Processing Letters*, 2023, **55**(4): 3863-3898.
- [16] LE Hoang, LIU Feng, ZHANG Shu, et al. Deep homography estimation for dynamic scenes[C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020, 7649-7658.
- [17] 郭璠, 李小龙, 刘文韬, 等. 基于参数回归的快速全景图像拼接算法[J]. *通信学报*, 2023, **44**(9): 36-47.  
GUO Fan, LI Xiaohu, LIU Wentao, et al. A fast panoramic image stitching algorithm based on parameter regression [J]. *Journal of Communications*, 2023, **44**(9): 36-47.
- [18] HONG Mingbo, LU Yuhang, YE Nianjin, et al. Unsupervised homography estimation with coplanarity-aware gan[C]//*Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2022: 17642-17651.
- [19] ZHOU Hongfei, ZHU Yuhe, LV Xiaoqian, et al. Rectangular-output image stitching[C]//*2023 IEEE International Conference on Image Processing (ICIP)*, 2023: 2800-2804.
- [20] 蓝朝桢, 卢万杰, 于君明, 等. 异源遥感影像特征匹配的深度学习算法[J]. *测绘学报*, 2021, **50**(2): 189-202.  
LAN Chaozhen, LU Wanjie, YU Junming, et al. Deep learning algorithm for feature matching of heterogeneous remote sensing images[J]. *Journal of Surveying and Mapping*, 2021, **50**(2): 189-202.
- [21] Dusmanu M, Rocco I, Pajdla T, et al. D2-Net: a trainable CNN for joint detection and description of local features[J]. *CoRR*, 2019, abs/1905.03561.
- [22] Cover Thomas M, Peter E. Hart. Nearest neighbor pattern classification[J]. *IEEE Trans. Inf. Theory*, 1967(13): 21-27.
- [23] LI Feifei, JIA Deng, KAI Li. ImageNet: constructing a large-scale image database[J]. *Journal of Vision*, 2010, **9**(8): 1037-1037.
- [24] Christopher W, Omar A, N D B, et al. Megadepth: efficient coverage quantification for BigWigs and BAMs[J]. *Bioinformatics*, 2021, **37**(18): 3014-3016.
- [25] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. *Communications of the ACM*, 1981, **24**(6): 381-395.