

# 空天异源图像匹配关键技术研究

刘庚辰<sup>1</sup>, 吴国强<sup>1,2,3,4</sup>, 黄坤<sup>1,3,4</sup>, 刘奇霏<sup>1</sup>, 黄蜀玲<sup>1</sup>

(1 航天时代飞鸿技术有限公司 北京 100094;

2 北京理工大学机电学院 北京 100081;

3 中国航天科技集团有限公司智能无人系统总体技术研发中心 北京 100094;

4 中国航天科技集团有限公司第九研究院无人体系中心 北京 100094)

**摘要:** 空天异源图像匹配技术是无人机的重要研究方向之一。本文系统构建了空天异源图像匹配框架,并深度剖析了其关键组成部分。基于文献综述,本文将空天异源图像匹配框架的关键技术归纳为三大类:图像质量评价技术、图像预处理技术与图像匹配技术。本文对各类技术的最新进展做了总结概括,重点分析了各方法应用于无人机领域的技术差异。在此基础上结合数据集,进行交叉对比实验,对各方法的具体效果进行分析。最后,针对空天异源图像匹配所面临的难点作出总结,并对未来研究方向和发展趋势进行了展望。

**关键词:** 空天异源; 图像质量评价; 图像预处理; 图像匹配

**中图分类号:** TP751 **文献标志码:** A **文章编号:** 2095-1000(2025)02-0001-11

**DOI:** 10.12347/j.ycyk.20241014001

**引用格式:** 刘庚辰, 吴国强, 黄坤, 等. 空天异源图像匹配关键技术研究[J]. 遥测遥控, 2025, 46(2): 1-11.

## Research on the Key Techniques of Space-Sky Heterogeneous Image Matching

LIU Gengchen<sup>1</sup>, WU Guoqiang<sup>1,2,3,4</sup>, HUANG Kun<sup>1,3,4</sup>, LIU Qifei<sup>1</sup>, HUANG Shuling<sup>1</sup>

(1. Aerospace Times Feihong Technology Limited Company, Beijing 100094, China;

2. School of Mechanical and Electrical Engineering, Beijing Institute of Technology, Beijing 100081, China;

3. China Aerospace Science and Technology Corporation Limited Intelligent Unmanned System Overall Technology Research and Development Center, Beijing 100094, China;

4. China Aerospace Science and Technology Corporation 9th Institute Unmanned System Center, Beijing 100094, China)

**Abstract:** Aerial-space image matching is one of the significant research directions of unmanned aerial vehicles. This paper systematically constructs a framework for matching heterogeneous aerial-space images and provides an in-depth analysis of its key components. Based on literature review, this paper categorizes the key technologies of the heterogeneous image matching framework into three major types: image quality assessment technology, image preprocessing technology, and image matching technology. It summarizes the latest advancements in each of these technologies, with a particular focus on analyzing the technical differences in their application to the UAV field. Based on this, cross-comparison experiments are conducted using datasets to analyze the specific effects of each method. Finally, the paper summarizes the challenges faced in matching heterogeneous aerial and space images and provides an outlook on future research directions and development trends.

**Keywords:** Space-sky heterogeneous; Image quality evaluation; Image preprocessing; Image matching

**Citation:** LIU Gengchen, WU Guoqiang, HUANG Kun, et al. Research on the Key Techniques of Space-Sky Heterogeneous Image Matching[J]. Journal of Telemetry, Tracking and Command, 2025, 46(2): 1-11.

## 0 引言

在现代战争中,无人机因其快速部署能力和对时敏目标长时间追踪定位优势,成为重要的侦

察工具。然而,在卫星导航拒止环境中,由于无人机侦察图像目标定位主要依靠姿态、载荷信息进行解算,导致自身定位信息获取困难,定位误差较大,难以精确解算目标定位信息。为解决这

一问题, 可采用空天异源图像匹配技术, 通过将无人机拍摄的图像与卫星图像进行匹配, 结合卫星图像的地理信息, 快速计算出无人机侦察图像中目标位置, 从而显著提高无人机目标定位精度, 获取目标准确坐标值。

实现空天异源图像精确匹配对于复杂环境目标定位至关重要, 能够显著提升无人机应用灵活性、泛化能力。可以预见, 全地形、全气候条件下的空天异源图像精确匹配将逐步成为无人机的重要功能。针对在实际应用中, 空天异源图像匹配可能遇到的图像质量下降, 图像干扰信息较多, 图像匹配结果不理想等问题, 本文介绍了一种可用于实际应用的空天异源图像匹配框架, 详细阐述了该框架的关键组成技术, 并对框架中各技术可采用的先进算法进行了比较分析。通过相关实验验证, 本文为

空天异源图像匹配提供了有力的技术参考。

### 1 空天异源图像匹配框架概述

空天异源图像匹配框架具体结构如图 1 所示, 主要由四部分组成: 图像质量评价技术, 图像预处理技术, 图像匹配及定位解算技术。对实时获取的无人机侦察影像, 首先进行图像质量评价, 筛选出由于链路问题造成的大面积失真或缺损图像。其次, 通过图像预处理技术, 解决无人机侦察图像由于噪声干扰等问题造成的图像纹理不清晰和图像目标轮廓模糊等问题。最后, 通过图像匹配算法, 将无人机侦察影像与卫星影像进行配准, 获取无人机图像变换到卫星图像的单应性矩阵, 并结合卫星图像地理信息与检测目标图像, 最终得到目标的高精度地理坐标。

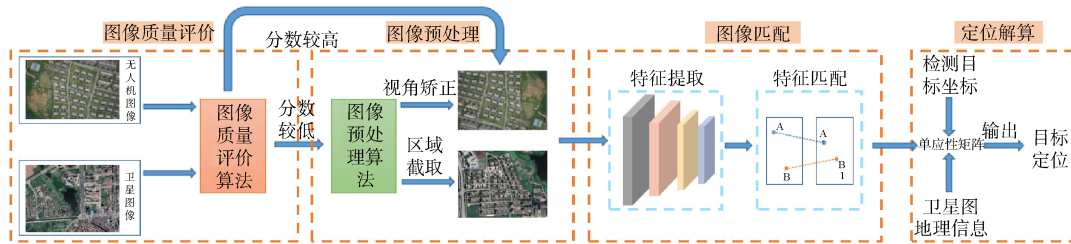


图 1 空天异源图像匹配框架

Fig. 1 The framework of space-sky heterogeneous image matching

### 2 图像质量评价

在无人机飞行过程中, 图像拍摄、信号传输、图像解析等环节可能导致图像信息丢失, 进而造成拍摄图像质量不均衡, 如图 2 所示。由于图像质量直接影响空天异源图像匹配的准确性, 因此需采用图像质量评价算法对无人机拍摄图像进行评分, 以筛选出信息损失较多的图像。

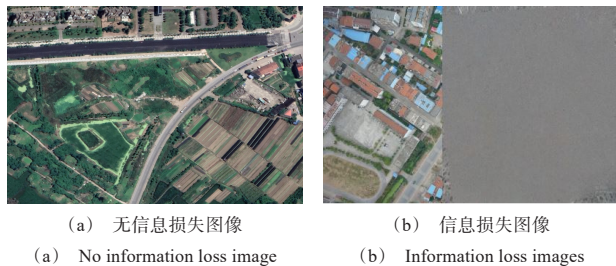


图 2 图像质量比较

Fig. 2 Image quality comparison

图像质量评价可以分为全参考图像质量评价(输入两张图, 通过对比两幅图像的信息量或特征相似

度来实现质量评价), 半参考图像质量评价(输入一张图, 结合另一张图的部分信息进行质量评价), 无参考图像质量评价(输入一张图直接评价)三种评价方式。

本课题在评价信息损失图像时无需参考原图, 因此可采用无参考图像质量评价方式。该方法主要分为两类: 传统图像质量评价和基于深度学习的图像质量评价。传统图像质量评价从数理角度出发, 通过数学运算与统计原理得出图像质量评分, 而基于深度学习的图像质量评价利用深度学习网络得出图像质量评分。

#### 2.1 传统图像质量评价

传统图像质量评价算法已在多个领域取得显著成果, 结合无人机实际应用情况, 可进行迁移应用。传统方法主要从数理角度、统计方法等角度出发对图像进行质量评价。从数学角度出发对图像质量做出评价的方法主要区别在于图像在不同域中分解方式不同。Moorthy 等人<sup>[1]</sup>在 2010 年提出了一种盲图像质量指数对图像质量进行评价, 2011 年, 他们提出了一种基于自然场景统计的图像质量评价

算法<sup>[2]</sup>。Saad<sup>[3]</sup>等人从频域角度出发，提取出图像的DCT(离散余弦变换)特征，提出了一种优化的基于频域的无参考图像质量评价方法。此外，Mittal<sup>[4]</sup>等人提出了一个基于空间域特征的算法NIQE(自然图像质量评价)。他们认为图像失真会改变自然图像归一化亮度值趋向于正态高斯分布的特点，所以通过统计空间域中的亮度特征可以计算出失真图像的质量分数。

统计方法在图像质量评价方法中也得到了广泛应用。例如，BRISQUE(无参考图像空间质量评价指标)<sup>[5]</sup>运算速度较快同时具备一定的准确性，适用于对无人机运算速度较高的场景中。BRISQUE首先通过均值减去对比度归一化在空间域中提取自然场景统计信息，接着将MSCN(移动文换中心节点)图像拟合到广义高斯分布并提取特征向量，最后利用支持向量机回归计算图像质量分数。然而，BRISQUE面对通用场景泛化性则需要进行实验验证。

### 2.2 基于深度学习图像质量评价

近年来，基于深度学习图像质量评价技术取得了显著进展，其中，代表性的深度学习卷积网络包括GIQA(通用图像质量评估)<sup>[6]</sup>、基于自适应超网络的图像质量盲评估<sup>[7]</sup>、TRIQ(时间参考图像质量评估方法)<sup>[8]</sup>和DIQA(去躁图像质量评估)<sup>[9]</sup>等。

DIQA主要通过生成客观误差图来增强训练数据集，有效解决了无人机训练数据集规模较小、不利于后续测试的问题。

基于自适应超网络的图像质量盲评估可以从多尺度层面预测无人机航摄图像质量，尤其适用于大视角无人机图像质量评价。该网络通过引入局部畸变感知模块，有效解决了无人机拍摄图像非均匀畸变情况，同时，它通过聚合局部失真特征和全局语义特征，综合细粒度细节与整体信息，实现对图像质量做出预测。该网络结构如图3所示。

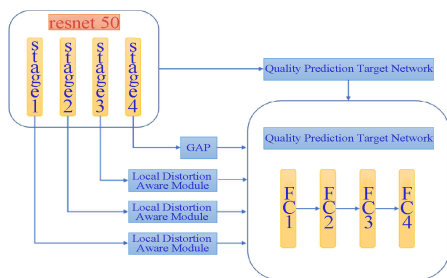


图3 基于自适应超网络的图像质量盲评估算法

Fig. 3 Blind evaluation algorithm of image quality based on adaptive hyper network

Transformer也可用于图像质量评估中。TRIQ提出了一种将Transformer应用于图像质量评估的思路，解决了Transformer不可以输入不同尺度图像的问题，同时保持较高的评估准确性。TRIQ结构如图4所示。在无人机实际应用中，若需要进行算法迁移等情况，可以利用此方法增加算法通用性，避免了二次消耗。

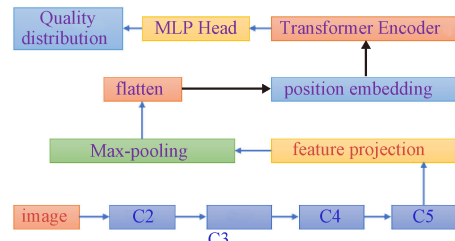


图4 TRIQ 结构图

Fig. 4 The network structure of TRIQ

采用hyper-IQA(超高性能图像质量评估), DIQA, BRISQE质量评价方法对图2进行质量评价，结果如表1所示。

表1 质量评价结果

Table 1 Results of quality evaluation

算法名称	图2(a)分数	图2(a)时间	图2(b)分数	图2(b)时间
hyper-IQA	56.42	0.99	42.58	0.86
DIQA	0.78	0.004	0.73	0.005
BRISQE	<b>32.98</b>	<b>0.012</b>	<b>21.93</b>	<b>0.007</b>

在本实验中，从表1可以得出，经BRISQE算法评价后，不同质量图片之间质量评价差异显著，同时BRISQE实时性强，不影响后续操作实时性，因此非常适用于无人机侦察影像图像质量评价。然而，在实验过程中，仍存在以下问题：

①基于深度学习图像质量评价算法效果不显著。这可能是由于训练集图像与实际应用场景相差较大所致。为了提高算法的准确性，在实际应用中，基于深度学习图像质量评价的方法应针对具体应用场景进行训练，以提高算法准确性。

②图像质量评价算法阈值难以确定。在实际应用中，通常需要量化图像质量的具体数值，但无人机拍摄的图像涵盖多种地形，如田野、山地、城市等。针对不同地形，需要进行大量实验得出具体阈值，也应考虑无人机在飞行过程中受到的

各种干扰对图像质量造成的影响。

### 3 图像预处理

由于无人机或者相机在拍摄过程中可能因本身的抖动或传输干扰引入较多噪声失真, 导致图像目标特征不清晰, 因此需要通过图像增强等技术对无人机拍摄图像进行预处理, 以提升图像纹理的准确性和特征的显著性。图像增强技术主要分为传统图像增强技术与基于深度学习的图像增强技术。

#### 3.1 传统图像增强技术

传统图像增强技术常从图像机理角度出发, 利用数学运算和统计方法对图像进行处理。部分研究从直方图角度入手, 采用直方图均衡 (HE)、自适应直方图均衡 (AHE) 和限制对比度自适应直方图均衡 (CLAHE) [10] 等。这些技术通过局部或全局进行直方图均衡, 有效改善局部图像过亮或过暗的问题, 并通过对比度限制抑制背景噪声的增强。此外, 还有研究利用卡尔曼滤波、空间滤波等方法来减小图像中噪声。

#### 3.2 基于深度学习的图像增强技术

尽管传统图像增强技术运行时间较短, 但机理单一, 增强效果受应用场景限制较大, 因此, 基于深度学习的图像增强技术逐渐成为主流。

DNCNN<sup>[11]</sup>深度卷积神经网络及其改进方法<sup>[12-13]</sup>是基于深度学习去噪网络的代表, 它们在 VGG<sup>[14]</sup>(Visual Geometry Group)网络的基础上, 结合残差学习、卷积层、BN层(批量归一化层)、ReLU层(修正线性单元层)等, 能够显著降低无人机图像中的噪声。然而, DNCNN及其改进方法网络结构较为复杂, 在运算资源有限情况下难以满足实时去噪需求。针对此情况, 另一些学者从实时性考虑, 提出一些轻量化去噪网络, 保证一定去噪效果的同时减少运算资源消耗。

##### 3.2.1 DNCNN及其改进方法

DNCNN采用残差学习策略, 直接从噪声观测中得出去噪后图像, FFDNet(快速灵活的去噪网络)在 DNCNN基础上进一步发展, 不仅能够更好地平衡去噪与细节保留, 还具有更快的运行速度, 并且能够处理大范围的噪声水平。因此, FFDNet在无人机航摄图像的实际去噪应用中表现更为理想。FFDNet结构如图5所示。此外, SRMD(基于图像先验最大后验估计的超分辨率重建方法)采用了与 DNCNN相似的非线性映射模块, 不同点在于在无人机航摄图像分辨率较低的情况下, 该方法使单个卷积超分辨率网络能够将退化过程的模糊核和噪声水平作为输入, 解决了低分辨率输入图像、模糊核和噪声水平之间的维度不匹配问题。SRMD结构如图6所示。



图5 FFDNET网络结构

Fig. 5 The network structure of FFDNet



图6 SRMD网络结构

Fig. 6 The network structure of SRMD

##### 3.2.2 轻量化去噪网络

为了满足无人机实际应用中计算效率的需求, 轻量化去噪网络逐渐成为研究热点。PMRID(实用移动设备原始图像去噪技术)<sup>[15]</sup>通过5×5内核增加接受域和减少网络深度, 并且采用 stride-2 卷积下采样特征映射, 仅在输入和输出阶段使用正态密集卷积层, 从而在保证去噪效果的同时降低了计算复杂度。

另一代表方法是 IMDN<sup>[16]</sup> (轻量级的信息多蒸

馏网络), 可以用于无人机实际应用中, 其核心组件是信息多蒸馏块 (IMDB)。IMDB包含蒸馏模块和选择融合模块。IMDN结构如图7所示, IMDN结构如图8所示。蒸馏模块逐步提取分层特征, 融合模块根据候选特征的重要程度对候选特征进行聚合, 并通过对比度感知通道关注机制对候选特征进行评估。

本文采用 UAV Visloc(无人机视觉定位)<sup>[17]</sup>数据集对算法进行测试。该数据集由多个子数据集构

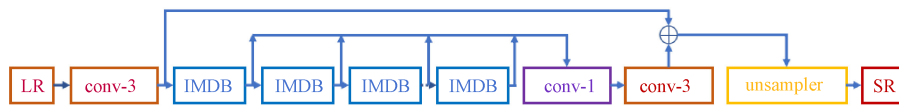


图 7 IMDN 网络结构

Fig. 7 The network structure of IMDN

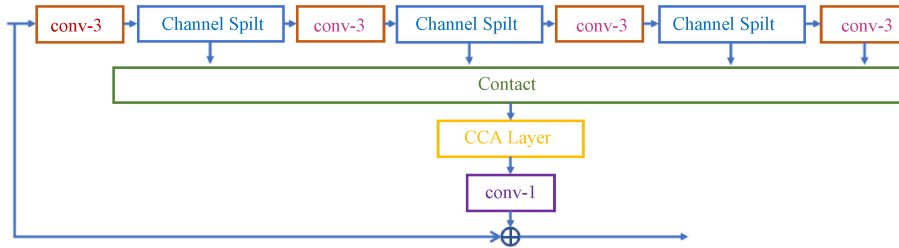
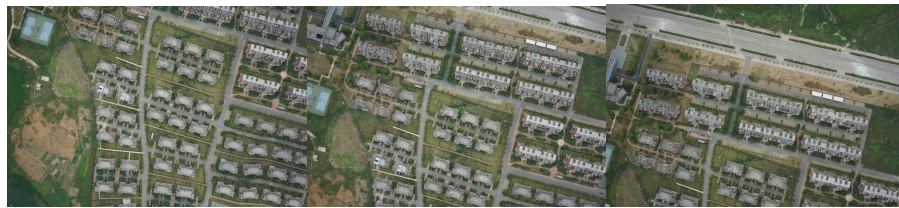


图 8 IMDB 网络结构

Fig. 8 The network structure of IMDB

成，每个子数据集包含了无人机侦察图像以及对应地区的卫星影像。无人机侦察图像涵盖了农村、城镇、农场、城市、河流、丘陵等多种场景。由于卫星影像包含实际地图范围较广，远大于无人机侦察图像包含的地图大小，因此在实际工程应用中为提升匹配效率，一般根据无人机侦察范围，

卫星影像略大于无人机侦察范围，同时基于无人机姿态信息、侦察载荷信息等能够初步计算无人机侦察图像朝向信息，基于上述条件，本文对数据集中图像进行旋转、裁剪等操作，以便于后续处理。无人机数据展示如图9 (a) 所示，卫星子图像展示如图9 (b) 所示：



(a) 无人机数据展示  
(a) Drone data display



(b) 卫星数据展示  
(b) Satellite data display

图9 无人机和卫星数据展示  
Fig.9 Drone and satellite data display

为验证预处理方法的有效性，首先对数据集中的无人机图像添加噪声，随后利用多种预处理方法对无人机噪声图像进行预处理，最后通过图像匹配方法得出匹配结果。无人机噪声图像经不同预处理方法处理后结果如图10所示。图像预处

理模块采取了IMDN(即时消息分发网络)、DNCNN(去噪卷积神经网络)、DPSR(数码像素结构重建)<sup>[18]</sup>、FFDNet(快速全卷积去噪网络)、MSRResNet(多尺度残差网络)<sup>[19]</sup>、RRDB(残差递归密集块)、SRMD(超分辨率多退化模型)等方法，而图像匹配模块采取

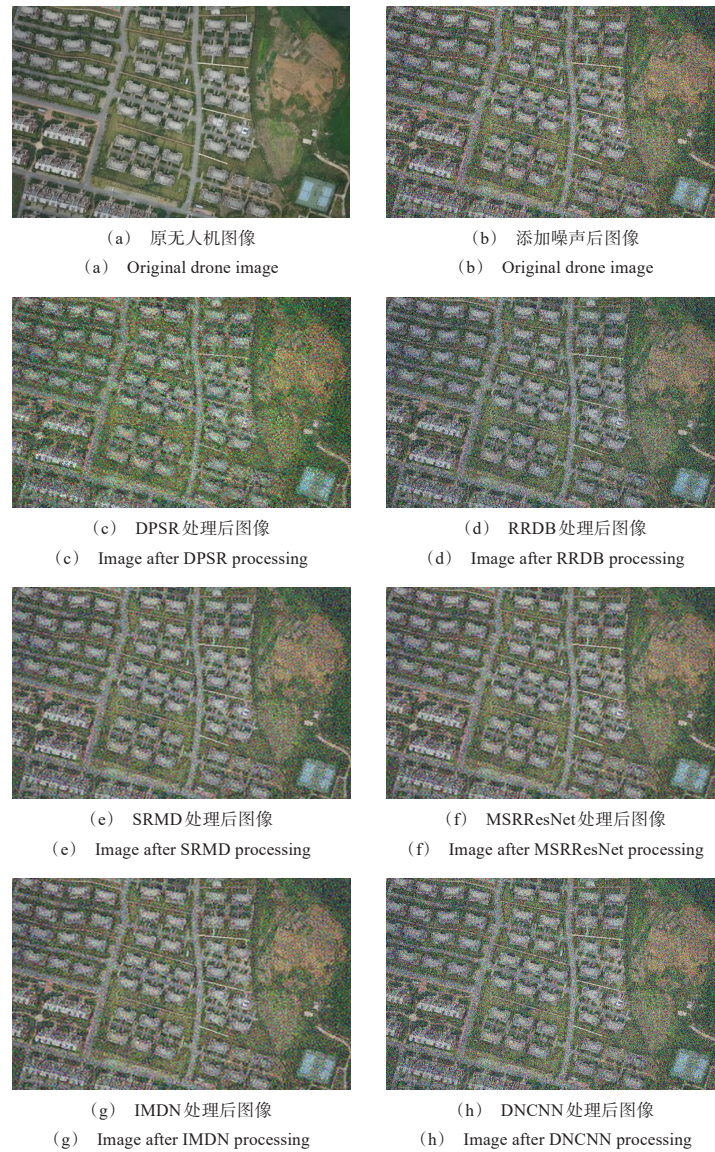


图 10 无人机噪声图像经不同预处理方法处理后结果

Fig.10 The UAV noise image is processed by different preprocessing methods

SuperPoint<sup>[20]</sup>+SuperGlue<sup>[21]</sup>方法(超级点+超级胶点)。在不同预处理情况下的 SuperPoint+SuperGlue 的匹配结果如表 2 所示。

表 2 不同预处理方法下 SuperPoint+SuperGlue 的匹配结果  
Table 2 The matching results of SuperPoint+SuperGlue under the different preprocessing method

算法名称	100 像素误差以内图片数量占比
SuperPoint+SuperGlue	25.5%
IMDN+SuperPoint+SuperGlue	<b>32.4%</b>
DNCNN+SuperPoint+SuperGlue	<b>25.9%</b>
DPSR+SuperPoint+SuperGlue	16.4%
RRDB+SuperPoint+SuperGlue	23.1%

续表

算法名称	100 像素误差以内图片数量占比
MSRResNet+SuperPoint+SuperGlue	<b>31.2%</b>
FFDNet+SuperPoint+SuperGlue	<b>26.8%</b>
SRMD+SuperPoint+SuperGlue	<b>32.8%</b>

从表 2 中可以看出, 在本实验环境中, 经 IMDN、DNCNN、MSRResNet、FFDNet 和 SRMD 算法处理后, 100 像素误差以内的图片数量均有显著提升, 其中, SRMD 算法效果最为突出, 100 像素误差以内图片数量达到了 32.8%, 相较于未预处理之前图片数量提高了 7.3%, 充分证明了预处理算法的有效性。然而, 并非所有的预处理方法都

能够有效提高匹配准确性。例如，DPSR与RRDB未能够提高100像素误差以内的图片数量。这可能是由于不同预处理方法针对的噪声种类不同，从而无法有效提升匹配结果。因此，在实际应用中，建议对可能出现的噪声进行提前调研，并基于实际应用场景图片进行测试，以选择合适的预处理方法。

## 4 图像匹配

空天异源图像匹配技术，现已有基于模板的图像匹配，主要包括基于特征的图像匹配等多种实现思路。基于模板的图像匹配主要通过计算模板图像与测试图像中目标相似度来实现匹配，其方法简单直观，但在复杂环境中，信息与图像畸变较大的情况下表现较差，并且需要根据具体的图像数据和匹配任务，选择合适的模板图像，泛化性不强。相比之下，基于特征的图像匹配可以更好地应对空天异源图像匹配中目标纹理改变、目标视角畸变等不利情况，可以分为传统特征匹配与基于深度学习的特征匹配。

### 4.1 传统特征匹配

传统特征匹配以SIFT<sup>[22]</sup>与SURF<sup>[23]</sup>为代表。SIFT算法又叫尺度不变特征变换匹配算法，SIFT提取出的特征对于无人机航摄图像中旋转，尺度和亮度等变化均具有不变性，计算速度较快，并且对于噪声、视角变化和光照变化具有良好的鲁棒性。SURF是对SIFT进一步改进，SURF特征也对于无人机航摄图像中旋转、尺度和亮度等变化均具有不变性，不同点在于：SURF使用Hessian矩阵来检测特征点并利用积分图加速计算但针对无人机航摄图像弱纹理场景准确率较低。

### 4.2 基于深度学习特征匹配

传统特征匹配方法虽然在一定程度上能满足空天异源图像特征匹配的要求，但无法很好地应对视角畸变、特征模糊等不利情况，特征提取效果不理想。而基于深度学习特征匹配方法通过卷积、池化等操作，能够提取传统特征的匹配方法无法提取到的深度特征，在多数应用场景中表现更为优异。基于深度学习特征匹配算法可以分为直接从原始图像中生成对应关系算法，如LoFTR<sup>[24]</sup>、Efficient-LoFTR<sup>[25]</sup>、Aspanformer<sup>[26]</sup>等，以及匹配稀疏关键点来建立跨图像对应关系算法，例如SuperPoint<sup>[22]</sup>与SuperGlue<sup>[23]</sup>，LightGlue<sup>[27]</sup>等。

与传统特征匹配方法相比，基于深度学习的特征匹配方法可以通过模型学习到图像的深层次特征，具有更强的描述能力和更高的泛化能力。

#### 4.2.1 LoFTR, Efficient-LoFTR, Aspanformer

LoFTR及其改进算法拥有较好的匹配效果与匹配效率，在空天异源图像匹配中得到广泛应用。LoFTR结构如图11所示，由四个部分组成。其首先在粗粒度上建立逐像素的密集匹配，随后在精粒度上进一步优化匹配结果，能够满足空天异源图像匹配中对局部精确匹配的高要求，并且具有较高的匹配准确度与较快的运算速度，但其匹配效率相对较低。

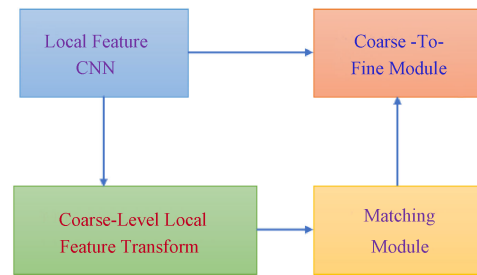


图11 LoFTR网络结构

Fig.11 The network structure of LoFTR

为了进一步提升性能，Efficient-LoFTR在LoFTR基础上进行改进，Efficient-LoFTR引入了自适应聚合注意力机制和两级相关层，显著提高了匹配效率和精度。Efficient-LoFTR结构如图12所示。此外，Aspanformer针对LoFTR缺乏局部关联性的问题，基于分层注意结构，采用了一种新颖的注意力模块，能够自适应地调整注意广度。Aspanformer结构如图13所示。

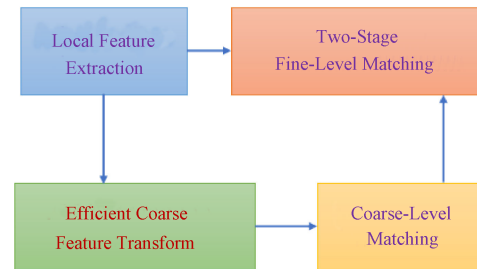


图12 Efficient-LoFTR网络结构

Fig.12 The network structure of Efficient-LoFTR

#### 4.2.2 SuperPoint与SuperGlue, LightGlue

SuperPoint与SuperGlue、LightGlue的结合运用是特征提取-特征匹配两阶段图像匹配方法的典

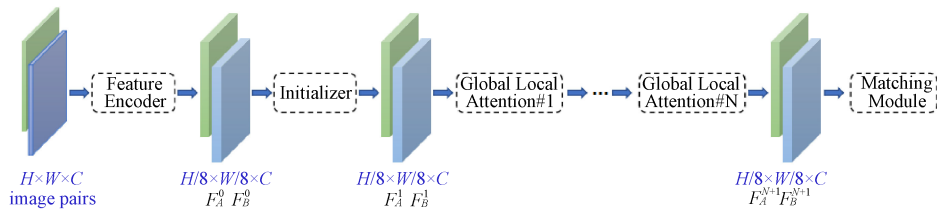


图 13 Aspanformer 网络结构

Fig.13 The network structure of Aspanformer

型代表, Superpoint 负责检测图像关键点并提取关键点的描述子, 而 SuperGlue 和 LightGlue 则基于这些描述子进行特征匹配, 这种方法已在多个实际应用场景取得显著效果。

在无人机应用场景下, SuperPoint 能够准确检测出关键点并生成高质量的特征描述子, 但

在光照变化较大或场景存在遮挡的情况下, 其性能可能有所下降。SuperPoint 网络结构如图 14 所示, 首先将图像输入到共享编码器中进行处理和降低图像维度。之后, 再输入到两个解码器模块, 一个用于兴趣点检测, 另一个用于兴趣点描述。

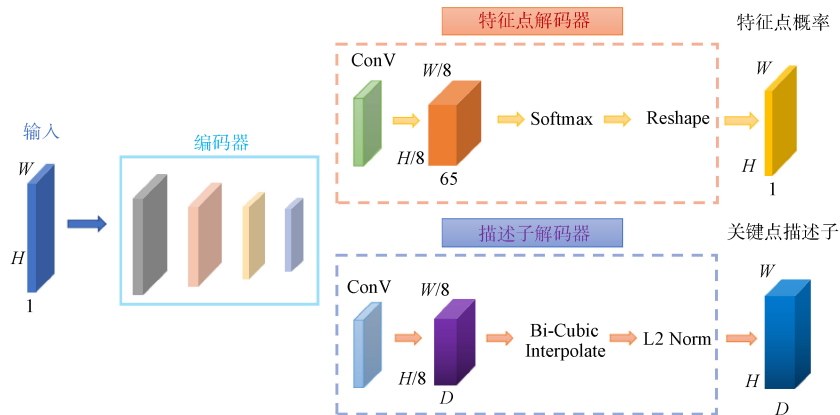


图 14 SuperPoint 网络结构

Fig.14 The network structure of SuperPoint

SuperGlue 是一种基于深度学习的局部特征匹配方法, 能够同时完成空天异源图像特征匹配和滤除外点的任务。然而, SuperGlue 在实时性、灵活性和效率等方面仍有提升空间。SuperGlue 流程图如图 15 所示。SuperGlue 引入了一种基于注意力的灵活上下文聚合机制, 通过联合查找对应点和

拒绝不匹配点来匹配两组局部特征。该方法通过求解一个可微最优运输问题来估计特征点之间的分配关系, 并利用图神经网络预测运输成本, 从而提高了匹配的精度和鲁棒性。

LightGlue 在 SuperGlue 的基础上更进一步, 能够更灵活地适应空天异源图像的匹配需求, 在准确

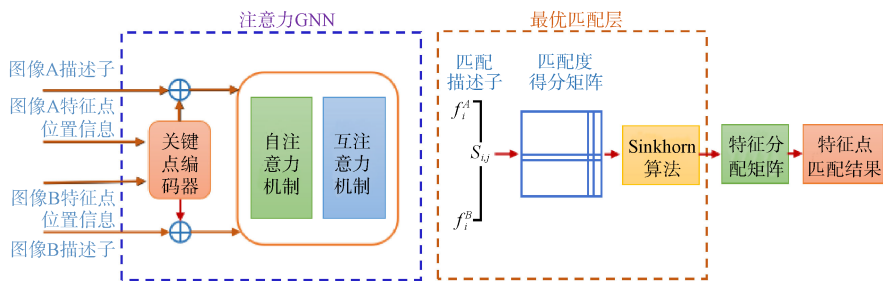


图 15 SuperGlue 网络结构

Fig.15 The network structure of SuperGlue

性、效率和训练易用性方面优于现有的 SuperGlue。其网络结构如图 16 所示，由  $L$  个相同的层堆叠而成。

每一层都由自注意单元和交叉注意单元组成。分类器在每一层决定是否停止推理，从而避免不必要的计算。

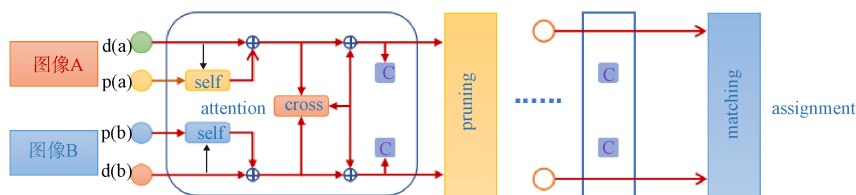


图 16 LightGlue 网络结构

Fig.16 The network structure of LightGlue

为验证匹配方法有效性，实验采用了 SRMD 方法进行图像预处理，并在图像匹配模块采取 LoFTR、Efficient-LoFTR、SuperPoint+SuperGlue、SuperPoint+LightGlue、Aspanformer、3dg-stfm 等方法。SRMD 预处理后不同图像匹配方法的 100 像素误差以内图片数量占比如表 3 所示。

表 3 SRMD 预处理后不同图像匹配方法的 100 像素误差以内图片数量

Table 3 The number of images within 100 pixel error of different image matching methods under SRMD

算法名称	预处理前	预处理后
SuperPoint+SuperGlue	25.5%	<b>32.8%</b>
SuperPoint+LightGlue	18.9%	<b>25.6%</b>
LoFTR	17.1%	<b>27.6%</b>
Efficient-LoFTR	32.6%	<b>44.4%</b>
3dg-stfm	32.3%	<b>37.1%</b>
Aspanformer	33.0%	<b>43.3%</b>

从表 3 可以得出，在本实验场景中，不同图像匹配算法经 SRMD 预处理后结果均有提升，验证了多算法融合的空天异源图像匹配框架合理性。然而，不同图像匹配方法的效果存在差异，其中 Efficient-LoFTR 表现最佳，最适用于本实验场景。但在实验过程中，仍存在以下问题：

①基于深度学习的图像匹配算法相较于传统图像匹配算法效果较好，但对算力要求比较高，很难在嵌入式的平台上部署。

②算法训练集图像与无人机侦察图像存在差异，算法训练集图像轮廓较为清晰，特征较为明显，无人机侦察图像存在物体阴影、轮廓不清晰等问题，导致特征提取困难。算法在实际应用之前如果先采用无人机航摄图像进行大量训练，可能进一步提高算法的准确率。

## 5 结束语

本文针对空天异源图像匹配技术，介绍了一个可用于实际应用的空天异源图像匹配框架，描述了框架的主要组成关键技术，给出了各关键技术可采用的典型方法的概念及应用情况，对各关键组成技术提及方法进行了实验对比，给出了实验结果与思考。

总体而言，空天异源图像匹配技术作为无人机的重要技术之一，虽然已经取得了一定成果，但涉及技术多为其他领域相关技术的迁移应用，针对空天异源图像匹配领域专用技术较少。空天异源图像匹配技术主要发展趋势如下：

①空天异源图像质量评价算法改进。目前，针对空天异源图像的图像质量评价方法较少，且无法很好地应对空天异源图像匹配实际应用中多场景、多分辨率、多干扰等情况。未来可以采用迁移学习与现有算法结合的策略，研究空天异源图像匹配领域专用算法，以去除无效图片、减少设备资源消耗并缩短任务执行时间。

②可用于工程实践的空天异源图像预处理算法。现有图像预处理算法难以有效处理如噪声、云雾、大目标遮挡等多种干扰同时存在的情况，需要设计可用于工程实践的空天异源图像预处理算法，有效地判别不同干扰情况，同时针对每种情况挑选多种算法进行实验，选出最优算法。

③实时鲁棒的空天异源图像匹配算法的研究。现有图像匹配算法由于无人机的接收数据量大与模型参数众多、体积庞大等原因耗时较长，无法满足实际工程需要。应进一步研究空天异源图像匹配算法，结合知识蒸馏、知识迁移等策略，同时兼顾算法实时性与准确性。

④多视角空天异源图像匹配算法的研究。利用多视角信息可以更好地辅助定位工作, 但现有算法无法充分利用多视角信息, 且多视角信道彼此之间的运算复杂独立、信息互通性不理想, 导致消耗很多运算资源。需在已有多视角算法基础上, 结合时空一致性、视角一致性和颜色一致性等策略进行空天异源图像多视角匹配研究。

### 参考文献

- [1] MOORTHY A K, BOVIK A C. A two-step framework for constructing blind image quality indices[J]. *IEEE Signal Processing Letters*, 2010, 7(5): 13-16.
- [2] MOORTHY ANUSH KRISHNA, BOVIK ALAN CONRAD. Blind image quality assessment: From natural scene statistics to perceptual quality[J]. *IEEE Transactions on Image Processing*, 2011, 20(12): 3350-3364. DOI: 10.1109/TIP.2011.2160799.
- [3] SAAD M A, BOVIK A C, CHARRIER C. Blind image quality assessment: A natural scene statistics approach in the DCT domain[J]. *IEEE Transactions on Image Processing*, 2012, 21(8): 3339-3352. DOI: 10.1109/TIP.2012.2197527.
- [4] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a "completely blind" image quality analyzer[J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212. DOI: 10.1109/LSP.2013.2248785.
- [5] MITTAL A K, MOORTHY A, BOVIK A C. No-reference image quality assessment in the spatial domain [J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695-4708. DOI: 10.1109/TIP.2012.2214050.
- [6] GU S Y, BAO J M, CHEN D, et al. GIQA: Generated image quality assessment[EB/OL]. [2022-11-08]. <https://arxiv.org/abs/2003.08932>.
- [7] SU S, et al. Blindly assess image quality in the wild guided by a self-adaptive hyper network[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA. New York: IEEE, 2020: 3664-3673. DOI: 10.1109/CVPR42600.2020.00372.
- [8] YOU J, KORHOVEN J. Transformer for image quality assessment[C]//Proceedings of the 2021 IEEE International Conference on Image Processing. Anchorage, AK, USA. New York: IEEE, 2021: 1389-1393. DOI: 10.1109/ICIP42928.2021.9506075.
- [9] KIM J, NGUYEN A D, LEE S. Deep CNN-based blind image quality predictor[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 30(1): 11-24. DOI: 10.1109/TNNLS.2018.2829819.
- [10] ZUIDERVELD K. Contrast limited adaptive histogram equalization[J]. *Graphics Gems*, 1994: 474-485. DOI: 10.1016/B978-0-12-336156-1.50061-6.
- [11] ZHANG K, ZUO W, CHEN D, et al. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising[J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142-3155. DOI: 10.1109/TIP.2017.2662206.
- [12] ZHANG K, ZUO W, ZHANG L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising[J]. *IEEE Transactions on Image Processing*, 2018, 27(9): 4608-4622. DOI: 10.1109/TIP.2018.2839891.
- [13] ZHANG K, ZUO W, ZHANG L. Learning a single convolutional super-resolution network for multiple degradations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2017. DOI: 10.48550/arXiv.1712.06116.
- [14] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. *Computer Science*, 2014. DOI: 10.48550/arXiv.1409.1556.
- [15] WANG Y, HUANG H, XU Q, et al. Practical deep raw image denoising on mobile devices[C]//Proceedings of the European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2020: 12351. DOI: 10.1007/978-3-030-58539-6\_1.
- [16] HUI Z, GAO X, YANG Y, et al. Lightweight Image Super-Resolution with Information Multi-distillation Network[J]. *ACM*, 2019. DOI: 10.1145/3343031.3351084.
- [17] XU W, YAO Y, CAO J, et al. UAV-VisLoc: A Large-scale Dataset for UAV Visual Localization[J]. 2024.
- [18] ZHANG K, ZUO W, ZHANG L. Deep Plug-And-Play Super-Resolution for Arbitrary Blur Kernels[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019. DOI: 10.1109/CVPR.2019.00177.
- [19] WANG X, YU K, WU S, et al. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks[J]. 2018. DOI: 10.1007/978-3-030-11021-5\_5.
- [20] Detone D, Malisiewicz T, Rabinovich A. SuperPoint:

- Self-Supervised Interest Point Detection and Description [J]. 2017. DOI:10.48550/arXiv.1712.07629.
- [21] SARLIN P E, DETONE D, MALISIEWICZ T, et al. SuperGlue: Learning feature matching with graph neural networks[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020. DOI: 10.1109/CVPR42600.2020.00499.
- [22] LOWE D. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2003, 20(91-110). DOI: 10.1023/B: AVISI.0000029664.99615.94.
- [23] BAY H, TUYTELAARS T, GOOLSVEN L V. SURF: Speeded up robust features[C]//Proceedings of the 9th European Conference on Computer Vision. Berlin: Springer-Verlag, 2006: 32. DOI: 10.1007/11744023\_32.
- [24] SUN J, SHEN Z, WANG Y, et al. LoFTR: Detector-free local feature matching with transformers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2021. DOI: 10.1109/CVPR46437.2021.00881.
- [25] WANG Y, HE X, PENG S, et al. Efficient LoFTR: Semi-dense local feature matching with sparse-like speed[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2024: 21666-21675.
- [26] CHEN H, LUO Z, ZHOU L, et al. Aspanformer: Detector-free image matching with adaptive span transformer[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 20-36.
- [27] LINDENBERGER P, SARLIN P E, POLLEFEYS M. LightGlue: Local feature matching at light speed[C]//Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision, Paris, France. New York: IEEE, 2023: 17581-17592. DOI: 10.1109/ICCV51070.2023.01616.

#### [作者简介]

- 刘庚辰 2000年生, 硕士研究生。  
吴国强 1984年生, 博士研究生。  
黄坤 1990年生, 硕士。  
刘奇霏 2000年生, 硕士研究生。  
黄蜀玲 1986年生, 博士。

(本文编辑: 赵尹默)