

面向非刚性形变图像的关键点匹配方法*

董家麟, 吴丽君

(福州大学 物理与信息工程学院, 福建 福州 350108)

摘要: 在图像匹配任务中, 非刚性变形会显著干扰传统关键点匹配算法的性能, 造成匹配精度下降。针对此问题, 提出一种基于概率建模与特征增强的匹配方法。首先, 设计自适应关键点采样策略, 通过动态空间约束与边界感知过滤, 解决复杂形变下的关键点分布冗余问题。其次, 针对确定性几何变换估计易陷入局部最优的缺陷, 引入 Student-t 分布建模几何变换参数的不确定性, 提升描述符在复杂形变下的适应能力。最后, 构建高斯先验的特征增强模块, 利用方向敏感算子强化关键结构信息, 有效抑制非刚性变形引发的特征偏移。实验结果表明, 该方法可有效提升非刚性图像的匹配精度, 在处理复杂非刚性形变场景时具有更强的鲁棒性。

关键词: 非刚性图像匹配; 关键点检测; 不确定性建模; 特征增强

中图分类号: TP391

文献标志码: A

DOI: 10.19358/j.issn.2097-1788.2026.03.005

中文引用格式: 董家麟, 吴丽君. 面向非刚性形变图像的关键点匹配方法 [J]. 网络安全与数据治理, 2026, 45(3): 33-39.

英文引用格式: Dong Jialin, Wu Lijun. Key point matching method for non-rigid deformed images [J]. Cyber Security and Data Governance, 2026, 45(3): 33-39.

Key point matching method for non-rigid deformed images

Dong Jialin, Wu Lijun

(College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China)

Abstract: Non-rigid deformations significantly impair the performance of traditional keypoint matching algorithms in image matching tasks, resulting in reduced matching accuracy. To address this issue, a matching method based on probabilistic modeling and feature enhancement is proposed. First, an adaptive keypoint sampling strategy is designed to resolve the redundancy in keypoint distribution under complex deformations by employing dynamic spatial constraints and boundary-aware filtering. Second, to overcome the limitation where deterministic geometric transformation estimation tends to fall into local optima, a Student-t distribution is introduced to model the uncertainty of geometric transformation parameters, thereby enhancing descriptor adaptability in complex deformation scenarios. Finally, a feature enhancement module based on Gaussian priors is constructed. This module utilizes direction-sensitive operators to reinforce key structural information, effectively mitigating feature drift caused by non-rigid deformations. Experimental results demonstrate that the proposed method effectively improves matching accuracy for non-rigid images and exhibits superior robustness in scenarios involving complex non-rigid deformations.

Key words: non-rigid image matching; keypoint detection; uncertainty modeling; feature enhancement

0 引言

图像匹配是计算机视觉领域的一项基础且关键的技术, 其核心任务是在不同视角、不同时间或不同传感器获取的图像之间建立可靠的对应关系。该技术广泛应用于三维重建^[1]、医学图像分析^[2]、目标跟踪^[3]、结构健康监测^[4]等众多领域。基于特征的匹配方法是

该领域的主流技术路线, 其通常包含关键点检测、特征描述和特征匹配三个核心步骤。以 SIFT (Scale-Invariant Feature Transform)、SURF (Speeded-Up Robust Features) 为代表的经典算法, 通过提取具有尺度和旋转不变性的局部特征, 在处理刚性或仿射变换的场景中取得了巨大成功。特别是随着深度学习的兴起, 基于学习的匹配方法已成为新的研究热点。早期工作如 SuperPoint^[5] 实现了关键点检测与描述的端到端联合学

* 基金项目: 国家自然科学基金 (62271151, W2421092)

习, 展现出超越传统方法的潜力。为了适应更具挑战性的环境, LiftFeat^[6]通过单目深度预测来引入三维几何线索, 有效增强了描述符在光照剧变和弱纹理区域的判别力。而 XFeat^[7]则专注于构建轻量化网络架构, 在保证高精度的同时实现了极高的运行速度, 满足了资源受限设备上的实时匹配需求。

然而, 在许多实际应用中, 图像内容会发生复杂的非刚性变形 (Non-rigid Deformation), 例如人体姿态变化、面部表情变化、衣物褶皱以及建筑结构或生物组织的形变等。这种不规则的几何形变严重破坏了传统特征匹配算法所依赖的局部几何结构不变性假设, 导致特征描述符的区分度下降、匹配歧义性增加, 最终造成匹配准确率和鲁棒性显著降低。因此, 如何有效处理非刚性变形, 实现高精度的图像匹配, 已成为计算机视觉领域备受关注的研究热点。

为应对非刚性变形挑战, 早期工作如 DaLi^[8]通过编码局部 3D 曲面提升了鲁棒性, 但计算与存储开销较大。GeoBit^[9]与 GeoPatch^[10]等方法虽引入了几何感知特征, 却严重依赖含噪的深度图像, 限制了应用。为减少对深度信息的依赖, DEAL^[11]在单目图像中隐式建模变形, 但其性能仍受制于传统关键点检测器缺乏形变等变性的约束。后续的 DALF^[12]虽通过空间变换网络 (STN) 联合学习形变特征取得了显著进展, 但其采用的确定性参数估计在面对复杂变形时易陷入局部最优, 并可能导致关键边缘特征的模糊与偏移。这

正是本研究旨在解决的核心问题。

针对上述局限, 本文提出一种面向非刚性图像的鲁棒关键点匹配方法, 主要贡献如下:

(1) 提出自适应关键点采样策略 (AKS): 突破了传统基于响应值的采样在形变区域分布不均的限制, 通过动态调整空间分布约束确保稀疏性, 提升了关键点在形变边缘的定位精度。

(2) 构建概率化薄板样条变换网络 (P-TPSNet): 摒弃了传统的确定性参数估计范式, 通过对变换参数进行概率分布建模与边际化处理, 有效规避了复杂形变估计中的局部最优解问题, 实现了对几何变换不确定性的显式建模。

(3) 设计高斯概率特征增强模块 (GPFE): 针对非刚性形变导致的特征漂移与高频信息丢失, 引入高斯空间先验强化关键结构表达, 在不增加显著计算开销的前提下提升了描述符的判别力。

1 非刚性图像的关键点检测与匹配方法

本文提出的非刚性图像关键点检测与匹配方法流程如图 1 所示。首先, 将两幅待匹配图像分别输入特征提取网络, 以获得其对应的特征表示以及关键点热图。随后, 基于热图执行自适应关键点采样以提取图像关键点, 同时对特征表示进行结构化增强, 并利用形变修正网络校正关键点图像块, 从而生成关键点描述符。最后, 通过描述符匹配与 RANSAC 几何校验, 筛选得到最终的匹配对。

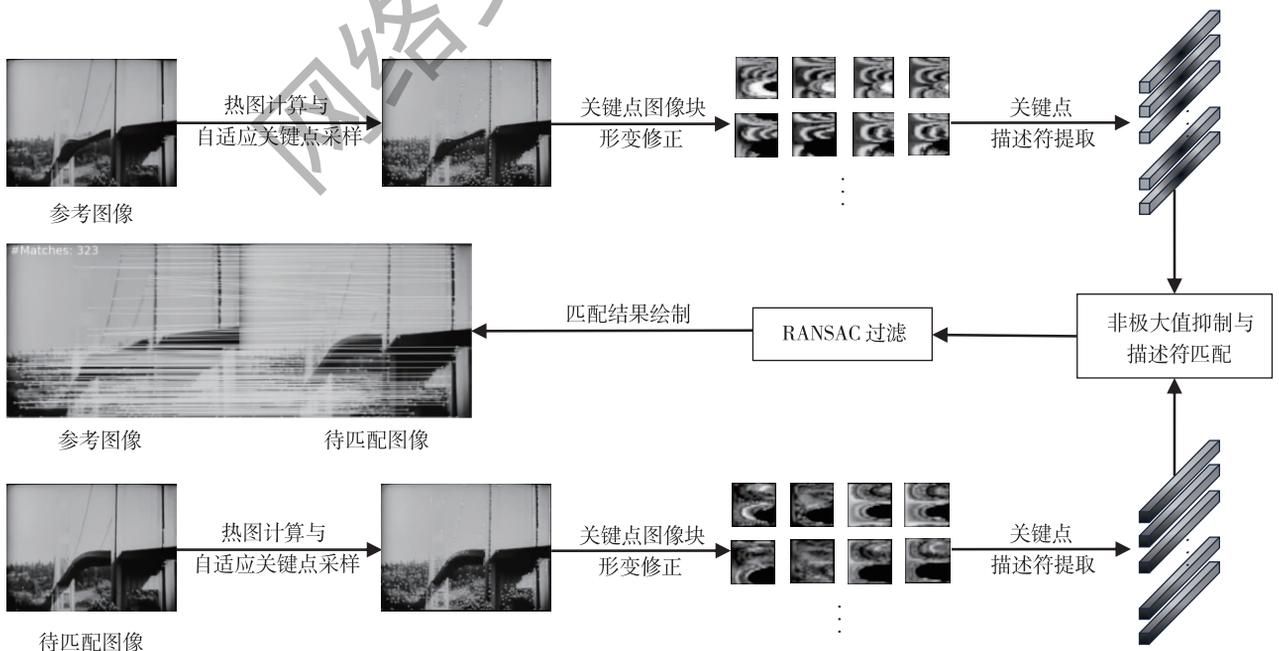


图 1 非刚性图像的关键点检测与匹配方法流程

具体网络架构如图 2 所示。首先，基于标准 U-Net 主干提取分层特征。其中，解码器输出热图 H 用于关键点采样；编码器瓶颈特征 F_{enc} 则输入后续模块。接

着，利用高斯概率建模增强特征，并经由概率化薄板样条网络 (P-TPSNNet) 计算非刚性描述符。最后，通过线性层融合通用描述符，获得最终的特征表示。

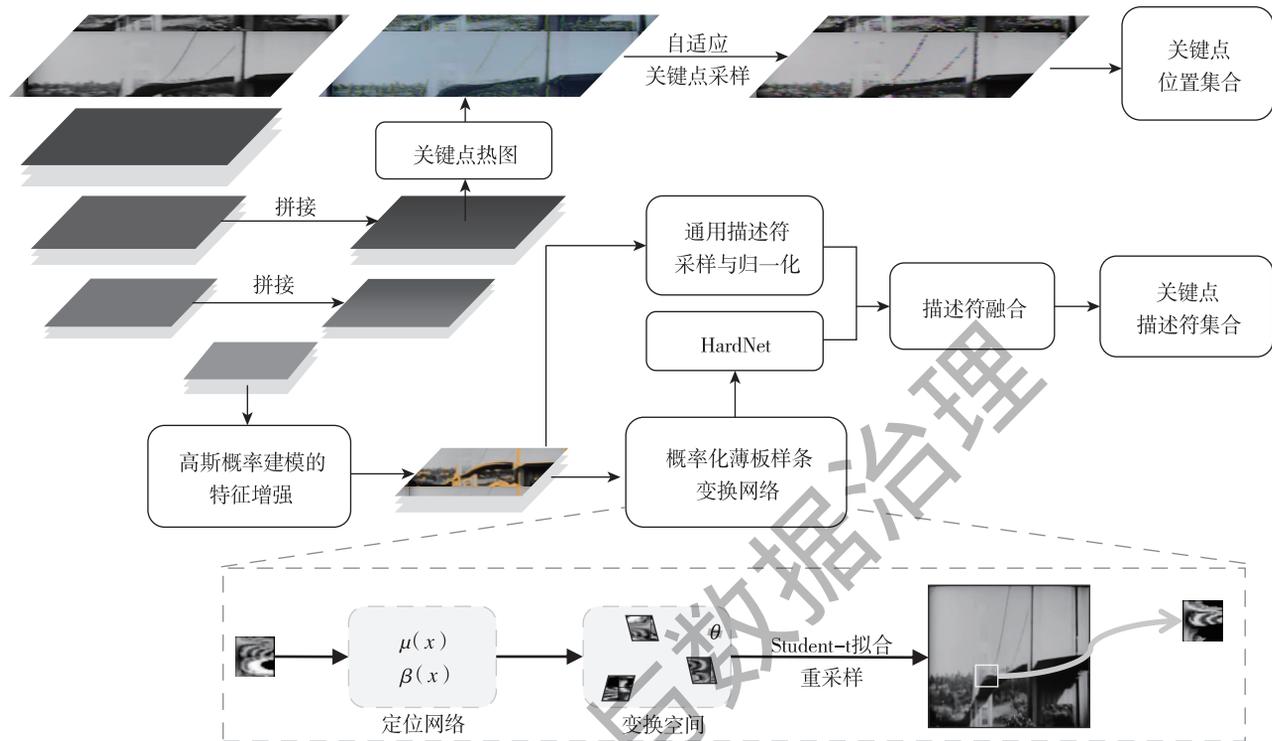


图 2 非刚性图像的关键点检测与匹配网络架构

1.1 自适应关键点采样策略

受 SuperPoint^[5] 与 DISK^[13] 的启发，本文的关键点提取基于对网络生成的响应热图进行采样。检测器模块与特征描述子模块进行协同端到端训练，旨在学习对常见视觉变换具有不变性的特征表示。为进一步提升所提取关键点的可重复性 (repeatability) 与匹配可靠性 (reliability)，引入了基于强化学习的梯度策略，以优化检测概率的分布，从而鼓励网络在不同视角、光照以及非刚性变换下检测出鲁棒的关键点。

在非刚性变形普遍存在的复杂场景中，稳定且精确的关键点采样尤为关键。首先构建基础概率采样框架，将解码器输出的热图 $H \in \mathbb{R}^{1 \times h \times w}$ 划分为 $m \times n$ 个非重叠的子块 $p_{i=1}^{m \times n}$ (窗口大小设为 8×8)。每个子块 p_i 代表局部关键点响应强度。随后，在每个子块 p_i 内采用两阶段概率采样，第一阶段按照 Softmax 归一化的响应值分布从 p_i 内采样像素位置 k ：

$$P(z_i = k) = \frac{\exp(p_i^k)}{\sum_{j=1}^{w \times h} w \times w \exp(p_i^j)} \quad (1)$$

其中， z_i 表示第 i 个子块中选中的像素索引， p_i^k 表示第

i 个子块中第 k 个位置的响应值。

第二阶段基于响应值通过伯努利分布判断是否保留：

$$P(a_i = 1 | z_i = k) = \sigma(p_i^k) \quad (2)$$

其中 $a_i \in \{0, 1\}$ 表示是否接受当前关键点， $\sigma(\cdot)$ 表示 Sigmoid 函数。仅 $a_i = 1$ 的候选点会反投影至原图坐标系。

为提升关键点可重复性，引入 DISK^[13] 中基于强化学习的训练策略，给定同场景图像对 (I, I') 以及真值流场 $GT(\cdot)$ ，定义奖励函数 $R(\cdot)$ ，用于度量一个关键点在另一幅图像中是否存在对应点：

$$R(p_j^i) = \begin{cases} 1, & \exists p_k^r \| GT(p_j^i) - p_k^r \| < \varepsilon \\ 0, & \text{其他} \end{cases} \quad (3)$$

其中 p_j^i 为图像 I 中检测到的关键点坐标， ε 为容忍的像素距离阈值。

此奖励函数仅当关键点能够在对应图像中找到可靠匹配时给予正反馈，进一步构造关键点对集合 $X = X_I \times X_{I'}$ ，其中 X_I 和 $X_{I'}$ 为图像 I 和 I' 中采样得到的关键点集合，优化目标是最大化期望奖励的对数似然，通

过策略梯度方法更新网络参数 θ :

$$\nabla_{\theta} E_K[R(\cdot)] = \sum_{(x,y) \in K} \nabla_{\theta} (\log P(x;\theta) + \log P(y;\theta)) R(x,y) + \lambda L_{reg} \quad (4)$$

其中 $P(x; \theta)$ 表示在当前参数下采样关键点 x 的概率, 采用梯度上升法来更新网络参数 θ , 以最大化该期望奖励。此外, 为了进一步提升描述符的判别力并稳定训练过程, 对那些判别性较差的关键点 (例如响应值过低或位于无纹理区域) 的奖励强制置零, 并引入一个正则化项 L_{reg} 来约束模型的复杂度。

然而, 基础策略在纹理密集区易产生冗余采样, 且在子块边界处定位精度较低, 增加了计算负担。为此, 本文提出自适应关键点采样策略 (Adaptive Keypoint Sampling, AKS)。该策略采用分阶段课程学习机制: 在训练初期 (前 30% 迭代), 移除稀疏性约束以充分探索概率分布; 在训练中期 (30% ~ 75%): 定义动态松弛因子 $f_{rlx} = tot_steps/cur_steps$ 和基础对数概率阈值 t_{base} (实验中设置为 -0.05) 以及最小距离阈值 $t_d = 1.5$ 像素, 保留对数概率高于动态阈值 $t_p = f_{rlx} \times t_{base}$ 的冲突点; 在训练后期 (75% 以后), 将最小距离提升至 2.0 像素, 并固定 f_{rlx} 为 1, 以强化空间分布的均匀性。

接着, 进行边界感知过滤, 将每个子块外围 1 像素设为“不安全边界区”, 其内点需更严格过滤。对边界候选集执行非极大值抑制, 即若两点距离小于当前阈值且概率劣于另一点, 则移除该点。最后, 重复采样点保留多次被选中的位置视为重要点, 即使与他点距离近也保留, 并将其对数得分更新为原始响应值的平均概率。通过动态空间正则、边界过滤与重复点保留, 在保证召回率的同时生成分布更均匀、冗余度更低的关键点集合, 提升匹配效率与精度。

1.2 关键点描述

为保证描述子在复杂变换下的鲁棒性, 本文在几何矫正中引入了概率化思想。具体地, 采用双分支网

络, 学习薄板样条 (TPS) 变换参数的 Student-t 分布, 取代了传统的确定性估计。通过在训练时对该分布进行采样, 模型能有效规避局部最优, 显著提升了描述符在复杂非刚性场景下的稳定性。此外, 为应对形变导致的边缘结构扭曲与高频信息丢失问题, 设计了基于高斯概率建模的特征增强模块。该模块在几何矫正后、生成描述子前对特征图进行强化, 以保留关键的结构信息。整体架构如图 2 所示。

1.2.1 基于高斯概率建模的特征增强

为弥补深度编码器下采样造成的精细结构信息损失, 本文设计了高斯概率建模特征增强模块 (Gaussian Probability-based Feature Enhancement, GPFE)。该模块利用关键点邻域特征呈空间高斯分布的先验, 对特征进行加权增强, 以强化关键结构并抑制噪声 (结构见图 3)。

首先, 利用一个固定的、不可训练的高斯核对输入特征进行深度卷积 (Depthwise Conv)。该操作在每个特征通道上独立进行, 能够模拟特征的自然空间分布, 从而在增强核心特征的同时平滑噪声, 得到空间增强特征 F_{gauss} :

$$F_{gauss} = \text{Gaussian}(F_{in}) \quad (5)$$

其中, $\text{Gaussian}(\cdot)$ 代表使用固定高斯核的深度卷积操作, 其权重由标准高斯函数生成。随后, 为了自适应地调整不同特征通道的重要性, 引入高效通道注意力 (Efficient Channel Attention, ECA) 机制。具体来说, 将原始特征 F_{in} 与空间增强特征 F_{gauss} 进行融合, 并通过全局平均池化 (Global Average Pooling, GAP) 和一维卷积来学习通道间的相关性, 生成通道注意力权重 F_{c_attn} :

$$F_{c_attn} = \text{Sigmoid}(\text{Conv}_{1D}(\text{GAP}(F_{fusion}))) \quad (6)$$

其中, F_{fusion} 是 F_{in} 和 F_{gauss} 的融合特征。最后, 将学习到的通道权重应用到融合特征上, 并通过残差连接和原始特征 F_{in} 整合, 得到最终的增强特征表示 F_{out} :

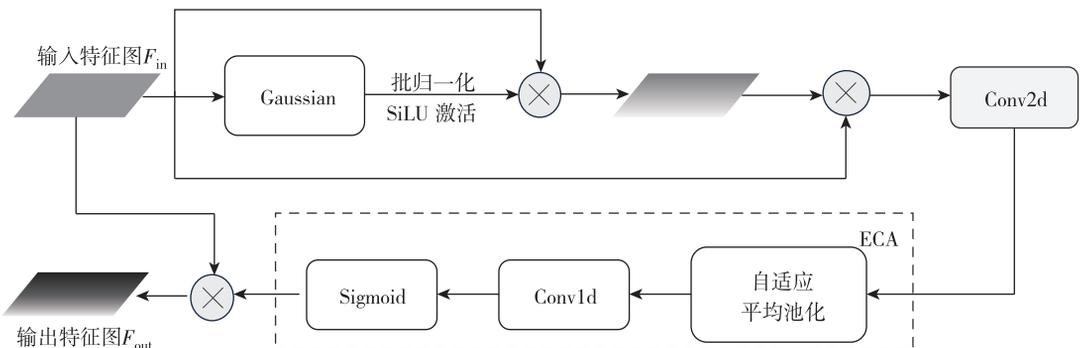


图 3 高斯概率建模特征增强模块结构

$$F_{\text{out}} = \text{Norm}((F_{\text{c_attn}} \odot F_{\text{fusion}}) + F_{\text{in}}) \quad (7)$$

通过这种方式，显著提升了描述子对边缘等核心结构信息的表达能力，为后续生成鲁棒的关键点描述符奠定了坚实的基础。

1.2.2 概率化薄板样条变换网络

传统的空间变换网络通常利用确定性映射函数来预测变换参数 $\theta = f(x)$ 。以 TPS 变换为例，这是一种基于径向基函数的非刚性形变变换方法，通过定义源控制点集 $\{p_j\}_{j=1}^N$ 到目标位置 $\{q_j\}_{j=1}^N$ 的映射来实现局部形变矫正。TPS 映射变换函数定义为：

$$T(x, y) = a_1 + a_2x + a_3y + \sum_{j=1}^N \omega_j U(\|(x, y) - p_j\|) \quad (8)$$

其中 $U(r) = r^2 \log(r)$ 为径向基函数，需要学习的变换参数为 $\{a_1, a_2, a_3, \omega_1, \dots, \omega_N\}$ 。这些参数通过求解线性方程组获得。然而，这种确定性估计方法存在显著局限^[14]：无法量化不确定性，单一的参数输出降低了模型在模糊区域的自适应能力；易陷入局部最优，尤其在变形边界模糊时，容易收敛到次优解；鲁棒性不足，预测误差会直接传递并影响后续特征提取质量。为克服上述限制，本文构建概率化薄板样条变换网络 (Probabilistic Thin-Plate Spline Transformation Network, P-TPSNet)。该网络不再预测一个确定的变换参数 θ ，而是学习其后验概率分布 $p(\theta | x)$ ，从而显式地对变换的不确定性进行建模。具体而言，选择 Student-t 分布对 TPS 参数进行概率化建模。相比于常见的高斯分布，Student-t 分布具有更重的尾部特性，这使其能够产生更多样化的变换样本，从而更好地捕捉和适应非刚性变形中可能出现的极端或异常情况。其概率密度函数定义如下：

$$p(\theta | x) = \text{StudentT}(\theta; \text{df} = 2\alpha, \text{loc} = \mu(x), \text{scale} = \sqrt{\beta(x)/\alpha}) \quad (9)$$

其中， α 是一个预设的、控制分布形状的超参数； $\mu(x)$ 和 $\beta(x)$ 分别是网络需要学习的位置（均值）参数和尺度相关参数。为了预测上述分布所需的参数，设计了一个双分支网络结构。给定输入特征 x ，网络并行地通过两个分支进行处理，得到预测参数的均值与方差，构建 Student-t 分布并从中多次采样得到最终的薄板样条参数，减少单次错误估计的影响。使用该参数在原图像关键点图像块上进行重采样以修复变形区域，送入 HardNet^[15] 网络得到描述符，此描述符专注于非刚性区域的不变形，代价是失去了部分描述符本身的区分性。因此，本文对通用描述符与非刚性描述

符通过全连接层进行融合得到最终描述符。

1.3 网络训练与损失函数

网络的训练分为两个阶段：首先，优化特征提取与增强模块以生成信息丰富的描述符；其次，训练概率化 TPS 网络以增强描述符的几何不变性。最终特征由两阶段结果融合而成。训练由一个复合损失函数驱动，该函数包含关键点检测损失与描述符损失。检测损失函数定义为期望奖励 $E_{\kappa}[R(\cdot)]$ 的负值，记作 L_{det} 。其中，奖励函数 $R(\cdot)$ 的具体定义和优化方式已在第 1.1 节中详细阐述。此损失项促使网络优先关注那些在不同视角下都能被稳定检测到的关键点。接着，为了提升描述符的判别能力，采用最难批内三元组损失 (Hardest-in-batch Triplet Loss, HTLoss)，记作 L_{desc} 。对于一个包含 N 组对应关键点对的批次，首先通过描述子网络 $H(\cdot)$ 提取其特征向量，构成两个特征矩阵 $A, B \in \mathbb{R}^{N \times D}$ ，其中 D 为描述符维度且所有向量均已 L2 归一化。对于第 i 对样本，其正样本对距离 $\delta_+^{(i)}$ 为对应描述符间的欧式距离：

$$\delta_+^{(i)} = \|A_i - B_i\|_2 \quad (10)$$

最难负样本距离 $\delta_h^{(i)}$ 则是在批内所有非对应匹配项中，与 A_i 距离最近的：

$$\delta_h^{(i)} = \min_{j \neq i} \|A_i - B_j\|_2 \quad (11)$$

最终，描述符三元组损失函数定义如下：

$$L_{\text{desc}} = \frac{1}{N} \sum_{i=1}^N \max(0, \mu + \delta_+^{(i)} - \delta_h^{(i)}) \quad (12)$$

其中， μ 是一个预设的间隔 (margin) 超参数。该损失函数旨在缩小正样本对间的距离，同时确保最难负样本对与正样本对的距离差至少为 μ ，从而增强描述符的判别力。网络总损失为 $L_{\text{total}} = L_{\text{det}} + \lambda L_{\text{desc}}$ ，其中， λ 是一个权重项，用于平衡检测损失和描述符损失之间的量级。

2 实验结果与分析

2.1 数据集与评价指标

本文采用无监督范式，基于 SFM 数据集^[16] 生成训练数据。通过对原始图像施加合成光度与几何形变生成训练图像对及真值光流场。为确保训练收敛的稳定性，本文实施课程学习策略：依照从易到难的原则，初始阶段仅使用弱形变样本，随后随迭代进程逐步增加形变幅度，直至训练进度的 60% 时达到最大形变难度。该方式有助于模型在早期充分学习关键点检测，并逐步适应复杂场景。本文在 GeoBit^[9] 提出的非刚性形变基准数据集上进行评估。该数据集包含 11 个真实形变物体的图像序列，并提供了可靠的真值对应关系。

性能评估采用标准的匹配得分 (Matching Score,

MS) 和平均匹配准确率 (Mean Matching Accuracy, MMA) 指标, 像素误差阈值统一设为 3。匹配得分是衡量在给定图像对 (I_i, I_j) 中正确匹配的关键点数量占总检测关键点数量的比例。

$$MS = \#correct / \min (\#keypints_i, \#keypints_j) \quad (13)$$

其中, #correct 是通过最近邻搜索得到的正确匹配数量, #keypints_i, #keypints_j 分别是图像 I_i 和 I_j 中检测到的关键点总数。MS 能够综合反映关键点的可重复性和描述符的判别性。

平均匹配准确率更专注于评估描述符的性能, 计算正确匹配数占所有可匹配关键点 (即存在真实对应点的关键点) 总数的比例。

$$MMA = \#correct / \#matchable \quad (14)$$

其中, #matchable 表示在两幅图像公共视域内存在真实对应关系的关键点总数。MMA 能更纯粹地反映描述符在找到正确对应关系时的准确度。

2.2 实施细节

在实验中, 使用以下超参数值: 描述符三元组损失中 $\mu = 0.5$; 平衡损失项 $\lambda = 0.005$; 关键点惩罚项 $c = -7 \times 10^{-5}$; 薄板样条空间变换网络控制点数为 64。模型训练遵循第 1.3 节所述的双阶段策略。训练阶段, 第一阶段训练 80k 次迭代, 第二阶段训练 100k 次迭代。网络在 PyTorch 框架下实现, 包含约 97 万可训练参数。使用单块 NVIDIA GeForce RTX 3080 Ti GPU, 两阶段总训练时长约为 20 h。在推理过程中, 从热图中提取关键点时, 采用 3×3 窗口的非极大值抑制 (NMS)。

2.3 消融实验

为验证各组件的有效性, 本文在 GeoBit^[9] 基准上进行了一系列消融实验。以完整模型为基准, 通过逐一移除以下关键组件来评估性能变化: 自适应关键点采样 (AKS)、高斯概率建模特征增强模块 (GPFE) 以及概率化薄板样条变换网络 (P-TPSNet)。消融实验的详细性能对比如表 1 所示, 涵盖了匹配得分与平均匹配准确率两项关键指标。

表 1 消融实验匹配得分与平均匹配准确率对比

本文方法	MS	MMA
完整模型	0.56	0.85
- AKS	0.55	0.85
- GPFE	0.55	0.83
- P-TPSNet	0.54	0.82

上述消融实验结果清晰地展示了各个改进模块的必要性。相比于缺失任意组件的变体, 完整模型在匹

配得分和平均匹配准确率上均取得了最优表现。这一结果有力地证明了本文提出的自适应采样、概率化建模及特征增强策略并非简单的叠加, 而是形成了有机的互补, 共同提升了算法在应对复杂非刚性形变时的鲁棒性与精确度。

2.4 对比实验

如表 2 所示, 在公开基准 GeoBit^[9] 上, 所提方法在匹配性能上优于现有最先进方法, 验证了其有效性与优越性。

表 2 GeoBit 公开基准数据集上各方法性能比较

算法模型	MS/MMA @ 3 pixels			
	Kinect1	Kinect2	Simulation	Total
DISK ^[13]	0.53/0.76	0.52/0.81	0.26/0.34	0.45/0.67
DALF ^[12]	0.54/0.82	0.62/0.90	0.42/0.69	0.55/0.82
XFeat ^[7]	0.40/0.90	0.23/0.84	0.20/0.44	0.26/0.76
LiftFeat ^[6]	0.50/ 0.91	0.30/0.80	0.18/0.29	0.31/0.72
本文	0.53/0.87	0.62/0.92	0.47/0.72	0.56/0.85

此外, 为了验证本文方法在实际工程中的应用潜力, 本文选取了四种极具挑战性的真实工况进行可视化评估, 并与最先进的方法进行直观对比, 结果如图 4 所示。由图 4 可以看出:

(1) 大型基础设施健康监测: 在塔科马大桥的卡门涡街震颤场景中 (图 4 (a)、(b)), 传统方法难以捕捉桥面剧烈的扭曲形变, 导致大量误匹配或漏匹配; 而本文方法成功建立了桥索与桥面在极端形变下的稠密对应关系。这表明该方法可为基于视觉的大型结构位移测量与振动分析提供可靠的关键点数据。

(2) 柔性材料与增强现实: 针对布料这种高频非刚性形变 (图 4 (c)、(d)), 本文方法有效克服了褶皱带来的纹理畸变, 保持了较高的匹配准确率。这一特性使其适用于虚拟试衣、柔性体追踪等 AR 应用场景。

(3) 工业结构探伤与测试: 在箱型梁交变载荷实验中 (图 4 (e)、(f)), 面对结构表面的四点弯曲形变, 本文方法能够精确追踪表面的拉伸与压缩区域, 展现了在工业零部件形变检测与材料力学测试中的应用潜力。

(4) 机器人复杂环境感知: 在非刚性多物体堆叠的复杂场景中 (图 4 (g)、(h)), 服务机器人常面临视角变化与物体形变的双重干扰。对比结果显示, 本文方法在杂乱、视点变换以及物体扭曲和重叠的背景下仍能保持鲁棒的特征关联, 有助于提升机器人在动态环境下的抓取成功率。

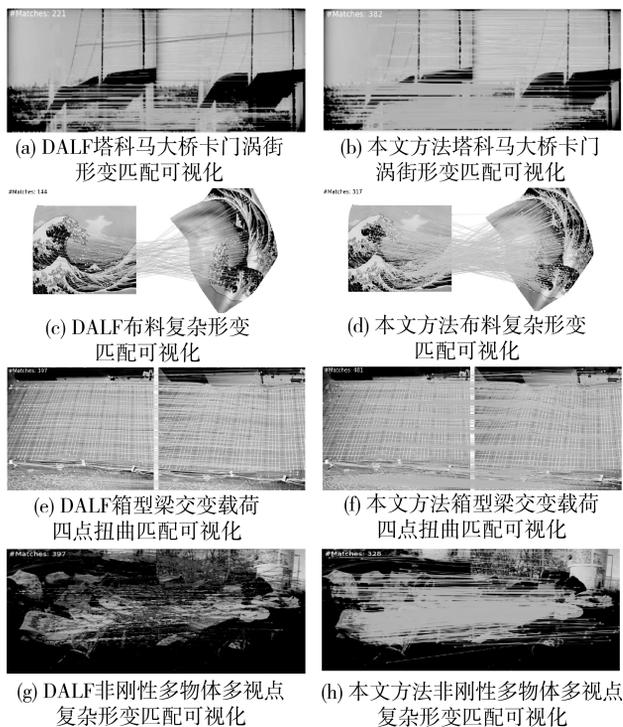


图4 实际应用场景下匹配效果可视化

3 结论

本文针对非刚性变形严重影响图像匹配性能的挑战,提出并实现了一种鲁棒的关键点匹配框架。该方法通过动态优化关键点的空间分布,并结合概率化思想对几何变换的不确定性进行建模,有效增强了特征描述符对复杂变形的适应能力。同时,通过强化对关键结构信息的鲁棒表达,缓解了由形变引起的局部特征偏移问题。在公开基准数据集上的实验结果表明,该方法在匹配得分与平均匹配准确率上均取得了具有竞争力的表现,验证了其在复杂非刚性场景下建立可靠图像对应关系的有效性,克服了传统方法在此类挑战下的性能瓶颈。

参考文献

[1] NIU Y, LIU L, HUANG F, et al. Overview of image-based 3D reconstruction technology [J]. Journal of the European Optical Society-Rapid Publications, 2024, 20 (1). DOI: 10.1051/jeos/2024018.

[2] SI S, LI Z, LIN Z, et al. 2D/3D medical image registration based on feature-point matching [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 5037609.

[3] WANG X, YU H, YU L, et al. Towards robust keypoint detection and tracking: a fusion approach with event-aligned image features [J]. IEEE Robotics and Automation Letters, 2024, 9 (9): 8059 - 8066.

[4] ZHAO W, SHI X, NI F, et al. Semi-dense sub-pixel displacement measurement for structural health monitoring: a framework of

deep learning-based detector-free feature matching [J]. Measurement, 2025, 254: 117899.

[5] DETONE D, MALISIEWICZ T, RABINOVICH A. SuperPoint: self-supervised interest point detection and description [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2018: 224 - 236.

[6] LIU Y, LAI W, ZHAO Z, et al. LiftFeat: 3D geometry-aware local feature matching [J]. arXiv preprint arXiv: 2505.03422, 2025.

[7] POTJE G, CADAR F, ARAUJO A, et al. XFeat: accelerated features for lightweight image matching [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 2682 - 2691.

[8] SIMO-SERRA E, TORRAS C, MORENO-NOGUER F. DaLI: deformation and light invariant descriptor [J]. International Journal of Computer Vision, 2015, 115 (2): 136 - 154.

[9] NASCIMENTO E R, POTJE G, MARTINS R, et al. GEOBIT: a geodesic-based binary descriptor invariant to non-rigid deformations for RGB-D images [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 10004 - 10012.

[10] POTJE G, MARTINS R, CADAR F, et al. Learning geodesic-aware local features from RGB-D images [J]. Computer Vision and Image Understanding, 2022, 219: 103409.

[11] POTJE G, MARTINS R, CHAMONE F, et al. Extracting deformation-aware local features by learning to deform [J]. Advances in Neural Information Processing Systems, 2021, 34: 10759 - 10771.

[12] POTJE G, CADAR F, ARAUJO A, et al. Enhancing deformable local features by jointly learning to detect and describe keypoints [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 1306 - 1315.

[13] TYSZKIEWICZ M, FUA P, TRULLS E. DISK: learning local features with policy gradient [J]. Advances in Neural Information Processing Systems, 2020, 33: 14254 - 14265.

[14] SCHWÖBEL P, WARBURG F R, JØRGENSEN M, et al. Probabilistic spatial transformer networks [C]//Uncertainty in Artificial Intelligence. PMLR, 2022: 1749 - 1759.

[15] MISHUK A, MISHKIN D, RADENOVIC F, et al. Working hard to know your neighbor's margins: local descriptor learning loss [J]. Advances in Neural Information Processing Systems, 2017, 30.

[16] WILSON K, SNAVELY N. Robust global translations with 1 DSFM [C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2014: 61 - 75.

(收稿日期: 2025 - 12 - 03)

作者简介:

董家麟 (2001 -), 男, 硕士研究生, 主要研究方向: 计算机视觉。

吴丽君 (1984 -), 女, 博士, 副教授, 主要研究方向: 计算机视觉、机器学习、智能信号处理等。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com