

非刚性点云配准综述

1. 引言

非刚性点云配准（Non-Rigid Point Cloud Registration, NRPCR）作为计算机视觉、机器人感知及医学图像分析中的核心问题，其本质是求解两个或多个发生形变的三维表面之间的对应关系与变换场。与寻找低维欧几里得变换 $SE(3)$ 的刚性配准不同，非刚性配准面临着求解高维位移场 $W: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 的挑战。这一过程在数学上是高度不稳定的，因为存在无穷多可能的形变解能够满足表面的对齐约束。在物理世界中，物体不仅发生刚体运动，还伴随着关节运动、软组织形变、拓扑撕裂等复杂的几何变化。此外，现实世界的数据采集往往伴随着传感器噪声、离群点、部分重叠以及由于遮挡导致的拓扑结构缺失，使得传统基于物理模型（如弹性体模型）或几何优化（如迭代最近点算法变体）的方法难以在精度、鲁棒性与计算效率之间取得平衡。

在过去的五年中，该领域经历了从手工设计特征与能量函数优化向深度学习驱动的范式转变。特别是近三年，随着隐式神经表示、扩散概率模型、以及功能映射理论的深入应用，非刚性配准的研究呈现出多模态融合、生成式建模和长序列时空一致性的新趋势。

本报告旨在对近五年发表于顶级会议及顶级期刊的非刚性点云配准文献进行详尽的综述与批判性分析。报告将深入剖析各类代表性方法的技术内核，特别是其损失函数设计的理论依据与数学机制，最终指出现有方法的局限性与潜在的研究空白。

2. 隐式神经表征与层级化形变建模

传统的非刚性配准方法通常依赖于离散的网格或体素表示，这不仅限制了形变场的分辨率，还增加了内存消耗，且难以处理拓扑结构变化的场景。近年来，隐式神经表征通过多层感知机将形变场参数化为连续函数 $f_\theta: (x, y, z) \rightarrow (\Delta x, \Delta y, \Delta z)$ ，引领了分辨率无关配准的新方向。

2.1 层级化运动分解：神经形变金字塔 NDP

(1) 核心问题与挑战

非刚性运动的高度复杂性是该任务的主要瓶颈。自然界的非刚性形变往往具有多尺度特性：宏观上表现为全局的刚性或关节运动（如人体的肢体摆动），微观上则表现为局部的精细形变（如衣物的褶皱、皮肤的拉伸）。标准的单一大容量 MLP 在拟合这种混合频率的信号时面临“频谱偏差”（Spectral Bias）问题，即网络倾向于优先学习低频分量，导致高频细节丢失；或者在强行拟合高频细节时，优化过程极易陷入局部极小值，导致整体几何结构的扭曲^[1]。

(2) 解法与假设

针对这一频域冲突问题，Neural Deformation Pyramid (NDP) 提出了基于层级化运动分解的解决方案。该方法的核心假设是：非刚性运动可以分解为一个金字塔结构，其中金字塔的高层级代表低频的全局运动，低层级代表高频的局部修正。通过显式地控制不同层级网络的频域响应，可以实现从粗糙到精细的解耦优化。

(3)技术路径

NDP 构建了一个由多级 MLP 组成的金字塔架构，每一级 MLP 接收经过正弦位置编码的三维坐标作为输入，并输出相对于上一级的运动增量。为了解决多尺度形变的频谱冲突，NDP 采用了分层频率控制机制，其中顶层网络使用极低频率的位置编码，在数学上限制了 MLP 的李普希茨常数，使其仅能拟合平滑的低频运动，从而快速收敛到全局大致对齐的状态；随后各层随着层级下降逐渐增加编码频率，逐步释放网络拟合高频细节的能力，最终的总形变场由各层增量累加而成。在推理优化过程中，NDP 通过最小化一个复合损失函数来求解形变场，该函数主要由衡量几何贴合度的倒角距离损失和防止非物理形变的刚性正则化项组成。其中，倒角距离采用单向或截断形式以增强对部分重叠的鲁棒性，而正则化项则通过惩罚形变梯度的雅可比矩阵与旋转矩阵 $SO(3)$ 的偏差，强制要求形变在局部保持刚性，这种约束被施加在金字塔的每一层，从而保证了从粗到细的每一步形变都具有物理可信度。

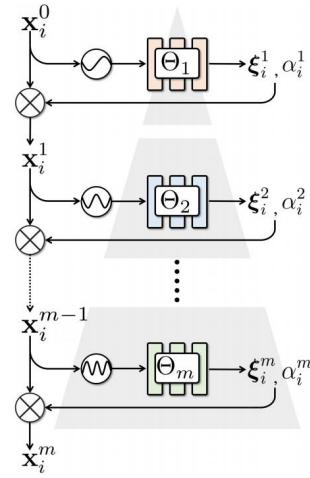


图 1 NDP 框架图

(4)总结与局限

NDP 成功地将信号处理中的频域分析引入到神经形变建模中，解决了多尺度优化的难题，但仍存在若干局限性。首先，该方法对正弦编码频率 (k_l) 等超参数极为敏感，若频率范围设定与实际形变的频谱特征不匹配，极易导致结果过度平滑或产生混叠效应。其次，作为一个基于测试时优化的方法，NDP 需要针对每一对点云进行耗时的迭代求解，尽管相比 Nerfies 等方法速度有所提升，但秒级甚至分钟级的推理延迟仍无法满足实时跟踪需求。此外，虽然 ARAP 正则化有效保持了局

部刚性，但在处理衣物撕裂等拓扑断裂场景时，强连续性假设会导致错误的“拉丝”现象，难以准确还原物理断裂边界。

2.2 各向异性混合场与辛几何约束：AniSym-Net

(1)核心问题与挑战

在处理存在大量遮挡的“近等距形变”，如人体运动时，现有的各向同性形变场方法（如基于 LBO 的方法）往往难以保持拓扑一致性。标准的欧几里得空间优化容易导致体积塌陷或自相交，特别是在观测数据稀缺的区域，网络倾向于通过非物理的捷径（直接穿过物体内部）来最小化数据项误差^[2]。

(2)解法与假设

AniSym-Net 引入了一种基于辛流形（Symplectic Manifold）的各向异性混合形状-运动场框架。其核心假设是：物理上合理的形变应当保持某种几何结构（如相空间中的体积或辛形式 ω ），这种约束可以通过辛几何中的全纯约束（Holonomic Constraints）来严格建模，从而在数学层面杜绝体积塌陷。

(3)技术路径

AniSym-Net 摒弃了单一的位移预测模式，转而构建了一个包含“子逆扩散场”和“子拉格朗日场”的混合系统，前者利用逆扩散过程从噪声观测中恢复静态几何分布以应对遮挡补全，后者则在拉格朗日视角下跟踪粒子轨迹以编码动量守恒。这两个场通过一个基于辛几何的 Transformer 模块进行融合，利用运动条件交叉注意力机制，依据运动场的速度边界信息来校准形状场的几何特征，确保形状重建不违反物理限制。为了训练这一复杂的混合场，AniSym-Net 设计了一个包含多重约束的协调损失函数：除了基础的几何重建误差外，最核心的是辛一致性损失（Symplectic Consistency Loss），它通过惩罚雅可比矩阵偏离辛群 $Sp(2n)$ 的程度，强制变换映射满足辛条件 $\Phi^*\omega = \omega$ ，从根本上保证了变换的保体积特性；同时引入协调损失，最小化由形状场和运动场分别导出的对应关系之间的差异，作为自监督的一致性约束，强迫静态形状理解与动态运动估计达成共识，从而在严重遮挡区域也能产生拓扑鲁棒的预测。

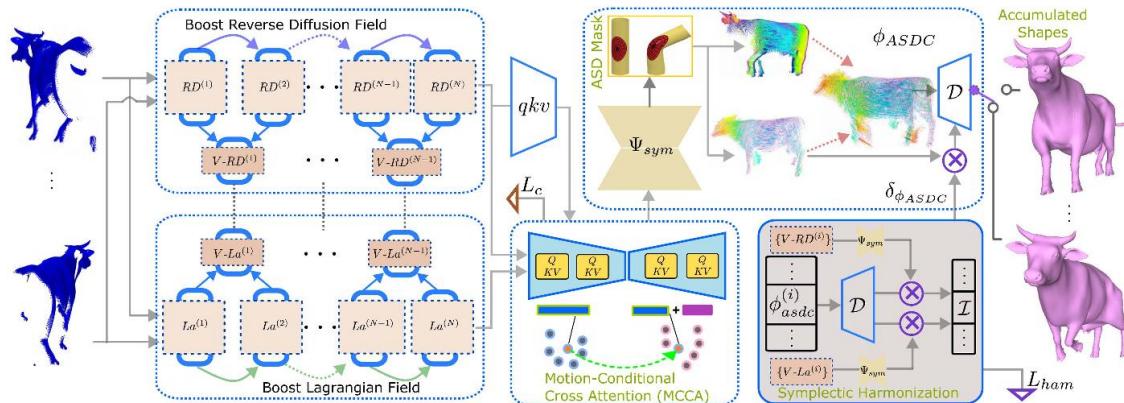


图 2 AniSym-Net 框架图

(4)总结与局限

AniSym-Net 将深奥的辛几何理论成功应用于点云配准，极大增强了模型的物理可解释性和拓扑保持能力，但这同时也带来了高昂的代价。其计算复杂度极高，在流形上求解辛流涉及复杂的积分器和高维张量运算，导致训练和推理的开销远超常规欧氏空间方法。同时，辛神经网络的架构设计与梯度反向传播涉及复杂的数学推导，复现与工程落地难度较大。此外，该方法对“近等距形变”的强假设限制了其通用性，在面对流体、烟雾或涉及手术切除等非等距、拓扑改变的场景时，其基于辛几何的体积守恒约束可能不再适用。

3.序列化配准与时空一致性

单帧对单帧的配准容易受到噪声和局部极小值的干扰，而在处理连续的点云序列（如 4D 扫描）时，利用时间上下文来消除歧义并防止误差累积成为了研究焦点。

3.1 高效前馈网络与变形图预测：ERNet

(1)核心问题与挑战

长序列非刚性配准面临两大核心难题：一是单帧配准目标函数的非凸性导致极易陷入局部极小值（特别是在大位移下）；二是长序列跟踪中的累积误差会导致跟踪漂移（Drifting），最终导致模型失效。现有的隐式方法（如 NDP）虽然精度高但速度慢；而传统的逐帧优化方法则缺乏全局的时间一致性^[3]。

(2)解法与假设

ERNet (Efficient Registration Network) 提出了一种高效的前馈网络，直接预测变形图（Deformation Graph）序列，而非密集的逐点流场。其核心假设是局部刚性（Local Rigidity）：即非刚性物体的形变可以近似为由一组稀疏节点驱动的局部刚性变换的组合。

(3)技术路径

ERNet 采用了一种“两阶段预测 + 显式变形图”的高效策略，首先通过源节点与当前帧点云的全局特征匹配来估计粗糙的节点位置，利用全局上下文解决大位移初始化问题；随后采用滑动窗口机制，利用时空 Transformer 在时间维度上聚合邻域帧信息以平滑轨迹，在空间维度上聚合邻域节点以保持结构。为了降低优化难度并保证旋转正交性，ERNet 并不直接回归旋转矩阵，而是预测节点对应关系后利用可微 SVD 解析求解每个节点的最优刚性变换。其训练过程受多项损失函数共同监督：流场损失直接约束预测点位置的准确性；图一致性损失确保变形图节点紧贴几何表面；而最为关键的 ARAP 正则化项则惩罚相邻节点间的距离变化，强制变形图骨架在运动过程中保持局部刚性，从而有效防止了非物理拉伸和长序列跟踪中的误差累积。

(4)总结与局限

ERNet 凭借高效的前馈架构和稀疏变形图表征，在保证精度的同时实现了极高的推理速度，是工程实用性极强的方法。然而，其对变形图节点采样的依赖是潜在软肋，通常使用的最远点采样（FPS）策略缺乏语义感知，若关键运动部件（如手指尖端）未被采样为节点，这些区域的配准精度将大幅下降。此外，基于变形图的线性混合蒙皮（LBS）机制虽然适用于有骨骼结构的生物体，但在处理缺乏骨骼支撑的柔性形变（如裙摆飘动、流体）时，可能会产生过于僵硬的伪影或“糖纸”效应，限制了其在非生物体配准中的表现。

3.2 4D 潜变量集合扩散模型：Motion2VecSets

(1)核心问题与挑战

从稀疏、噪声或部分缺失的点云中重建 4D 动态表面是一个高度不确定的多模态问题（*Ill-posed multimodal problem*）。确定性的前馈网络（如 ERNet）倾向于输出所有可能解的平均值，导致结果模糊或过度平滑。此外，使用单一的全局潜变量（*Global Latent Code*）难以捕捉局部的复杂运动细节，且泛化到未见过的物体类别能力差^[4]。

(2)解法与假设

Motion2VecSets 将 4D 重建重构为一个条件生成问题。它假设形状和运动的分布可以通过 扩散概率模型（Diffusion Probabilistic Model, DPM）来学习，从而能够从噪声中采样出高频细节。

(3)技术路径

Motion2VecSets 摒弃了单一全局编码，转而使用一组局部潜变量集合来参数化 4D 动力学，每个潜变量负责物体的一个局部区域，以此提升对细节的捕捉能力。为了保证时间一致性，该方法并非逐帧去噪，而是引入了“同步形变向量集扩散”机制，将整个序列的形变潜变量作为一个整体进行联合去噪。在网络架构上，采用了交错时空注意力机制，交替聚合空间内的局部信息和时间轴上的演化信息，有效降低了计算复杂度。作为扩散模型，其训练目标遵循标准的去噪得分匹配（Denoising Score Matching）范式，即最小化预测噪声与添加噪声之间的均方误差；同时，为了规范化潜空间分布以便于推理采样，还引入了 KL 散度正则化项，确保潜变量分布接近标准正态分布。这种生成式设计使得模型能够在输入稀疏或缺失时，基于学习到的先验“脑补”出合理的几何细节。

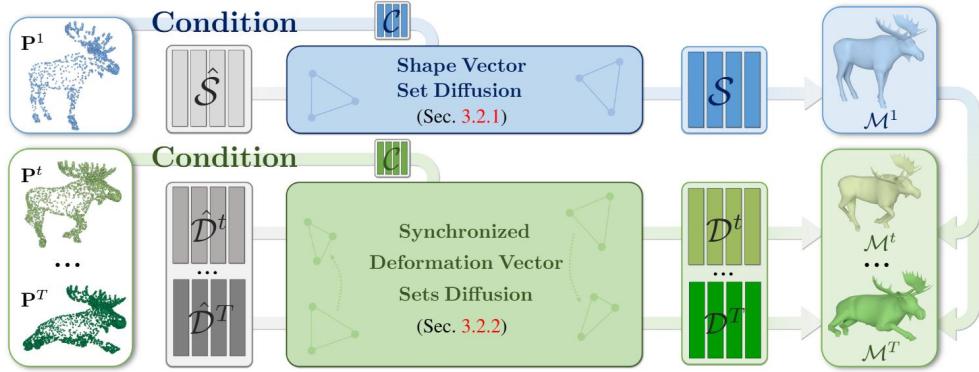


图 3 Motion2VecSets 框架图

(4) 总结与局限

Motion2VecSets 展示了生成式模型在非刚性配准中的强大潜力，尤其是在处理稀疏输入和重建精细几何细节方面表现卓越。然而，作为一种生成式方法，它面临着“幻觉”风险，即生成的逼真细节可能并非源自真实观测，而是基于训练数据分布的猜测，这在医疗等严谨场景中可能构成安全隐患。此外，扩散模型的推理本质上是一个迭代去噪过程（通常需要 50-100 步），导致其推理延迟相比 ERNet 等单步前馈网络高出数个数量级，这极大地限制了其在实时交互式应用中的部署可能性。

4. 谱方法与功能映射

除了在空间域（Spatial Domain）直接求解形变，另一类方法尝试在谱域（Spectral Domain）解决配准问题。功能映射（Functional Maps）理论将点对点的对应关系转化为函数空间之间的线性映射，这使得处理多形状配准和拓扑噪声成为可能。

4.1 多路配准与映射同步：SyNoRiM

(1) 核心问题与挑战

成对（Pairwise）配准方法在处理多个形状集合（如一个很长的视频序列或多视角的扫描）时，往往无法保证全局的一致性。即 $A \rightarrow B \rightarrow C$ 的变换组合，往往不等于 $A \rightarrow C$ 的直接变换。这种循环一致性的缺失会导致累积误差^[5]。

(2) 解法与假设

SyNoRiM 提出了基于功能映射同步的多路非刚性配准框架。其核心假设是：形状间的对应关系可以由一组基函数上的线性变换矩阵（功能映射 C ）紧凑地表示，且这些矩阵在全局范围内应当满足循环一致性约束 ($C_{ij}C_{jk} \approx C_{ik}$)。

(3) 技术路径

SyNoRiM 首先通过神经网络（如 KPConv）直接从点云数据中学习一组基函数，这组基函数允许是非正交的，相比传统的拉普拉斯算子特征函数对部分缺失和遮挡具有更强的鲁棒性。在推理阶段，该方法不仅计算成对的功能映射，还求解一个全局同步优化问题，试图在所有形状对之间找到一组满足循环一致性的绝对映射。其损失函数设计紧扣这一目标：除了包含最小化特征投影误差的成对功能映射损失外，还显式引入了循环一致性损失，惩罚三元组闭环误差（即要求 $A \rightarrow B \rightarrow A$ 回归单位阵），并辅以正则化项以鼓励映射与 LBO 的可交换性，从而在全局范围内协调所有形状的配准关系，消除长序列漂移。

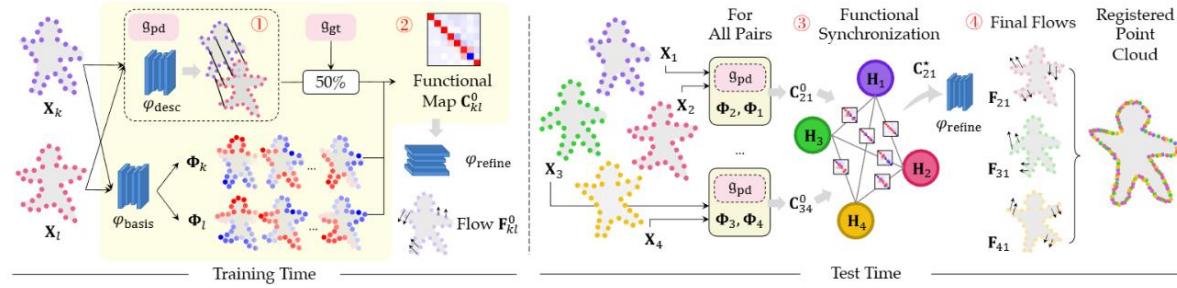


图 4 SyNoRiM 框架图

(4) 总结与局限

SyNoRiM 通过功能映射同步有效解决了长序列和多视角的全局一致性问题，消除了漂移现象，但仍受限于“转换瓶颈”。即从低秩的功能映射矩阵恢复到高分辨率的点对点对应关系时，由于功能映射的低通滤波特性，往往会丢失高频几何细节，导致精细结构模糊。此外，全局同步步骤高度依赖于初始成对映射图的质量，如果初始图中包含大量错误的离群值，全局同步优化可能会发生坍塌，导致所有形状收缩到一点或产生退化解，这限制了其在极端噪声环境下的鲁棒性。

4.2 混合特征引导配准：NFR

(1) 核心问题与挑战

纯学习型的特征嵌入缺乏直观的几何意义，且在未见过的形状上泛化能力差；而纯几何优化的方法在面对大形变时容易陷入局部最优。NFR (Neural Feature-Guided Registration) 试图结合二者的优点^[6]。

(2) 解法与假设

NFR 提出了一种 师生蒸馏 (Teacher-Student Distillation) 框架。它假设：可以利用在高质量网格上预训练的深度功能映射 (DFM) 网络作为“教师”，来指导点云特征提取器 (学生) 的学习。

(3) 技术路径

NFR 采用师生蒸馏策略，利用在高质量全网格数据上预训练的深度功能映射 (DFM) 网络作为教师，指导仅以稀疏点云为输入的学生网络 (DGCNN) 学习具

有全局几何感知的“伪谱特征”。在配准过程中，NFR 实施混合优化方案：第一阶段利用学习到的特征对应进行全局粗配准，解决大形变下的初始化问题；第二阶段则切换到环境空间，利用 ARAP 能量项和加权倒角距离进行精细化几何优化。其训练损失主要由蒸馏损失构成，强制学生网络的特征分布逼近教师网络的谱特征，同时辅以无监督损失，利用教师网格上的测地线距离一致性作为监督信号，从而使模型兼具学习型方法的泛化能力和几何方法的精度。

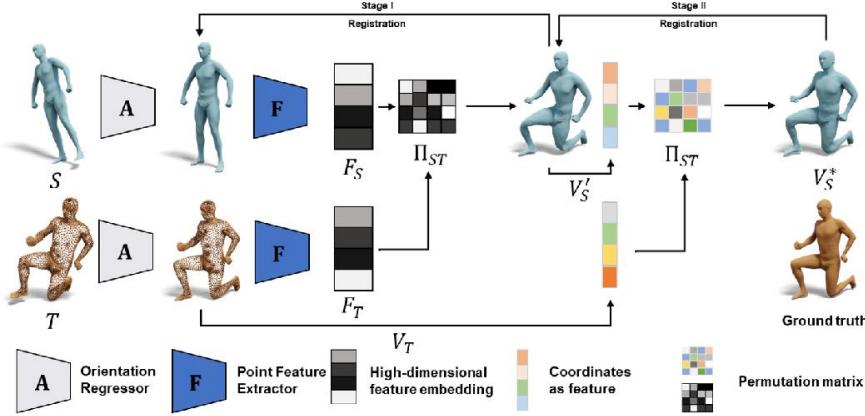


图 5 NFR 框架图

(4) 总结与局限

NFR 虽然巧妙结合了谱特征的全局感知能力和几何优化的精度，但其对高质量模板网格的依赖构成了主要瓶颈。该方法必须依赖在完整网格上预训练的 DFM 教师网络来指导学习，如果应用场景中仅有稀疏点云而缺乏可用的网格拓扑先验（例如未知的工业零件或新发现的生物组织），这种“师生蒸馏”策略便无法实施。此外，分阶段的“先特征后几何”配准策略切断了端到端的梯度回传，使得后续几何优化的误差无法反馈用于修正特征提取网络，限制了模型在极端情形下的自适应能力。

5. 对应关系估计与扩散匹配

配准的核心往往在于寻找可靠的对应关系。近年来，直接对匹配矩阵进行生成式建模成为新趋势。

5.1 扩散匹配模型：Diff-Reg

(1) 核心问题与挑战

现有的对应关系估计方法（如基于 Transformer 的方法）多依赖于单次前馈预测（Single-pass Prediction）。在面对复杂形变、尺度不一致和对称性歧义（如左手配右手）时，单次预测很难捕捉到多模态的匹配分布，往往输出模糊的平均值。此外，理想的匹配矩阵应当是双随机矩阵（Doubly Stochastic Matrix），但强制这一约束在传统网络中较难实现^[7]。

(2) 解法与假设

Diff-Reg 提出将对应关系估计视为双随机矩阵空间内的去噪扩散过程。其核心假设是：真实的匹配矩阵可以通过从高斯噪声矩阵开始，逐步去噪并投影到双随机流形上来获得。

(3) 技术路径

Diff-Reg 将对应关系估计建模为 $N \times M$ 匹配矩阵空间中的扩散生成过程。在前向过程中，向真实匹配矩阵逐步添加高斯噪声；在反向过程中，训练神经网络逐步去除噪声以恢复匹配矩阵。该方法的核心创新在于引入了 Sinkhorn 投影机制：在反向去噪的每一步，网络输出的矩阵被显式地应用 Sinkhorn 算法投影回双随机矩阵流形（即行和列之和均为 1），这相当于在流形上进行受约束的随机游走，保证了中间结果始终满足物理上的双射约束。为了提高效率，特征提取仅进行一次，扩散过程在紧凑的矩阵空间迭代。其损失函数包含标准的 MSE 去噪损失（或变分下界），并引入 Focal Loss 来处理匹配矩阵极度稀疏（绝大多数元素为 0）导致的类别不平衡问题，防止网络倾向于预测全零矩阵。

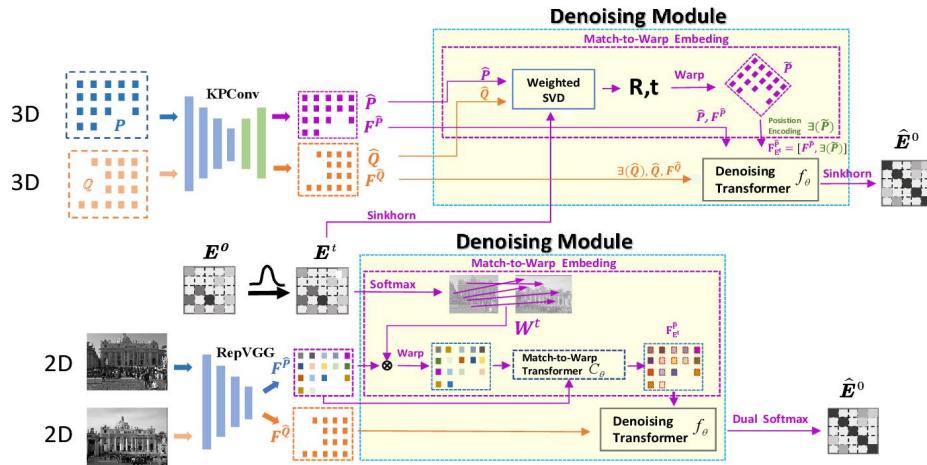


图 6 Diff-Reg 框架图

(4) 总结与局限

Diff-Reg 虽然在解决大形变和多模态匹配歧义上表现出色，但其计算成本极为高昂。由于在扩散生成的每一步（通常需 20-50 步）都需要嵌入迭代式的 Sinkhorn 算法进行投影，这构成了“迭代中的迭代”，使得推理过程涉及数千次矩阵运算，难以处理大规模点云。此外，其核心的双随机矩阵假设隐含了严格的双射前提（即源点云与目标点云一一对应），而在现实中普遍存在的部分重叠（Partial Overlap）场景中，这一假设失效，强制的 Sinkhorn 投影可能会错误地将离群点强制匹配，反而引入噪声干扰配准结果。

5.2 3D 位置编码与部分匹配：Lepard

(1) 核心问题与挑战

KPConv 等主流点云特征提取器具有平移不变性。虽然这对于刚性物体识别有利，但在非刚性物体中，几何结构往往具有重复性（如左手和右手、重复的纹理），导致特征空间中的歧义。仅靠局部几何特征无法区分这些部位^[8]。

(2) 解法与假设

Lepard 提出显式地利用 3D 位置知识来消除歧义。它假设：相对位置关系是区分重复几何结构的关键线索。

(3) 技术路径

Lepard 引入了解耦表征，将点云特征空间与 3D 位置空间分离，并将 NLP 中的旋转位置编码（Rotary Positional Encoding, RoPE）推广到 3D 空间，使特征点积注意力能够显式感知相对距离和方向信息，从而打破对称性歧义。此外，Lepard 采用了一种迭代式的“重定位”（Repositioning）机制，利用当前的粗糙变换调整源点云位置后重新计算位置编码，实现特征匹配的逐步细化。其训练损失由两部分组成：一是基于 Focal Loss 的分类损失，用于监督匹配矩阵的置信度并解决正负样本不平衡；二是几何扭曲损失（Warping Loss），直接优化匹配点对在经过几何变换后的欧氏距离，确保匹配结果在几何上的精确对齐。

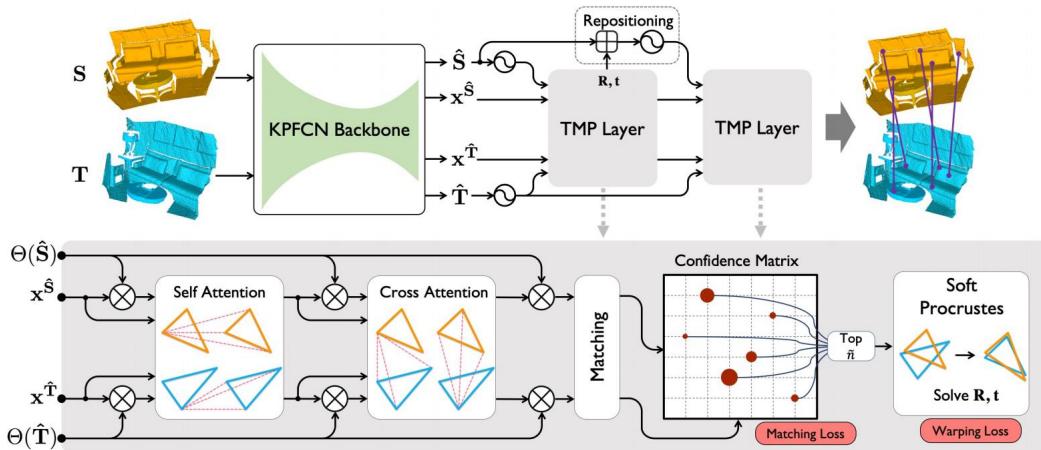


图 7 Lepard 框架图

(4) 总结与局限

Lepard 通过显式引入 3D 旋转位置编码有效解决了非刚性物体中的几何对称性歧义，显著提升了匹配召回率。然而，其基于 Transformer 的全局注意力机制导致内存消耗随点数呈二次方增长 ($O(N^2)$)，这限制了其直接应用于高分辨率点云，通常需要进行降采样处理。同时，显式的位置编码可能导致模型对训练数据的空间分布产生过拟合（例如训练数据物体多居中），当测试数据的空间位置发生剧烈变化时，位置编码可能会产生误导，影响配准性能。

6. 无监督与聚类范式

鉴于真实世界非刚性形变真值获取的极度困难，通常需要昂贵的动作捕捉系统，无监督学习成为了重要的研究方向。

6.1 基于聚类的对应无关配准：CluReg

(1)核心问题与挑战

大多数现有方法（如 Lepard, Diff-Reg）采用“先匹配后对齐”的两步策略。这种策略严重依赖于初始匹配的质量，一旦匹配中存在大量离群点，后续的 SVD 或优化就会彻底失败。CluReg 试图回答：能否跳过显式的点对点匹配，直接求解形变^[9]？

(2)解法与假设

CluReg 将配准重构为无监督聚类分析问题。它假设：源点云是“聚类中心”，目标点云是“样本数据”。配准的过程，就是移动聚类中心（源点云）使其覆盖样本数据（目标点云）分布的过程。

(3)技术路径

CluReg 采用期望最大化（EM）算法框架，在 E-步计算目标点隶属于当前变形后源点的软分配概率，在 M-步更新源点位置（形变场）以最大化目标数据的对数似然。为了克服传统高斯核对大形变过度平滑的问题，CluReg 创新性地引入了 L1 诱导的拉普拉斯核作为正则化项，利用拉普拉斯分布的长尾特性来容忍关节处等剧烈的运动跳变。为了处理稠密点云的计算瓶颈，该方法利用 Nyström 方法对 Gram 矩阵进行低秩逼近，将复杂度从平方级降低到线性级。其目标函数由负对数似然函数（数据项）和拉普拉斯核正则项（平滑项）组成，整个过程无需任何对应关系标注，完全以无监督方式驱动源点云覆盖目标分布。

(4)总结与局限

CluReg 作为一种纯几何优化的无监督方法，摆脱了对标注数据的依赖，且计算效率极高，但其本质上缺乏语义感知能力。由于仅依赖欧氏距离和局部几何相似性进行“盲”配准，当目标形状发生拓扑结构改变（如手掌张开变握拳）或存在显著的语义部分缺失时，聚类算法容易将源点云错误地“拉扯”到最近的几何区域，导致非物理的扭曲。此外，该方法对正则化参数 λ 极其敏感，不同的形变程度往往需要手动微调该参数，缺乏深度学习方法在不同场景下的自适应能力。

7.总结与未来展望

7.1 方法论比较与趋势综合

下表总结了近五年核心方法的关键特性：

表 1 近五年核心方法的关键特性

方法	类型	核心机制	优势	主要局限
NDP	隐式/优化	正弦 MLP 金字塔	细节保持好, 无监督, 避免局部极小值	推理慢 (需逐个优化), 受限于正弦编码的频谱偏差
SyNoRIM	谱方法	学习型基函数/同步	多体一致性, 对拓扑噪声和遮挡鲁棒	高频几何细节易在谱变换中丢失, 全局同步优化容易坍塌
ERNet	图/前馈	变形图 + SVD + 混合蒙皮	推理极快 ($>4x$), 时序稳定性好, 显式保持局部刚性	依赖节点采样 (FPS) 质量, 若关键部位采样不足会导致局部配准失效
Motion2VecSets	生成式	潜变量集扩散	处理模糊、稀疏及部分重叠输入能力强, 捕捉非线性运动	存在生成“幻觉”风险, 推理延迟高 (需迭代去噪)
Lepard	特征学习	3D 旋转位置编码	显式消除对称性歧义, 提升部分重叠下的匹配召回率	可能对绝对空间位置过拟合, Transformer 带来的 $O(N^2)$ 复杂度限制了规模
Diff-Reg	扩散匹配	双随机矩阵扩散 + Sinkhorn	对大形变和离群点鲁棒性强, 能探索组合匹配空间	计算成本极其昂贵 (Sinkhorn 投影嵌套在扩散步中), 难以处理大规模点云
CluReg	无监督	拉普拉斯核聚类 + EM 算法	无需对应关系标注, 线性计算复杂度, 对大形变鲁棒	缺乏语义感知能力, 对正则化参数敏感, 难以处理拓扑结构剧烈变化

7.2 存在的空白与未来方向

尽管取得了显著进展, 非刚性点云配准领域仍面临严峻挑战:

(1)泛化性的鸿沟: 目前的监督方法 (如 ERNet, Lepard) 大多在合成数据集 (如 DeformingThings4D) 上训练。当应用到真实世界的扫描数据 (如 Kinect 深度图) 时, 由于传感器噪声模式和几何特性的差异 (Domain Gap), 性能往往大幅下降。未来的研究需要关注域适应 (Domain Adaptation) 和自监督学习技术。

(2)物理约束的缺失: 目前的形变场大多仅受几何平滑项约束 (ARAP, Laplacian), 缺乏真实的物理属性 (如弹性模量、质量守恒、碰撞检测)。

(3)AniSym-Net 引入辛几何是一个好的开始, 但未来的方向应是将物理引擎的可微模拟直接嵌入到配准网络的损失函数中, 从而彻底杜绝自相交。

(4)大模型与零样本能力: 随着 NLP 和 2D 视觉进入大模型时代, 3D 几何领域也呼唤“几何基础模型”。利用海量 3D 数据预训练一个通用的形变先验模型, 使其能够在零样本条件下适应未见过的物体类别, 将是颠覆性的研究方向。

(5)实时性与精度的二律背反: 目前的方法要么快但不准 (前馈网络), 要么准但极慢 (扩散模型、测试时优化)。未来的突破点可能在于“预测-优化”混合架构, 即利用轻量级网络快速初始化, 再利用高效的 GPU 并行优化器 (如可微渲染器) 进行微调。

综上所述, 2021-2025 年间的非刚性点云配准研究展示了从手工几何特征向深度几何学习的深刻演变。尽管在处理大形变、部分重叠和时序一致性方面取得了长足进步, 但在物理真实性、计算效率和跨域泛化性方面仍面临严峻挑战。

参考文献

- [1] Li, Yang and Tatsuya Harada. “Non-rigid Point Cloud Registration with Neural Deformation Pyramid.” ArXiv abs/2205.12796 (2022): n. pag.
- [2] Wang, Jinyang et al. “Non-Rigid Point Cloud Registration via Anisotropic Hybrid Field Harmonization.” IEEE Transactions on Pattern Analysis and Machine Intelligence 47 (2025): 7898-7915.
- [3] He, Guangzhao et al. “ERNet: Efficient Non-Rigid Registration Network for Point Sequences.” ArXiv abs/2510.15800 (2025): n. pag.
- [4] Cao, Wei et al. “Motion2VecSets: 4D Latent Vector Set Diffusion for Non-Rigid Shape Reconstruction and Tracking.” 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024): 20496-20506.
- [5] Huang, Jiahui et al. “Multiway Non-Rigid Point Cloud Registration via Learned Functional Map Synchronization.” IEEE Transactions on Pattern Analysis and Machine Intelligence 45 (2021): 2038-2053.
- [6] Jiang, Puhua et al. “NFR: Neural Feature-Guided Non-Rigid Shape Registration.” ArXiv abs/2505.22445 (2025): n. pag.
- [7] Wu, Qianliang et al. “Diff-Reg: Diffusion Model in Doubly Stochastic Matrix Space for Registration Problem.” European Conference on Computer Vision (2024).
- [8] Li, Yang and Tatsuya Harada. “Lepard: Learning partial point cloud matching in rigid and deformable scenes.” 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021): 5544-5554.
- [9] Zhao, Mingyang et al. “Correspondence-Free Non-Rigid Point Set Registration Using Unsupervised Clustering Analysis.” 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024): 21199-21208.