RINO: Rotation-Invariant Non-Rigid Correspondences

Maolin Gao[1,3] Shao Jie Hu-Chen[1,3] Congyue Deng[2,4]
Riccardo Marin[1,3] Leonidas Guibas[2] Daniel Cremers[1,3]
[1] TUM [2] Stanford University [3] MCML [4] MIT

Figure 1. We propose **RINO**, an unsupervised framework unifying rigid and non-rigid shape matching. Our end-to-end point feature extractor, the *RINONet*, is invariant to extrinsic rotations and robust to intrinsic symmetries, without relying on handcrafted descriptors (shape orientations indicated by RGB frames). It handles challenging non-rigid matching cases with arbitrary poses, non-isometry, partiality, non-manifoldness, and noise. From left to right we show pairs of humanoid, quadruped, airplane, partial and non-manifold shapes.

Abstract

Dense 3D shape correspondence remains a central challenge in computer vision and graphics as many deep learning approaches still rely on intermediate geometric features or handcrafted descriptors, limiting their effectiveness under non-isometric deformations, partial data, and non-manifold inputs. To overcome these issues, we introduce RINO, an unsupervised, rotation-invariant dense correspondence framework that effectively unifies rigid and non-rigid shape matching. The core of our method is the novel RINONet, a feature extractor that integrates vector-based SO(3)-invariant learning with orientation-aware complex functional maps to extract robust features directly from raw geometry. This allows for a fully end-to-end, data-driven approach that bypasses the need for shape pre-alignment or handcrafted features. Extensive experiments show unprecedented performance of RINO across challenging non-rigid matching tasks, including arbitrary poses, non-isometry, partiality, non-manifoldness, and noise.

1. Introduction

Establishing correspondence between 3D shapes is a foundational and challenging problem in computer vision and graphics. While deep learning has enhanced classical shape matching by replacing handcrafted descriptors with learned features, most approaches still depend on intermediate geometric representations rather than operating directly on raw shape data [3, 11, 22, 30, 34, 37]. This reliance on human-designed features, which are often specialized (e.g., for isometric shapes), fundamentally limits robustness and generalization across diverse, challenging conditions like non-isometric deformations, partiality, and non-manifold structures. Hence, a compelling alternative is to start learning directly from the raw data — namely, using the vertex coordinates as input. Such an approach appears appealing for several reasons: it does not require preprocessing, it does not rely on theoretical assumptions, and it has been the most successful approach in other domains (e.g., CNNs generally take RGB pixel values as input directly, LLMs start from words or sub-words tokens). In accordance with Sutton's bitter lesson [58], we believe that the path toward robust shape correspondence lies in scalable end-to-end learning, with little or no human-encoded sophistication.

However, directly processing raw 3D geometry introduces the challenge of *shape–pose entanglement*, which has traditionally led to splitting the problem into two separate subproblems: rigid matching, estimating global rotations and translations; and non-rigid matching, assuming pose alignment to infer dense correspondences. This separation, however, is inherently ill-posed, as global transformations and local deformations are tightly coupled. The distinction becomes especially ambiguous when canonical poses are hard to define, for example, between a running and a sitting human, or in cases of partial shape matching.

To eliminate the need for rigid pose alignment, fully intrinsic shape matching methods [11, 22] leverage handcrafted intrinsic features (e.g., wks [6]) and intrinsic architectures (e.g., DiffusionNet [54]). Yet, reliance solely on

intrinsic information fundamentally limits their scope: it cannot theoretically distinguish shape symmetries, and it yields severely degraded and incompatible features for non-smooth or irregular shapes [26, 27, 35, 67, 68].

In this work, we propose a paradigm that goes beyond the need of manual pose alignment and does not suffer from any of the previous limitations. Our insight stems from three properties which we believe coexist in the ideal shape descriptor: first, it should be robust to noise and artifacts; second, it incorporates the geometric prior provided by the underlying surface; finally, it characterizes shape vertices regardless of the extrinsic orientation of the shape. To our knowledge, no available methods respect all these points. To this end, our idea is to develop a rotation-invariant architecture, which factors out the global shape orientation from the articulations' non-rigid deformation. In particular, we revise the components of the popular DiffusionNet [54] to be SO(3)-equivariant end-to-end, using a vector neuron representation. Such modifications are well-motivated by theoretical derivations that we report in the appendix. Our network can learn SO(3)-invariant, geometric features directly from raw 3D geometry. The feature gradients are further integrated with the orientation-aware Complex Functional Maps (CFMaps) [21, 22] to map tangent bundles between shapes. The entire system is then trained using a novel unsupervised loss that unifies Functional Maps (FMaps), CFMaps, and pointwise maps, creating a unified framework that leverages their complementary strengths for achieving high-quality, robust features and correspondences (Fig. 1-3). To summarize, our key contributions are:

- We introduce the first unsupervised rotation-invariant dense correspondence method, **RINO**, that unifies rigid and non-rigid matching.
- Our novel feature extractor, **RINONet**, achieves direct, end-to-end, and robust feature extraction from raw geometry, and discovers complex, scalable features without the need for handcrafted descriptors or shape pre-alignment.
- We integrate SO(3)-invariant learning with orientation-aware complex FMaps on tangent vector fields, yielding a coherent network architecture and unsupervised loss.
- Extensive experiments demonstrate superior performance across challenging non-rigid matching scenarios, including arbitrary poses, non-isometric deformations, partiality, non-manifold structures, and noise perturbations.

2. Related Works

Non-rigid shape matching aims to find correspondences between points on two shapes. We focus on the most relevant works below and refer to [51, 61] for an extensive review.

Learning Non-Rigid Correspondences. The Deep Functional Maps framework [34, 43] enabled deep learning in non-rigid matching. It uses a learnable feature extractor to refine initial descriptors (e.g., shot [52]) before computing a functional map. This framework has since been extended to zero-shot [2, 4], weakly supervised [53] and unsupervised settings [10, 30], incorporating more advanced objectives [20, 33, 57]. While some explored learning bases [28, 32, 39], the theoretical strength of the Laplace-Beltrami Operator (LBO) bases remains preferred. As a feature extractor, initial works used ResNet with shot or PointNet++/KPConv with xyz [20, 53], but lately DiffusionNet [54] has emerged as the standard for learning robust descriptors [3, 5, 10, 22, 57] among other networks [36, 65]. Notably, the focus of these works has been mainly on designing losses. We instead propose a novel feature extractor that learns SO(3)-invariant features directly from raw 3D geometry.

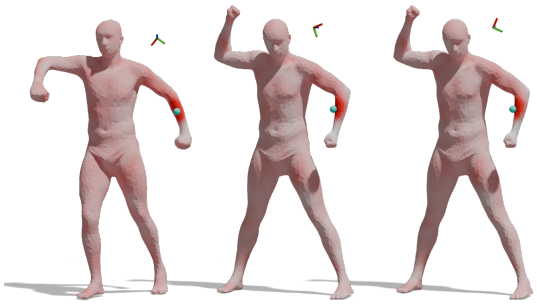

Figure 2. **Our learned SO(3)-invariant features.** We visualize the Euclidean distance to the blue surface point in the learned feature space (darker red means higher similarity). Semantic correspondences have a similar feature similarity pattern (left & mid.). The learned feature is invariant to rotations (mid. & right). Shape orientations are depicted by RGB frames.

Shape Descriptors in 3D Matching. Historically, 3D shape analysis relied on handcrafted descriptors like hks [56], wks [6], and shot [52] to establish canonical feature spaces, which modern deep learning models now use as refined inputs. However, shot is highly sensitive to discretization and mesh connectivity [5], while wks is limited by an isometric assumption (Sec. 6) that fails under significant non-rigid deformation and cannot theoretically disambiguate intrinsic symmetry. This has led to an unprincipled reliance on pre-aligned xyz coordinates in challenging datasets [7, 11, 31, 37, 45]. While recent foundational 2D features (e.g., DINOv3) offer high expressivity [15, 23, 66], they lack geometric understanding of shape orientation, require upright alignment for multi-view aggregation, and struggle to resolve rotational equivariance from data alone. To address these gaps, we propose a principled SO(3)-invariant feature extractor that directly processes raw xyz coordinates, unifying the input feature choice and enabling genuine data-driven learning from 3D shapes.

Intrinsic Shape Symmetry. Current approaches either use non-rigorous orientation-aware features [27, 47] or rely on impractical extrinsic pre-alignment [45, 53] to resolve

symmetries. While DiffusionNet combined with intrinsic descriptors attempts to learn anisotropic diffusion, it assumes consistently oriented normals, which rarely applies in real data. This leads to a lack of performance guarantees, especially for partial shapes. The use of SO(3)-*variant* DiffusionNet with xyz further hinders symmetry disambiguation due to poor handling of shape rotations (cf. Sec. 6.3). We build upon CFMaps [21], which theoretically encodes only orientation-preserving maps, and combine it with our learned SO(3)-invariant features from arbitrarily oriented raw geometry. This eliminates intrinsic symmetry flips without requiring extrinsic pre-alignment.

Equivariant Learning in 3D Understanding. Equivariant neural networks (ENN) inject group symmetry into the learning process [9, 16, 18, 46]. Architectures like Equivariant Point Network [14] offer approximate SE(3)-equivariance, while Tensor Field Networks [59] and e3nn [29] achieve strict continuous SO(3)-equivariance using steerable convolutions, often at a high computational cost. While successful in rigid tasks [44, 69], applying SO(3)-equivariant networks to fine-grained, non-rigid correspondence remains a less-explored challenge.

3. Background

Functional Maps & Complex Functional Maps. The functional maps framework [43] encodes shape matching as a compact function-to-function mapping $\mathbf{C}$ between scalar-valued features $\mathbf{F}_\mathcal{X}$ and $\mathbf{F}_\mathcal{Y}$ on two shapes $\mathcal{X}$ and $\mathcal{Y}$. Using the LBO eigenfunctions $\mathbf{\Phi}_.$ as a basis, the features are represented by coefficients $\mathbf{\Phi}_.^\dagger \mathbf{F}_.$. The functional map $\mathbf{C}$ is recovered by minimizing an energy that enforces feature preservation and structural regularity:

$$\mathcal{E}_{\text{data}}(\mathbf{C}) = \|\mathbf{C}\mathbf{\Phi}_\mathcal{X}^\dagger \mathbf{F}_\mathcal{X} - \mathbf{\Phi}_\mathcal{Y}^\dagger \mathbf{F}_\mathcal{Y}\|_F^2, \tag{1}$$

$$\mathbf{C}_{\mathcal{XY}} = \arg\min_{\mathbf{C}} \mathcal{E}_{\text{data}}(\mathbf{C}) + \lambda \mathcal{E}_{\text{reg}}(\mathbf{C}) \tag{2}$$

where $\mathcal{E}_{\text{data}}$ promotes feature preservation, and $\mathcal{E}_{\text{reg}}$ encourages structural properties, implicitly promoting isometry [22, 48]. Subscripts $_{\mathcal{X}/\mathcal{Y}}$ are omitted for general shapes.

The *complex* functional maps $\mathbf{Q}$ extend the FMaps concept to estimate correspondences between tangent vector fields (versus scalar functions) [21]. The optimization structure remains similar to Eq. 2, but it uses vector-valued features (versus scalar-valued) and the complex eigenbasis of the Connection Laplacian (versus the LBO). The key benefit of $\mathbf{Q}$ is its integration of surface orientation, making it suitable for disambiguating intrinsic shape symmetries. Both $\mathbf{C}$ and $\mathbf{Q}$ are small matrices, providing a highly compact representation compared to pointwise permutation matrices.

Equivariance and Vector Neurons. Equivariance is a fundamental property where a function $\mathcal{F}$ maintains the structure of a transformation $\mathbf{R}$ applied to its input $\mathbf{u}$: $\mathcal{F}(\mathbf{uR}) = \mathcal{F}(\mathbf{u})\mathbf{R}, \quad \forall \mathbf{R} \in \mathcal{T}$, where $\mathcal{T}$ is a transformation group (e.g. 3D rotations). Our work focuses on $\mathrm{SO}(3)$-equivariance due to its relevance in 3D vision. The vector neuron (VN) framework [19] provides a general method for constructing SO(3)-equivariant networks. The core idea is to extend traditional scalar neurons to 3D vectors, enabling a simple, explicit mapping of SO(3) actions to the hidden feature space. This allows for straight-forward generalization of classic operations, such as the VN-Linear layer:

$$\begin{aligned}\text{VN-linear}(\mathbf{uR}) &= \mathbf{M}(\mathbf{uR}) \\ &= (\mathbf{Mu})\mathbf{R} = \text{VN-linear}(\mathbf{u})\mathbf{R},\end{aligned}$$

where $\mathbf{u} \in \mathbb{R}^{c\times 3}$ is the input with c feature channels and $\mathbf{M} \in \mathbb{R}^{c\times c}$ are the learnable parameters. The additional dimension of size 3 is termed the *VN dimension*. Since invariance is a specific case of equivariance where $\mathcal{F}(\mathbf{uR}) = \mathcal{F}(\mathbf{u})$, our network learns $\mathrm{SO}(3)$-equivariant latent features that can be easily converted into invariant features to predict rotation-invariant correspondences.

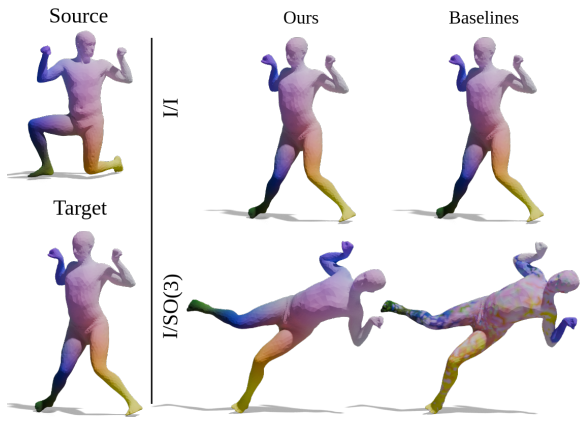

Figure 3. **SO(3)-invariant correspondences.** Our method is unaffected by shape orientations. In contrast, baselines (represented by DUOFM) perform well only when train and test shapes are pre-aligned (I/I), failing dramatically on unseen rotations (I/SO(3)).

4. SO(3)-Invariant Manifold Learning

Motivation. We seek an ideal 3D shape descriptor characterized by three properties: invariance to extrinsic shape embedding; effective exploitation of surface geometric structure, and robustness to noise and topological variations. The current standard, DiffusionNet [54], fails to satisfy all these criteria simultaneously. Its reliance on the extrinsic xyz coordinates requires either pre-alignment or the use of manually designed (intrinsic) descriptors as input. To overcome these limitations and realize a purely data-driven, geometric feature extractor, we propose RINONet. Our approach maintains feature diffusion on the surface as in [54], but integrates VN representations to enforce SO(3)-equivariance (cf. Fig. 2). This integration required a care-

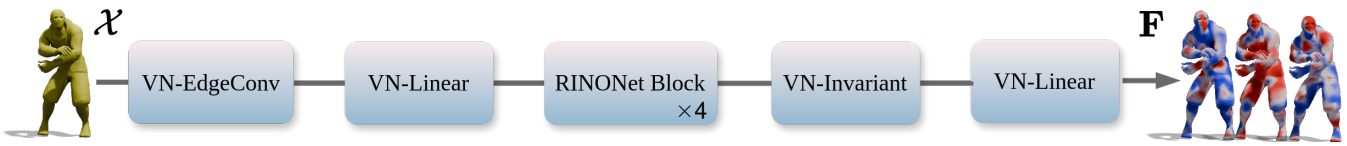

Figure 4. We propose **RINONet**, which learns SO(3)-invariant features $\mathbf{F}$ from input shape $\mathcal{X}$. Our novel network inherits all nice properties from DiffusionNet [54], and learns smooth, high-quality per-point features, while additionally retaining invariant to SO(3) actions applied to input shapes by employing *vector neurons* as the feature representation in our hidden layers. The RINONet has a simple structure and contains four consecutive **RINONet Blocks** at its core, which is combined with VN-EdgeConv and VN-linear layers to achieve the desired I/O dimensionality. The VN-invariant layer is employed to convert SO(3)-equivariant features to invariant ones.

ful architectural redesign to ensure rotational equivariance across all spatial diffusion and gradient operations.

Overview. RINONet starts with a VN-EdgeConv layer [19, 64], converting the input xyz coordinates $\mathbf{V} \in \mathbb{R}^{n\times 3}$ into features $\mathbb{R}^{n\times c\times 3}$ which serves as an aggregation of neighborhood information. The core structure is composed of four consecutive RINONet blocks, after which we append a VN-invariant layer ([19, Sec.3.5]) to obtain SO(3)-invariant features and, finally, a VN-linear layer to obtain output features with desired dimensionality. Notice that internally our RINONet remains SO(3)-equivariant until the VN-invariant layer (cf. Fig. 4). In the next, we present the design insights of our novel *RINONet block* and state its property of in-baked SO(3)-equivariance in Thm. 2. The full architecture is illustrated in Fig. 4. For details please refer to the supp.mat..

4.1. RINONet Block

Our RINONet block consists of three main modules: a diffusion layer, a gradient layer, and an MLP (cf. Fig. 5). All modules are designed to be SO(3)-equivariant by harnessing the VN representation. The hidden states of our network, $\mathbf{u}$ and $\mathbf{d}$, become vector-valued with dimensionality $\mathbb{R}^{n\times c\times 3}$ (instead of scalar-valued $\mathbb{R}^{n\times c}$), with n vertices and c feature channels. Next, we explain these modules.

VN-Diffusion Layer. The goal of this layer is to perform feature diffusion using a learned diffusion time t for each feature channel, which governs the size of the spatial support of the resulting features. Naively, feature diffusion is analogous to applying a linear operator $\mathcal{H}_t$ on an input feature $\mathbf{u}$, yielding $\mathbf{h} \coloneqq \mathcal{H}_t(\mathbf{u})$. However, to ensure SO(3)-equivariance, we must apply the same diffusion time t across all three VN dimensions of the input feature $\mathbf{u}$. This crucial design, together with the linearity of the diffusion operator, guarantees the SO(3)-equivariance, namely $\mathcal{H}_t(\mathbf{uR}) = \mathcal{H}_t(\mathbf{u})\mathbf{R}$. Furthermore, to enhance the model's ability to adjust the size of its spatial support, the diffusion time t is learned per feature channel. Note that although t is a learnable parameter updated during back-propagation, it remains fixed during any single feed-forward pass.

VN-Gradient Layer. The idea of this layer is to compute spatial gradient features $\mathbf{e}$ from a diffused input feature $\mathbf{h}$ to learn radially asymmetric filters, a crucial capability for expressive geometric representation. Similar to [54], we begin by computing the intrinsic spatial gradient $\mathbf{w} \coloneqq \mathcal{G}(\mathbf{h})$, which is then expressed in the local frame of each vertex as complex numbers. The gradient is transformed by a learnable (complex) matrix $\mathbf{A} \in \mathbb{C}^{c\times c}$, which is further aggregated by $\mathrm{Re}(\overline{\mathbf{w}} \odot \mathbf{Aw})$. However, this naive approach breaks the SO(3)-equivariance, as the aggregation operation does not commute with rotation, meaning $(\overline{\mathbf{w}}\mathbf{R} \odot \mathbf{AwR}) \neq (\overline{\mathbf{w}} \odot \mathbf{Aw})\mathbf{R}$. Our idea is an additional summation along the VN dimension to obtain the aggregated feature $\mathbf{f}$:

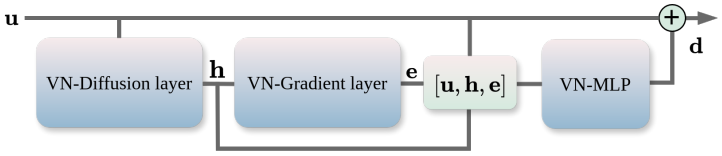

Figure 5. **Our RINONet block** is the core of our RINONet, and it consists of three main modules: a VN-Diffusion layer, a VN-Gradient layer, and a VN-MLP. In contrast to the original DiffusionNet block [54], ours is SO(3)-equivariant by design.

$$\mathbf{f} \coloneqq \mathrm{sum}(\mathrm{Re}(\overline{\mathbf{w}} \odot \mathbf{Aw}), \mathrm{dim} = 1). \tag{3}$$

where Re extracts the real part of a complex number. In the supp.mat., we demonstrate another equivalent formulation of Eq. 3 and prove the following result:

Theorem 1. *The feature $\mathbf{f} \in \mathbb{R}^{n\times c}$ computed as in Eq. 3 is* SO(3)*-invariant.*

After stabilizing training by passing $\mathbf{f}$ through a $\tanh$ layer $\mathbf{g} \coloneqq \tanh(\mathbf{f})$, we obtain the final equivariant features $\mathbf{e}$ by an element-wise multiplication (with broadcasting) of the invariant feature $\mathbf{g} \in \mathbb{R}^{n\times c}$ with the normalized equivariant feature $\mathbf{h} \in \mathbb{R}^{n\times c\times 3}$: $\mathbf{e} \coloneqq \mathbf{g} \odot \mathbf{h}$. This formulation maintains the network equivariance as much as possible.

VN-MLP. An MLP consists of linear layers followed by ReLU non-linearities, we disable the bias terms in the linear layers, since $\mathbf{M}(\mathbf{uR}) + \mathbf{b} \neq (\mathbf{Mu} + \mathbf{b})\mathbf{R}$ and swap the ReLU with its VN counterpart, to enable equivariance.

Remark. A RINONet block can be summarized as:

$$\mathbf{d} = \text{VN-MLP}([\mathbf{u}, \mathbf{h}, \mathbf{e}]) + \mathbf{u}, \tag{4}$$

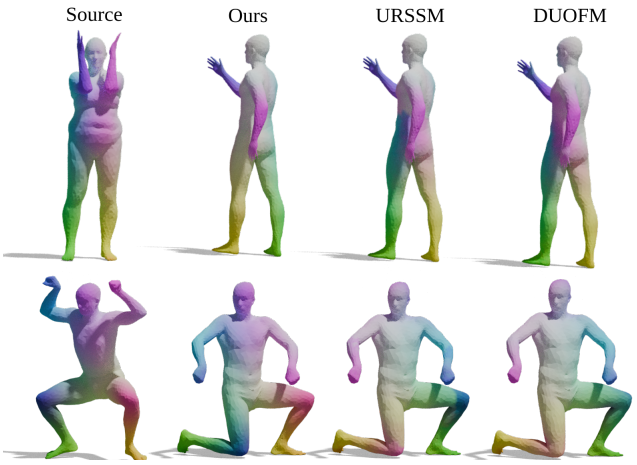

Figure 6. **Intrinsic symmetry.** Ours fully resolves symmetry, while baselines suffer. **Top:** partial left-right (legs) and front-back (belly-back) flips. **Bottom:** full left-right flips (arms and legs).

where the input feature $\mathbf{u}$, the equivariant diffusion feature $\mathbf{h}$ and the equivariant gradient feature $\mathbf{e}$ are concatenated together along the c dimension before going through a VN-MLP.

Theorem 2. *The RINONet block is* $\mathrm{SO}(3)$*-equivariant and the whole RINONet is* $\mathrm{SO}(3)$*-invariant.*

Please refer to the supp.mat. for a proof.

As discussed in [54], a complex matrix $\mathbf{A}$ is beneficial for disambiguating intrinsic symmetry due to its awareness of the shape orientation. In our experiments, we found that symmetrically flipped correspondences still appear when the xyz coordinates of unaligned shapes are fed into the network (cf. Fig. 6 & Tab. 3). Hence, we propose to exploit CFMaps, which can only encode orientation-preserving maps in theory, to fully eliminate intrinsic symmetry and regularize learned features.

5. Unsupervised SO(3)-Invariant Matching

Overview. We integrate our RINONet into a novel triple branch matching framework and propose RINO, the first unsupervised SO(3)-invariant approach for non-rigid shape matching. The process starts by passing two input shapes, $\mathcal{X}$ and $\mathcal{Y}$ (with $n_\mathcal{X}$ and $n_\mathcal{Y}$ vertices, respectively), through the network to obtain their (flattened) SO(3)-invariant features $\mathbf{F}_\mathcal{X} \in \mathbb{R}^{n_\mathcal{X} \times 3c}$ and $\mathbf{F}_\mathcal{Y} \in \mathbb{R}^{n_\mathcal{Y} \times 3c}$ in a Siamese fashion (cf. Fig. 2). These expressive features are then leveraged to compute three distinct map representations for matching: a (soft) pointwise map $\mathbf{\Pi}_{\mathcal{XY}}$, an FMap $\mathbf{C}_{\mathcal{XY}}$, and a CFMap $\mathbf{Q}_{\mathcal{XY}}$. All of them provide important structural properties for the matching that we aim to exploit with a careful coupling of the losses. During *inference*, RINO can match shapes in a single forward pass by Euclidean nearest neighbor search in the learned feature space. Our full pipeline is shown in Fig. 7.

Train/Test	I/I	I/SO(3)	SO(3)/SO(3)	Y/Y
CnsFM	5.4	58.7	9.1	5.4
DUOFM	11.7	41.4	25.2	16.5
URSSM	4.8	62.1	24.8	7.9
SMS	**4.6**	58.5	57.6	39.3
SmpFM	5.6	62.6	9.9	7.3
HbrFM	5.6	61.8	26.4	5.3
Ours	**4.6**	**4.6**	**4.6**	**4.6**

Table 1. **SO(3)-invariance on SMAL.** The rotation type is indicated by: I (aligned), SO(3) (fully random rotation), and Y (y-axis rotation). *e.g.* I/SO(3) means aligned shapes for training and fully randomly rotated shapes for testing. We report mGeoErr (↓) and ours consistently outperforms baselines, especially on unseen rotations. Best results are **bold**, and highlighted cells indicate the experiment settings used in baseline papers.

Map Representations. Our method derives three complementary map representations from the $\mathrm{SO}(3)$-invariant features $\mathbf{F}_\mathcal{X}$ and $\mathbf{F}_\mathcal{Y}$ generated by RINONet. The soft pointwise map is computed via softmax on the feature similarity: $\mathbf{\Pi}_{\mathcal{XY}} = \mathrm{Softmax}(\mathbf{F}_\mathcal{X}\mathbf{F}_\mathcal{Y}^T/\tau)$, where τ is a temperature parameter controlling the match entropy. Simultaneously, $\mathbf{F}_\mathcal{X}$ and $\mathbf{F}_\mathcal{Y}$ are fed into two non-learnable, differentiable blocks: the *FMaps* and the *CFMaps Solver*. These blocks efficiently solve convex problems to estimate $\mathbf{C}_{\mathcal{XY}}$ and $\mathbf{Q}_{\mathcal{XY}}$ (cf. Eq. 2). Our insight is twofold: first, $\mathbf{Q}_{\mathcal{XY}}$ inherently encodes only orientation-preserving maps, which is essential for disambiguating intrinsic shape symmetry [21]. Second, while $\mathbf{C}_{\mathcal{XY}}$ enforces feature consistency, $\mathbf{Q}_{\mathcal{XY}}$ additionally enforces consistency of the features' first-order derivatives (obtained via the gradient operator ∇), leading to more accurate features and correspondences.

Unsupervised Loss. Our unsupervised objective L_{total} is composed of three terms: a novel coupling loss L_{couple}, a structural loss L_{struct}, and a contrastive loss L_{contr}. The terms L_{struct} and L_{contr} are inspired by [13, 21, 22, 50], which encourage key properties of $\mathbf{C}$ and $\mathbf{Q}$, and discriminative per-vertex features (see supp.mat. for details).

Our contribution is the L_{couple}, which links the three map representations: $\mathbf{\Pi}, \mathbf{C}, \mathbf{Q}$, derived from our invariant features. We enforce consistency between the soft pointwise map $\mathbf{\Pi}$ and its functional pullbacks $\mathbf{C}$ and $\mathbf{Q}$ respectively. The coupling loss is defined as:

$$L_{\text{couple}} = L_{\mathbf{\Pi C}} + L_{\mathbf{\Pi Q}} \tag{5}$$

where $L_{\mathbf{\Pi C}}$ is the consistency between the pointwise map and $\mathbf{C}$, and $L_{\mathbf{\Pi Q}}$ is the proposed term enforcing consistency between the pointwise map and the orientation-aware complex functional map $\mathbf{Q}$. This targeted coupling is crucial for simultaneously achieving highly accurate features and robust symmetry disambiguation without requiring excessive coupling that can jeopardize training. Note that although

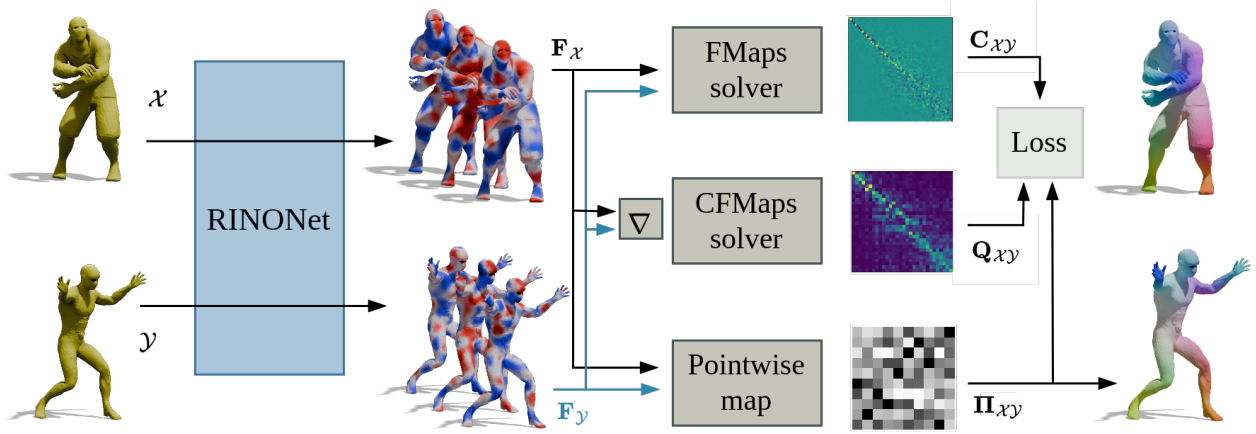

Figure 7. **Overview of RINO.** Our novel RINONet is used to extract SO(3)-invariant features $\mathbf{F}_\mathcal{X}$ and $\mathbf{F}_\mathcal{Y}$ from input shapes $\mathcal{X}$ and $\mathcal{Y}$. The features are further processed in three non-learnable but differentiable blocks: the FMaps solver to compute the functional map $\mathbf{C}_{\mathcal{XY}}$, the CFMaps solver to compute the complex functional map $\mathbf{Q}_{\mathcal{XY}}$ (based on feature gradients), and the pointwise map block to compute the soft pointwise map $\mathbf{\Pi}_{\mathcal{XY}}$ based on feature similarity. Finally, our unsupervised loss prompts structural properties of $\mathbf{C}_{\mathcal{XY}}$, $\mathbf{Q}_{\mathcal{XY}}$ and consistency between different map representations.

	wks			xyz		
	SMAL	**DT4D**	**FSCAN**	**SMAL**	**DT4D**	**FSCAN**
CnsFM	23.4	9.3	3.6	9.1	7.4	50.5
DUOFM	6.7	15.8	-	25.2	55.8	-
URSSM	29.6	8.4	3.0	24.8	59.3	24.8
SMS	30.6	58.6	12.1	57.6	64.9	54.0
SmpFM	29.9	15.3	5.7	9.9	5.5	24.3
HbrFM	34.5	5.4	3.5	26.4	33.8	22.1
Ours	28.7	9.8	2.7	**4.6**	**5.3**	**2.5**

Table 2. **Non-isometric and raw scan matching.** We report mGeoErr (↓) and highlight the **best** and 2nd best results. Ours outperforms all baselines regardless of the input types. Highlighted cells indicate the experiment settings used in baseline papers.

	FAUST			SCAPE		
	E	**ES**	**#Flips**	**E**	**ES**	**#Flips**
CnsFM	4.4	2.8	25	27.2	7.4	249
DUOFM	28.0	7.2	204	28.3	8.9	232
URSSM	30.0	7.8	240	26.9	7.7	245
SmpFM	3.0	2.7	1	29.9	7.9	63
HbrFM	29.9	7.2	227	27.2	8.1	236
Ours	**1.6**	**1.5**	**0**	**2.0**	**2.0**	**0**

Table 3. **Symmetry analysis**. We compute mGeoErr (↓) by allowing only symmetry-free correspondences (col. **E**) and additionally the symmetric ones (col. **ES**), and report the number of symmetrically flipped estimations out of 400 test pairs (col. **#Flips**).

the coupling of $\mathbf{\Pi}$ and $\mathbf{C}$ has been explored in [3, 11, 57], we are the first to study the coupling of all three map representations. Following our novel pipeline and loss, our estimated $\mathbf{C}$ is an orientation-preserving isometry akin to the results in [22] (cf. supp.mat.).

6. Experiments

Baselines and Metrics The proposed method is extensively compared against state-of-the-art unsupervised shape matching techniques, including the deep complex functional maps method DUOFM[22], the deep map coupling method URSSM[11], the cycle-consistent multi-matching method CnsFM[57], the deep non-rigid ICP method SMS[12], the hybrid functional maps method HbrFM[7], and the memory-efficient functional maps method SmpFM[37]. The focus on competitive unsupervised baselines is due to their superior performance over axiomatic approaches and their practical relevance given the scarcity of 3D assets with ground truth annotations. For a fair evaluation, all shapes are subject to random rotation during both training and testing under xyz as input unless otherwise mentioned, and no post-processing [11, 42, 62, 63] is applied, as all compared methods can be post-processed.

For all experiments, we adhere to the Princeton protocol [55], where the standard evaluation metric reported is the mean geodesic error (**mGeoErr**) with respect to the ground truth correspondence.

6.1. SO(3)-Invariant Matching

To assess SO(3)-invariant features, we utilized the remeshed SMAL dataset [47, 70] across four train/test configurations based on shape rotation: I/I (aligned), Y/Y (y-axis rotated), SO(3)/SO(3) (fully randomly rotated), and I/SO(3) (aligned training, fully rotated testing). Due to the non-isometry presented in SMAL, all baselines (except DUOFM) use xyz coordinates as input. As shown in Tab. 1, while all baselines perform well in the I/I setting, their performance significantly degrades in the Y/Y and SO(3)/SO(3) cases, and crucially, all baselines fail severely in the challenging I/SO(3) configuration (unseen rotation at test time). This confirms that existing methods lack robustness to arbitrary rotational changes and rely on

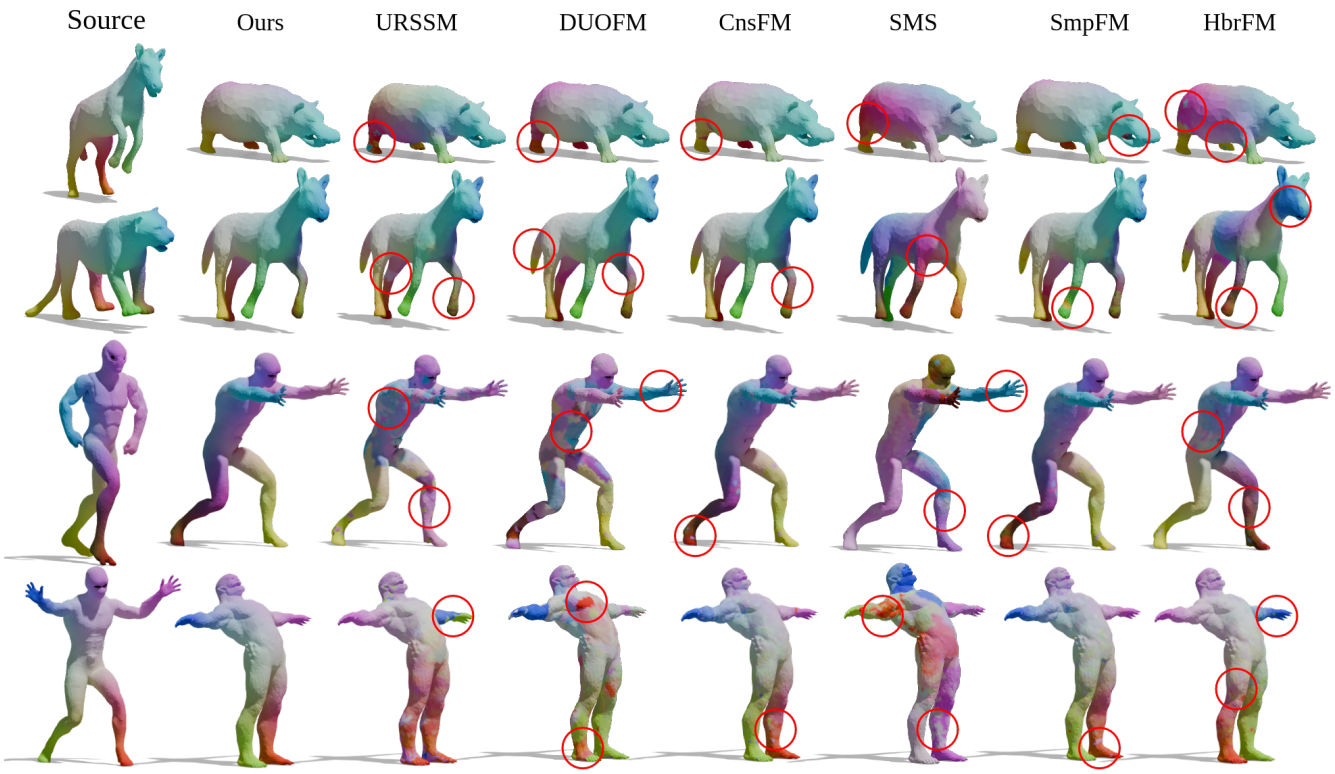

Figure 8. **Non-isometric matching on SMAL and DT4D.** Our method outperforms all baselines; it is invariant to extrinsic rigid poses (no pre-alignment) and robust to intrinsic shape symmetry by directly processing raw shape geometry.

shape alignment. While data augmentation can partially address this at the cost of training effort, our method achieves SO(3)-invariant features by design, thus completely eliminating the need for cumbersome and inefficient rotational augmentation. See Fig. 3 for visual examples.

6.2. Non-Isometric Matching

To evaluate our method's robustness to non-isometric deformations, we utilize the SMAL and DT4D [38, 47, 70] datasets, with quantitative results reported in Tab. 2. RINO outperforms all baselines, regardless of their input type. This superior performance directly addresses the known limitation of spectral descriptors like wks, whose near-isometric dependency makes them unstable under non-isometric deformation. By leveraging SO(3)-invariance, RINO is able to directly process raw 3D geometry. Qualitative results in Fig. 8 further show that our approach successfully yields high-quality, flip-free correspondences, whereas baselines suffer from non-smoothness and intrinsic left-right symmetry. We also notice the effectiveness of SmpFM when using xyz, which is undiscovered in its original paper. This is likely due to its effective architecture, which limits the use of SO(3)-variant features only in the differentiable Zoomout module, thereby avoiding detrimental coupling with low-quality pointwise maps. Methods like DUOFM and URSSM struggle to learn meaningful correspondences under challenging random rotations. Additionally, results on the FAUSTSCAN [8] (denoted as FSCAN) are reported, with further discussion in the supp.mat..

6.3. Disambiguating Intrinsic Symmetry

This experiment analyzes robustness against intrinsic symmetry using SCAPE and FAUST [1, 8], with results in Tab. 3. Allowing for symmetrically flipped matches significantly improves baseline accuracy, indicating that a substantial portion of their correspondence error is due to symmetry confusion (symmetric flips in $\sim$200 of 400 test pairs). Our method inherently avoids this issue due to the integration of CFMaps. While DUOFM also uses CFMaps, it still confuses symmetry because its features are not SO(3)-invariant, making them dependent on extrinsic embeddings. Although DiffusionNet's spatial gradient features can offer some aid via anisotropic diffusion [54, Fig. 5], we found it insufficient to fully resolve intrinsic symmetry, especially under random shape rotations.

6.4. Partial Shape Matching

The severe non-isometry inherent in partial shapes often forces baselines to switch their input features to xyz coordinates. Following the training strategy of [11], we first pretrain RINO on a collection of four full shape datasets (DT4D, SMAL, FAUST, SCAPE) for two epochs. The model is then fine-tuned on the SHREC16-Partiality (CUTS and HOLES) [17, 24] for 500 epochs, without any post-

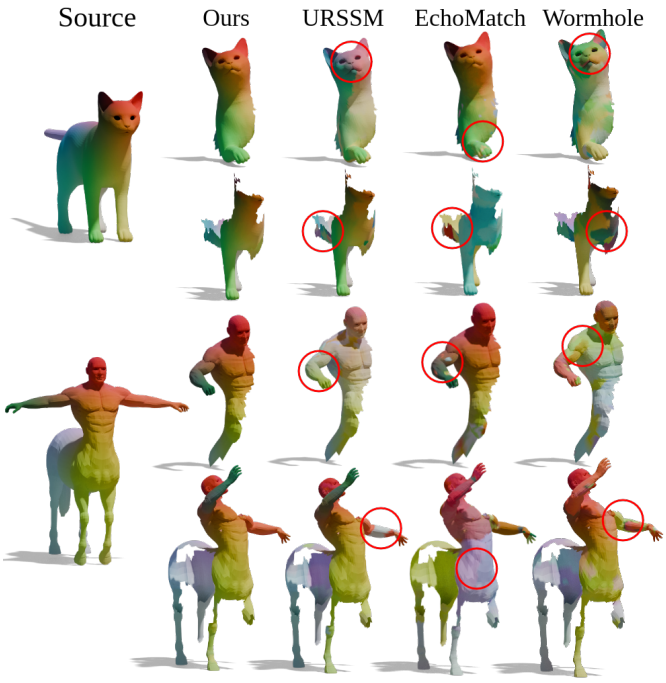

Figure 9. **Qualitative results on SHREC16-Partiality.**

	URSSM	EchoMatch	Wormhole	Ours
CUTS	42.1	12.7	48.2	**6.7**
HOLES	29.8	57.1	26.9	**12.09**

Table 4. **mGeoErr (↓) on SHREC16-Partiality.** Ours outperforms all baselines, including the supervised method EchoMatch.

processing for evaluation. As shown in Tab. 4, RINO outperforms strong partial matching baselines, even the supervised method EchoMatch. This superior performance is a strong indication that the ability to handle unaligned partial shapes, a feature built into our design, is crucial for successful partial shape matching (cf. Fig. 9).

6.5. Robustness under Noise

To analyze robustness to noise, we evaluate ours and baselines on ten SCAPE shapes corrupted by Gaussian noise with variance $\sigma \in [1\text{e-}3, 1\text{e-}2]$. As shown in Fig. 11, while all baselines perform acceptably up to $\sigma = 4\text{e-}3$, ours demonstrates superior robustness, with performance noticeably degrading only after $\sigma = 6\text{e-}3$, beyond which the degradation rate is approximately linear. At this high noise level, ours maintains accurate correspondences where baselines yield spurious results (cf. Fig. 20). We attribute this advantage to two factors: 1) our design processes xyz coordinates directly, avoiding noise-sensitive shape descriptors like wks; and 2) our RINONet has a significantly reduced number of trainable parameters, which acts as a powerful regularization effect, enhancing generalization to imperfect shapes.

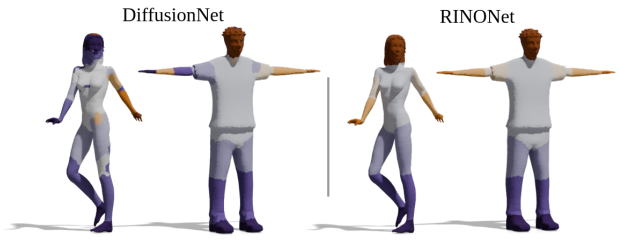

Figure 10. **Human Segmentation.** We show the segmentation results of DiffusionNet and ours on the composite dataset[40]. Our RINONet produces sharper segments.

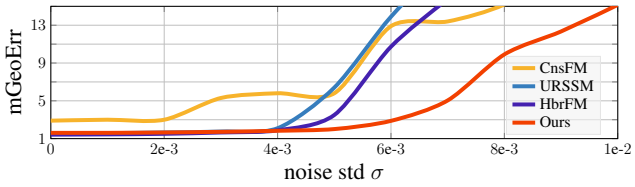

Figure 11. **Robustness under noise perturbations.**

6.6. Segmentation & More

As a proof-of-concept, we employ RINONet for human segmentation tasks using the composite dataset introduced in [40]. We compare against the SoTA DiffusionNet [54] and Fig. 10 shows the superior performance in the segmentation task, hence the promising potential of RINONet in a wider range of 3D understanding tasks.

Please refer to the supp.mat. for further studies on computational complexity, an ablation analysis of our loss and architecture, and experimental results on both isometric and non-manifold shapes.

7. Conclusion

We introduce **RINO**, the first unsupervised rotation-invariant dense correspondence method that unifies rigid and non-rigid matching. At its core is the novel feature extractor **RINONet**, that enables end-to-end SO(3)-invariant feature learning on surfaces. The integration with orientation-aware CFMaps further ensures robustness to intrinsic shape symmetries. Our method takes a significant step toward shape matching methods that are applicable to real-world data and achieves new SoTA across challenging non-rigid matching scenarios, including arbitrary poses, non-isometric deformations, partiality, non-manifold structures, and noise perturbations, thereby validating the feasibility and full potential of a genuinely data-driven non-rigid matching paradigm. As promising future work, we will explore matching methods that can directly consume the learned vector-valued features (avoiding feature flattening). We hope our work can inspire further advancements in rotation-invariant and data-driven 3D shape understanding.

Acknowledgment

We thank Vladimir Golkov, Thomas Dagès, Viktoria Ehm, and Linus Härenstam-Nielsen for proofreading the early version of the manuscript and providing valuable feedback, which significantly improved the work. We are also grateful to Nicolas Donati and Simone Melzi for the fruitful and inspiring discussions. We express our gratitude to the respective funding agencies. MG, SH, RM, and DC are supported by the ERC Advanced Grant SIMULACRON and the Munich Center for Machine Learning. MG also acknowledges support from the Bavarian Californian Technology Center. CD and LG gratefully acknowledge support from the Toyota Research Institute University 2.0 Program, a Vannevar Bush Faculty Fellowship, and a gift from the Flexiv Corporation. Additionally, CD is supported by a Tayebati Postdoctoral Fellowship.

References

[1] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. SCAPE: shape completion and animation of people. In *TOG*, 2005. 7, 3

[2] Souhaib Attaiki and Maks Ovsjanikov. NCP: Neural correspondence prior for effective unsupervised shape matching. In *NeurIPS*, 2022. 2

[3] Souhaib Attaiki and Maks Ovsjanikov. Understanding and improving features learned in deep functional maps. In *CVPR*, 2023. 1, 2, 6

[4] Souhaib Attaiki and Maks Ovsjanikov. Shape non-rigid kinematics (SNK): a zero-shot method for non-rigid shape matching via unsupervised functional map regularized reconstruction. In *NeurIPS*, 2023. 2

[5] Souhaib Attaiki, Gautam Pai, and Maks Ovsjanikov. DPFM: Deep partial functional maps. In *3DV*, 2021. 2

[6] Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *ICCV Workshops*, 2011. 1, 2

[7] Lennart Bastian, Yizheng Xie, Nassir Navab, and Zorah Lähner. Hybrid functional maps for crease-aware non-isometric shape matching. *CVPR*, 2024. 2, 6

[8] Federica Bogo, Javier Romero, Matthew Loper, and Michael J Black. FAUST: Dataset and evaluation for 3D mesh registration. In *CVPR*, 2014. 7, 3

[9] Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: Going beyond Euclidean data. *Signal Processing Magazine*, 2017. 3

[10] Dongliang Cao and Florian Bernard. Unsupervised deep multi-shape matching. In *ECCV*, 2022. 2

[11] Dongliang Cao, Paul Roetzer, and Florian Bernard. Unsupervised learning of robust spectral shape matching. In *TOG*, 2023. 1, 2, 6, 7, 3

[12] Dongliang Cao, Marvin Eisenberger, Nafie El Amrani, Daniel Cremers, and Florian Bernard. Spectral meets spatial: Harmonising 3d shape matching and interpolation. In *CVPR*, 2024. 6

[13] Dongliang Cao, Paul Roetzer, and Florian Bernard. Revisiting map relations for unsupervised non-rigid shape matching. In *3DV*, 2024. 5, 2

[14] Haiwei Chen, Shichen Liu, Weikai Chen, Hao Li, and Randall Hill. Equivariant point network for 3d point cloud analysis. *CVPR*, 2021. 3

[15] Zhangquan Chen, Puhua Jiang, and Ruqi Huang. DV-Matcher: Deformation-based non-rigid point cloud matching guided by pre-trained visual features. In *CVPR*, 2025. 2

[16] Taco S. Cohen and Max Welling. Group equivariant convolutional networks. In *ICML*, 2016. 3

[17] Luca Cosmo, Emanuele Rodola, Michael M Bronstein, Andrea Torsello, Daniel Cremers, Y Sahillioglu, et al. SHREC'16: Partial matching of deformable shapes. *3DOR*, 2016. 7, 3

[18] Pim de Haan, Maurice Weiler, Taco Cohen, and Max Welling. Gauge equivariant mesh cnns: Anisotropic convolutions on geometric graphs. In *ICLR*, 2021. 3

[19] Congyue Deng, Or Litany, Yueqi Duan, Adrien Poulenard, Andrea Tagliasacchi, and Leonidas Guibas. Vector neurons: A general framework for SO(3)-equivariant networks. In *ICCV*, 2021. 3, 4, 1

[20] Nicolas Donati, Abhishek Sharma, and Maks Ovsjanikov. Deep geometric functional maps: Robust feature learning for shape correspondence. In *CVPR*, 2020. 2, 1

[21] Nicolas Donati, Etienne Corman, Simone Melzi, and Maks Ovsjanikov. Complex functional maps: a conformal link between tangent bundles. In *CGF*, 2022. 2, 3, 5, 1

[22] Nicolas Donati, Etienne Corman, and Maks Ovsjanikov. Deep orientation-aware functional maps: Tackling symmetry issues in shape matching. In *CVPR*, 2022. 1, 2, 3, 5, 6

[23] Niladri Shekhar Dutt, Sanjeev Muralikrishnan, and Niloy J. Mitra. Diffusion 3d features (diff3f): Decorating untextured shapes with distilled semantic features. In *CVPR*, 2024. 2

[24] Viktoria Ehm, Paul Roetzer, Marvin Eisenberger, Maolin Gao, Florian Bernard, and Daniel Cremers. Geometrically consistent partial shape matching. In *3DV*, 2023. 7

[25] Viktoria Ehm, Maolin Gao, Paul Roetzer, Marvin Eisenberger, Daniel Cremers, and Florian Bernard. Partial-to-partial shape matching with geometric consistency. In *CVPR*, 2024. 3

[26] Maolin Gao, Zorah Lähner, Johan Thunberg, Daniel Cremers, and Florian Bernard. Isometric multi-shape matching. In *CVPR*, 2021. 2

[27] Maolin Gao, Paul Roetzer, Marvin Eisenberger, Zorah Lähner, Michael Moeller, Daniel Cremers, and Florian Bernard. SIGMA: Quantum scale-invariant global sparse shape matching. In *ICCV*, 2023. 2

[28] Maolin Gao, Huajian Zeng, and Daniel Cremers. CoE: Deep coupled embedding for non-rigid point cloud correspondences. In *3DV*, 2025. 2

[29] Mario Geiger and Tess Smidt. e3nn: Euclidean neural networks. *arXiv preprint arXiv:2207.09453*, 2022. 3

[30] Oshri Halimi, Or Litany, Emanuele Rodolà Rodolà, Alex M. Bronstein, and Ron Kimmel. Unsupervised learning of dense shape correspondence. In *CVPR*, 2019. 1, 2

[31] Qi-Xing Huang and Leonidas Guibas. Consistent shape maps via semidefinite programming. In *CGF*, 2013. 2

[32] Puhua Jiang, Mingze Sun, and Ruqi Huang. Neural intrinsic embedding for non-rigid point cloud matching. In *CVPR*, 2023. 2

[33] Tung Le, Khai Nguyen, Shanlin Sun, Nhat Ho, and Xiaohui Xie. Integrating efficient optimal transport and functional maps for unsupervised shape correspondence learning. In *CVPR*, 2024. 2

[34] Or Litany, Tal Remez, Emanuele Rodolà, Alex Bronstein, and Michael Bronstein. Deep functional maps: Structured prediction for dense shape correspondence. In *ICCV*, 2017. 1, 2

[35] Hsueh-Ti Derek Liu, Alec Jacobson, and Keenan Crane. A Dirac operator for extrinsic shape analysis. In *CGF*, 2017. 2

[36] Arman Maesumi, Tanish Makadia, Thibault Groueix, Vladimir Kim, Daniel Ritchie, and Noam Aigerman. PoissonNet: A local-global approach for learning on surfaces. *SIGGRAPH Asia*, 2025. 2

[37] Robin Magnet and Maks Ovsjanikov. Memory-scalable and simplified functional map learning. In *CVPR*, 2024. 1, 2, 6

[38] Robin Magnet, Jing Ren, Olga Sorkine-Hornung, and Maks Ovsjanikov. Smooth non-rigid shape matching via effective dirichlet energy optimization. In *3DV*, 2022. 7, 3

[39] Riccardo Marin, Marie-Julie Rakotosaona, Simone Melzi, and Maks Ovsjanikov. Correspondence learning via linearly-invariant embedding. In *NeurIPS*, 2020. 2

[40] Haggai Maron, Meirav Galun, Noam Aigerman, Miri Trope, Nadav Dym, Ersin Yumer, Vladimir G. Kim, and Yaron Lipman. Convolutional neural networks on surfaces via seamless toric covers. *TOG*, 2017. 8, 5, 6

[41] S. Melzi, R. Marin, E. Rodolà, U. Castellani, J. Ren, A. Poulenard, P. Wonka, and M. Ovsjanikov. Matching Humans with Different Connectivity. In *3DOR*, 2019. 3

[42] Simone Melzi, Jing Ren, Emanuele Rodolà, Abhishek Sharma, Peter Wonka, and Maks Ovsjanikov. ZoomOut: Spectral upsampling for efficient shape correspondence. In *TOG*, 2019. 6

[43] Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: A flexible representation of maps between shapes. In *TOG*, 2012. 2, 3

[44] Chunghyun Park, Seungwook Kim, Jaesik Park, and Minsu Cho. Learning SO(3)-invariant semantic correspondence via local shape transform. In *CVPR*, 2024. 3

[45] Emery Pierson, Lei Li, Angela Dai, and Maks Ovsjanikov. DiffuMatch: Category-agnostic spectral diffusion priors for robust non-rigid shape matching. *ICCV*, 2025. 2

[46] Adrien Poulenard, Marie-Julie Rakotosaona, Yann Ponty, and Maks Ovsjanikov. Effective rotation-invariant point cnn with spherical harmonics kernels. *3DV*, 2019. 3, 5

[47] Jing Ren, Adrien Poulenard, Peter Wonka, and Maks Ovsjanikov. Continuous and orientation-preserving correspondences via functional maps. In *TOG*, 2018. 2, 6, 7

[48] Jing Ren, Mikhail Panine, Peter Wonka, and Maks Ovsjanikov. Structured regularization of functional map computations. In *CGF*, 2019. 3

[49] Jing Ren, Simone Melzi, Peter Wonka, and Maks Ovsjanikov. Discrete optimization for shape matching. In *CGF*, 2021. 2, 6

[50] Jean-Michel Roufosse, Abhishek Sharma, and Maks Ovsjanikov. Unsupervised deep learning for structured shape matching. In *ICCV*, 2019. 5, 1, 6

[51] Yusuf Sahillioğlu. Recent advances in shape correspondence. *Visual Computer*, 2020. 2

[52] Samuele Salti, Federico Tombari, and Luigi Di Stefano. SHOT: Unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding*, 2014. 2

[53] Abhishek Sharma and Maks Ovsjanikov. Weakly supervised deep functional maps for shape matching. In *NeurIPS*, 2020. 2

[54] Nicholas Sharp, Souhaib Attaiki, Keenan Crane, and Maks Ovsjanikov. DiffusionNet: Discretization agnostic learning on surfaces. In *TOG*, 2022. 1, 2, 3, 4, 5, 7, 8

[55] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The princeton shape benchmark. In *Proceedings Shape Modeling Applications*, 2004. 6

[56] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *CGF*, 2009. 2

[57] Mingze Sun, Shiwei Mao, Puhua Jiang, Maks Ovsjanikov, and Ruqi Huang. Spatially and spectrally consistent deep functional maps. In *ICCV*, 2023. 2, 6

[58] Richard S. Sutton. The bitter lesson. *Incomplete Ideas*, 2019. 1

[59] Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation-and translation-equivariant neural networks for 3D point clouds. *arXiv:1802.08219*, 2018. 3

[60] Greg Turk and Marc Levoy. Zippered polygon meshes from range images. In *Conference on Computer Graphics and Interactive Techniques*. Association for Computing Machinery, 1994. 5

[61] Oliver Van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. A survey on shape correspondence. In *CGF*, 2011. 2

[62] Matthias Vestner, Zorah Lähner, Amit Boyarski, Or Litany, Ron Slossberg, Tal Remez, Emanuele Rodola, Alex Bronstein, Michael Bronstein, Ron Kimmel, et al. Efficient deformable shape correspondence via kernel matching. In *3DV*, 2017. 6

[63] Giulio Viganò, Maks Ovsjanikov, and Simone Melzi. NAM: Neural adjoint maps for refining shape correspondences. *TOG*, 2025. 6

[64] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. Dynamic graph CNN for learning on point clouds. In *TOG*, 2019. 4

[65] Ruben Wiersma, Ahmad Nasikun, Elmar Eisemann, and Klaus Hildebrandt. DeltaConv: anisotropic operators for geometric deep learning on point clouds. *TOG*, 2022. 2

[66] Yizheng Xie, Viktoria Ehm, Paul Roetzer, Nafie El Amrani, Maolin Gao, Florian Bernard, and Daniel Cremers. EchoMatch: Partial-to-partial shape matching via correspondence reflection. In *CVPR*, 2025. 2

[67] Zi Ye. *Discrete Spin Geometry for Surfaces*. PhD thesis, Technische Universität München, 2020. 2

[68] Zi Ye, Olga Diamanti, Chengcheng Tang, Leonidas Guibas, and Tim Hoffmann. A unified discrete framework for intrinsic and extrinsic Dirac operators for geometry processing. In *CGF*, 2018. 2

[69] Jia-Xing Zhong, Ta-Ying Cheng, Yuhang He, Kai Lu, Kaichen Zhou, Andrew Markham, and Niki Trigoni. Multibody SE(3) equivariance for unsupervised rigid segmentation and motion estimation. *NeurIPS*, 2023. 3

[70] Silvia Zuffi, Angjoo Kanazawa, David Jacobs, and Michael J. Black. 3D menagerie: Modeling the 3D shape and pose of animals. In *CVPR*, 2017. 6, 7, 3