

Hyper-Network Neural Functional Maps for Unsupervised Robust 3D Shape Matching

Dongliang Cao¹  and Florian Bernard¹ 

University of Bonn, Germany

Abstract. Functional maps are the cornerstone of recent non-rigid 3D shape matching methods due to their efficiency and performance. However, existing methods struggle with challenging scenarios, such as partiality, topological noise, and raw point clouds. A primary bottleneck is that significant intrinsic distortion prevents truncated spectral bases from being accurately aligned via linear transformations (i.e., functional maps). To address this, we introduce a *hyper-network that predicts non-linear neural functional maps (NFM)*, learned in an unsupervised manner, to better align spectral bases. Specifically, we model the NFM as an MLP with skip-connection to refine standard FM and employ a hyper-network to predict its weights, conditioned on standard FM. Our framework is trained using a novel unsupervised spectral alignment loss. Experiments demonstrate that our approach can be seamlessly integrated into state-of-the-art unsupervised deep functional map pipelines, substantially improving matching accuracy in demanding scenarios.

Keywords: 3D shape matching · Functional maps

1 Introduction

Establishing correspondences between non-rigid 3D shapes is a fundamental problem in computer vision and computer graphics. Its utility spans a diverse range of applications, from texture and pose transfer [20, 70, 71] to statistical shape modeling [25, 48, 50]. Despite being a long-standing challenge with decades of extensive study [19, 65], achieving accurate vertex-wise correspondence still remains a challenging task. This is particularly evident when dealing with topological noise and partiality. Such distortions are nearly ubiquitous in real-world 3D scanning data but difficult for traditional algorithms [54, 55] to tackle.

Recent advancements in deep functional map frameworks have led to significant breakthroughs across various non-rigid 3D shape matching tasks. These include robust solutions for near-isometric [2] and non-isometric matching [21, 47], as well as progress in multi-shape [14, 30, 72], multi-modal [15, 40] and partial shape matching [3, 11, 81]. Notably, state-of-the-art frameworks [6, 16, 17] have demonstrated remarkable versatility across diverse settings. However, despite these gains in standard scenarios, existing methods [16, 17] still fall short in more

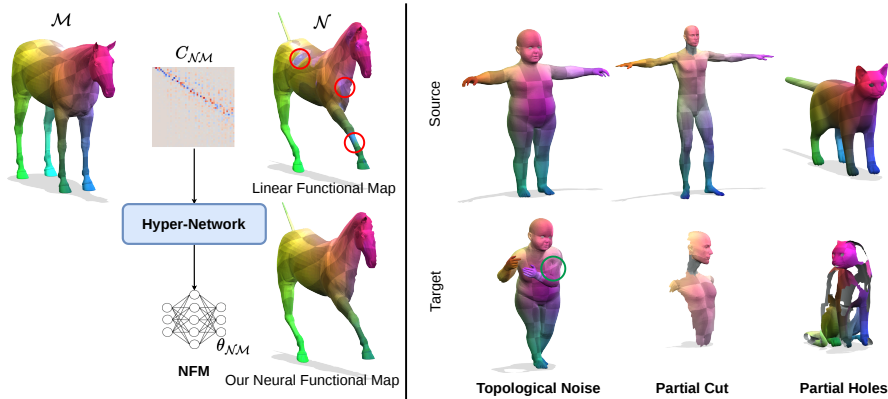


Fig. 1: Standard functional maps leverage a *linear* transformation and often fail to establish accurate correspondences when faced with significant intrinsic distortions – this is because a linear transformation is insufficient to align truncated spectral bases under severe partiality or topological noise. To overcome this, we introduce *non-linear neural functional maps (NFM)*, which are dynamically predicted by a hyper-network conditioned on standard FM. Our framework enables a more flexible and robust basis alignment, allowing our method to recover high-quality vertex-wise correspondences even in the presence of topological noise and extreme missing parts.

demanding conditions. Achieving accurate matching results remains a struggle when faced with significant partiality, topological noise, or the unstructured nature of raw point clouds. The failure of current functional map methods in these demanding scenarios is primarily rooted in a fundamental spectral misalignment: significant intrinsic distortion prevents the spectral bases of the two shapes from being accurately aligned via simple *linear* transformations, see Fig. 1. Because the standard functional map framework [54] relies on the assumption that a linear operator can effectively map one basis to another, it becomes inherently ill-equipped to reconcile the non-linear basis shifts caused by partiality or topological noise as shown in Fig. 2. To overcome the aforementioned bottleneck, we draw inspiration from recent work on map refinement [77] and propose an unsupervised learning framework for *non-linear neural functional maps (NFM)*.

In our approach we specifically substitute the standard linear functional map [54] with a non-linear neural functional map, which is modeled as an MLP, to facilitate more flexible spectral basis alignment. Unlike the previous refinement method [77] that requires accurate initial vertex-wise correspondences to optimize neural functional maps individually, we propose an unsupervised framework that learns a generalizable hyper-network using a training set to learn priors over the training distribution. At inference time, our hyper-network dynamically predicts the MLP weights conditioned on standard functional maps. To facilitate training, we introduce a novel spectral basis alignment loss, ensuring that the spectral bases are accurately aligned via the resulting non-linear transformation. Due to its design simplicity and flexibility, our framework can be

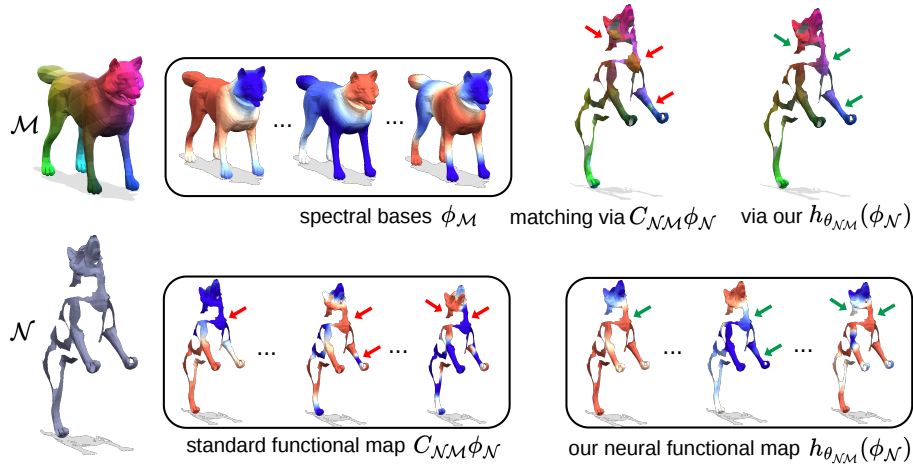


Fig. 2: Comparison between standard and neural functional maps on a challenging partial shape matching task. We visualize the source spectral bases $\phi_{\mathcal{M}}$ (top) and the corresponding aligned target bases $\phi_{\mathcal{N}}$ (bottom). Due to the significant intrinsic distortion induced by partiality, the standard linear functional map $C_{\mathcal{NM}}$ fails to accurately align the spectral bases (highlighted in **red arrows**). In contrast, our neural functional map $h_{\theta_{\mathcal{NM}}}(\cdot)$ facilitates a more flexible, non-linear alignment (highlighted in **green arrows**), resulting in superior matching accuracy and robustness under geometric irregularities.

seamlessly integrated into state-of-the-art deep functional map pipelines [15, 16]. We demonstrate through extensive experiments that our method substantially improves matching accuracy in demanding scenarios. Our primary contributions are summarized as follows:

- We propose the first unsupervised learning framework for *non-linear neural functional maps* (NFM), enabling flexible spectral basis alignment in challenging scenarios.
- We introduce a generalizable hyper-network that dynamically predicts the weights of these neural functional maps conditioned on standard functional maps, alongside a novel unsupervised spectral alignment loss to facilitate training.
- We demonstrate that our framework can be seamlessly integrated into existing state-of-the-art deep functional map pipelines, yielding substantial performance gains, particularly in the presence of partiality and topological noise.

2 Related Work

Non-rigid 3D shape matching is an established research area in computer vision and graphics, with a rich history of development [73, 75]. Rather than provid-

ing an exhaustive literature review, we focus our discussion on contemporary approaches, specifically those built upon the functional map framework [54].

2.1 Deep Functional Map Methods

Rather than directly finding vertex-wise correspondences, which is often formulated as an expensive combinatorial optimization problem [7, 34, 37, 63, 80], the functional map framework operates in the spectral domain. This paradigm encodes correspondence relationships into a compact matrix, termed the functional map, by projecting them onto a truncated set of basis functions [54] (typically the first k Laplace-Beltrami eigenfunctions [45]). Due to its mathematical elegance and computational efficiency, the functional map framework has been extensively integrated into modern deep learning pipelines. Instead of using traditional handcrafted features [4, 13, 66], deep learning approaches learn robust descriptors directly from training data to establish correspondences. Deep functional map methods are generally categorized as either supervised or unsupervised, depending on whether they require ground-truth correspondences for training. Supervised methods leverage ground-truth correspondences to optimize feature descriptors, primarily differing in their loss domain. While some minimize vertex-wise mismatches in the spatial domain [49], others operate on the spectral domain by regularizing the functional map itself [3, 22]. More recent hybrid approaches [46] combine both constraints to achieve superior robustness. Unsupervised methods typically rely on the near-isometric deformation assumption to optimize correspondences without ground truth. Early approaches [5, 36] penalize spatial mismatches by enforcing geodesic distance preservation. Subsequent works [64, 67] shift focus to the spectral domain, directly regularizing the structural properties (e.g., orthogonality and commutativity) of the functional map. Most recently, state-of-the-art performance has been achieved by frameworks [2, 16, 72] that couple functional maps with vertex-wise correspondences [59]. Despite these advances, the large intrinsic distortions inherent in partial or topologically noisy shapes prevent simple linear transformations from accurately aligning spectral bases. This fundamental limitation hinders the performance of current functional map methods in such demanding scenarios. To overcome this, we propose a novel unsupervised framework that learns *non-linear neural functional maps*, enabling a more flexible and robust alignment of bases.

2.2 Functional Map Refinement Methods

A common strategy to enhance matching accuracy is to iteratively refine functional maps as a post-processing step. The foundational ICP method iteratively converts functional maps to vertex-wise mappings via nearest-neighbor search [54]. Subsequent methods [31, 61, 62] optimize this relationship by minimizing specific energy functions. For instance, PMF [76] enforces bijectivity as a hard constraint, while ZoomOut [53] progressively increases spectral resolution during conversion. Other techniques focus on map properties: RHM [32] promotes reversible harmonic maps to reduce conformal distortion, and FSF [56]

utilizes theoretical spectral alignment for vertex-wise map recovery. For shape collections, another line of work leverages cycle-consistency [8, 33, 38, 39, 78]. The method most closely related to ours is NAM [77], which replaces standard linear functional maps with neural adjoint maps (NAM). By extending ZoomOut [53] into a Neural ZoomOut framework, it achieves superior refinement, particularly for non-isometric shapes. However, as a two-stage "match-and-refine" pipeline, its performance remains heavily dependent on the quality of initial correspondences; erroneous starts often lead to sub-optimal local minima. In contrast, our approach introduces a generalizable hyper-network that directly predicts the neural functional map weights for any given shape pair. This shift from per-pair optimization to a predictive learning framework and thus avoids the need for careful initialization and eliminates the time-consuming optimization required by traditional refinement techniques.

3 Background: Unsupervised Deep Functional Maps

In this section, we provide the background for the contemporary unsupervised deep functional map framework [16, 17] and a more comprehensive discussion is provided in the supplementary materials.

Consider a pair of 3D shapes, $\mathcal{M}$ and $\mathcal{N}$, represented as triangle meshes with $n_{\mathcal{M}}$ and $n_{\mathcal{N}}$ vertices, respectively. We compute the associated positive semi-definite Laplace-Beltrami operators $L_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times n_{\mathcal{M}}}$ and $L_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}}$ using the standard cotangent discretization $L_{\mathcal{M}} = A_{\mathcal{M}}^{-1}W_{\mathcal{M}}$, where $A_{\mathcal{M}}$ denotes the diagonal matrix of lumped vertex areas and $W_{\mathcal{M}}$ represents the cotangent weight matrix [58]. The spectral bases for each shape are then defined by the k most dominant eigenfunctions, $\Phi_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times k}$ and $\Phi_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times k}$, obtained via the generalized eigenvalue decomposition of the respective Laplacians. A Siamese feature extractor (e.g., DiffusionNet [68]) processes both shapes to extract c -dimensional vertex-wise features, denoted as $F_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times c}$ and $F_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times c}$. Given the extracted features, the functional map $C_{\mathcal{MN}} \in \mathbb{R}^{k \times k}$ (relating the two spectral bases) is computed by solving a least-squares optimization problem

$$C_{\mathcal{MN}} = \operatorname{argmin}_C E_{\text{data}}(C) + \lambda E_{\text{reg}}(C). \quad (1)$$

Here, minimizing $E_{\text{data}} = \left\| C \Phi_{\mathcal{M}}^{\dagger} F_{\mathcal{M}} - \Phi_{\mathcal{N}}^{\dagger} F_{\mathcal{N}} \right\|_F^2$ enforces descriptor preservation, while minimizing the regularization term E_{reg} imposes some form of structural properties (e.g., commutativity [54]). The operator $\dagger$ denotes the Moore-Penrose inverse, which projects the features onto the spectral bases. Meanwhile, soft vertex-wise correspondences can be obtained via feature similarity

$$\Pi_{\mathcal{NM}} = \operatorname{Softmax}(F_{\mathcal{N}} F_{\mathcal{M}}^T / \tau), \quad (2)$$

where τ is the scaling factor to determine the softness of the correspondence matrix. During training, structural regularization (e.g. orthogonality, bijectivity [64]) is imposed on the computed functional maps $C_{\mathcal{MN}}$. Moreover, to encourage that the computed functional map is associated with a valid vertex-wise

map, an additional coupling loss is applied between the computed functional map and soft vertex-wise correspondences, i.e.,

$$L_{\text{couple}} = \left\| C_{\mathcal{M}\mathcal{N}} - \phi_{\mathcal{N}}^{\dagger} \Pi_{\mathcal{N}\mathcal{M}} \phi_{\mathcal{M}} \right\|_F^2. \quad (3)$$

During inference, the final vertex-wise correspondences can be obtained based on feature similarities using

$$\Pi_{\mathcal{N}\mathcal{M}} = \text{NN}(F_{\mathcal{N}}, F_{\mathcal{M}}), \quad (4)$$

where NN denotes nearest neighbor search in $F_{\mathcal{N}}$ for each entry in $F_{\mathcal{M}}$.

Despite its mathematical elegance, the standard deep functional map framework remains limited by its reliance on *linear* transformations. This linear dependency proves insufficient when handling the significant intrinsic distortions present in partial or topologically noisy shapes, as a single matrix C cannot accurately model the complex spectral shifts induced by such irregularities. This limitation motivates the proposal of *non-linear neural functional maps*, which facilitate a more flexible and robust alignment of spectral bases. By transcending the limitations of linear transformations, it effectively accommodates the complex intrinsic distortions that standard functional maps fail to resolve.

4 Our Unsupervised Neural Functional Map Method

We propose an unsupervised framework that replaces the traditional linear spectral transformation under truncated bases with non-linear neural functional maps to compensate for this truncation-induced misalignment. As illustrated in Fig. 3, our pipeline consists of three core stages: (a) an initial deep functional map block, (b) neural functional map prediction via a hyper-network, and (c) spectral alignment for final correspondence recovery.

First, we utilize a standard unsupervised functional map pipeline (*cf.* Sec. 3) to estimate an initial functional map $C_{\mathcal{N}\mathcal{M}}$ and soft vertex-wise correspondences $\Pi_{\mathcal{N}\mathcal{M}}$. To overcome the inherent rigidity of this linear estimate, we introduce a hyper-network H that takes $C_{\mathcal{N}\mathcal{M}}$ to predict the parameters $\theta_{\mathcal{N}\mathcal{M}}$ of a neural functional map $h_{\theta_{\mathcal{N}\mathcal{M}}}$. This neural functional map (NFM) performs a non-linear alignment of the source spectral bases $\phi_{\mathcal{N}}$ to the target bases $\phi_{\mathcal{M}}$, effectively compensating for complex intrinsic distortions that a linear functional map cannot resolve. Finally, refined vertex-wise correspondences are established via a nearest-neighbor search in the resulting aligned spectral space. In the following sections, we detail the architectures of the hyper-network and neural functional map, alongside our unsupervised spectral alignment loss.

4.1 Neural Functional Map

Standard functional maps rely on a linear transformation matrix $C_{\mathcal{N}\mathcal{M}} \in \mathbb{R}^{k \times k}$ to align the truncated spectral bases of two shapes. While computationally efficient, this linear formulation assumes that the relationship between spectral

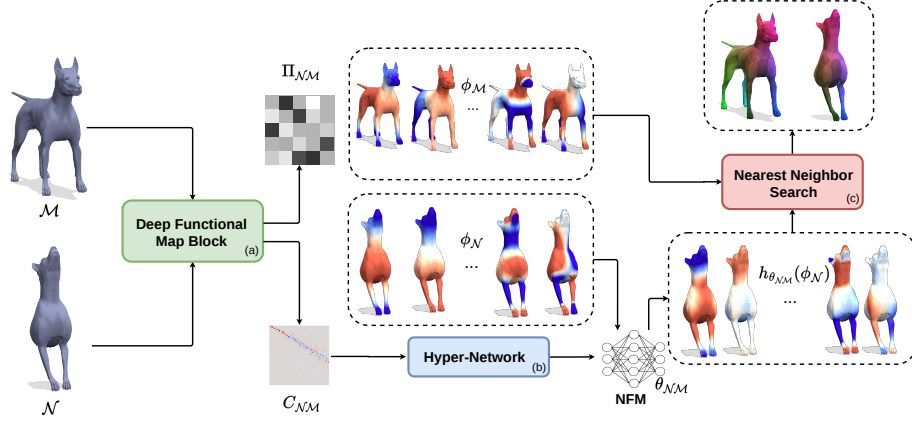


Fig. 3: Our pipeline for unsupervised learning of a hyper-network to predict neural functional maps. Given a pair of 3D shapes, we first employ a standard deep functional map block to estimate an initial functional map and its corresponding vertex-wise correspondences. A hyper-network then predicts the weights of a neural functional map (i.e., $\theta_{N,M}$), conditioned on the initial functional map (i.e., $C_{N,M}$). This neural functional map facilitates a non-linear alignment of the target spectral bases with the source bases. Finally, refined vertex-wise correspondences are established via a nearest-neighbor search on this newly aligned spectral domain. Our unsupervised spectral alignment loss utilizes the initial vertex-wise correspondences $\Pi_{N,M}$ to train the hyper-network.

coefficients is strictly additive and proportional, which often fails under non-isometric deformations, topological noise or partiality. To address this, we define a neural functional map $h_{\theta_{N,M}} : \mathbb{R}^k \rightarrow \mathbb{R}^k$, a non-linear function parameterized by a neural network.

Instead of a global matrix multiplication, the mapping is defined as:

$$\hat{\phi}_N(x) = h_{\theta_{N,M}}(\phi_N(x)), \forall x \in V_N, \quad (5)$$

where $\phi_N(x)$ represents the spectral coefficients of the vertex x on the source shape N . The network $h_{\theta_{N,M}}$ acts as a learnable transformation between spectral bases that can align the basis functions in a high-dimensional feature space. This formulation allows for a more flexible alignment, as the neural network can learn to ignore noisy spectral components or compensate for non-linear shifts in the truncated spectral bases.

In our implementation, $h_{\theta_{N,M}}$ is designed as a lightweight multi-layer perceptron (MLP) with residual connections. Specifically, we use a three layer MLP (similar to NAM [77]) with a skip-connection of the standard FM $C_{N,M}$, i.e.,

$$h_{\theta_{N,M}}(\phi_N(x)) = C_{N,M}\phi_N(x) + W^3\sigma(W^2\sigma(W^1\phi_N(x))), \quad (6)$$

where σ is a non-linear activation function (i.e., ReLU), $C_{N,M} \in \mathbb{R}^{k \times k}$, $W^1 \in \mathbb{R}^{k \times k}$, $W^2 \in \mathbb{R}^{k \times k}$, $W^3 \in \mathbb{R}^{k \times k}$ are linear transformations.

In contrast to NAM [77], our parameters $\theta_{\mathcal{NM}}$ are not fixed during training but are instead dynamically generated by the hyper-network based on the initial standard functional map. By operating directly in the spectral domain, the neural functional map maintains the advantages of low-dimensional representation while gaining the expressiveness of deep neural architectures. By isolating the residual component in Eq. (6), it becomes evident that the neural functional map recovers the standard linear transformation when the non-linear residuals are zero. Consequently, this formulation can be viewed as a generalized functional map framework: it preserves the structural benefits of classical spectral alignment while providing the necessary degrees of freedom to model the non-linear deformations that a fixed matrix $C_{\mathcal{NM}}$ cannot capture.

4.2 Hyper-Network for Neural Functional Map Prediction

To achieve a generalized matching framework that avoids expensive per-pair optimization, we introduce a hyper-network H to dynamically generate the parameters of the neural functional map. Unlike traditional refinement methods [53, 77] that treat map parameters as variables to be optimized at test-time, our hyper-network learns a mapping from the spectral domain representation (i.e., functional maps) to the parameter space of the MLP by exploiting priors over the training distribution.

The hyper-network takes the initial functional map $C_{\mathcal{NM}} \in \mathbb{R}^{k \times k}$, which encapsulates the global geometric relationship between the two shapes, as input. It then predicts the weights $\theta_{\mathcal{NM}}$ for the neural functional map h_θ , i.e.,

$$\theta_{\mathcal{NM}} = H(C_{\mathcal{NM}}), \quad (7)$$

where $\theta_{\mathcal{NM}} = \{W_{\mathcal{NM}}^1, W_{\mathcal{NM}}^2, W_{\mathcal{NM}}^3\}$. By conditioning the neural mapping on $C_{\mathcal{NM}}$, the hyper-network allows the model to adapt its non-linear alignment strategy based on the specific distortion profile of the input pair characterized by the initial functional map (e.g., slanted-diagonal structure under partiality).

Given that the initial functional map $C_{\mathcal{NM}}$ is a structured 2D representation of spectral correlations, we interpret it as a single-channel image as commonly done in prior functional map generation methods [57, 83]. We then employ a Vision Transformer (ViT) architecture [23] for the hyper-network to exploit its self-attention mechanism, which is uniquely suited for capturing the global dependencies across different frequency bands within the functional map. Specifically, H treats the $k \times k$ functional map as input and produces a $3 \times k \times k$ output. This output is interpreted as a three-channel feature map, where each channel corresponds to a specific linear layer for the neural functional map h_θ as defined in Eq. (6). By formulating the parameter prediction as an image-to-image translation task, the ViT can learn to recognize and correct complex patterns of spectral misalignment induced by partiality and topological changes, such as those appearing as off-diagonal noise or block-diagonal shifts [60].

4.3 Unsupervised Spectral Alignment Loss

To train the hyper-network without ground-truth supervision, we introduce a spectral alignment loss that encourages the neural functional map to find a non-linear mapping consistent with the estimated spatial correspondences. Let $\Pi_{\mathcal{N}\mathcal{M}} \in [0, 1]^{n_{\mathcal{N}} \times n_{\mathcal{M}}}$ represent the soft vertex-wise correspondence matrix obtained from the initial deep functional map block (*cf.* Eq. (2)). This matrix represents the probability of a vertex on the target shape $\mathcal{M}$ being mapped to a vertex on the source shape $\mathcal{N}$.

The unsupervised spectral alignment loss is defined as the discrepancy between the target bases and the non-linearly transformed source bases in the spectral domain, i.e.,

$$\mathcal{L}_{\text{align}} = \|\Pi_{\mathcal{N}\mathcal{M}}\phi_{\mathcal{M}} - h_{\theta_{\mathcal{N}\mathcal{M}}}(\phi_{\mathcal{N}})\|_F^2. \quad (8)$$

Here, the soft correspondence matrix $\Pi_{\mathcal{N}\mathcal{M}}$ acts as a spatial bridge, transporting the per-vertex spectral signals from the target domain to the source domain. The neural functional map $h_{\theta_{\mathcal{N}\mathcal{M}}}$ is then tasked with non-linearly warping the source bases to match this transported target bases. By minimizing this objective, the hyper-network learns to predict parameters $\theta_{\mathcal{N}\mathcal{M}}$ that correct for the misalignment presents in the initial linear functional map.

5 Experimental Results

In this section, we evaluate the performance of our unsupervised neural functional map framework against state-of-the-art baselines across several challenging 3D shape matching scenarios, including topological noise, partiality, and multi-modal data. To demonstrate the versatility and robustness of our approach, we integrate our non-linear neural functional map learning module into two leading unsupervised frameworks: the state-of-the-art deep functional map pipeline [16] for topologically noisy and partial shape matching, and the self-supervised functional map learning framework [15] for multi-modal matching between triangle meshes and point clouds. By maintaining consistent experimental settings with these baselines, we highlight the specific performance gains obtained by our unsupervised neural functional map learning framework.

5.1 Matching with Topological Noise

Datasets. Real-world 3D scans often exhibit significant topological inconsistencies, such as self-intersections of distinct surfaces. These irregularities pose a severe challenge for existing matching methods: they distort the intrinsic geometry, causing spectral shifts that compromise functional map methods [43], and they introduce non-manifold artifacts that disrupt spatial registration approaches [30]. To evaluate the robustness of our neural functional map method in the presence of such noise, we conduct experiments on the **TOPKIDS** dataset [43], which contains non-rigid shapes with varying levels of topological changes. Here

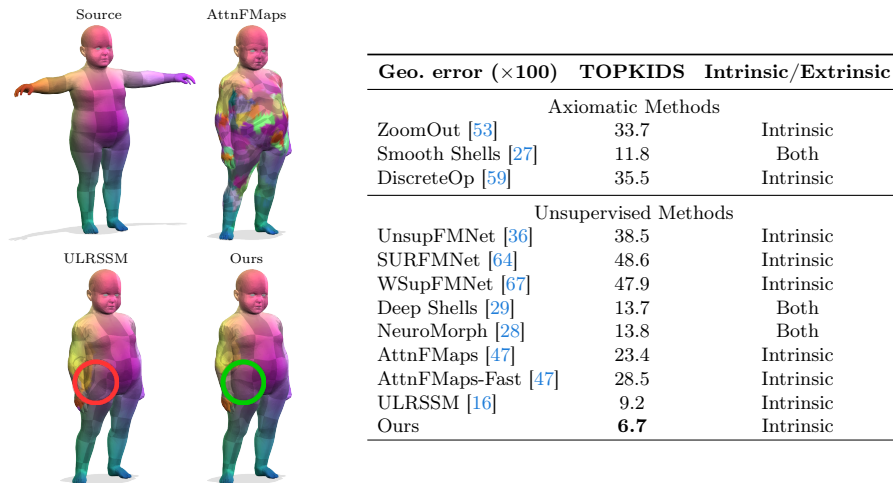


Fig. 4: Results on the TOPKIDS dataset. We categorize existing approaches into fully intrinsic methods (based strictly on functional maps) and extrinsic-aware methods that utilize additional spatial information (e.g., rigid alignment). Our method achieves a substantial performance margin over all baselines, notably even outperforming even those methods that rely on extrinsic priors.

the shapes with topological noise are expected to be matched to a clean template shape with T pose, as shown in Fig. 4 left.

Results. The results on the **TOPKIDS** dataset demonstrate the superior robustness of our neural functional map framework. We use the mean geodesic error [42] as our evaluation metric. As shown in Fig. 4, our method achieves a significant reduction in mean geodesic error compared to both functional map methods and methods based on spatial registration. Notably, our approach maintains high correspondence accuracy even in cases of severe self-intersection where standard linear methods suffer from failure. The performance gain can be attributed to the non-linear flexibility of the neural functional map. In the presence of topological noise, the spectral bases undergo localized shifts that cannot be modeled by a linear matrix (i.e., functional map). In contrast, our neural functional map learns to warp the spectral domain non-linearly. By effectively re-aligning these distorted basis functions, the neural functional map preserves the underlying geometric structure, allowing for accurate pointwise recovery where linear methods typically fail.

5.2 Partial Shape Matching

Datasets. For partial shape matching, we use the **SHREC16** dataset [18], which is based on the synthetic shapes of TOSCA [12]. It contains 76 nearly-isometric shapes from eight different classes of humans and animals. Each class has one template mesh in a neutral pose. The partial shapes are divided into CUTS and HOLES. The CUTS dataset contains partial shapes cut by a plane, whereas the HOLES dataset contains shapes with irregular holes that are produced by an

erosion process. We rely on the CUTS24 split [26] to ensure that no shapes from the test set are used for training. We also test on the **PFAUST** dataset [10], a modification of the FAUST dataset [9], where 10 different humans (8 are used for training and 2 for testing) are seen in 10 different poses. PFAUST processes the dataset by adding artificial holes of varying difficulty, including PFAUST-M (medium difficulty) and PFAUST-H (hard difficulty). Following recent advancements in hybrid geometric-visual descriptors [24, 79], we incorporate DINOv3 features [69] to further improve matching performance. Specifically, we extract high-level semantic features from multi-view rendered images of each shape and lift them back to the shape vertices [24]. This allows our model to leverage robust visual priors alongside geometric information. To ensure a fair and transparent evaluation, we report matching accuracy using two distinct feature configurations: first, a version with pure vertex coordinates to maintain consistency with prior literature [3, 16]; second, a version enhanced by DINOv3 visual descriptors. In this way, it allows us to isolate the gains of our neural functional map from the improvements attributed to high-level semantic feature representations.

Geo. error ($\times 100$)	SHREC16		PFAUST	
	CUTS	HOLES	Medium	Hard
Supervised Methods				
GeomFMaps [22]	12.8	15.3	4.3	6.6
+ <i>DINOv3</i>	3.2	6.8	2.8	3.8
DPFM-sup [3]	3.2	13.1	3.0	6.8
+ <i>DINOv3</i>	3.2	6.6	2.8	3.8
Unsupervised Methods				
DPFM-unsup [3]	9.0	20.5	9.3	12.7
+ <i>DINOv3</i>	10.8	10.4	3.8	6.5
ULRSSM [16]	3.2	8.2	7.5	9.2
+ <i>DINOv3</i>	4.3	9.7	1.7	5.2
Ours	2.8	7.6	7.0	7.9
+ <i>DINOv3</i>	1.9	4.7	1.7	2.3

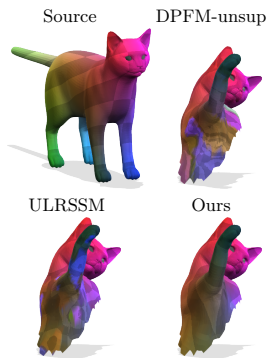


Fig. 5: Results on partial shape matching benchmarks. Our method demonstrates superior robustness across diverse partiality patterns, significantly outperforming all unsupervised baselines in most cases. Notably, on the most challenging benchmarks (e.g., SHREC16 HOLES and PFAUST Hard), our unsupervised approach even significantly exceeds the performance of state-of-the-art supervised methods. This highlights the efficacy of our non-linear neural functional map, which accurately aligns spectral bases despite severe boundary artifacts and surface incompleteness.

Results. The results on partial shape matching benchmarks confirm that our method sets a new state-of-the-art for unsupervised partial shape matching. As indicated in Fig. 5, our method significantly outperforms all existing baseline methods in nearly every test scenario. Most notably, our unsupervised approach even surpasses the accuracy of established supervised methods on the most challenging benchmarks. This performance gap is particularly evident in datasets characterized by extreme missing geometry, where the inherent incompleteness of the shapes makes finding a stable basis alignment traditionally difficult. By enabling a flexible, non-linear transformation in the spectral domain based on

Source



Fig. 6: Qualitative results on the SHREC16 HOLES benchmark. Even under conditions of extreme partiality, our method recovers accurate, semantically consistent correspondences. This success is driven by the non-linear neural functional map, which provides the flexibility to align spectral bases that have been significantly distorted by the presence of large holes that traditional linear approximations typically collapse.

our neural functional maps, our method can accurately correlate the relationship between partial and complete spectral bases, ensuring robust pointwise recovery even when a large portion of the shape is absent, as shown in Fig. 6.

5.3 Multi-Modal Shape Matching

Datasets. For multi-modal shape matching, we evaluate on the **FAUST** [9], **SCAPE** [1], and **SHREC19** [52] datasets. In alignment with prior works [15, 40, 41], we assess the model’s ability to establish correspondences across different shape representations, specifically between triangle meshes and unstructured point clouds. The FAUST dataset consists of 100 shapes (10 people in 10 poses), where the evaluation is performed on the last 20 shapes. The SCAPE dataset contains 71 different poses of the same person, where the last 20 shapes are used for evaluation. The SHREC19 dataset is a more challenging benchmark dataset due to its significant variations in mesh connectivity and irregular sampling. It comprises 44 shapes and a total of 430 evaluation pairs. To be consistent with prior works [15, 40, 41], we use vertex coordinate as the input for all datasets.

Results. Across all evaluated benchmarks, our framework sets the new state of the art, consistently surpassing both supervised and self-supervised baselines (*cf.* Tab. 1) and shows remarkable cross-dataset generalization ability for both mesh and point cloud. In the multi-modal setting, the neural functional map proves particularly effective at bridging the spectral gap between triangle meshes and point clouds, where discretization differences often lead to basis misalignment. By outperforming existing self-supervised and supervised methods on FAUST, SCAPE, and the challenging SHREC’19 datasets, our approach demonstrates superior generalization and robustness. These results validate that our non-linear refinement is not only more expressive than standard linear transformations but also more resilient to the varied sampling densities and connectivity artifacts inherent in multi-modal data.

Table 1: Results on multi-modal shape matching datasets. We evaluate all methods on individual dataset for shapes represented as triangle meshes and point clouds. The first row indicates the training dataset and the second row indicates the test datasets. The numbers in parentheses indicates the mean geodesic error for point clouds. The best results in each column are **highlighted**.

Geo. error ($\times 100$)	FAUST			SCAPE		
	FAUST	SCAPE	SHREC19	SCAPE	FAUST	SHREC19
Supervised Methods						
3D-CODED [35]	2.5 (2.5)	31.0 (31.0)	-	31.0 (31.0)	33.0 (33.0)	-
DiffFMaps [51]	3.6 (3.6)	19.0 (19.0)	16.4 (16.4)	12.0 (12.0)	12.0 (12.0)	17.6 (17.6)
TransMatch [74]	2.7 (2.7)	33.6 (33.6)	21.0 (21.0)	18.6 (18.6)	18.3 (18.3)	38.8 (38.8)
Unsupervised/Self-Supervised Methods						
CorrNet3D [82]	26.5 (26.5)	38.0 (38.0)	44.0 (44.0)	37.3 (37.3)	36.2 (36.2)	42.6 (42.6)
DCP [44]	11.1 (11.1)	17.5 (17.5)	31.0 (31.0)	17.3 (17.3)	11.2 (11.2)	28.7 (28.7)
NIE [40]	5.5 (5.5)	15.0 (15.0)	15.1 (15.1)	11.0 (11.0)	8.7 (8.7)	15.6 (15.6)
DFR [41]	3.0 (3.0)	6.3 (6.3)	5.9 (5.9)	2.9 (2.9)	4.0 (4.0)	4.8 (4.8)
SSMSM [15]	2.0 (2.4)	9.2 (11.0)	7.5 (9.0)	3.1 (4.1)	7.2 (8.5)	6.8 (7.3)
Ours	1.8 (2.2)	4.6 (6.0)	5.0 (5.7)	2.3 (2.8)	3.2 (3.7)	4.3 (4.7)

6 Ablation Study

6.1 Component Analysis

To evaluate the contribution of each component within our framework, we conduct a series of ablation experiments on the TOPKIDS dataset, as its significant topological noise provides a rigorous test for method robustness. Our investigation focuses on two primary aspects: the architecture of the hyper-network and the depth (capacity) of the neural functional map.

First, we evaluate the impact of the hyper-network by replacing our default Vision Transformer (ViT) backbone with CNN and MLP alternatives. This comparison demonstrates that the global receptive field and self-attention mechanism of the ViT are essential for capturing long-range spectral dependencies required to predict coherent map parameters. Second, we ablate the complexity of the neural functional map itself by varying the number of hidden layers in its MLP structure, seeking the optimal balance between expressivity and regularization.

(a) Hyper-network as MLP. We replace the ViT with a standard multi-layer perceptron. This setting lacks the spatial and structural priors of the ViT, treating the input functional map as a flattened vector, which limits its ability to recognize localized spectral patterns.

(b) Hyper-network as CNN. We utilize a convolutional neural network as the hyper-network. While this introduces inductive biases for local feature extraction, it lacks the global context provided by self-attention, which is crucial for understanding the overall functional map structure.

(c) 1-hidden layer. We restrict the neural functional map to a single hidden layer. This shallow configuration serves as a "near-linear" baseline to test if minimal non-linearity is sufficient to correct spectral distortions.

(d) 3-hidden layers. We increase the depth of the neural functional map to three hidden layers. This setting evaluates whether more non-linear capacity can capture the complex warping required for topologically noisy shapes, compared to our 2-hidden layers default setting.

(e) 4-hidden layers. We test a deeper four-layer configuration for the neural functional map. This allows us to observe if excessive depth leads to overfitting or if the added capacity further aids in basis alignment.

Table 2: Qualitative results of our ablation study on TOPKIDS dataset. Compared to different settings, our default setting obtains the best matching performance.

TOPKIDS	(a) MLP	(b) CNN	(c) 1-hidden layer	(d) 3-hidden layers	(e) 4-hidden layers	(f) Ours
Geo. error	10.8	9.6	10.3	8.6	8.9	6.7

Results. As shown in Tab. 2, our default configuration (ViT hyper-network with a 2-hidden-layer neural functional map) achieves the lowest geodesic error. The significant performance gap between the ViT and the CNN/MLP backbones (settings (a) and (b)) highlights that global context is vital for parameter prediction. Furthermore, the results for settings (c), (d), and (e) suggest that a 2-layer MLP provides the ideal zone for expressivity; fewer layers lack the flexibility to model distortions, while deeper layers do not yield additional gains, likely due to the difficulty of optimizing highly non-linear spectral transformations.

6.2 Well-Posedness of Our Method

To evaluate whether our hyper-network can exploit priors over the training distribution to recover high-frequency non-linear residual from standard functional alone, we conduct the following ablative experiments, as shown in Tab. 3.

(a) Swap C : swap original C with another pair’s C . The high errors indicate that the hyper-network *does not* learn a constant dataset-specific correction.

(b) Rnd-init: randomly initialized hyper-network weights with per-pair test-time optimization. High errors indicate that per-pair test-time optimization does not work, confirming that the hyper-network learns a generalized mapping from linear FM to NFM from the *collective information* in the training dataset.

(c) Cross-dataset: train on CUTS, test on HOLES (and vice-versa). The results show that our hyper-network does not only learn dataset-specific bias but generalizes across datasets.

(d) C -corruption: degradation of C by Gaussian noise $\sigma = 0.1$ during inference on test datasets. The results demonstrate the robustness of our hyper-network and that it does not overfit to training data bias.

Results. Our method successfully learns generalizable priors from training data. By construction (Eq. (6)), our NFM reduces to C when the residual vanishes and operates entirely in the spectral domain. Overall, our hyper-network’s task is to enrich a coarse spectral alignment with a non-linear correction, conditioned on the distortion profile that C already encodes. To this end, our method can effectively exploit priors over the training distribution and generalize to other datasets.

Table 3: Qualitative results on SHREC16 dataset. Results demonstrate our method successfully learns generalizable priors from training data distribution.

Geo. error ($\times 100$)	Ours	Swap C	Rnd-init	Cross-dataset	C -corruption
CUTS	1.9	21.7	57.3	3.0	3.1
HOLES	4.7	32.8	61.9	5.7	6.2

7 Limitation and Future Work

Despite its robustness, our framework is limited by its fundamental reliance on spectral alignment. In cases of extreme geometric distortion where Laplace-Beltrami eigenfunctions become significantly degraded, the non-linear mapping may fail to bridge the resulting spectral gap. Future work could address this, for example by extending the framework to multi-shape matching through the enforcement of cycle-consistency as regularization. Specifically, learning a set of virtual universal spectral bases to serve as a canonical reference domain would allow the hyper-network to align an entire collection to a shared latent space, ensuring globally consistent correspondences for advanced joint shape analysis.

8 Conclusion

In this paper, we introduced a novel framework for 3D shape matching that moves beyond the limitations of standard functional maps by incorporating non-linear neural functional maps. By utilizing a ViT-based hyper-network, our method dynamically predicts the parameters of a non-linear functional map based on the standard functional map, providing the flexibility required to align spectral bases in the presence of significant distortions. Our extensive experiments across diverse benchmarks, including topologically noisy, partial, and multi-modal datasets, demonstrate that our approach consistently sets the new state of the art. Notably, our unsupervised framework often surpasses the accuracy of fully supervised alternatives, particularly in challenging scenarios where standard linear approximations collapse. By bridging the gap between spectral processing and deep non-linear architectures, this work provides a robust and versatile foundation for high-quality correspondence for 3D shape analysis and beyond. Ultimately, this paradigm shift toward adaptive spectral alignment paves the way for more robust matching algorithm capable of handling the inherent geometric inconsistencies of real-world data in fields such as robotics, medical imaging and digital twins.

9 Acknowledgment

This work is supported by the ERC starting grant no. 101160648 (Harmony).

References

1. Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J.: Scape: shape completion and animation of people. In: ACM SIGGRAPH (2005) [12](#)
2. Attaiki, S., Ovsjanikov, M.: Understanding and improving features learned in deep functional maps. In: CVPR (2023) [1](#), [4](#)
3. Attaiki, S., Pai, G., Ovsjanikov, M.: Dpfm: Deep partial functional maps. In: International Conference on 3D Vision (3DV) (2021) [1](#), [4](#), [11](#)
4. Aubry, M., Schlickewei, U., Cremers, D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In: ICCV (2011) [4](#)
5. Aygün, M., Lähner, Z., Cremers, D.: Unsupervised dense shape correspondence using heat kernels. In: Conference on 3D Vision (3DV) (2020) [4](#)
6. Bastian, L., Xie, Y., Navab, N., Lähner, Z.: Hybrid functional maps for crease-aware non-isometric shape matching. In: CVPR (2024) [1](#)
7. Bernard, F., Suri, Z.K., Theobalt, C.: Mina: Convex mixed-integer programming for non-rigid shape alignment. In: CVPR (2020) [4](#)
8. Bernard, F., Thunberg, J., Swoboda, P., Theobalt, C.: Hipp: Higher-order projected power iterations for scalable multi-matching. In: ICCV (2019) [5](#)
9. Bogo, F., Romero, J., Loper, M., Black, M.J.: Faust: Dataset and evaluation for 3d mesh registration. In: CVPR (2014) [11](#), [12](#)
10. Bracha, A., Dagès, T., Kimmel, R.: On unsupervised partial shape correspondence. ACCV (2024) [11](#)
11. Bracha, A., Dagès, T., Kimmel, R.: Wormhole loss for partial shape matching. In: NIPS (2024) [1](#)
12. Bronstein, A., Bronstein, M., Kimmel, R.: Numerical Geometry of Non-Rigid Shapes. Springer Publishing Company, Incorporated, 1 edn. (2008) [10](#)
13. Bronstein, M.M., Kokkinos, I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In: CVPR (2010) [4](#)
14. Cao, D., Bernard, F.: Unsupervised deep multi-shape matching. In: ECCV (2022) [1](#)
15. Cao, D., Bernard, F.: Self-supervised learning for multimodal non-rigid 3d shape matching. In: CVPR (2023) [1](#), [3](#), [9](#), [12](#), [13](#)
16. Cao, D., Roetzer, P., Bernard, F.: Unsupervised learning of robust spectral shape matching. ACM Transactions on Graphics (ToG) (2023) [1](#), [3](#), [4](#), [5](#), [9](#), [10](#), [11](#)
17. Cao, D., Roetzer, P., Bernard, F.: Revisiting map relations for unsupervised non-rigid shape matching. In: 2024 International Conference on 3D Vision (3DV). IEEE (2024) [1](#), [5](#)
18. Cosmo, L., Rodola, E., Bronstein, M.M., Torsello, A., Cremers, D., Sahillioglu, Y.: Shrec’16: Partial matching of deformable shapes. Proc. 3DOR **2**(9), 12 (2016) [10](#)
19. Deng, B., Yao, Y., Dyke, R.M., Zhang, J.: A survey of non-rigid 3d registration. In: Computer Graphics Forum. vol. 41, pp. 559–589. Wiley Online Library (2022) [1](#)
20. Dinh, H.Q., Yezzi, A., Turk, G.: Texture transfer during shape transformation. ACM Transactions on Graphics (ToG) **24**(2), 289–310 (2005) [1](#)
21. Donati, N., Corman, E., Ovsjanikov, M.: Deep orientation-aware functional maps: Tackling symmetry issues in shape matching. In: CVPR (2022) [1](#)
22. Donati, N., Sharma, A., Ovsjanikov, M.: Deep geometric functional maps: Robust feature learning for shape correspondence. In: CVPR (2020) [4](#), [11](#)
23. Dosovitskiy, A.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020) [8](#)

24. Dutt, N.S., Muralikrishnan, S., Mitra, N.J.: Diffusion 3d features (diff3f): Decorating untextured shapes with distilled semantic features. In: CVPR (2024) [11](#)
25. Egger, B., Smith, W.A., Tewari, A., Wuhler, S., Zollhoefer, M., Beeler, T., Bernard, F., Bolkart, T., Kortylewski, A., Romdhani, S., et al.: 3d morphable face models—past, present, and future. *ACM Transactions on Graphics (ToG)* **39**(5), 1–38 (2020) [1](#)
26. Ehm, V., Gao, M., Roetzer, P., Eisenberger, M., Cremers, D., Bernard, F.: Partial-to-partial shape matching with geometric consistency. In: CVPR (2024) [11](#)
27. Eisenberger, M., Lahner, Z., Cremers, D.: Smooth shells: Multi-scale shape registration with functional maps. In: CVPR (2020) [10](#)
28. Eisenberger, M., Novotny, D., Kerchenbaum, G., Labatut, P., Neverova, N., Cremers, D., Vedaldi, A.: Neuromorph: Unsupervised shape interpolation and correspondence in one go. In: CVPR (2021) [10](#)
29. Eisenberger, M., Toker, A., Leal-Taixé, L., Cremers, D.: Deep shells: Unsupervised shape correspondence with optimal transport. *NIPS* (2020) [10](#)
30. Eisenberger, M., Toker, A., Leal-Taixé, L., Cremers, D.: G-msm: Unsupervised multi-shape matching with graph-based affinity priors. In: CVPR (2023) [1](#), [9](#)
31. Ezuz, D., Ben-Chen, M.: Deblurring and denoising of maps between shapes. In: *Computer Graphics Forum*. Wiley Online Library (2017) [4](#)
32. Ezuz, D., Solomon, J., Ben-Chen, M.: Reversible harmonic maps between discrete surfaces. *ACM Transactions on Graphics (ToG)* **38**(2), 1–12 (2019) [4](#)
33. Gao, M., Lahner, Z., Thunberg, J., Cremers, D., Bernard, F.: Isometric multi-shape matching. In: CVPR (2021) [5](#)
34. Gao, M., Roetzer, P., Eisenberger, M., Lähner, Z., Moeller, M., Cremers, D., Bernard, F.: Sigma: Scale-invariant global sparse shape matching. In: ICCV (2023) [4](#)
35. Groueix, T., Fisher, M., Kim, V.G., Russell, B.C., Aubry, M.: 3d-coded: 3d correspondences by deep deformation. In: ECCV (2018) [13](#)
36. Halimi, O., Litany, O., Rodola, E., Bronstein, A.M., Kimmel, R.: Unsupervised learning of dense shape correspondence. In: CVPR (2019) [4](#), [10](#)
37. Holzschuh, B., Lähner, Z., Cremers, D.: Simulated annealing for 3d shape correspondence. In: 2020 International Conference on 3D Vision (3DV) (2020) [4](#)
38. Huang, Q., Wang, F., Guibas, L.: Functional map networks for analyzing and exploring large shape collections. *ACM Transactions on Graphics (ToG)* **33**(4), 1–11 (2014) [5](#)
39. Huang, R., Ren, J., Wonka, P., Ovsjanikov, M.: Consistent zoomout: Efficient spectral map synchronization. In: *Computer Graphics Forum*. Wiley Online Library (2020) [5](#)
40. Jiang, P., Sun, M., Huang, R.: Neural intrinsic embedding for non-rigid point cloud matching. In: CVPR (2023) [1](#), [12](#), [13](#)
41. Jiang, P., Sun, M., Huang, R.: Non-rigid shape registration via deep functional maps prior. In: NIPS (2023) [12](#), [13](#)
42. Kim, V.G., Lipman, Y., Funkhouser, T.: Blended intrinsic maps. *ACM Transactions on Graphics (ToG)* **30**(4), 1–12 (2011) [10](#)
43. Lähner, Z., Rodola, E., Bronstein, M.M., Cremers, D., Burghard, O., Cosmo, L., Dieckmann, A., Klein, R., Sahillioglu, Y.: Shrec’16: Matching of deformable shapes with topological noise. *Proc. 3DOR* **2**(10.2312) (2016) [9](#)
44. Lang, I., Ginzburg, D., Avidan, S., Raviv, D.: Dpc: Unsupervised deep point correspondence via cross and self construction. In: International Conference on 3D Vision (3DV). IEEE (2021) [13](#)

45. Lévy, B.: Laplace-beltrami eigenfunctions towards an algorithm that "understands" geometry. In: IEEE International Conference on Shape Modeling and Applications 2006 (SMI'06). pp. 13–13. IEEE (2006) [4](#)
46. Li, L., Attaiki, S., Ovsjanikov, M.: Srfeat: Learning locally accurate and globally consistent non-rigid shape correspondence. In: International Conference on 3D Vision (3DV). IEEE (2022) [4](#)
47. Li, L., Donati, N., Ovsjanikov, M.: Learning multi-resolution functional maps with spectral attention for robust shape matching. NIPS (2022) [1](#), [10](#)
48. Li, T., Bolkart, T., Black, M.J., Li, H., Romero, J.: Learning a model of facial shape and expression from 4D scans. ACM Transactions on Graphics (ToG) **36**(6), 194:1–194:17 (2017) [1](#)
49. Litany, O., Remez, T., Rodola, E., Bronstein, A., Bronstein, M.: Deep functional maps: Structured prediction for dense shape correspondence. In: ICCV (2017) [4](#)
50. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: Smpl: A skinned multi-person linear model. ACM Transactions on Graphics (ToG) **34**(6), 1–16 (2015) [1](#)
51. Marin, R., Rakotosaona, M.J., Melzi, S., Ovsjanikov, M.: Correspondence learning via linearly-invariant embedding. In: NeurIPS (2020) [13](#)
52. Melzi, S., Marin, R., Rodolà, E., Castellani, U., Ren, J., Poulénard, A., Wonka, P., Ovsjanikov, M.: Shrec 2019: Matching humans with different connectivity. In: Eurographics Workshop on 3D Object Retrieval (2019) [12](#)
53. Melzi, S., Ren, J., Rodolà, E., Sharma, A., Wonka, P., Ovsjanikov, M.: Zoomout: spectral upsampling for efficient shape correspondence. ACM Transactions on Graphics (ToG) **38**(6), 1–14 (2019) [4](#), [5](#), [8](#), [10](#)
54. Ovsjanikov, M., Ben-Chen, M., Solomon, J., Butscher, A., Guibas, L.: Functional maps: a flexible representation of maps between shapes. ACM Transactions on Graphics (ToG) **31**(4), 1–11 (2012) [1](#), [2](#), [4](#), [5](#)
55. Ovsjanikov, M., Mérigot, Q., Mémoli, F., Guibas, L.: One point isometric matching with the heat kernel. In: Computer Graphics Forum. Wiley Online Library (2010) [1](#)
56. Pai, G., Ren, J., Melzi, S., Wonka, P., Ovsjanikov, M.: Fast sinkhorn filters: Using matrix scaling for non-rigid shape correspondence with functional maps. In: CVPR (2021) [4](#)
57. Pierson, E., Li, L., Dai, A., Ovsjanikov, M.: Diffumatch: Category-agnostic spectral diffusion priors for robust non-rigid shape matching. In: ICCV (2025) [8](#)
58. Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. *Experimental mathematics* **2**(1), 15–36 (1993) [5](#)
59. Ren, J., Melzi, S., Wonka, P., Ovsjanikov, M.: Discrete optimization for shape matching. In: Computer Graphics Forum. Wiley Online Library (2021) [4](#), [10](#)
60. Rodolà, E., Cosmo, L., Bronstein, M.M., Torsello, A., Cremers, D.: Partial functional correspondence. In: Computer Graphics Forum. Wiley Online Library (2017) [8](#)
61. Rodolà, E., Moeller, M., Cremers, D.: Point-wise map recovery and refinement from functional correspondence. arXiv preprint arXiv:1506.05603 (2015) [4](#)
62. Rodola, E., Moeller, M., Cremers, D.: Regularized pointwise map recovery from functional correspondence. In: Computer Graphics Forum. Wiley Online Library (2017) [4](#)
63. Roetzer, P., Swoboda, P., Cremers, D., Bernard, F.: A scalable combinatorial solver for elastic geometrically consistent 3d shape matching. In: CVPR (2022) [4](#)
64. Roufousse, J.M., Sharma, A., Ovsjanikov, M.: Unsupervised deep learning for structured shape matching. In: ICCV (2019) [4](#), [5](#), [10](#)

65. Sahillioglu, Y.: Recent advances in shape correspondence. *The Visual Computer* **36**(8), 1705–1721 (2020) [1](#)
66. Salti, S., Tombari, F., Di Stefano, L.: Shot: Unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding* **125**, 251–264 (2014) [4](#)
67. Sharma, A., Ovsjanikov, M.: Weakly supervised deep functional maps for shape matching. *NIPS* (2020) [4](#), [10](#)
68. Sharp, N., Attaiki, S., Crane, K., Ovsjanikov, M.: Diffusionnet: Discretization agnostic learning on surfaces. *arXiv preprint arXiv:2012.00888* (2020) [5](#)
69. Siméoni, O., Vo, H.V., Seitzer, M., Baldassarre, F., Oquab, M., Jose, C., Khali-dov, V., Szafraniec, M., Yi, S., Ramamonjisoa, M., et al.: Dinov3. *arXiv preprint arXiv:2508.10104* (2025) [11](#)
70. Song, C., Wei, J., Li, R., Liu, F., Lin, G.: 3d pose transfer with correspondence learning and mesh refinement. *NIPS* (2021) [1](#)
71. Song, C., Wei, J., Li, R., Liu, F., Lin, G.: Unsupervised 3d pose transfer with cross consistency and dual reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023) [1](#)
72. Sun, M., Mao, S., Jiang, P., Ovsjanikov, M., Huang, R.: Spatially and spectrally consistent deep functional maps. In: *ICCV* (2023) [1](#), [4](#)
73. Tam, G.K., Cheng, Z.Q., Lai, Y.K., Langbein, F.C., Liu, Y., Marshall, D., Martin, R.R., Sun, X.F., Rosin, P.L.: Registration of 3d point clouds and meshes: A survey from rigid to nonrigid. *IEEE transactions on visualization and computer graphics* **19**(7), 1199–1217 (2012) [3](#)
74. Trappolini, G., Cosmo, L., Moschella, L., Marin, R., Melzi, S., Rodolà, E.: Shape registration in the time of transformers. *NIPS* (2021) [13](#)
75. Van Kaick, O., Zhang, H., Hamarneh, G., Cohen-Or, D.: A survey on shape correspondence. In: *Computer Graphics Forum*. Wiley Online Library (2011) [3](#)
76. Vestner, M., Litman, R., Rodola, E., Bronstein, A., Cremers, D.: Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space. In: *CVPR* (2017) [4](#)
77. Viganò, G., Ovsjanikov, M., Melzi, S.: Nam: Neural adjoint maps for refining shape correspondences. *ACM Transactions on Graphics (ToG)* **44**(4), 1–15 (2025) [2](#), [5](#), [7](#), [8](#)
78. Wang, F., Huang, Q., Guibas, L.J.: Image co-segmentation via consistent functional maps. In: *ICCV* (2013) [5](#)
79. Wang, W., Weißberg, T., El Amrani, N., Bernard, F.: kh: Symmetry understanding of 3d shapes via chirality disentanglement. In: *ICCV* (2025) [11](#)
80. Windheuser, T., Schlickewei, U., Schmidt, F.R., Cremers, D.: Geometrically consistent elastic matching of 3d shapes: A linear programming solution. In: *ICCV* (2011) [4](#)
81. Xie, Y., Ehm, V., Roetzer, P., El Amrani, N., Gao, M., Bernard, F., Cremers, D.: Echomatch: Partial-to-partial shape matching via correspondence reflection. In: *CVPR* (2025) [1](#)
82. Zeng, Y., Qian, Y., Zhu, Z., Hou, J., Yuan, H., He, Y.: Corrnnet3d: Unsupervised end-to-end learning of dense correspondence for 3d point clouds. In: *CVPR* (2021) [13](#)
83. Zhuravlev, A., Lähner, Z., Golyanik, V.: Denoising functional maps: Diffusion models for shape correspondence. In: *CVPR* (2025) [8](#)