Coarse-to-Fine: A Hybrid Self-Supervised Method for Non-rigid 3D Shape Matching

Feifan Luo[1,2], Ting Li[3], Zhao Li[2], and Hongyang Chen[2]*

[1] Zhejiang University, China
[2] Zhejiang Lab, China
[3] Wenzhou University of Technology, China
luoff@zju.edu.cn dr.h.chen@ieee.org

Abstract. Non-rigid 3D shape matching is a fundamental task in computer vision and graphics. In this paper, we propose a hybrid self-supervised method based on a coarse-to-fine strategy, which ensures consistency between the coarse mapping and the refined correspondence produced by our refinement module. The architecture features a dual-branch design, consisting of two symmetric functional map learning streams: one based on the Laplacian basis and the other utilizing the elastic basis. Extensive experiments show that our approach not only maintains computational efficiency, but also achieves state-of-the-art performance across a variety of challenging scenarios, including non-isometric deformations and topological noise. Finally, we rigorously demonstrate that contrastive energies promote feature discrimination. Furthermore, integrating these energies with existing methods yields consistent improvements, validating the overall efficacy of our approach. Our code is available at https://github.com/LuoFeifan77/Coarse-to-Fine-Hybrid-Self-Supervised-Matching.

Keywords: 3D shape matching · Functional maps · Self-supervised learning · Coarse-to-fine

1 Introduction

Establishing accurate correspondences between non-rigid 3D shapes is a long-standing challenge in computer vision and graphics, which has a wide range of applications, including shape analysis [45], shape generation [21], deformation transfer [67], and shape interpolation [25].

Deep functional maps [42] have emerged as a preferred solution for non-rigid shape matching, paving the way for significant advancements in near-isometric shape matching [5, 20, 30, 62], non-isometric shape matching [19, 25, 26, 34, 39], multi-shape matching [11, 27, 69], and partial shape matching [3, 22, 23].

However, these methods [5, 19, 20, 34, 39, 62] primarily rely on geometric prior constraints (e.g., orthogonality) to supervise pointwise or functional maps, while overlooking self-supervised prior constraints, leading to suboptimal performance.

* Corresponding author.

Although incorporating the coupling loss [15] to promote functional map properness [58] enhances performance, it inevitably incurs substantial computational overhead. This is because these methods either necessitate solving linear optimization problems to estimate functional maps [2, 6, 15, 19, 48, 69] or rely on fine-tuning techniques [6, 13, 15, 38] during inference to bolster results.

To address these limitations, we propose a novel and efficient hybrid self-supervised method for non-rigid shape matching. We first introduce *a coarse-to-fine framework* for functional map learning. The core idea is to align the coarse map with the refined map, and it consists of two main components: (1) the proposal of three different self-supervised prior constraints, which facilitate the alignment of the coarse map and the refined map from spatial, spectral, and spatial-spectral perspectives; (2) the introduction of a general refinement method, which can be applied not only to generate both refined pointwise and functional maps but also to non-orthogonal settings. Based on the aforementioned framework, *a hybrid self-supervised method* is composed of two fully symmetric branches: the intrinsic branch, which employs a functional map learning module based on the orthonormal Laplacian basis, and the extrinsic branch, which uses a functional map learning module based on the non-orthogonal elastic basis, while eliminating the time-consuming least-squares solvers and fine-tuning techniques. Extensive experiments demonstrate that our method outperforms existing state-of-the-art methods in both matching accuracy and runtime. Finally, we rigorously prove that contrastive energies promote feature discrimination and lead to consistent improvements when integrated with other methods, thereby validating the effectiveness of our proposed framework.

The main contributions are summarized as follows:

- We propose a general coarse-to-fine strategy, which establishes a unified bridge between functional map learning frameworks and map refinement methods.
- We present a robust and efficient hybrid self-supervised method for non-rigid 3D shape matching.
- Extensive experiments demonstrate that our approach achieves excellent performance across challenging benchmarks, while maintaining remarkable computational efficiency.
- We provide a rigorous proof that contrastive energies induce feature discrimination. Moreover, our proposed energies deliver consistent performance gains when integrated into baseline methods.

2 Related Work

Non-rigid 3D shape matching is a basic problem that has been extensively studied. For a comprehensive overview of the field, readers are referred to surveys such as [17, 44, 63, 78].

2.1 Axiomatic Functional Maps

Numerous axiomatic approaches [1,7,10,54,70,74] have been proposed to address the shape matching problem. Among these, the functional map framework [54] has gained significant prominence due to its inherent flexibility and efficiency, which has inspired extensive subsequent works [18,28,29,32,33,35,49,51,53,55,59, 60,71]. However, axiomatic functional map methods rely heavily on the quality of handcrafted features [4,40,43,64,68], resulting in unsatisfactory performance in challenging settings.

2.2 Deep Functional Maps

FMNet [42], also known as deep functional maps, effectively overcomes the bottlenecks of axiomatic functional map approaches [54] by learning local features via neural networks. Subsequent approaches [5,20,30] followed this principle, constraining pointwise or functional maps to enhance feature representation [62,65], where the geometric prior constraints like bijectivity and orthogonality [59] have been widely adopted. Recently, the remarkable success of functional maps in ensuring properness [15,69] has brought dual-branch functional map learning architectures [2,6,13,16,38,48] and a single-branch self-supervised network [50]. Moreover, approaches based on hybrid bases [6,41], diffusion models [56,79], neural adjoint maps [71], unsupervised contrastive learning [47], basis learning [46], and other studies [8,9,61,72,75,77] have gradually been applied to the field of shape matching.

Despite advancements, existing methods either demonstrate poor matching performance [5,19,20,34,39,62,65] in challenging scenarios, such as non-isometric deformations and cross-dataset generalization, or rely on time-consuming linear system solvers [2,6,11–16,19,20,39,69], or fine-tuning techniques [6,13–16]. In contrast, we propose an effective and efficient hybrid self-supervised method for non-rigid shape matching, which outperforms these methods in both matching accuracy and computational efficiency. Furthermore, we demonstrate that the self-supervised losses introduced in previous works [16,50] are special cases of the contrastive energy proposed in our study. Finally, the contrastive energies introduced in our method consistently improve the performance of other functional map learning techniques, such as ULRSSM [15].

3 Background

We begin by providing a brief overview of the basic pipeline for deep functional maps, directing interested readers to relevant literature [6,15,54,78] for further details. The notation used throughout the paper is summarized in Table 1.

3.1 Functional Map Learning Pipeline

Given a pair of non-rigid 3D shapes denoted as $\mathcal{X}$ and $\mathcal{Y}$, the common functional map learning frameworks mainly consist of five main stages:

Table 1: Summary of the symbol description.

Symbol	Description
$\mathcal{X}, \mathcal{Y}$	3D shapes (triangle mesh) with $\lvert V_{\mathcal{X}}\rvert$, $\lvert V_{\mathcal{Y}}\rvert$ vertices
$\Lambda_{\mathcal{X}}$	$\mathbb{R}^{k\times k}$ eigenvalue matrix of shape $\mathcal{X}$
$\Phi_{\mathcal{X}}$	$\mathbb{R}^{\lvert V_{\mathcal{X}}\rvert\times k}$ eigenfunctions matrix
$\Phi_{\mathcal{X}}^{\dagger}$	$\mathbb{R}^{k\times\lvert V_{\mathcal{X}}\rvert}$ Moore-Penrose inverse of $\Phi_{\mathcal{X}}$
$A_{\mathcal{X}}$	$\mathbb{R}^{\lvert V_{\mathcal{X}}\rvert\times\lvert V_{\mathcal{X}}\rvert}$ mass matrix.
$A_{\mathcal{X},k}$	$\mathbb{R}^{k\times k}$ reduced mass matrix.
$F_{\mathcal{X}}$	$\mathbb{R}^{\lvert V_{\mathcal{X}}\rvert\times d}$ vertex-wise features of shape $\mathcal{X}$, d denotes the feature dimension
$\Pi_{\mathcal{Y}\mathcal{X}}$	$\mathbb{R}^{\lvert V_{\mathcal{Y}}\rvert\times\lvert V_{\mathcal{X}}\rvert}$ point-wise map between shapes $\mathcal{Y}$ and $\mathcal{X}$
$C_{\mathcal{X}\mathcal{Y}}$	$\mathbb{R}^{k\times k}$ (spectral) functional maps matrix obtained by solving a linear system Eq.(1)
$\hat{C}_{\mathcal{X}\mathcal{Y}}$	$\mathbb{R}^{k\times k}$ spatial functional maps matrix obtained from $\Pi_{\mathcal{Y}\mathcal{X}}$
$\lvert\lvert\cdot\rvert\rvert_{\mathrm{F}}$	Frobenius norm for Laplacian eigensystem
$\lvert\lvert\cdot\rvert\rvert_{\mathrm{HS}}$	Hilbert–Schmidt norm for elastic eigensystem.
$\{g_s(\cdot)\}_{s=1}^{S}$	a set of filter functions

(1) **Precomputation.** Compute the first k eigenfunctions $\Phi_{\mathcal{X}}, \Phi_{\mathcal{Y}}$, eigenvalues $\Lambda_{\mathcal{X}}, \Lambda_{\mathcal{Y}}$ in matrix notation via generalized eigendecomposition, respectively. For example, the orthonormal Laplacian eigensystem $\{\Phi_{\mathcal{X}}^{LB}, \Lambda_{\mathcal{X}}^{LB}\}$ and the non-orthogonal elastic eigensystem $\{\Phi_{\mathcal{X}}^{EL}, \Lambda_{\mathcal{X}}^{EL}\}$ can be obtained by decomposing the Laplacian operator [57] and elastic thin-shell energy [73].

(2) **Feature extraction.** The vertex features $F_{\mathcal{X}}$ and $F_{\mathcal{Y}}$ are extracted by a feature extractor network $\mathcal{F}_{\Theta}$, respectively, with shared parameters Θ. Among numerous feature extraction methods, DiffusionNet [66] remains the predominant choice for the majority of functional map learning approaches.

(3) **Functional map computation.** Functional maps have two computation methods. The first is the *definition-based* approach [54], which is obtained by solving a linear optimization problem, specifically:

$$C_{\mathcal{X}\mathcal{Y}} = \underset{C_{\mathcal{X}\mathcal{Y}}}{\arg\min} \left\| C_{\mathcal{X}\mathcal{Y}} \Phi_{\mathcal{X}}^{\dagger} F_{\mathcal{X}} - \Phi_{\mathcal{Y}}^{\dagger} F_{\mathcal{Y}} \right\| + \mu E_{reg}(C_{\mathcal{X}\mathcal{Y}}), \tag{1}$$

where $E_{reg} = \|C_{\mathcal{X}\mathcal{Y}}\Lambda_{\mathcal{X}} - \Lambda_{\mathcal{Y}} C_{\mathcal{X}\mathcal{Y}}\|$ denotes the regularization term, μ is a hyperparameter. Here, the choice of norm is determined by the type of basis functions Φ employed. For the orthonormal Laplacian eigensystem, Eq. (1) is computed using the Frobenius norm $||\cdot||_{\mathrm{F}}$, whereas for the non-orthogonal elastic eigensystem, the Hilbert–Schmidt norm $||\cdot||_{\mathrm{HS}}$ is adopted [6].

The second is the *projection-based* approach [54], where the pointwise map is projected onto the spectral domain, expressed as:

$$\hat{C}_{\mathcal{X}\mathcal{Y}} = \Phi_{\mathcal{Y}}^{\dagger} \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}, \tag{2}$$

where the pointwise map satisfies $F_{\mathcal{Y}} = \Pi_{\mathcal{Y}\mathcal{X}} F_{\mathcal{X}}$.

To avoid ambiguity, the functional map $C_{\mathcal{X}\mathcal{Y}}$ derived from Eq.(1) is termed the (spectral) functional map, while that explicitly computed via pointwise map, i.e., Eq. (2), is designated as the *spatial* functional map [69, 76] (or the proper functional map in some references [50, 56, 58]).

(4) **Loss functions.** During the training phase, axiomatic prior constraints, such as bijectivity and orthogonality [59], are typically applied as structural

penalties, namely,

$$L_{bi} = \|C_{\mathcal{X}\mathcal{Y}}C_{\mathcal{Y}\mathcal{X}} - I\|, L_{or} = \left\|C_{\mathcal{X}\mathcal{Y}}C_{\mathcal{X}\mathcal{Y}}^{\mathrm{T}} - I\right\|, \tag{3}$$

or the coupling [15] between spatial and spectral functional maps, i.e.,

$$L_{co} = \left\|C_{\mathcal{X}\mathcal{Y}} - \hat{C}_{\mathcal{X}\mathcal{Y}}\right\|. \tag{4}$$

Additionally, the aforementioned loss terms can be extended to bidirectional constraints by incorporating reverse mapping terms.

However, most existing functional map learning methods [5,19,39,79] mostly rely on geometric prior constraints while ignoring self-supervised prior constraints, leading to subpar performance in challenging scenarios, such as non-isometric deformations and matching with topological noise. On the other hand, they rely on time-consuming computational modules, such as the linear system solver [2,20] in Eq. (1) and fine-tuning techniques [6,13,15], which significantly increase the computational cost. In contrast, we propose a novel, efficient, and effective hybrid self-supervised method that achieves outstanding performance in both matching accuracy and computational cost.

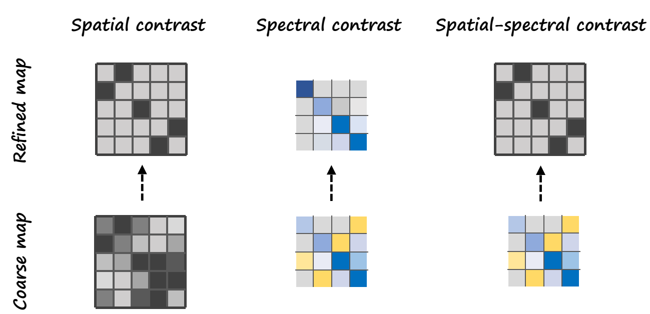

Fig. 1: Comparison of the three coarse-to-fine contrastive energy terms. From left to right: spatial contrastive energy, spectral contrastive energy, and spatial-spectral contrastive energy.

4 Coarse-to-Fine Guided Functional Maps

This section introduces the coarse-to-fine strategy tailored for the functional map framework, which aims to align the coarse map more closely with the refined map. We begin by defining three types of self-supervised prior constraint terms: spatial, spectral, and spatial-spectral contrastive energies, each of which promotes the alignment of the coarse and refined maps from different directions in Section 4.1. We then present a general refinement method that generates refined functional and pointwise maps, applicable to both orthogonal and non-orthogonal basis functions in Section 4.2.

4.1 Contrastive Energy

The essence of the coarse-to-fine strategy lies in enforcing the coarse maps to converge toward the refined maps, where the latter serves as a high-fidelity guidance. To this end, we formulate a series of contrastive energy constraints based on this coarse-to-fine hierarchy, see Fig. 1.

Spatial contrast. Given a coarse pointwise map $\Pi_{\mathcal{Y}\mathcal{X}}$, the pointwise map contrastive energy is defined as:

$$\|\Pi_{\mathcal{Y}\mathcal{X}} - \Pi_{\mathcal{Y}\mathcal{X}}^{ref}\|, \tag{5}$$

where $\Pi_{\mathcal{Y}\mathcal{X}}^{ref}$ represents the refined pointwise map. Since the refined pointwise maps $\Pi_{\mathcal{Y}\mathcal{X}}^{ref}$ encode higher-quality correspondences, they can be regarded as a proxy for ground truth, guiding the optimization of the coarse pointwise maps $\Pi_{\mathcal{Y}\mathcal{X}}$.

If the pointwise map represents a self-mapping relationship within the shape itself, denoted as $\Pi_{\mathcal{Y}\mathcal{Y}}$, then the refined map $\Pi_{\mathcal{Y}\mathcal{Y}}^{ref}$ reduces to the identity mapping $I_{\mathcal{Y}} \in \mathbb{R}^{|V_{\mathcal{Y}}|\times|V_{\mathcal{Y}}|}$ due to the self-evident correspondence of vertex features, leading to

$$\|\Pi_{\mathcal{Y}\mathcal{Y}} - I_{\mathcal{Y}}\|. \tag{6}$$

Interestingly, the self-supervised constraints introduced in earlier works [16, 50] are special cases of our spatial contrastive energies, with additional details provided in the supplementary materials.

Spectral contrast. Parallel to the spatial contrastive regularization, we incorporate a coarse-to-fine representation for the spectral functional map $C_{\mathcal{X}\mathcal{Y}}$, denoted as

$$\|C_{\mathcal{X}\mathcal{Y}} - C_{\mathcal{X}\mathcal{Y}}^{ref}\|. \tag{7}$$

Unlike the self spatial contrastive energy Eq. (6), since spectral functional maps are computed from a linear system Eq. (1) via the least-square method, the resulting $C_{\mathcal{Y}\mathcal{Y}}$ is necessarily equal to I. Consequently, we omit this trivial contrastive constraint.

Spatial-spectral contrast. The coupling constraint Eq. (4), which promotes properness [15,69], has proven highly effective in challenging shape matching tasks. However, it specifically aligns the coarse functional maps with the coarse pointwise maps, thereby ensuring what we term as *coarse properness*. In contrast, we propose a novel spatial-spectral contrastive energy to promote *fine properness*, namely,

$$\left\|C_{\mathcal{X}\mathcal{Y}} - \Phi_{\mathcal{Y}}^{\dagger}\Pi_{\mathcal{Y}\mathcal{X}}^{ref}\Phi_{\mathcal{X}}\right\|. \tag{8}$$

The above equation can be interpreted as aligning spectral functional maps with refined pointwise maps.

On the other hand, we also leverage the refined spectral functional map $C_{\mathcal{X}\mathcal{Y}}^{ref}$ to facilitate the optimization of the coarse pointwise maps, i.e.,

$$\left\|\Pi_{\mathcal{Y}\mathcal{X}} - \Phi_{\mathcal{Y}}C_{\mathcal{X}\mathcal{Y}}^{ref}\Phi_{\mathcal{X}}^{\dagger}\right\|. \tag{9}$$

Our coarse-to-fine strategy establishes a unified bridge between deep functional map frameworks and map refinement methods, by employing the aforementioned contrastive energies as self-supervised losses to enhance map quality. On the other hand, this integration allows various established axiomatic refinement techniques [33, 53, 60, 70] to generate refined maps.

4.2 Map Refinement

Despite the proposal of numerous axiomatic map refinement techniques [32, 53], they lack a unified framework capable of refining both pointwise and functional maps simultaneously. Inspired by [48], we introduce a more generalized refinement paradigm that is equally effective for both pointwise and functional map representations, even under non-orthogonal basis functions.

Functional map refinement. We extend the refinement method proposed in [48]—specifically, the general multi-scale Laplacian commutativity term—to the context of non-orthogonal basis functions. For additional background and detailed information, we strongly recommend that readers refer to the original paper. The general multi-scale Laplacian commutativity term is as follows:

$$E(C_{\mathcal{X}\mathcal{Y}}) = \sum_{s=1}^{S} \|C_{\mathcal{X}\mathcal{Y}} g_s(\Lambda_{\mathcal{X}}) - g_s(\Lambda_{\mathcal{Y}}) C_{\mathcal{X}\mathcal{Y}}\|_{\mathrm{HS}}^2, \tag{10}$$

where $\{g_s\}_{s=1}^{S}$ denote the filter functions.

Although one could compute this following the approach for solving standard Laplacian commutativity in the previous study [6], such a computation is complex and incurs significant computational overhead. Instead, we recommend adopting a strategy similar to the prior work [48]. Specifically, we assume that one of the functional maps is already obtained (denoted as $C_{\mathcal{X}\mathcal{Y}}^{init}$), then,

$$E(C_{\mathcal{X}\mathcal{Y}}) = \sum_{s=1}^{S} \|C_{\mathcal{X}\mathcal{Y}} g_s(\Lambda_{\mathcal{X}}) - g_s(\Lambda_{\mathcal{Y}}) C_{\mathcal{X}\mathcal{Y}}^{init}\|_{\mathrm{HS}}^2. \tag{11}$$

Consequently, we only need to solve for the other functional map.

Lemma 1 *The regularization term $E(C_{\mathcal{X}\mathcal{Y}})$ can be formulated as:*

$$C_{\mathcal{X}\mathcal{Y}}^{ref} = \arg\min_{C_{\mathcal{X}\mathcal{Y}}} \sum_{s=1}^{S} \|\sqrt{A_{\mathcal{Y},k}}\left(C_{\mathcal{X}\mathcal{Y}} g_s(\Lambda_{\mathcal{X}}) - g_s(\Lambda_{\mathcal{Y}}) C_{\mathcal{X}\mathcal{Y}}^{init}\right)\sqrt{A_{\mathcal{X},k}^{-1}}\|_{\mathrm{F}}^2,$$

where $A_{\mathcal{X},k} = \Phi_{\mathcal{X}}^{\mathrm{T}} A_{\mathcal{X}} \Phi_{\mathcal{X}}$.

Moreover, if the set of filter functions $\{g_s(\lambda)\}_{s=1}^{S}$ satisfy Parseval condition that $\sum_s g_s^2(\lambda) \equiv 1, \forall \lambda$, then the functional map in Eq. (11) can be obtained via

$$C_{\mathcal{X}\mathcal{Y}}^{ref} = \sum_s g_s(\Lambda_{\mathcal{Y}}) C_{\mathcal{X}\mathcal{Y}}^{init} g_s(\Lambda_{\mathcal{X}}). \tag{12}$$

Proof. The first statement expresses the HS-norm in terms of the Frobenius norm. For the second claim, the formulation of Eq. (12) is analogous to that in reference [48]. Detailed proofs are provided in supplementary materials.

Clearly, Eq. (12) can be interpreted as refining the coarse functional map $C_{\mathcal{X}\mathcal{Y}}^{init}$ using multiple sets of filters, analogous to signal denoising.

Pointwise map refinement. Similarly, we employ the same strategy to generate refined pointwise maps. First, we compute the spatial functional map $\hat{C}_{\mathcal{X}\mathcal{Y}} = \Phi_{\mathcal{Y}}^{\dagger}\Pi_{\mathcal{Y}\mathcal{X}}\Phi_{\mathcal{X}}$ and then refine it via Eq. (12), yielding the refined spatial functional map $\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref}$. Then, we utilize a common strategy originating from ZoomOut [53], which converts the refined spatial functional map $\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref}$ into a refined pointwise map by minimizing the following energy:

$$\Pi_{\mathcal{Y}\mathcal{X}}^{ref} = \arg\min_{\Pi_{\mathcal{Y}\mathcal{X}}} \left\| \Phi_{\mathcal{Y}} - \Pi_{\mathcal{Y}\mathcal{X}}\Phi_{\mathcal{X}}(\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref})^{*} \right\|_{\mathrm{HS}}, \tag{13}$$

where $(\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref})^{*} = A_{\mathcal{X},k}^{-1}(\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref})^{\mathrm{T}}A_{\mathcal{Y},k}$ denotes the *adjoint operator* of $\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref}$. The above equation can be solved via nearest neighbor search [32], i.e.,

$$\Pi_{\mathcal{Y}\mathcal{X}}^{ref} = NN\left((A_{\mathcal{Y},k}^{1/2}\Phi_{\mathcal{Y}}^{\dagger}A_{\mathcal{Y}}^{-1})^{\mathrm{T}}, (A_{\mathcal{Y},k}^{1/2}\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref}\Phi_{\mathcal{X}}^{\dagger}A_{\mathcal{X}}^{-1})^{\mathrm{T}}\right). \tag{14}$$

In the orthonormal setting, since $A_{\mathcal{X},k} = I$ and $A_{\mathcal{Y},k} = I$, the adjoint operator of $\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref}$ reduces to its transpose, i.e., $(\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref})^{*} = (\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref})^{\mathrm{T}}$. Additionally, with $\Phi_{\mathcal{Y}}^{\dagger}A_{\mathcal{Y}}^{-1} = \Phi_{\mathcal{Y}}^{\mathrm{T}}$, Eq. (14) is simplified to

$$\Pi_{\mathcal{Y}\mathcal{X}}^{ref} = NN(\Phi_{\mathcal{Y}}, \Phi_{\mathcal{X}}(\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref})^{\mathrm{T}}). \tag{15}$$

Algorithm 1 Map refinement

Input: $C_{\mathcal{X}\mathcal{Y}}$, $\Pi_{\mathcal{Y}\mathcal{X}}$, $\{g_s(\Lambda_{\mathcal{X}})\}_{s=1}^{S}$, $\{g_s(\Lambda_{\mathcal{Y}})\}_{s=1}^{S}$
Output: $C_{\mathcal{X}\mathcal{Y}}^{ref}$ and $\Pi_{\mathcal{Y}\mathcal{X}}^{ref}$
L_2-normalization ensures $\sum_s g_s^2(\Lambda_{\mathcal{X}}) \equiv 1$ and $\sum_s g_s^2(\Lambda_{\mathcal{Y}}) \equiv 1$.
Estimate refined functional map: $C_{\mathcal{X}\mathcal{Y}}^{ref} = \sum_s g_s(\Lambda_{\mathcal{Y}})\, C_{\mathcal{X}\mathcal{Y}}\, g_s(\Lambda_{\mathcal{X}})$.
Estimate refined pointwise map: (1): $\hat{C}_{\mathcal{X}\mathcal{Y}}^{ref} = \sum_s g_s(\Lambda_{\mathcal{Y}})\, \hat{C}_{\mathcal{X}\mathcal{Y}}\, g_s(\Lambda_{\mathcal{X}})$ with $\hat{C}_{\mathcal{X}\mathcal{Y}} = \Phi_{\mathcal{Y}}^{\dagger}\Pi_{\mathcal{Y}\mathcal{X}}\Phi_{\mathcal{X}}$. (2) Recover $\Pi_{\mathcal{Y}\mathcal{X}}^{ref}$ from Eq. (13).

Filter normalization. If the set of filters does not satisfy the Parseval framework [31] (i.e., $\sum_s g_s^2(\lambda) \neq 1, \exists\lambda$), the refinement process necessitates matrix inversion [48]. To circumvent numerical instability arising from matrix inversion, we impose an L_2 norm constraint on the filters $\{g_s(\lambda)\}_{s=1}^{S}$ to ensure they satisfy the Parseval condition. This approach not only facilitates computation but also enhances numerical robustness.

We summarize the generalized refinement strategy for both functional and pointwise maps in Algorithm 1.

5 Hybrid Self-Supervised Learning Shape Matching

In this section, we propose a novel and efficient hybrid self-supervised learning method for non-rigid 3D shape matching, a dual-branch framework that bypasses the computationally expensive least-squares solver. This framework consists of an *intrinsic* branch, which constructs functional maps using the Laplacian basis in Section 5.3, and an *extrinsic* branch, which utilizes an Elastic basis for functional map learning in Section 5.4. Each branch incorporates two distinct computational pipelines: a refinement stream and unsupervised losses. The network architecture is illustrated in Fig. 2.

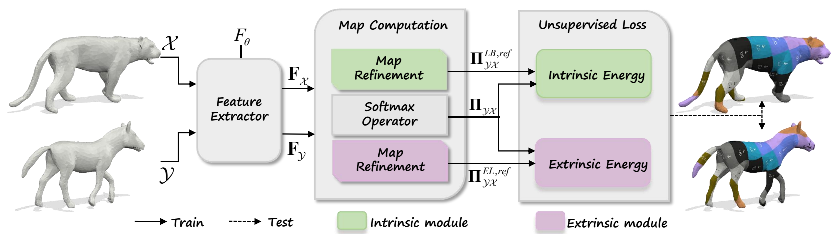

Fig. 2: An overview of our method. (1) Feature extraction: Utilize DiffusionNet [66] to extract vertex representations $F_\mathcal{X}$ and $F_\mathcal{Y}$ from $\mathcal{X}$ and $\mathcal{Y}$, respectively. (2) Map computation: Employ the softmax operator and the refinement Algorithm 1 to generate the coarse soft map $\Pi_{\mathcal{YX}}$ and the refined hard maps $\Pi_{\mathcal{YX}}^{LB,ref}$ and $\Pi_{\mathcal{YX}}^{EL,ref}$, respectively. (3) Unsupervised loss: Apply intrinsic and extrinsic energies (Eq. (22) and Eq. (31)) to supervise the network learning. (4) Map recovery. Use Eq. (17) and Eq. (32) to recover the pointwise maps during inference.

5.1 Eliminating Least-Squares Optimization

Solving the functional map optimization problem via least-squares, as shown in Eq. (1), has become a standard component in many state-of-the-art methods [6, 11, 13–16, 19, 38, 39, 48]. However, the least-squares solver introduces significant computational overhead. For instance, HybridFMaps [6] requires solving a dense $k^2 \times k^2$ linear system (where k is the number of eigenvectors, typically between 100 and 200); detailed runtimes are provided in supplementary materials. Furthermore, least-squares solvers can suffer from numerical instability and require higher-dimensional feature inputs [50]. Consequently, although we have designed a series of regularization terms Eq. (7), Eq. (8), and Eq. (9) for spectral functional maps that could potentially enhance performance, we deliberately bypass these potential gains in favor of computational efficiency. Instead, we establish a highly efficient shape matching framework based *solely* on pointwise map constraints.

5.2 Coarse Map

The coarse map refers to the initial mapping generated prior to any refinement operations. To ensure both differentiability and computational efficiency, after extracting shape features $F_{\mathcal{X}}$ and $F_{\mathcal{Y}}$, we follow several state-of-the-art methods [15, 25, 48] by employing a softmax operator to compute the differentiable pointwise map matrix, defined as:

$$\Pi_{\mathcal{Y}\mathcal{X}} = \mathrm{Softmax}(F_{\mathcal{Y}} F_{\mathcal{X}}^{\mathrm{T}} / \tau), \tag{16}$$

where the element at position (j, i) represents the probability of correspondence between the j-th point on $\mathcal{Y}$ and the i-th point on $\mathcal{X}$, and τ is the temperature factor to determine the softness of the matrix.

5.3 Intrinsic Branch

The Laplacian basis encodes essential intrinsic spectral information and has been widely adopted in functional map-based methods. Motivated by its efficacy, we utilize the spectral basis to construct the intrinsic self-supervised learning branch of our framework.

Non-differentiable refined map. The refined map is designed to provide optimization guidance for the coarse soft map, therefore it does not strictly require differentiability. On the other hand, an ideal soft correspondence matrix $\Pi_{\mathcal{Y}\mathcal{X}}$ should approximate a binary correspondence matrix. This implies achieving optimal discriminativity, where the mapping is concentrated on the most similar candidate points while maintaining maximum separation from the remaining low-similarity or dissimilar points.

We first compute the initial binary pointwise map matrix based on a nearest neighbor search, i.e.,

$$\Pi_{\mathcal{Y}\mathcal{X}}^{init} = NN(F_{\mathcal{Y}}, F_{\mathcal{X}}), \tag{17}$$

where $\Pi_{\mathcal{Y}\mathcal{X}}^{init} \in \{0,1\}^{|V_{\mathcal{Y}}| \times |V_{\mathcal{X}}|}$. Subsequently, we apply the Algorithm 1 to derive the refined binary pointwise map $\Pi_{\mathcal{Y}\mathcal{X}}^{LB,ref}$.

Unsupervised loss. Unsupervised loss functions are essential for learning functional maps. Our unsupervised loss design consists of two components: the widely used geometric prior constraints [58, 59] and the self-supervised constraints proposed previously.

Geometric prior constraints impose structural penalties on the mapping, such as bijectivity and orthogonality. However, these constraints fail to penalize the image of $\Pi_{\mathcal{Y}\mathcal{X}}$ that lies outside the span of $\Phi_{\mathcal{Y}}$. Inspired by [58], we employ variants of orthogonality and bijectivity penalties to constrain the coarse map $\Pi_{\mathcal{Y}\mathcal{X}}$, formulated as:

$$L_{or}^{LB} = \left\| \Phi_{\mathcal{Y}}^{LB} - \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{LB} (\hat{C}_{\mathcal{X}\mathcal{Y}}^{LB})^{\mathrm{T}} \right\|_{\mathrm{F}} \tag{18}$$

and

$$L_{bi}^{LB} = \left\| \Phi_{\mathcal{Y}}^{LB} - \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{LB} \hat{C}_{\mathcal{Y}\mathcal{X}}^{LB} \right\|_{\mathrm{F}}, \tag{19}$$

respectively, where $\Phi_{\mathcal{X}}^{LB} \in \mathbb{R}^{|V_{\mathcal{X}}| \times k^{LB}}$.

However, directly imposing these constraints on coarse maps would lead to a significant numerical discrepancy between different losses, as the orthogonality Eq. (18) and bijectivity terms Eq. (19) are defined in $\mathbb{R}^{N_{\mathcal{Y}} \times k}$ while the spatial contrastive losses Eq. (5) and Eq. (6) reside in $\mathbb{R}^{N_{\mathcal{Y}} \times N_{\mathcal{X}}}$ and $\mathbb{R}^{N_{\mathcal{Y}} \times N_{\mathcal{Y}}}$, respectively. This imbalance could cause network training to fail. To address this, we multiply by the corresponding Laplacian basis, which not only ensures numerical consistency across loss terms but also has no adverse impact on the optimization of the coarse map, namely,

$$L_{cross}^{LB} = \left\| \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{LB} - \Pi_{\mathcal{Y}\mathcal{X}}^{LB,ref} \Phi_{\mathcal{X}}^{LB} \right\|_{\mathrm{F}}. \tag{20}$$

On the other hand, for self-spatial contrast energy, we also leverage the fine pointwise map $I_{\mathcal{Y}}$ to facilitate the optimization of the coarse maps, i.e.,

$$L_{self}^{LB} = \left\| \Pi_{\mathcal{Y}\mathcal{Y}} \Phi_{\mathcal{Y}}^{LB} - \Phi_{\mathcal{Y}}^{LB} \right\|_{\mathrm{F}}. \tag{21}$$

The unsupervised losses of the intrinsic branch are expressed as:

$$L_{total}^{LB} = \theta_{cross}^{LB} L_{cross}^{LB} + \theta_{self}^{LB} L_{self}^{LB} + \theta_{bi}^{LB} L_{bi}^{LB} + \theta_{or}^{LB} L_{or}^{LB}, \tag{22}$$

where θ^{LB} denotes the corresponding weight.

5.4 Extrinsic Branch

As the functional map learning pipeline on the extrinsic branch mirrors the intrinsic branch in its underlying formulation, we omit repetitive details for the sake of brevity and focus on the primary distinctions.

Non-differentiable refined map. Similarly, by employing Algorithm 1 with the same initial pointwise map, we can obtain the refined binary pointwise maps, denoted as $\Pi_{\mathcal{Y}\mathcal{X}}^{EL,ref}$.

Unsupervised loss. To impose a structured penalty on elastic maps, we introduce, for the first time, novel orthogonality and bijectivity loss terms, denoted as:

$$L_{or}^{EL} = \left\| \Phi_{\mathcal{Y}}^{EL} - \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{EL} (\hat{C}_{\mathcal{X}\mathcal{Y}}^{EL})^{*} \right\|_{\mathrm{HS}} \tag{23}$$

and

$$L_{bi}^{EL} = \left\| \Phi_{\mathcal{Y}}^{EL} - \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{EL} \hat{C}_{\mathcal{Y}\mathcal{X}}^{EL} \right\|_{\mathrm{HS}}, \tag{24}$$

where $\Phi_{\mathcal{X}}^{EL} \in \mathbb{R}^{|V_{\mathcal{X}}| \times k^{EL}}$.

These are reformulated in terms of the standard Frobenius norm, denoted as:

$$L_{or}^{EL} = \left\| A_{\mathcal{Y}}^{1/2} \left(\Phi_{\mathcal{Y}}^{EL} - \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{EL} (\hat{C}_{\mathcal{X}\mathcal{Y}}^{EL})^{*} \right) A_{\mathcal{Y},k}^{-1/2} \right\|_{\mathrm{F}} \tag{25}$$

and

$$L_{bi}^{EL} = \left\| A_{\mathcal{Y}}^{1/2} \left(\Phi_{\mathcal{Y}}^{EL} - \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{EL} (\hat{C}_{\mathcal{Y}\mathcal{X}}^{EL}) \right) A_{\mathcal{Y},k}^{-1/2} \right\|_{\mathrm{F}}, \tag{26}$$

respectively.

Analogously, for the spatial contrastive regularizer defined on elastic bases, we have

$$L_{cross}^{EL} = \left\| \Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{EL} - \Pi_{\mathcal{Y}\mathcal{X}}^{EL,ref} \Phi_{\mathcal{X}}^{EL} \right\|_{\mathrm{HS}} \tag{27}$$

and

$$L_{self}^{EL} = \left\| \Pi_{\mathcal{Y}\mathcal{Y}}^{EL} \Phi_{\mathcal{Y}}^{EL} - \Phi_{\mathcal{Y}}^{EL} \right\|_{\mathrm{HS}}. \tag{28}$$

Expressing them by the Frobenius norm, namely,

$$L_{cross}^{EL} = \left\| A_{\mathcal{Y}}^{1/2} \left(\Pi_{\mathcal{Y}\mathcal{X}} \Phi_{\mathcal{X}}^{EL} - \Pi_{\mathcal{Y}\mathcal{X}}^{EL,ref} \Phi_{\mathcal{X}}^{EL} \right) A_{\mathcal{X},k}^{-1/2} \right\|_{\mathrm{F}} \tag{29}$$

and

$$L_{self}^{EL} = \left\| A_{\mathcal{Y}}^{1/2} \left(\Pi_{\mathcal{Y}\mathcal{Y}}^{EL} \Phi_{\mathcal{Y}}^{EL} - \Phi_{\mathcal{Y}}^{EL} \right) A_{\mathcal{Y},k}^{-1/2} \right\|_{\mathrm{F}}, \tag{30}$$

respectively. Then, the unsupervised losses of the extrinsic branch are represented as:

$$L_{total}^{EL} = \theta_{cross}^{EL} L_{cross}^{EL} + \theta_{self}^{EL} L_{self}^{EL} + \theta_{bi}^{EL} L_{bi}^{EL} + \theta_{or}^{EL} L_{or}^{EL}. \tag{31}$$

Finally, overall losses $L_{total} = L_{total}^{LB} + L_{total}^{EL}$ and the detailed proofs of Eq. (25), (26), (29), and (30) are provided in supplementary materials.

5.5 Inference

Inspired by [15], employing distinct recovery strategies during the inference stage yields more robust results, a practice widely adopted in existing state-of-the-art approaches [6, 14, 16, 38]. For non-isometric matching scenarios, we directly recover pointwise maps from the learned feature embedding space using the nearest neighbor algorithm, namely, $\Pi_{\mathcal{Y}\mathcal{X}} = NN(F_{\mathcal{Y}}, F_{\mathcal{X}})$. For near-isometric matching scenarios, we recover pointwise maps from the hybrid basis space by applying the proposed map refinement to the obtained $\Pi_{\mathcal{Y}\mathcal{X}}$, i.e.,

$$\Pi_{\mathcal{Y}\mathcal{X}}^{ref} = NN\left(\Phi_{\mathcal{Y}}^{LB} \| (A_{\mathcal{Y},k}^{1/2} \Phi_{\mathcal{Y}}^{EL\dagger} A_{\mathcal{Y}}^{-1})^{\mathrm{T}}, \Phi_{\mathcal{X}}^{LB} (\hat{C}_{\mathcal{X}\mathcal{Y}}^{LB,ref})^{\mathrm{T}}) \| (A_{\mathcal{Y},k}^{1/2} \hat{C}_{\mathcal{X}\mathcal{Y}}^{EL,ref} \Phi_{\mathcal{X}}^{EL,\dagger} A_{\mathcal{X}}^{-1})^{\mathrm{T}} \right), \tag{32}$$

where $\|$ denotes matrix concatenation.

6 Experimental Results

6.1 Datasets

We extensively evaluate our method on widely used challenging benchmarks, including

- *Near-isometry.* The near-isometric datasets comprise remeshed FAUST(F for short) [60], SCAPE [60], and challenging SHREC'19 [52]. Moreover, anisotropic meshing datasets FAUST [15](F_a for short) and SCAPE [15] were used to evaluate robustness across different discretizations.

- *Non-isometry.* Commonly employed non-isometric datasets include SMAL [80] and DT4D-H [51]. DT4D-H offers two configurations: intra-class and inter-class. We omit the intra-class setting because it is overly simple, thereby precluding further evaluation of the method's performance. Consequently, we retain SMAL and the most challenging DT4D-H inter-class setting.
- *Topological noise.* The TOPKIDS dataset [37] is employed to evaluate our method's robustness to topological noise.

The training and testing protocols on SMAL follow the same procedure as previous related studies [19, 79], and the remaining datasets use the same splits as those in many earlier works [15, 39].

6.2 Baselines

We extensively compare our method with existing non-rigid deformable shape matching methods, which we categorize as follows:

- *Axiomatic approaches*, including ZoomOut [53], Smooth Shells [24], DiscreteOp [58], and MWP [33].
- *Supervised approaches*, FMNet [42], GeomFMaps [20].
- *Unsupervised approaches*, including Deep Shells [26], DUO-FMNet [19], AttentiveFMaps [39], RFMNet [34], ULRSSM [15], DiffZO [50], HybridFMaps [6], DeepFAFM [48], and DenoiseFMaps [79].

Excluding test-time fine-tuning. Employing a fine-tuning technique during the inference stage can significantly enhance method performance [15], particularly in terms of generalization capability. However, except for ULRSSM [15] and HybridFMaps [6], the remaining baselines do not adopt any fine-tuning strategies during inference to improve matching performance. Utilizing that refinement technique involving parameter updates during the inference stage would lead to unfair competition with the other unsupervised methods. Therefore, we omit the fine-tuning technique from ULRSSM and HybridFMaps.

6.3 Results

The mean geodesic error [36] is adopted to evaluate correspondence accuracy, with all results multiplied by 100 for improved readability.

Near-isometric matching. These results are summarized in Table 2. We achieve competitive performance in near-isometric matching scenarios, surpassing many axiomatic, supervised, and unsupervised methods. For instance, on the FAUST dataset, our method achieves the second-best performance.

Cross-dataset generalization. Cross-dataset generalization is a critical metric for evaluating a method's practical effectiveness. We assess this by training our model on one dataset and testing it on an unseen one, with quantitative results reported in Table 2. Our method achieves highly competitive performance across most settings. Notably, even on the more challenging SHREC'19 dataset [52] (i.e., training on FAUST or SCAPE and testing on SHREC'19),

Table 2: Evaluating the matching results across various benchmarks, including near-isometric shape matching, cross-dataset generalization (FAUST, SCAPE, and SHREC'19), anisotropic meshing (F_a and S_a), non-isometric shape matching (SMAL and DT4D-H), and matching with topological noise (TOPKIDS), respectively. The numbers in the table are mean geodesic errors (×100). **Bold**: Best. Underline: Runner-up.

Train	FAUST			SCAPE			FAUST		SCAPE		SMAL	DT4D-H inter	TOPKIDS
Test	F	S	S19	F	S	S19	F_a	S_a	F_a	S_a			
Axiomatic Methods													
ZoomOut [53]	6.1	7.5	-	6.1	7.5	-	8.7	15.0	8.7	15.0	47.7	29.0	33.7
SmoothShells [24]	2.5	4.7	-	2.5	4.7	-	5.4	5.0	5.4	5.0	34.9	6.3	35.5
DiscreteOp [58]	5.6	13.1	-	5.6	13.1	-	6.2	14.6	6.2	14.6	36.1	27.6	10.8
MWP [33]	3.1	4.1	-	3.1	4.1	-	8.2	8.7	8.2	8.7	20.9	25.4	5.7
Supervised Methods													
FMNet [42]	11.1	30.0	-	33.0	17.0	-	42.0	43.0	43.0	41.0	-	38.0	-
GeomFMaps [20]	2.6	3.4	9.9	3.0	3.0	12.2	3.2	3.8	8.4	3.1	4.3	4.1	-
Unsupervised Methods													
Deep Shells [26]	1.7	5.4	27.4	2.7	2.5	23.4	12.0	16.0	15.0	10.0	21.4	31.1	13.7
DUO-FMNet [19]	2.5	4.2	6.4	2.7	2.6	8.4	3.0	4.4	3.1	2.7	4.8	15.8	-
AttentiveFMaps [39]	1.9	2.6	6.4	2.2	2.2	9.9	2.4	<u>2.8</u>	2.5	2.3	4.4	11.6	23.4
RFMNet [34]	1.7	**2.3**	6.3	<u>1.7</u>	2.1	6.9	3.6	2.6	3.6	3.9	4.4	13.9	-
ULRSSM [15]	1.6	6.7	14.5	4.8	1.9	18.5	2.5	8.9	7.0	1.9	4.5	5.2	9.4
DiffZO [50]	1.9	2.4	**4.2**	1.9	2.4	<u>6.9</u>	2.2	3.8	2.7	2.4	4.3	4.1	-
HybridFMaps [6]	**1.4**	4.2	9.5	2.3	**1.8**	13.0	2.0	4.6	3.4	**1.8**	<u>3.5</u>	<u>3.9</u>	<u>5.0</u>
DeepFAFM [48]	1.6	2.7	7.0	1.9	<u>1.9</u>	7.9	2.0	2.9	**2.6**	1.9	3.9	4.2	6.3
DenoiseFMaps [79]	1.8	-	-	-	2.3	-	<u>2.0</u>	-	-	2.3	4.3	12.8	43.6
Ours	<u>1.6</u>	<u>2.4</u>	<u>4.4</u>	**1.6**	2.0	**4.0**	**1.9**	**2.4**	**2.4**	<u>1.9</u>	**2.9**	**3.5**	**4.9**

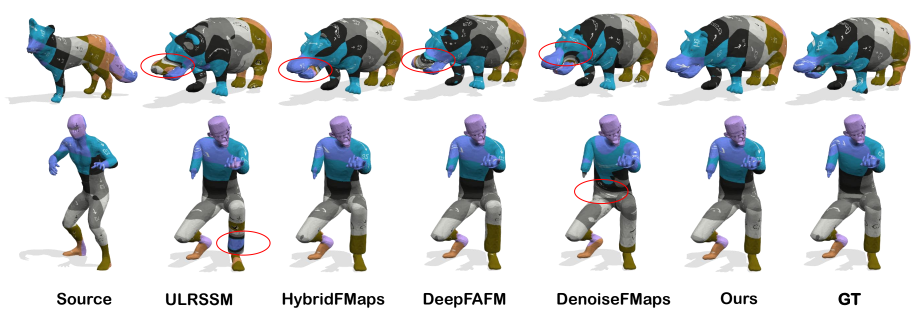

Fig. 3: Qualitative non-isometric matching results on the SMAL (top) [80] and DT4D-H (bottom) [51] datasets. Our method consistently suppresses matching errors and preserves texture coherence better than alternative methods, underscoring its efficacy in challenging non-isometric scenarios.

our performance remains robust. In contrast, existing state-of-the-art methods—such as DUO-FMNet [19], ULRSSM [15], and HybridFMaps [6]—exhibit varying degrees of degradation in relatively simpler cross-dataset scenarios (e.g., training on FAUST, testing on SCAPE) and suffer even steeper performance drops on SHREC'19. These quantitative findings confirm that our approach generalizes substantially better than current baselines.

Matching with anisotropic meshing. To evaluate the robustness of the method under different surface discretizations, we trained on remeshed data and tested on anisotropic remeshed data. The quantitative results are shown in Table 2. Many state-of-the-art methods exhibited varying degrees of fluctuation. For example, the accuracy of DiffZO significantly dropped when training on FAUST and testing on anisotropic remeshed SCAPE (S_a). However, our method achieved outstanding performance in most scenarios, demonstrating the robustness of our approach to discretization variations and showing that it consistently outperforms current supervised and unsupervised methods.

Non-isometric matching. We evaluate our method on the highly challenging non-isometric SMAL and DT4D-H datasets, with quantitative results presented in Table 2. Across all non-isometric settings, our method consistently outperforms all baselines—even fully supervised ones—firmly validating its effectiveness and reliability. In contrast, existing methods such as DenoiseFMaps exhibit a significant drop in accuracy on the particularly demanding DT4D-H dataset. Furthermore, the qualitative results in Fig. 3 demonstrate that our approach yields significantly lower matching errors and more coherent texture transfers compared to competitive baselines. Together, these results underscore our framework's superior performance and robustness when handling large-scale non-isometric deformations.

Matching with topological noise. We conducted experiments on the SHREC'16 TOPKIDS dataset [37] to evaluate robustness against topological perturbations. The results are summarized in Table 2. Many functional map learning methods [39,79] struggle to produce satisfactory results, as the intrinsic geometric structure is distorted. In contrast, our method still outperforms all baselines and achieves superior performance under such conditions.

For additional comparisons and discussions, please refer to the supplementary materials.

7 Conclusion

We propose a novel and effective hybrid self-supervised method for non-rigid shape matching, based on the proposed coarse-to-fine strategy. Extensive experiments demonstrate that our method outperforms existing state-of-the-art methods in terms of matching accuracy, generalization capability, and computational efficiency. Finally, we rigorously demonstrate that contrastive energies promote feature discrimination and yield consistent improvements when integrated into existing architectures, further confirming the validity of our approach.

8 Acknowledgments

We extend our gratitude to Yizheng Xie for his generous assistance. This work was supported by the National Natural Science Foundation of China under Grant No. 62271452.

References

1. Aflalo, Y., Dubrovina, A., Kimmel, R.: Spectral generalized multi-dimensional scaling. International Journal of Computer Vision **118**(3), 380–392 (2016) 3
2. Attaiki, S., Ovsjanikov, M.: Understanding and improving features learned in deep functional maps. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1316–1326 (2023) 2, 3, 5
3. Attaiki, S., Pai, G., Ovsjanikov, M.: Dpfm: Deep partial functional maps. In: International Conference on 3D Vision (3DV). pp. 175–185 (2021) 1
4. Aubry, M., Schlickewei, U., Cremers, D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In: IEEE International Conference on Computer Vision Workshops. pp. 1626–1633 (2011) 3
5. Aygün, M., Lähner, Z., Cremers, D.: Unsupervised dense shape correspondence using heat kernels. In: International Conference on 3D Vision (3DV). pp. 573–582 (2020) 1, 3, 5
6. Bastian, L., Xie, Y., Navab, N., Lähner, Z.: Hybrid functional maps for crease-aware non-isometric shape matching. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3313–3323 (2024) 2, 3, 4, 5, 7, 9, 12, 13, 14
7. Besl, P.J., McKay, N.D.: Method for registration of 3-d shapes. In: Sensor fusion IV: control paradigms and data structures. vol. 1611, pp. 586–606. Spie (1992) 3
8. Bracha, A., Dagès, T., Kimmel, R.: On unsupervised partial shape correspondence. In: Proceedings of the Asian Conference on Computer Vision. pp. 4488–4504 (2024) 3
9. Bracha, A., Dagès, T., Kimmel, R.: Wormhole loss for partial shape matching. In: Proceedings of the 38th International Conference on Neural Information Processing Systems. pp. 131247–131277 (2024) 3
10. Bronstein, A.M., Bronstein, M.M., Kimmel, R.: Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. Proceedings of the National Academy of Sciences **103**(5), 1168–1172 (2006) 3
11. Cao, D., Bernard, F.: Unsupervised deep multi-shape matching. In: European Conference on Computer Vision. pp. 55–71. Springer (2022) 1, 3, 9
12. Cao, D., Bernard, F.: Self-supervised learning for multimodal non-rigid 3d shape matching. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17735–17744 (2023) 3
13. Cao, D., Eisenberger, M., El Amrani, N., Cremers, D., Bernard, F.: Spectral meets spatial: Harmonising 3d shape matching and interpolation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3658–3668 (2024) 2, 3, 5, 9
14. Cao, D., Lähner, Z., Bernard, F.: Synchronous diffusion for unsupervised smooth non-rigid 3d shape matching. In: European Conference on Computer Vision. pp. 262–281 (2024) 3, 9, 12
15. Cao, D., Roetzer, P., Bernard, F.: Unsupervised learning of robust spectral shape matching. ACM Transactions on Graphics **42**, 1–15 (2023) 2, 3, 5, 6, 9, 10, 12, 13, 14
16. Cao, D., Roetzer, P., Bernard, F.: Revisiting map relations for unsupervised non-rigid shape matching. In: 2024 International Conference on 3D Vision (3DV). pp. 1371–1381. IEEE (2024) 3, 6, 9, 12
17. 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 (2022) 2

18. Donati, N., Corman, E., Melzi, S., Ovsjanikov, M.: Complex functional maps: A conformal link between tangent bundles. Computer Graphics Forum **41**, 317–334 (2022) 3
19. Donati, N., Corman, E., Ovsjanikov, M.: Deep orientation-aware functional maps: Tackling symmetry issues in shape matching. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 732–741 (2022) 1, 2, 3, 5, 9, 13, 14
20. Donati, N., Sharma, A., Ovsjanikov, M.: Deep geometric functional maps: Robust feature learning for shape correspondence. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8589–8598 (2020) 1, 3, 5, 13, 14
21. Egger, B., Smith, W.A., Tewari, A., Wuhrer, 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
22. Ehm, V., Gao, M., Roetzer, P., Eisenberger, M., Cremers, D., Bernard, F.: Partial-to-partial shape matching with geometric consistency. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 27488–27497 (2024) 1
23. Ehm, V., Roetzer, P., Eisenberger, M., Gao, M., Bernard, F., Cremers, D.: Geometrically consistent partial shape matching. In: 2024 International Conference on 3D Vision (3DV). pp. 914–922. IEEE Computer Society (2024) 1
24. Eisenberger, M., Lähner, Z., Cremers, D.: Smooth shells: Multi-scale shape registration with functional maps. In: Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition. pp. 12262–12271 (2020) 13, 14
25. Eisenberger, M., Novotny, D., Kerchenbaum, G., Labatut, P., Neverova, N., Cremers, D., Vedaldi, A.: Neuromorph: Unsupervised shape interpolation and correspondence in one go. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7473–7483 (2021) 1, 10
26. Eisenberger, M., Toker, A., Leal-Taixé, L., Cremers, D.: Deep shells: Unsupervised shape correspondence with optimal transport. Advances in Neural information processing systems **33**, 10491–10502 (2020) 1, 13, 14
27. Eisenberger, M., Toker, A., Leal-Taixé, L., Cremers, D.: G-msm: Unsupervised multi-shape matching with graph-based affinity priors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 22762–22772 (2023) 1
28. Fan, A., Ma, J., Tian, X., Mei, X., Liu, W.: Coherent point drift revisited for non-rigid shape matching and registration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1424–1434 (2022) 3
29. Gao, M., Lahner, Z., Thunberg, J., Cremers, D., Bernard, F.: Isometric multi-shape matching. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14183–14193 (2021) 3
30. Halimi, O., Litany, O., Rodolà, E.R., Bronstein, A.M., Kimmel, R.: Unsupervised learning of dense shape correspondence. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4365–4374 (2019) 1, 3
31. Hammond, D.K., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Applied and computational harmonic analysis **30**(2), 129–150 (2011) 8

32. Hartwig, F., Sassen, J., Azencot, O., Rumpf, M., Ben-Chen, M.: An elastic basis for spectral shape correspondence. In: ACM SIGGRAPH conference proceedings. pp. 1–11 (2023) 3, 7, 8
33. Hu, L., Li, Q., Liu, S., Liu, X.: Efficient deformable shape correspondence via multiscale spectral manifold wavelets preservation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14531–14540 (2021) 3, 7, 13, 14
34. Hu, L., Li, Q., Liu, S., Yan, D.M., Xu, H., Liu, X.: Rfmnet: Robust deep functional maps for unsupervised non-rigid shape correspondence. Graphical Models **129**, 1–11 (2023) 1, 3, 13, 14
35. Huang, R., Ren, J., Wonka, P., Ovsjanikov, M.: Consistent zoomout: Efficient spectral map synchronization. Computer Graphics Forum **39**, 265–278 (2020) 3
36. Kim, V.G., Lipman, Y., Funkhouser, T.: Blended intrinsic maps. ACM transactions on graphics **30**, 1–12 (2011) 13
37. Lähner, Z., Rodolà, E., Bronstein, M.M., Cremers, D., Burghard, O., Cosmo, L., Dieckmann, A., Klein, R., Sahillioğlu, Y., et al.: Shrec'16: Matching of deformable shapes with topological noise. In: Eurographics Workshop on 3D Object Retrieval. pp. 55–60 (2016) 13, 15
38. Le, T., Nguyen, K., Sun, S., Ho, N., Xie, X.: Integrating efficient optimal transport and functional maps for unsupervised shape correspondence learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 23188–23198 (2024) 2, 3, 9, 12
39. Li, L., Donati, N., Ovsjanikov, M.: Learning multi-resolution functional maps with spectral attention for robust shape matching. Advances in Neural Information Processing Systems **35**, 29336–29349 (2022) 1, 3, 5, 9, 13, 14, 15
40. Li, Q., Hu, L., Liu, S., Yang, D., Liu, X.: Anisotropic spectral manifold wavelet descriptor. In: Computer Graphics Forum. vol. 40, pp. 81–96. Wiley Online Library (2021) 3
41. Li, Q., Meng, J., Wang, H., Liu, S.: Mdnd: Unsupervised learning guided by non-differentiable refinement for shape correspondence. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 40, pp. 6351–6359 (2026) 3
42. Litany, O., Remez, T., Rodolà, E., Bronstein, A., Bronstein, M.: Deep functional maps: Structured prediction for dense shape correspondence. In: IEEE International Conference on Computer Vision. pp. 5660–5668 (2017) 1, 3, 13, 14
43. Liu, S., Luo, F., Li, Q., Liu, X., Hu, L.: Awedd: a descriptor simultaneously encoding multiscale extrinsic and intrinsic shape features. The Visual Computer **40**(4), 2537–2554 (2024) 3
44. Liu, S., Wang, H., Yan, D.M., Li, Q., Luo, F., Teng, Z., Liu, X.: Spectral descriptors for 3d deformable shape matching: A comparative survey. IEEE transactions on visualization and computer graphics **31**(3), 1677–1697 (2024) 2
45. Loncaric, S.: A survey of shape analysis techniques. Pattern recognition **31**(8), 983–1001 (1998) 1
46. Luo, F., Chen, H.: From feature learning to spectral basis learning: A unifying and flexible framework for efficient and robust shape matching. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 31377–31388 (2026) 3
47. Luo, F., Chen, H.: Unsupervised contrastive learning for efficient and robust spectral shape matching. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 40, pp. 7662–7670 (2026) 3

48. Luo, F., Li, Q., Hu, L., Wang, H., Xu, H., Liu, X., Liu, S., Chen, H.: Deep frequency awareness functional maps for robust shape matching. IEEE Transactions on Visualization and Computer Graphics **31**, 7781–7794 (2025) 2, 3, 7, 8, 9, 10, 13, 14
49. Maggioli, F., Baieri, D., Rodolà, E., Melzi, S.: Rematching: Low-resolution representations for scalable shape correspondence. In: European Conference on Computer Vision. pp. 183–200. Springer (2024) 3
50. Magnet, R., Ovsjanikov, M.: Memory-scalable and simplified functional map learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4041–4050 (2024) 3, 4, 6, 9, 13, 14
51. Magnet, R., Ren, J., Sorkine-Hornung, O., Ovsjanikov, M.: Smooth non-rigid shape matching via effective dirichlet energy optimization. In: International Conference on 3D Vision (3DV). pp. 495–504 (2022) 3, 13, 14
52. Melzi, S., Marin, R., Rodolà, E., Castellani, U., Ren, J., Poulenard, A., Ovsjanikov, P., et al.: Shrec'19: matching humans with different connectivity. In: Eurographics Workshop on 3D Object Retrieval. pp. 1–8 (2019) 12, 13
53. Melzi, S., Ren, J., Rodolà, E., Sharma, A., Wonka, P., Ovsjanikov, M.: Zoomout: Spectral upsampling for efficient shape correspondence. ACM Transactions on Graphics **38**, 1–14 (2019) 3, 7, 8, 13, 14
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 **31**, 1–11 (2012) 3, 4
55. 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: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 384–393 (2021) 3
56. Pierson, E., Li, L., Dai, A., Ovsjanikov, M.: Diffumatch: Category-agnostic spectral diffusion priors for robust non-rigid shape matching. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5745–5756 (2025) 3, 4
57. Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. Experimental mathematics **2**(1), 15–36 (1993) 4
58. Ren, J., Melzi, S., Wonka, P., Ovsjanikov, M.: Discrete optimization for shape matching. In: Computer Graphics Forum. vol. 40, pp. 81–96 (2021) 2, 4, 10, 13, 14
59. Ren, J., Panine, M., Wonka, P., Ovsjanikov, M.: Structured regularization of functional map computations. In: Computer Graphics Forum. vol. 38, pp. 39–53 (2019) 3, 4, 10
60. Ren, J., Poulenard, A., Wonka, P., Ovsjanikov, M.: Continuous and orientation-preserving correspondences via functional maps. ACM Transactions on Graphics **37**, 1–16 (2018) 3, 7, 12
61. Roetzer, P., Bernard, F.: Spidermatch: 3d shape matching with global optimality and geometric consistency. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14543–14553 (2024) 3
62. Roufosse, J.M., Sharma, A., Ovsjanikov, M.: Unsupervised deep learning for structured shape matching. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1617–1627 (2019) 1, 3
63. Sahillioğlu, Y.: Recent advances in shape correspondence. The Visual Computer **36**, 1705–1721 (2020) 2
64. 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) 3

65. Sharma, A., Ovsjanikov, M.: Weakly supervised deep functional maps for shape matching. Advances in Neural Information Processing Systems **33**, 19264–19275 (2020) 3
66. Sharp, N., Attaiki, S., Crane, K., Ovsjanikov, M.: Diffusionnet: Discretization agnostic learning on surfaces. ACM Transactions on Graphics (TOG) **41**(3), 1–16 (2022) 4, 9
67. Sumner, R.W., Popović, J.: Deformation transfer for triangle meshes. ACM Transactions on graphics **23**, 399–405 (2004) 1
68. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. Computer Graphics Forum **28**, 1383–1392 (2009) 3
69. Sun, M., Mao, S., Jiang, P., Ovsjanikov, M., Huang, R.: Spatially and spectrally consistent deep functional maps. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 14497–14507 (2023) 1, 2, 3, 4, 6
70. Vestner, M., Lähner, Z., Boyarski, A., Litany, O., Slossberg, R., Remez, T., Rodola, E., Bronstein, A., Bronstein, M., Kimmel, R., et al.: Efficient deformable shape correspondence via kernel matching. In: 2017 International Conference on 3D Vision (3DV). pp. 517–526. IEEE (2017) 3, 7
71. Viganò, G., Ovsjanikov, M., Melzi, S.: Nam: Neural adjoint maps for refining shape correspondences. ACM Transactions on Graphics (TOG) **44**(4), 1–15 (2025) 3
72. Weber, S., Dages, T., Gao, M., Cremers, D.: Finsler-laplace-beltrami operators with application to shape analysis. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 3131–3140 (2024) 3
73. Wirth, B., Bar, L., Rumpf, M., Sapiro, G.: A continuum mechanical approach to geodesics in shape space. International Journal of Computer Vision **93**(3), 293–318 (2011) 4
74. Xiang, R., Lai, R., Zhao, H.: Efficient and robust shape correspondence via sparsity-enforced quadratic assignment. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9513–9522 (June 2020) 3
75. 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: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11665–11675 (2025) 3
76. Xu, H., Li, Q., Hu, L., Liu, S., Wang, H., Liu, X.: Sedfmnet: A simple and efficient unsupervised functional map for shape correspondence based on deconstruction. Graphical Models **139**, 101270 (2025) 4
77. Yona, G., Velich, R., Rivlin, E., Kimmel, R.: Neural descriptors: Self-supervised learning of robust local surface descriptors using polynomial patches. In: International Conference on Scale Space and Variational Methods in Computer Vision. pp. 218–230. Springer (2025) 3
78. Zhuravlev, A., Bastian, L., Cao, D., Amrani, N.E., Roetzer, P., Ehm, V., Marin, R., Nishizawa, H., Morishima, S., Theobalt, C., et al.: Non-rigid 3d shape correspondences: From foundations to open challenges and opportunities. In: Computer Graphics Forum. p. e70397. Wiley Online Library (2026) 2, 3
79. Zhuravlev, A., Lähner, Z., Golyanik, V.: Denoising functional maps: Diffusion models for shape correspondence. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 26899–26909 (2025) 3, 5, 13, 14, 15
80. Zuffi, S., Kanazawa, A., Jacobs, D.W., Black, M.J.: 3d menagerie: Modeling the 3d shape and pose of animals. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 6365–6373 (2017) 13, 14