

Unsupervised Learning of Robust Spectral Shape Matching

DONGLIANG CAO, PAUL ROETZER, and FLORIAN BERNARD, University of Bonn, Germany

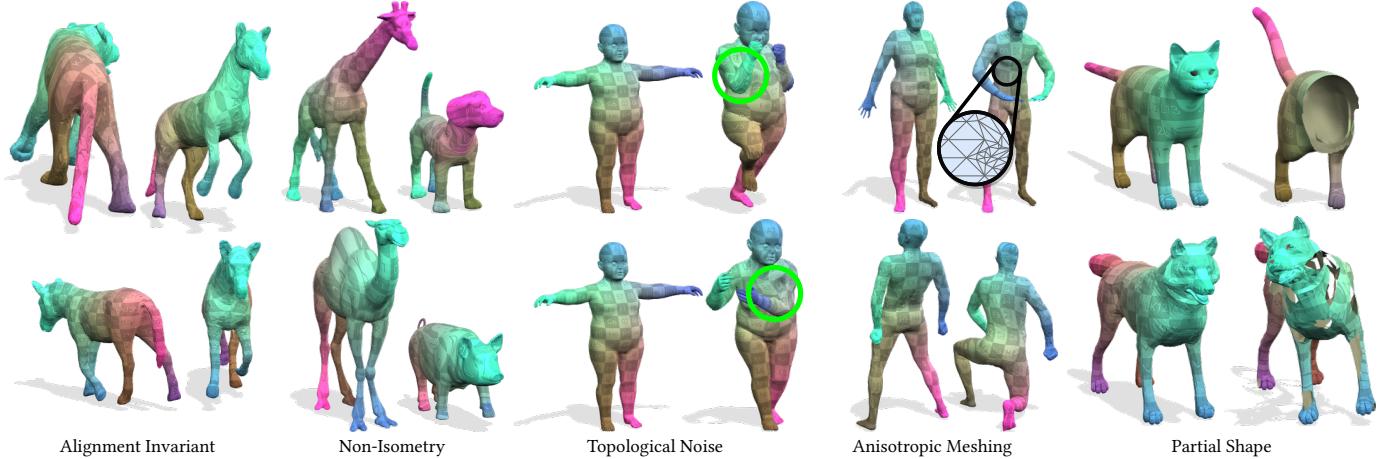


Fig. 1. We propose the first **unsupervised spectral shape matching** approach that is **robust across a broad range of challenging settings**: shape matching without initial alignment due to the intrinsic formulation, shape matching of non-isometric shape pairs, shape matching with topological noise, shape matching with anisotropic meshing, and partial shape matching.

We propose a novel learning-based approach for robust 3D shape matching. Our method builds upon deep functional maps and can be trained in a fully unsupervised manner. Previous deep functional map methods mainly focus on predicting optimised functional maps alone, and then rely on off-the-shelf post-processing to obtain accurate point-wise maps during inference. However, this two-stage procedure for obtaining point-wise maps often yields sub-optimal performance. In contrast, building upon recent insights about the relation between functional maps and point-wise maps, we propose a novel unsupervised loss to couple the functional maps and point-wise maps, and thereby directly obtain point-wise maps without any post-processing. Our approach obtains accurate correspondences not only for near-isometric shapes, but also for more challenging non-isometric shapes and partial shapes, as well as shapes with different discretisation or topological noise. Using a total of nine diverse datasets, we extensively evaluate the performance and demonstrate that our method substantially outperforms previous state-of-the-art methods, even compared to recent supervised methods. Our code is available at <https://github.com/dongliangcao/Unsupervised-Learning-of-Robust-Spectral-Shape-Matching>.

CCS Concepts: • Computing methodologies → Shape analysis.

Authors' address: Dongliang Cao, dcao@uni-bonn.de; Paul Roetzer, paul.roetzer@uni-bonn.de; Florian Bernard, fb@uni-bonn.de, University of Bonn, Friedrich-Hirzebruch-Allee 5, Bonn, North Rhine-Westphalia, Germany, D-53115.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2023/8-ART \$15.00
<https://doi.org/10.1145/3592107>

Additional Key Words and Phrases: Shape matching, deep learning, functional maps

ACM Reference Format:

Dongliang Cao, Paul Roetzer, and Florian Bernard. 2023. Unsupervised Learning of Robust Spectral Shape Matching. *ACM Trans. Graph.* 42, 4 (August 2023), 15 pages. <https://doi.org/10.1145/3592107>

1 INTRODUCTION

Shape matching is a long-standing problem in computer graphics and shape analysis with a diverse range of applications, including texture transfer [Dinh et al. 2005], deformation transfer [Sumner and Popović 2004], and parametric model construction [Egger et al. 2020; Loper et al. 2015]. Even though shape matching has been studied for decades [Tam et al. 2012; Van Kaick et al. 2011], finding correspondences between non-rigidly deformed shapes still remains highly challenging, especially in the presence of large non-isometric deformation, partiality, or topological inconsistency.

With recent developments in deep learning, numerous learning-based approaches were proposed for non-rigid 3D shape matching, including both supervised [Donati et al. 2020; Groueix et al. 2018; Litany et al. 2017a; Trappolini et al. 2021; Wiersma et al. 2020] and unsupervised methods [Eisenberger et al. 2020b; Halimi et al. 2019; Roufosse et al. 2019; Sharma and Ovsjanikov 2020]. While most of them mainly focus on near-isometric shape matching (such as the recent unsupervised work by Eisenberger et al. [2020b] which demonstrated near-perfect performance), only few works explicitly consider more challenging scenarios, such as non-isometric shape matching [Donati et al. 2022a; Li et al. 2022] and partial shape

Table 1. Method comparison. Our method is the first unsupervised learning approach that operates in the intrinsic/spectral domain and achieves accurate matching performance for both non-isometric and partial shapes.

Methods	Unsup.	Fully intrinsic	Non-iso.	Part.
[Litany et al. 2017a]	✗	✓	✗	✗
[Groueix et al. 2018]	✗	✗	✗	✗
[Halimi et al. 2019]	✓	✓	✗	✗
[Roufosse et al. 2019]	✓	✓	✗	✗
[Wiersma et al. 2020]	✗	✗	✗	✗
[Li et al. 2020]	✗	✗	✗	✗
[Donati et al. 2020]	✗	✓	✗	✗
[Eisenberger et al. 2020b]	✓	✗	✓	✗
[Attaiki et al. 2021]	✗	✓	✗	✓
[Eisenberger et al. 2021]	✓	✗	✓	✗
[Trappolini et al. 2021]	✗	✗	✗	✗
[Cao and Bernard 2022]	✓	✓	✗	✓
[Donati et al. 2022a]	✓	✓	✓	✗
[Li et al. 2022]	✓	✓	✓	✗
Ours	✓	✓	✓	✓

matching [Attaiki et al. 2021; Cao and Bernard 2022]. For such settings, the performance gap between the state-of-the-art supervised methods [Attaiki et al. 2021; Donati et al. 2020] and unsupervised methods [Cao and Bernard 2022; Li et al. 2022] is still large, while supervised methods at the same time suffer from overfitting.

In this work we propose an unsupervised learning framework to address these shortcomings. Our universal framework can be applied to a broad range of shape matching settings, including near-isometric shapes and more challenging non-isometric shapes, as well as partial shapes. Our method builds upon the powerful functional map framework [Ovsjanikov et al. 2012] integrated into a deep neural network [Litany et al. 2017a]. While previous unsupervised deep functional map approaches (e.g. [Halimi et al. 2019; Roufosse et al. 2019]) only regularise the functional map alone and ignore relations between the functional map and the underlying point-wise maps, we propose to explicitly exploit this relation during neural network training. Specifically, inspired by the recent axiomatic (i.e. non-learning-based) method by Ren et al. [2021], we introduce a novel differentiable unsupervised coupling loss, which accounts for the functional map being associated to a point-wise map. Furthermore, we propose a test-time adaptation strategy to simultaneously optimise the functional map and the corresponding point-wise map to further improve the matching performance. Overall, we demonstrate that this substantially improves upon previous shape matching approaches on a diverse and extensive selection of datasets. We summarise our main contributions as follows:

- For the first time we propose a *universal unsupervised learning* framework that allows to obtain accurate correspondences not only for near-isometric shapes but also for more challenging non-isometric and partial shapes, see Tab. 1.
- We introduce a *novel unsupervised loss* to enforce the functional map to be associated with a point-wise map. Together with our *test-time adaptation*, our method is much more robust to the choice of the spectral resolution compared to existing deep functional map methods.

- We demonstrate *state-of-the-art* performance on numerous challenging benchmarks in diverse settings, including near-isometric, non-isometric and partial shape matching, even in comparison to recent supervised methods.

2 RELATED WORK

In the following we will focus on reviewing those methods that are most relevant to our work. A more comprehensive overview can be found in [Sahillioğlu 2020; Tam et al. 2012; Van Kaick et al. 2011].

2.1 Axiomatic Functional Map Methods

Several shape matching approaches [Holzschuh et al. 2020; Huang et al. 2008; Ovsjanikov et al. 2010; Roetzer et al. 2022; Windheuser et al. 2011] directly establish correspondences between a given pair of shapes, while some other methods [Bernard et al. 2020; Eisenberger et al. 2019; Ezuz et al. 2019a] attempt to solve the problem by finding a non-rigid deformation to align them. Nevertheless, they solve the correspondence problem by minimising an energy function defined on the shape surface and thus often result in complex optimisation problems. In contrast, the functional map framework encodes the correspondence relationship into a small matrix (i.e. the functional map) that can be efficiently computed [Ovsjanikov et al. 2012]. Due to its simplicity and efficiency, the functional map framework has been extended in numerous works, e.g., in terms of improving the matching accuracy and robustness [Eynard et al. 2016; Ren et al. 2019], or extending it to non-isometric shape matching [Eisenberger et al. 2020a; Kovnatsky et al. 2013; Magnet et al. 2022; Ren et al. 2021, 2018], partial shape matching [Litany et al. 2017b; Rodolà et al. 2017], multi-shape matching [Cohen and Ben-Chen 2020; Gao et al. 2021; Huang et al. 2014, 2020], and matching with multiple solutions [Ren et al. 2020]. Nevertheless, axiomatic functional map methods heavily rely on handcrafted features [Aubry et al. 2011; Bronstein and Kokkinos 2010; Salti et al. 2014], which limits their matching performance, especially in the presence of non-isometries and partiality. In contrast, our method directly learns robust features from training data and achieves more accurate and robust matching performance on challenging non-isometric and partial settings.

2.2 Deep Functional Map Methods

Unlike axiomatic approaches that use handcrafted features, deep functional map methods attempt to learn features directly from training data. The supervised FMNet [Litany et al. 2017a] was first proposed to learn a non-linear transformation of SHOT descriptors [Salti et al. 2014]. Later works [Halimi et al. 2019; Roufosse et al. 2019] enable unsupervised training of FMNet by introducing isometry losses in the spatial and spectral domain, respectively. To improve the matching performance, several works [Donati et al. 2020; Sharma and Ovsjanikov 2020] replace FMNet by point-based networks [Qi et al. 2017; Thomas et al. 2019]. In order to extend the learning framework for point cloud matching, DiffFMaps [Marin et al. 2020] learns both the basis functions and features together. A similar idea was used in Azencot and Lai [2021]. Recently, DiffusionNet [Sharp et al. 2020] introduces an implicit diffusion process and has led to state-of-the-art matching performance [Attaiki

and Ovsjanikov 2022; Attaiki et al. 2021; Cao and Bernard 2022; Donati et al. 2022a; Li et al. 2022]. DPFM [Attaiki et al. 2021] enables supervised partial shape matching building upon a cross-attention mechanism and outlier detection. ConsistFMaps [Cao and Bernard 2022] introduces a virtual universe shape to enable cycle-consistent multi-shape matching. DUO-FMNet [Donati et al. 2022a] learns orientation-aware features by considering complex functional maps [Donati et al. 2022b]. AttentiveFMaps [Li et al. 2022] introduces a spectral attention mechanism to combine functional maps of different spectral resolutions. Despite the rapid progress of deep functional map methods, existing unsupervised approaches mostly focus on near-isometric shape matching, or they suffer from large performance gaps compared to supervised methods in the context of non-isometric or partial shape matching. In this work, we close this gap by introducing a universal unsupervised learning framework that achieves state-of-the-art performance in diverse settings, including near-isometric, non-isometric and partial shape matching.

2.3 Functional Map Refinement

A common strategy to improve the final matching performance is to iteratively refine functional maps as a post-processing step. As the simplest refinement technique, ICP iteratively converts the functional map to a point-wise map via nearest neighbour search [Ovsjanikov et al. 2012]. Follow-up works [Ezuz and Ben-Chen 2017; Rodolà et al. 2015; Rodola et al. 2017] iteratively optimise the functional map and point-wise map by minimising an energy function. PMF [Vestner et al. 2017] introduces bijectivity as a hard constraint for point-wise map conversion. ZoomOut [Melzi et al. 2019b] increases the spectral resolution of the functional map during the map conversion. RHM [Ezuz et al. 2019b] prompts reversible harmonic maps, thereby resulting in maps with lower conformal distortion. FSF [Pai et al. 2021] theoretically analyses the relationship between functional maps and point-wise maps and proposes a point-wise map recovery method based on spectral alignment. Another line of works encourages cycle-consistency [Bernard et al. 2019; Huang et al. 2014, 2020; Wang et al. 2013] in the case of processing a collection of shapes. Nevertheless, two-stage ‘match+refine’ approaches often yield sub-optimal performance, especially given erroneous initial functional maps. In contrast, our approach directly enforces the functional map to be associated with a point-wise map and optimises both of them simultaneously.

3 BACKGROUND

In this section we explain the background and introduce the notation used throughout the rest of the paper, see Tab. 2.

3.1 Functional Map Pipeline

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. The associated positive semi-definite Laplacian matrices $L_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times n_{\mathcal{M}}}, L_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times n_{\mathcal{N}}}$ [Pinkall and Polthier 1993] are computed as $L_{\mathcal{M}} = A_{\mathcal{M}}^{-1} W_{\mathcal{M}}$, where $A_{\mathcal{M}}$ is the diagonal matrix of lumped area elements and $W_{\mathcal{M}}$ is the cotangent weight matrix. The first k eigenfunctions $\Phi_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times k}, \Phi_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times k}$ of the respective Laplacian matrices are

Table 2. Summary of the **notation** used in this paper.

Symbol	Description
\mathcal{M}, \mathcal{N}	3D shapes (triangle mesh) with $n_{\mathcal{M}} (n_{\mathcal{N}})$ vertices
$X_{\mathcal{M}}$	vertex positions of shape \mathcal{M}
$F_{\mathcal{M}}$	vertex-wise features of shape \mathcal{M}
$L_{\mathcal{M}}$	Laplacian matrix associated with shape \mathcal{M}
$\Phi_{\mathcal{M}}$	eigenfunctions of Laplacian matrix $L_{\mathcal{M}}$
$C_{\mathcal{MN}}$	functional map between shapes \mathcal{M} and \mathcal{N}
$\Pi_{\mathcal{MN}}$	point-wise map between shapes \mathcal{M} and \mathcal{N}
$E(\cdot)$	energy of the functional map solver
$L(\cdot)$	loss function used during neural network training

used as the spectral embedding of each shape. Given c -dimensional features defined on each shape $F_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times c}, F_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}} \times c}$, either computed from handcrafted feature descriptors or extracted from learnable feature extractor, the functional map $C_{\mathcal{MN}} \in \mathbb{R}^{k \times k}$ associated with the spectral embedding can be computed by solving the continuous optimisation problem

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

Here, minimising $E_{\text{data}} = \|C\Phi_{\mathcal{M}}^{\dagger} F_{\mathcal{M}} - \Phi_{\mathcal{N}}^{\dagger} F_{\mathcal{N}}\|_F^2$ enforces descriptor preservation, while minimising the regularisation term E_{reg} imposes some form of structural properties (see e.g. [Ovsjanikov et al. 2012]). The operator \dagger denotes the Moore-Penrose inverse. From the optimal $C_{\mathcal{MN}}$, the point-wise map $\Pi_{\mathcal{NM}} \in \{0, 1\}^{n_{\mathcal{N}} \times n_{\mathcal{M}}}$ can be recovered based on the relationship $C_{\mathcal{MN}} = \Phi_{\mathcal{N}}^{\dagger} \Pi_{\mathcal{NM}} \Phi_{\mathcal{M}}$, e.g. either by nearest neighbour search or by other post-processing techniques [Ezuz et al. 2019b; Melzi et al. 2019b; Vestner et al. 2017].

3.2 Deep Functional Maps

With the recent progress in deep learning, many deep functional map methods [Roufosse et al. 2019; Sharma and Ovsjanikov 2020] have been proposed and lead to state-of-the-art matching performance. The common pipeline of those methods is shown in Fig. 2.

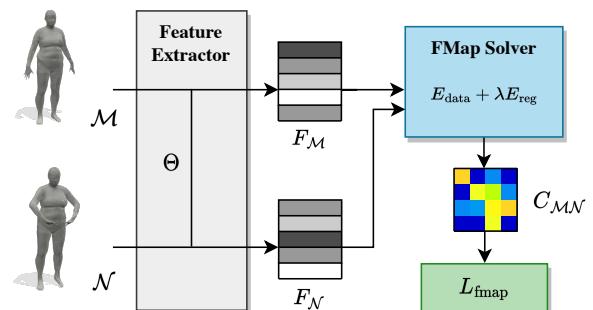


Fig. 2. **Common pipeline of deep functional map methods.** First, the feature extractor computes per-vertex features for each of the two input shapes. Then the functional map solver is used to compute the (bidirectional) functional map based on the extracted features. To train the feature extractor, structural regularisation is imposed on the computed functional maps.

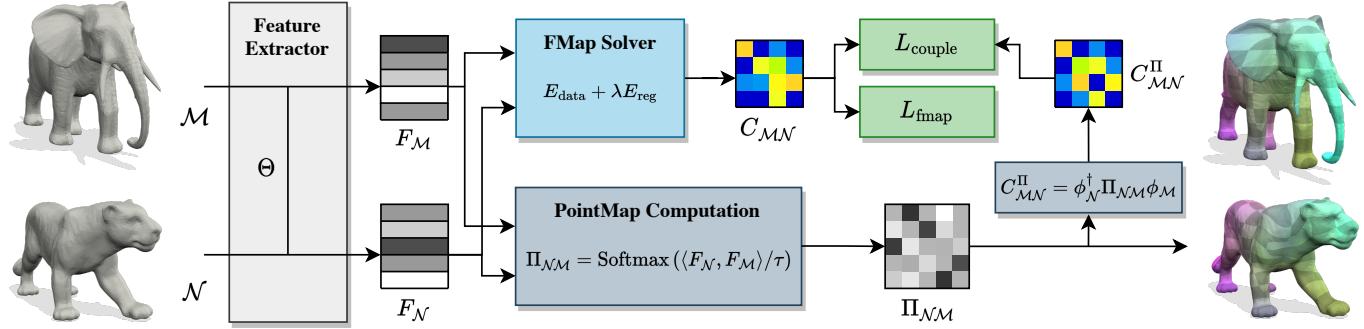


Fig. 3. Overview of our unsupervised robust spectral shape matching method. First, the feature extractor with shared weights Θ takes a pair of shapes M and N and extracts vertex-wise features F_M and F_N , respectively. Afterwards, the (non-trainable but differentiable) functional map solver is used to compute the functional map C_{MN} given extracted features. At the same time, the point-wise map Π_{NM} is obtained based on the similarity of the features F_M and F_N . Eventually, a coupling loss term is used to regularise the functional map C_{MN} to be associated to the point-wise map Π_{NM} .

3.2.1 Feature Extraction. Deep functional map methods use a neural network to learn features directly from training data, rather than using handcrafted features. Commonly, the feature extractor is used in a Siamese way, i.e. the same network with shared weights processes both shapes M and N . The extracted features F_M and F_N are then used in a differentiable functional map pipeline, cf. Sec. 3.1 and Fig. 4.

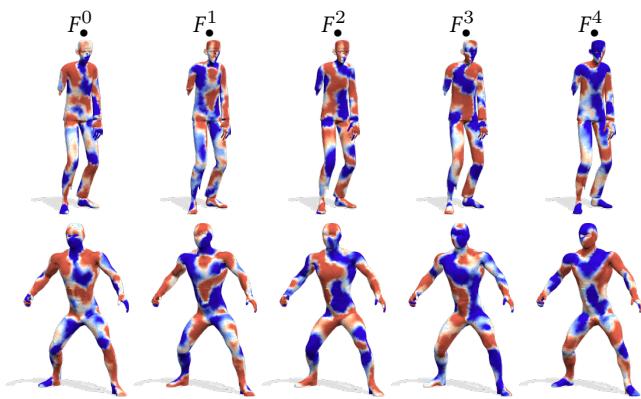


Fig. 4. Visualisation of the first five channels of the extracted features from the **feature extractor trained by our method** for a shape pair from DT4D-H test dataset. Smaller values are in blue, white represents values in-between and larger values are in red. Values outside of 25%–75% quantiles are clipped for improved visuals. Despite non-isometry and partiality (upper shape misses one arm) we can see that our network is able to predict consistent features across the shape.

3.2.2 Structural Regularisation. To train the feature extractor in an unsupervised way, a common strategy is to impose structural regularisation on the computed functional maps. Following Roufosse et al. [2019], structural functional map regularisation can be expressed in the form

$$L_{\text{fmap}} = \lambda_{\text{bij}} L_{\text{bij}} + \lambda_{\text{orth}} L_{\text{orth}}. \quad (2)$$

The bijectivity regularisation L_{bij} enforces the map from M through N back to M to be the identity map (and vice versa), i.e.

$$L_{\text{bij}} = \|C_{MN}C_{NM} - I\|_F^2 + \|C_{NM}C_{MN} - I\|_F^2. \quad (3)$$

The orthogonality regularisation L_{orth} prompts locally area-preserving matchings for both matching directions, i.e.

$$L_{\text{orth}} = \|C_{MN}^\top C_{MN} - I\|_F^2 + \|C_{NM}^\top C_{NM} - I\|_F^2. \quad (4)$$

In the context of partial shape matching, the functional map C_{MN} from complete shape M to partial shape N becomes a slanted diagonal matrix [Rodolà et al. 2017]. As a result, the terms L_{bij} and L_{orth} can be modified based on this property [Attaiki et al. 2021].

Most existing deep functional map methods follow the above formulation and only consider functional maps while *ignoring the relation between the functional map and the point-wise map*. In contrast, our approach explicitly considers this relation, which makes it possible to abandon the use of additional post-processing techniques to obtain point-wise maps during inference.

4 UNSUPERVISED ROBUST SPECTRAL SHAPE MATCHING

Our work aims to robustly estimate functional maps and point-wise maps for shape pairs in diverse settings, including near-isometric and non-isometric deformations, as well as shapes with partiality. The key insight of our method is based on the observation that existing learning-based methods solely focus on optimising functional maps and rely on some non-learnable post-processing techniques to obtain the point maps. However, the two-stage procedure often yields sub-optimal matching results, especially under challenging scenarios. In order to directly obtain point maps without post-processing, we explicitly consider the map relationship during training. To this end, we introduce a novel coupling loss to enable the optimisation of the functional map and the point-wise map simultaneously (unlike existing axiomatic methods, which often update them alternatingly). The whole framework of our approach is depicted in Fig. 3. Our framework has four main components:

feature extraction (Sec. 3.2.1), functional map solver (Sec. 3.1), differentiable point-wise map computation (Sec. 4.1) and unsupervised loss (Sec. 4.2).

4.1 Differentiable Point-Wise Map Computation

As illustrated in Sec. 3.2, most existing deep functional map methods use extracted features to compute functional maps and ignore the associated point-wise map. As a consequence, the existing functional map methods rely on off-the-shelf post-processing techniques during inference to obtain the final point-wise maps. When doing so, the relation between the functional and point-wise maps is underexplored, which in turn hampers the final matching quality as we demonstrate in Sec. 6. While recent work by Cao and Bernard [2022] introduced a universe classifier to compute *shape-to-universe* point-wise maps, it requires the knowledge of the number of universe vertices, which is typically unknown in practice. In contrast, our method directly computes the *pairwise* point-wise map based on the similarity measurement between features F_N and F_M defined on each shape. In theory, the point-wise map Π_{NM} should be a (partial) permutation matrix, i.e.

$$\left\{ \Pi \in \{0, 1\}^{n_N \times n_M} : \Pi \mathbf{1}_{n_M} = \mathbf{1}_{n_N}, \mathbf{1}_{n_N}^\top \Pi \leq \mathbf{1}_{n_M}^\top \right\}, \quad (5)$$

where the element at position (i, j) indicates whether the i -th point in N corresponds to the j -th point in M . In practice, we use the softmax operator to produce a soft correspondence matrix to make the computation differentiable, i.e.

$$\Pi_{NM} = \text{Softmax} \left(F_N F_M^T / \tau \right), \quad (6)$$

where τ is the scaling factor to determine the softness of the correspondence matrix. The softmax operator is applied in each row to ensure non-negativity and $\Pi_{NM} \mathbf{1}_{n_M} = \mathbf{1}_{n_N}$. Moreover, we convert the point-wise map to a functional map (via $\phi_N^\dagger \Pi_{NM} \phi_M$) for our unsupervised loss computation and test-time adaptation, which we introduce in the following.

4.2 Unsupervised Loss

Our unsupervised loss can be divided into two parts, structural regularisation (see Sec. 3.2.2) and a coupling between functional and point-wise maps.

4.2.1 Coupling Functional and Point-Wise Maps. The structural regularisation terms are not sufficient to obtain accurate matching results for shapes undergoing non-isometric deformations. The recent axiomatic approach by Ren et al. [2021] optimises for functional maps in the (discrete) constraint set

$$\left\{ C_{MN} \mid \exists \Pi_{NM}, \text{ s.t. } C_{MN} = \phi_N^\dagger \Pi_{NM} \phi_M \right\}, \quad (7)$$

which leads to more robust and accurate matchings. Inspired by this insight, we introduce our coupling loss term to ensure that the functional map is associated with a valid point-wise map, i.e.

$$L_{\text{couple}} = \left\| C_{MN} - \phi_N^\dagger \Pi_{NM} \phi_M \right\|_F^2. \quad (8)$$

Unlike the axiomatic approach by Ren et al. [2021], which formulates the relation between functional and point-wise maps as a discrete constraint, our approach allows to incorporate a soft correspondence

matrix as the point-wise map. To this end, the correspondence can be interpreted as a probabilistic matching over all vertices rather than the severely restricted case of vertex-to-vertex matching, see Fig. 5 for a conceptual 2D illustration. In practice, this is important for shape pairs with inconsistent discretisation, as we experimentally demonstrate in Sec. 6.1. In contrast to existing deep functional map methods, our approach uses the extracted features to compute both, functional maps and point-wise maps, in a differentiable way during training, and thus explicitly models their relationship and allows for their simultaneous optimisation.

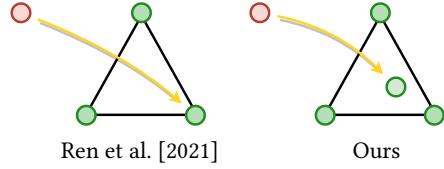


Fig. 5. Hard vs. soft correspondences in a conceptual 2D case. The method proposed by Ren et al. [2021] only allows for discrete matchings, i.e. the red vertex can only be matched to one of the green vertices of the triangle (left). In contrast, our method allows for a smooth matching, i.e. the red vertex can also be matched to points in the interior as a convex combination of the three vertices (right).

4.2.2 Total Unsupervised Loss. The overall unsupervised loss for training is a weighted sum of the individual losses, i.e.

$$L_{\text{total}} = L_{\text{fmap}} + \lambda_{\text{couple}} L_{\text{couple}}. \quad (9)$$

4.3 Inference and Test-Time Adaptation

In the following we explain how we obtain our final point-wise maps during inference.

4.3.1 Test-Time Adaptation. To improve the final matching performance, we propose a test-time adaptation strategy for each test shape pair individually. In contrast to post-processing techniques required by existing works, our test-time adaptation can be considered as a differentiable refinement process that directly optimises the feature extractor and seamlessly integrates into the learning paradigm. Specifically, we compute L_{total} defined in Eq. 9 and update the feature extractor in an iterative way during inference. The effect of this is that both the functional map and the point-wise map are updated simultaneously, since both of them are computed based on the extracted features. In contrast, most existing post-processing techniques optimise them either separately [Melzi et al. 2019b], or require an additional coupling term [Ren et al. 2018].

To address *non-isometric* matching, we introduce an additional smoothness penalty term for the point-wise map, which is based on the Dirichlet energy [Ezuz et al. 2019b; Magnet et al. 2022], i.e.

$$L_{\text{dirichlet}} = \left\| \Pi_{NM} X_M \right\|_{L_N}^2, \quad (10)$$

where $\|X\|_L^2 := \text{Trace}(X^\top L X)$. The smoothness penalty term encourages neighbouring vertices on shape M to be matched to neighbouring vertices on shape N . Imposing this smoothness penalty term is crucial for non-isometric shape matching, since a smooth

point-wise map is a desirable property for non-isometric matching. To avoid degenerate matchings (such as all vertices from shape \mathcal{M} being matched to the same vertex of shape \mathcal{N} , for which the Dirichlet energy becomes zero), we only apply this smoothness term during test-time adaptation.

4.3.2 Point-Wise Maps at Inference Time. Existing deep functional map methods mainly focus on computing accurate functional maps and either use nearest neighbour search or other post-processing techniques [Eisenberger et al. 2020a; Melzi et al. 2019b; Vestner et al. 2017] to obtain the final point-wise maps. In contrast, our method explicitly models the relation between the functional map and the point-wise map, and is thereby capable of directly obtaining the point-wise map without a functional map solver or other post-processing. During inference, our method directly obtains the point-wise map based on feature similarities using

$$\Pi_{\mathcal{NM}} = \text{NN}(\mathcal{F}_{\mathcal{N}}, \mathcal{F}_{\mathcal{M}}), \quad (11)$$

where NN denotes nearest neighbour search in $\mathcal{F}_{\mathcal{N}}$ for each entry in $\mathcal{F}_{\mathcal{M}}$.

Furthermore, in the near-isometric setting, our method obtains more robust and accurate correspondences by first converting the point-wise map to a functional map and then recovering the final point-wise map via nearest neighbour search in the spectral domain, i.e.

$$\Pi_{\mathcal{NM}}^{\text{iso}} = \text{NN}\left(\phi_{\mathcal{N}}\phi_{\mathcal{N}}^\dagger \Pi_{\mathcal{NM}}\phi_{\mathcal{M}}, \phi_{\mathcal{M}}\right), \quad (12)$$

where $\Pi_{\mathcal{NM}}$ is the soft correspondence matrix defined in Eq. 6. On the one hand, this inference strategy explicitly enforces the functional map $C_{\mathcal{MN}}$ to be associated with a valid point-wise map (by replacing it with $\phi_{\mathcal{N}}^\dagger \Pi_{\mathcal{NM}}\phi_{\mathcal{M}}$). On the other hand, this strategy can be understood as a low-pass filter on the computed point-wise map $\Pi_{\mathcal{NM}}$, thereby making it more robust against mismatches. Considering the complete set of eigenfunctions (i.e. $\phi_{\mathcal{N}}$ is of size $n_{\mathcal{N}} \times n_{\mathcal{N}}$, and $\phi_{\mathcal{N}}\phi_{\mathcal{N}}^\dagger = I$), Eq. 12 turns to be $\Pi_{\mathcal{NM}}^{\text{iso}} = \text{NN}(\Pi_{\mathcal{NM}}\phi_{\mathcal{M}}, \phi_{\mathcal{M}}) = \Pi_{\mathcal{NM}}$. By using the first k eigenfunctions, it is equivalent to applying a low-pass filter on the Fourier domain in signal processing [Ezuz and Ben-Chen 2017].

4.4 Implementation Details

We use DiffusionNet as our feature extractor and wave kernel signatures (WKS) [Aubry et al. 2011] as input features. Following Li et al. [2022], the dimension of the WKS is 128. For the spectral resolution, we choose the first 200 eigenfunctions of the Laplacian matrices as the spectral embedding. The dimension of the output features from DiffusionNet is 256 rather than 128 in the original implementation. The total number of learnable parameters is 510,336. In the context of the functional map computation, we use the regularised functional map solver proposed by Ren et al. [2019] and set $\lambda = 100$ in Eq. 1. For point-wise map computation, we set $\tau = 0.07$ in Eq. 6. In terms of our unsupervised loss, we empirically set $\lambda_{\text{bij}} = 1.0$, $\lambda_{\text{orth}} = 1.0$ in Eq. 2, and $\lambda_{\text{couple}} = 1.0$ in Eq. 9. For training and test-time adaptation, we use the Adam optimiser [Kingma and Ba 2015] with learning rate equal to 10^{-3} . For test-time adaptation, we choose the number of iterations equal to 15. In the context of non-isometric matching, the loss weight for the smoothness penalty term $L_{\text{dirichlet}}$ is empirically

set to be 5.0. We emphasise that throughout all experiments we use the same set of parameters.

5 EXPERIMENTAL RESULTS

In this section we evaluate our method and compare it to previous approaches in a broad range of different challenging scenarios. We refer the reader to our supplementary material, where we provide our complete shape matching results on all datasets, as well as our video, which shows additional animated results.

5.1 Near-Isometric Shape Matching

Datasets. We evaluate our method on three common near-isometric benchmark datasets: FAUST [Bogo et al. 2014], SCAPE [Anguelov et al. 2005] and SHREC’19 [Melzi et al. 2019a]. Instead of the original datasets, we choose the more challenging remeshed versions from [Donati et al. 2020; Ren et al. 2018]. The FAUST dataset consists of 100 human shapes (10 persons in 10 poses), where the training and testing split is 80/20. The SCAPE dataset contains 71 different poses of the same person, which is split into 51 and 20 shapes for training and testing. The SHREC’19 dataset is an even more challenging dataset with 44 human shapes and is used only as a test set.

Baselines. We perform an extensive comparison with existing non-rigid shape matching methods that are categorised as follows:

- (1) **Axiomatic approaches**, including BCICP [Ren et al. 2018], ZoomOut [Melzi et al. 2019b], Smooth Shells [Eisenberger et al. 2020a], and DiscreteOp [Ren et al. 2021];
- (2) **Supervised approaches**, including FMNet [Litany et al. 2017a], 3D-CODED [Groueix et al. 2018], HSN [Wiersma et al. 2020], ACSCNN [Li et al. 2020], GeomFMaps [Donati et al. 2020], and TransMatch [Trappolini et al. 2021]; and
- (3) **Unsupervised approaches**, including UnsupFMNet [Halimi et al. 2019], SURFMNet [Roufosse et al. 2019], WSupFMNet [Sharma and Ovsjanikov 2020], Deep Shells [Eisenberger et al. 2020b], NeuroMorph [Eisenberger et al. 2021], ConsistFMaps [Cao and Bernard 2022], DUO-FMNet [Donati et al. 2022a], and AttentiveFMaps [Li et al. 2022].

3D-CODED and TransMatch are template-based methods that require a large amount of training data. Therefore, the networks pre-trained on the large SURREAL dataset [Varol et al. 2017] are used. GeomFMaps and WSupFMNet are strong baselines related to our work. For a fair comparison, we use them with DiffusionNet as feature extractor, as it significantly improves shape matching accuracy [Sharp et al. 2020]. As for axiomatic approaches, they are initialised as originally proposed by respective authors.

Results. The mean geodesic error [Kim et al. 2011] is used for method evaluation. The results are summarised in Tab. 3. Our method outperforms the previous state-of-the-art in most settings, even in comparison to the supervised methods without relying on any additional post-processing techniques. Moreover, our method demonstrates substantially better cross-dataset generalisation ability compared to existing learning-based methods. Fig. 6 shows the proportion of correct keypoints (PCK) curves of our method in comparison to the existing state-of-the-art method [Li et al. 2022] for

Table 3. Near-isometric shape matching and cross-dataset generalisation on FAUST, SCAPE and SHREC’19. The numbers in parentheses show refined results using the indicated post-processing technique. The **best** results in each column are highlighted. Our method outperforms previous axiomatic, supervised and unsupervised methods in most settings without any post-processing techniques and demonstrates better cross-dataset generalisation ability (see columns in which *Train* and *Test* are different).

Train	FAUST			SCAPE			FAUST + SCAPE		
Test	FAUST	SCAPE	SHREC’19	FAUST	SCAPE	SHREC’19	FAUST	SCAPE	SHREC’19
Axiomatic Methods									
BCICP	6.1	11.0	-	6.1	11.0	-	6.1	11.0	-
ZoomOut	6.1	7.5	-	6.1	7.5	-	6.1	7.5	-
Smooth Shells	2.5	4.7	-	2.5	4.7	-	2.5	4.7	-
DiscreteOp	5.6	13.1	-	5.6	13.1	-	5.6	13.1	-
Supervised Methods									
FMNet (+ pmf)	11.0 (5.9)	30.0 (11.0)	-	33.0 (14.0)	17.0 (6.3)	-	-	-	-
3D-CODED	2.5	31.0	-	33.0	31.0	-	-	-	-
HSN	3.3	25.4	-	16.7	3.5	-	-	-	-
ACSCNN	2.7	8.4	-	6.0	3.2	-	-	-	-
GeomFMaps (+ zoomout)	2.6 (1.9)	3.4 (2.4)	9.9 (7.9)	3.0 (1.9)	3.0 (2.4)	12.2 (9.8)	2.6 (1.9)	2.9 (2.4)	7.9 (7.5)
TransMatch	1.7	30.4	14.5	15.5	12.0	37.5	1.6	11.7	10.9
Unsupervised Methods									
SURFMNet (+ icp)	15.0 (7.4)	32.0 (19.0)	-	32.0 (23.0)	12.0 (6.1)	-	33.0 (23.0)	29.0 (17.0)	-
UnsupFMNet (+ pmf)	10.0 (5.7)	29.0 (12.0)	-	22.0 (9.3)	16.0 (10.0)	-	11.0 (6.7)	13.0 (8.3)	-
WSupFMNet (+ zoomout)	3.8 (1.9)	4.8 (2.7)	-	3.6 (1.9)	4.4 (2.6)	-	3.6 (1.9)	4.5 (2.6)	-
Deep Shells	1.7	5.4	27.4	2.7	2.5	23.4	1.6	2.4	21.1
NeuroMorph	8.5	28.5	26.3	18.2	29.9	27.6	9.1	27.3	25.3
ConsistFMaps	1.5	3.2	19.7	3.2	2.0	28.3	1.7	3.2	17.8
DUO-FMNet	2.5	4.2	6.4	2.7	2.6	8.4	2.5	4.3	6.4
AttentiveFMaps	1.9	2.6	6.4	2.2	2.2	9.9	1.9	2.3	5.8
AttentiveFMaps-Fast	1.9	2.6	5.8	1.9	2.1	8.1	1.9	2.3	6.3
Ours	1.6	2.2	5.7	1.6	1.9	6.7	1.6	2.1	4.6

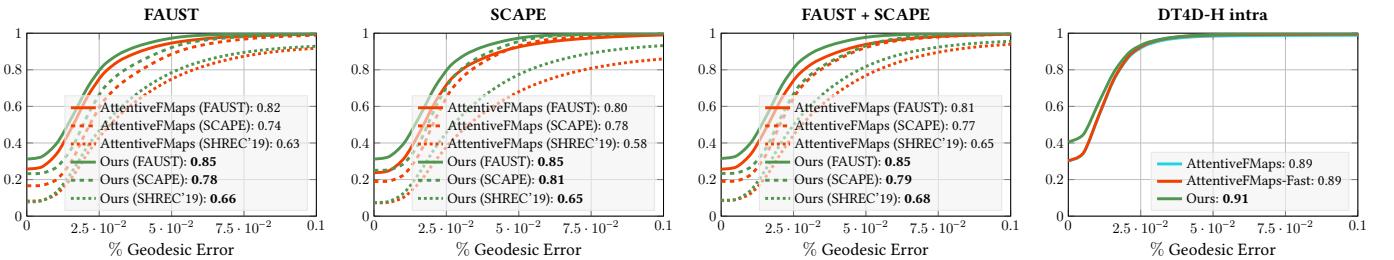


Fig. 6. Near-isometric shape matching on FAUST, SCAPE, SHREC’19 and DT4D-H intra and cross-dataset generalisation on FAUST, SCAPE and SHREC’19. Proportion of correct keypoints (PCK) curves and corresponding area under curve (scores in the legend) of our method in comparison to the existing state-of-the-art method. The title of each figure indicates the used training dataset and the names in parentheses shown in the legend correspond to the test dataset.

near-isometric shape matching. Fig. 7 shows some qualitative results on the SHREC’19 dataset trained on FAUST and SCAPE datasets.

5.2 Matching with Anisotropic Meshing

Datasets. To evaluate the robustness to different meshing, we follow Donati et al. [2022a] and use an anisotropic remeshed version of the FAUST and SCAPE datasets (denoted F_a and S_a, respectively). In this anisotropic remeshing the triangle scale is coordinate-dependent, so that methods that overfit to mesh connectivity are likely to fail to predict the correct maps. For a fair comparison, all

methods are trained on the original datasets and evaluated on the anisotropic remeshed ones.

Results. Tab. 4 summarises the matching performance. We observe that existing state-of-the-art methods [Cao and Bernard 2022; Eisenberger et al. 2020b] overfit to mesh connectivity and thus suffer from huge performance drops when testing on the anisotropic remeshed datasets. In contrast, our method is much more robust to varying mesh connectivity and outperforms the state of the art in most settings.



Fig. 7. **Cross-dataset generalisation on SHREC'19 (trained on FAUST and SCAPE).** Our method demonstrates previously unseen cross-dataset generalisation ability.

Table 4. **Anisotropic meshing on FAUST, SCAPE and their remeshed versions.** Our method outperforms previous state-of-the-art methods in most settings and is much more robust to varying mesh connectivity.

Train	FAUST				SCAPE			
Test	F	F_a	S	S_a	F	F_a	S	S_a
Axiomatic Methods								
BCICP	6.1	8.5	-	-	-	-	11.0	14.0
ZoomOut	6.1	8.7	-	-	-	-	7.5	15.0
Smooth Shells	2.5	5.4	-	-	-	-	4.7	5.0
DiscreteOp	5.6	6.2	-	-	-	-	13.1	14.6
Supervised Methods								
FMNet	11.0	43.0	30.0	44.0	33.0	43.0	17.0	41.0
GeomFMaps	2.6	3.2	3.4	3.8	3.0	8.4	3.0	3.1
Unsupervised Methods								
UnsupFMNet	10.0	42.0	29.0	43.0	22.0	44.0	16.0	44.0
Deep Shells	1.7	12.0	5.4	16.0	2.7	15.0	2.5	10.0
ConsistFMaps	1.5	15.3	3.2	14.1	3.2	23.3	2.0	4.9
DUO-FMNet	2.5	3.0	4.2	4.4	2.7	3.1	2.6	2.7
AttentiveFMaps	1.9	2.4	2.6	2.8	2.2	2.5	2.2	2.3
AttentiveFMaps-Fast	1.9	2.3	2.6	2.8	1.9	2.4	2.1	2.2
Ours	1.6	1.9	2.2	2.4	1.6	2.1	1.9	1.9

5.3 Matching with Topological Noise

Datasets. When working with data acquired from real scans, the mesh topology is often corrupted due to self-intersections of separate parts of the scanned objects. Such topological noise presents an additional challenge to shape matching methods as it distorts the intrinsic shape geometry non-isometrically. To evaluate our method’s robustness to topological noise, the TOPKIDS dataset [Lähner et al. 2016] is used, which contains synthetic shapes of children with topological merging based on the outer hull of intersecting shape parts. Due to the small amount of training data (26 shapes), in our comparison we consider only axiomatic and unsupervised methods. The ground truth correspondences are provided by matching the other shapes to one selected reference shape.

Results. The matching results are summarised in Tab. 5. Our method achieves the best matching performance and is much more robust against topological noise, in particular compared to existing methods based on functional maps. As shown in both Fig. 8 (middle) and Fig. 9, due to the topological degradation, none of the existing functional map methods provide satisfactory matching results (see the column *Fully intrinsic* in Tab. 5). In comparison to Deep Shells and NeuroMorph, our method does not rely on shape alignment in the spatial domain and thus does not require shape rigid alignment as pre-processing.

Table 5. **Topological noise on TOPKIDS.** Our method is more robust to topological noise compared to existing methods.

	Geo. error ($\times 100$)	TOPKIDS	Fully intrinsic
Axiomatic Methods			
ZoomOut	33.7	✓	
Smooth Shells	11.8	✗	
DiscreteOp	35.5	✓	
Unsupervised Methods			
UnsupFMNet	38.5	✓	
SURFMNet	48.6	✓	
WSupFMNet	47.9	✓	
Deep Shells	13.7	✗	
NeuroMorph	13.8	✗	
ConsistFMaps	39.3	✓	
AttentiveFMaps	23.4	✓	
AttentiveFMaps-Fast	28.5	✓	
Ours	9.2	✓	

5.4 Non-isometric Shape Matching

Datasets. Regarding non-isometric shape matching, we first consider the SMAL [Zuffi et al. 2017] dataset for inter-class shape matching. The SMAL dataset contains 49 four-legged animal shapes of eight species. Following Donati et al. [2022a], we use five species for training and three unseen species for testing, resulting in a 29/20 split of the dataset. To this end, no shapes similar to the test data are seen during training, and the deformation across species is highly non-isometric, presenting great challenges to existing approaches [Donati et al. 2022a]. We further test on a recent non-isometric shape matching benchmark introduced by Magnet et al. [2022], which is based on the large-scale animation dataset DeformingThings4D [Li et al. 2021]. Following Donati et al. [2022a], nine classes of humanoid shapes are used for evaluation, resulting in 198/95 shapes for training/testing. The dataset is denoted as DT4D-H.

Results. Tab. 6 summarises the matching results on the SMAL and DT4D-H datasets. Our approach outperforms the existing state-of-the-art on the challenging SMAL dataset even in comparison to supervised methods. Meanwhile, our method demonstrates near-perfect matching results for intra-class matching on the DT4D-H dataset. In the context of inter-class matching, our method outperforms existing axiomatic and unsupervised methods by a large margin and shows comparable matching performance for inter-class matching compared to the state-of-the-art supervised method.

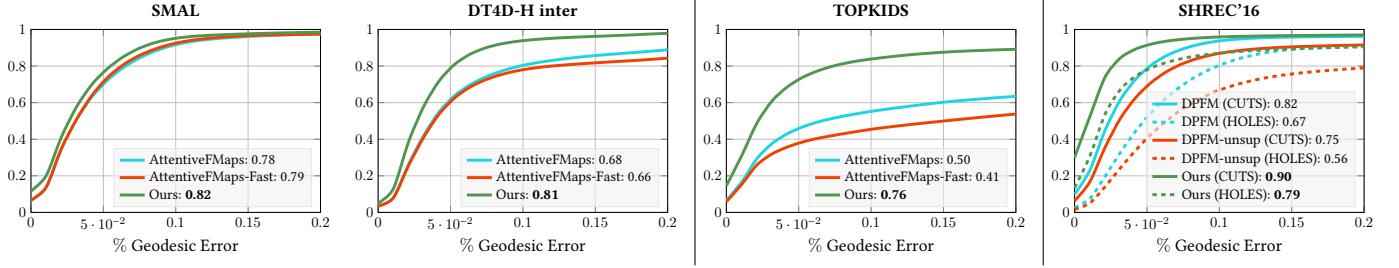


Fig. 8. (Left) **Non-isometric matching on SMAL and DT4D-H.** (Middle) **Matching with topological noise on TOPKIDS.** (Right) **Partial shape matching on SHREC'16.** Our method substantially outperforms existing state-of-the-art methods.

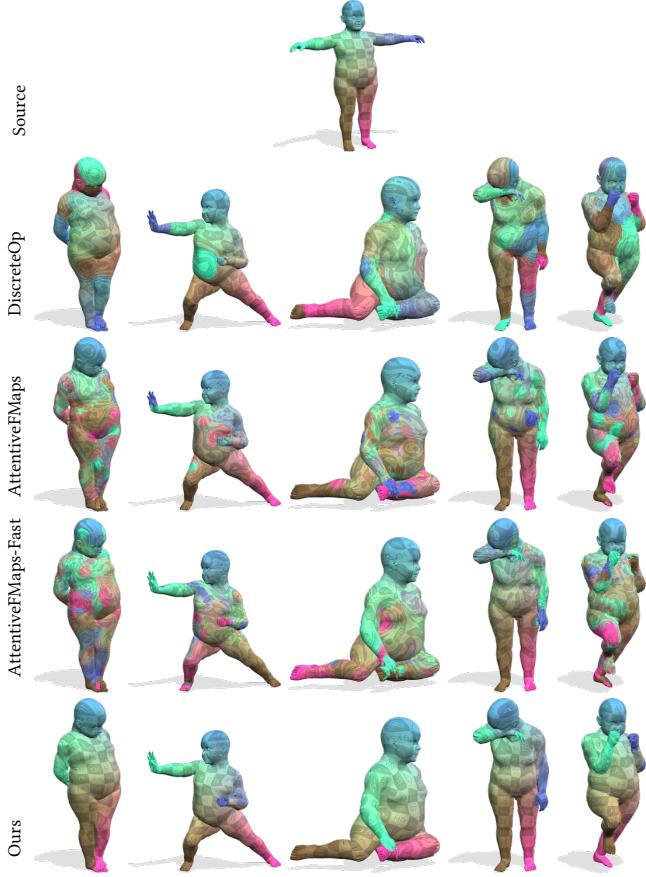


Fig. 9. **Topological noise on TOPKIDS.** Ours is the first deep functional map method that properly handles matching under topological noise.

In Fig. 8 (left), we summarise the PCK curves of our method in comparison to AttentiveFMaps [Li et al. 2022] on the SMAL and DT4D-H datasets. Fig. 11 provides some qualitative results of our method in comparison to recent state-of-the-art methods on both SMAL and DT4D-H datasets. We observe that our method consistently outperforms existing approaches even in comparison to supervised methods. To further demonstrate the superior performance of our method, we also consider the highly non-isometric SHREC'20

Table 6. **Non-isometric matching on SMAL and DT4D-H.** Our method sets to new state of the art on the SMAL dataset by a large margin. For DT4D-H inter-class matching, our method is the first unsupervised method that shows comparable performance to the state-of-the-art supervised method.

	Geo. error ($\times 100$)	SMAL		DT4D-H	
		intra-class	inter-class	intra-class	inter-class
Axiomatic Methods					
ZoomOut	38.4	4.0	29.0		
Smooth Shells	36.1	1.1	6.3		
DiscreteOp	38.1	3.6	27.6		
Supervised Methods					
FNet	42.0	9.6	38.0		
GeomFMaps	8.4	2.1	4.1		
Unsupervised Methods					
WSupFNet	7.6	3.3	22.6		
Deep Shells	29.3	3.4	31.1		
DUO-FNet	6.7	2.6	15.8		
AttentiveFMaps	5.4	1.7	11.6		
AttentiveFMaps-Fast	5.8	1.2	14.6		
Ours	3.9	0.9	4.1		

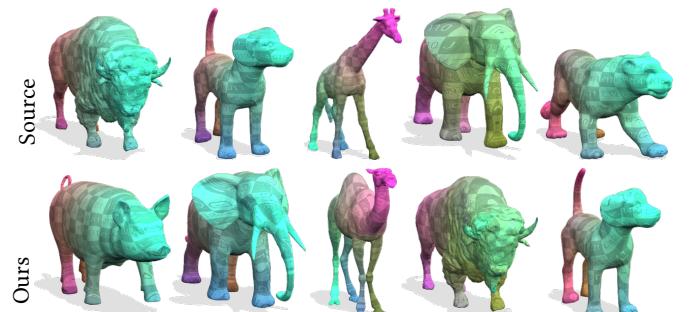


Fig. 10. **Non-isometric matching on SHREC'20.** Our method obtains smooth and accurate matching results even in the presence of large non-isometric shape deformations.

dataset [Dyke et al. 2020], which contains animals with extremely different appearance (e.g. giraffe and dog). Since SHREC'20 only contains 10 shapes, we directly use the training data for qualitative evaluation. The qualitative results are shown in Fig. 10. We



Fig. 11. Non-isometric matching on SMAL and DT4D-H. Comparing to existing methods, our approach demonstrates superior matching performance for both isometric and non-isometric shapes. For example, for shape matching between the lion and the hippo (the third column from the right), our method is the only one that provides smooth and accurate correspondences, while all competitors match the tail of the lion wrongly to the mouth of the hippo.

observe that our method is capable of obtaining smooth and accurate matching results even in the presence of large non-isometric deformations.

5.5 Partial Shape Matching

Datasets. We evaluate our method in the context of partial shape matching on the challenging SHREC’16 partial dataset [Cosmo et al. 2016]. The dataset contains 200 training shapes, categorised into 8 classes (humans and animals). Each class has a complete shape to which the other partial shapes are matched. The dataset is divided into two subsets, namely CUTS (missing a large part) with 120 pairs, and HOLES (missing many small parts) with 80 pairs. Following Attaiki et al. [2021], we train our method for each subset individually and evaluate it on the corresponding test set (200 shapes for each subset).

Baseline. Only a few existing shape matching approaches can be applied to partial shape matching. We categorise them as follows:

- (1) **Axiomatic approaches**, including PFM [Rodolà et al. 2017], FSP [Litany et al. 2017b];
- (2) **Supervised approaches**, including GeomFMaps [Donati et al. 2020], DPFM [Attaiki et al. 2021]; and
- (3) **Unsupervised approaches**, including DPFM-unsup [Attaiki et al. 2021], ConsistFMaps [Cao and Bernard 2022].

Results. Tab. 7 summarises the partial shape matching results on the SHREC’16 dataset. Our approach outperforms existing axiomatic and unsupervised methods by a large margin without relying on any post-processing techniques. In comparison to the supervised DPFM, our method achieves better performance on the more challenging HOLES subset. In the context of generalisation ability, our method substantially outperforms existing methods even in comparison to the supervised DPFM. Fig. 8 (right) summarises the PCK curves on the SHREC’16 partial datasets. Fig. 12 shows qualitative results on the SHREC’16 partial shape datasets of our method in comparison to other learning-based approaches. Compared to other

Table 7. Partial shape matching on SHREC’16. The numbers in parentheses show refined results using the indicated post-processing technique. Our method is the first unsupervised approach that bridges the huge performance gap between supervised and unsupervised methods.

Train	CUTS		HOLES	
Test	CUTS	HOLES	CUTS	HOLES
Axiomatic Methods				
PFM (+zoomout)	9.7 (9.0)	23.2 (22.4)	9.7 (9.0)	23.2 (22.4)
FSP (+zoomout)	16.1 (15.2)	33.7 (32.7)	16.1 (15.2)	33.7 (32.7)
Supervised Methods				
GeomFMaps (+zoomout)	12.8 (10.4)	20.6 (17.4)	19.8 (16.7)	15.3 (13.0)
DPFM (+zoomout)	3.2 (1.8)	15.8 (13.9)	8.6 (7.4)	13.1 (11.9)
Unsupervised Methods				
DPFM-unsup (+zoomout)	9.0 (7.8)	22.8 (20.0)	16.5 (14.6)	20.5 (18.4)
ConsistFMaps	8.4	23.7	15.7	17.9
Ours	3.3	13.7	5.2	9.1

learning-based approaches, we observe that our approach is capable of obtaining more accurate and smoother correspondences, even for partial shapes with several large missing parts. Overall, our method is the first unsupervised approach that bridges the huge performance gap between supervised and unsupervised methods.

5.6 Runtime Comparison

We evaluate the runtime of our method and compare it to the state-of-the-art supervised method GeomFMaps [Donati et al. 2020] (with and without ZoomOut [Melzi et al. 2019b]). In Fig. 13 (left) we show the runtime for shapes with a different number of vertices ranging from 1k to 20k when using the default setting of the number of eigenfunctions for our method ($k = 200$) and GeomFMaps ($k = 30$), respectively. We observe that our method is slower compared to GeomFMaps for shapes with a relatively small number of vertices. Nevertheless, when increasing the shape resolution, the runtime difference between GeomFMaps and our method is decreasing. In Fig. 13 (middle) we show the runtime for a different number of eigenfunctions ranging from 30 to 200. Here, we fix the number of vertices of the input shapes to 10k. We observe that our approach requires less computational time compared to GeomFMaps when the number of eigenfunctions is the same for both methods. The main reason is that our method does not compute functional maps (i.e. solving Eq. 1) or rely on any off-the-shelf post-processing techniques during inference. Overall, we consider matching a pair of high-resolution shapes within a few seconds as reasonable, especially in light of the remarkable performance of our method.

6 ABLATION STUDIES

In the following we perform ablative experiments to validate the merits of the individual components of our approach.

6.1 Coupling Loss, Inference and Test-Time Adaptation

We evaluate the importance of our introduced loss term L_{couple} in Eq. 8, as well as the proposed inference strategy and test-time adaptation described in Sec. 4.3. Additionally, we investigate the

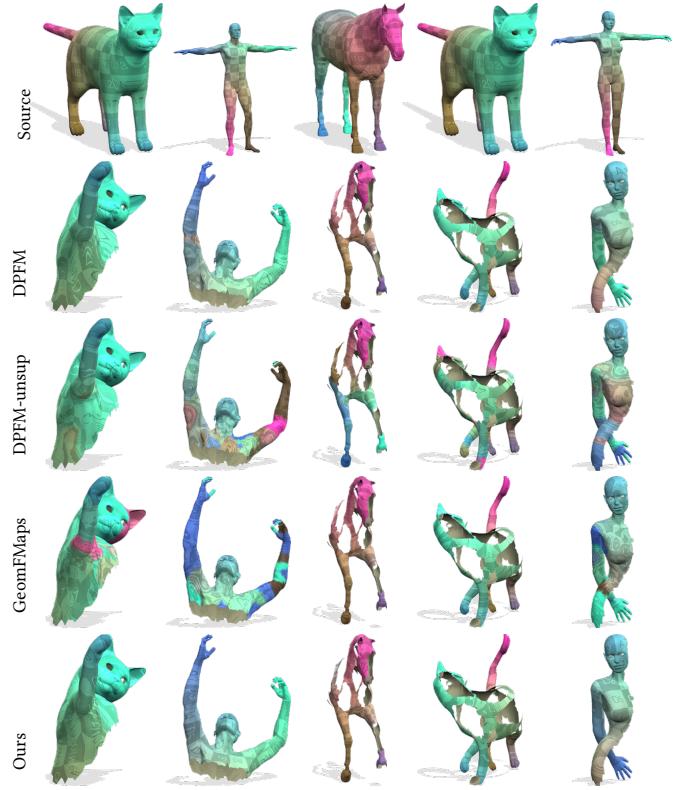


Fig. 12. Partial matching on SHREC’16. Our method achieves near-perfect matching performance, even for shapes with several large missing parts, and outperforms existing (supervised and unsupervised) methods.

advantage of using soft point-wise maps instead of hard ones. To this end, we consider the Gumbel-trick [Jang et al. 2016] to obtain hard point-wise maps during training. For all ablative experiments, we consider challenging non-isometric shape matching on the SMAL dataset.

Table 8. Ablation study on SMAL. The first row shows the network trained only with L_{fmap} in Eq. 2. In the second row we obtain the point-wise map based on the functional map computed by functional map solver. The third row represents inference without using test-time adaptation. The fourth row indicates not using E_d in Eq. 10 during test-time adaptation. In the fifth row we obtain the hard point-wise maps by using Gumbel trick.

Ablation Setting	SMAL
w.o. L_{couple}	10.3
w.o. our inference strategy	5.8
w.o. test-time adaptation	5.5
w.o. $L_{\text{dirichlet}}$ in test-time adaptation	4.3
w.o. soft point-wise maps	4.4
Ours	3.9

Results. Tab. 8 summarises the quantitative results. By comparing the first row and the last row, we can conclude that L_{couple} plays an important role for accurate matching. By comparing the second

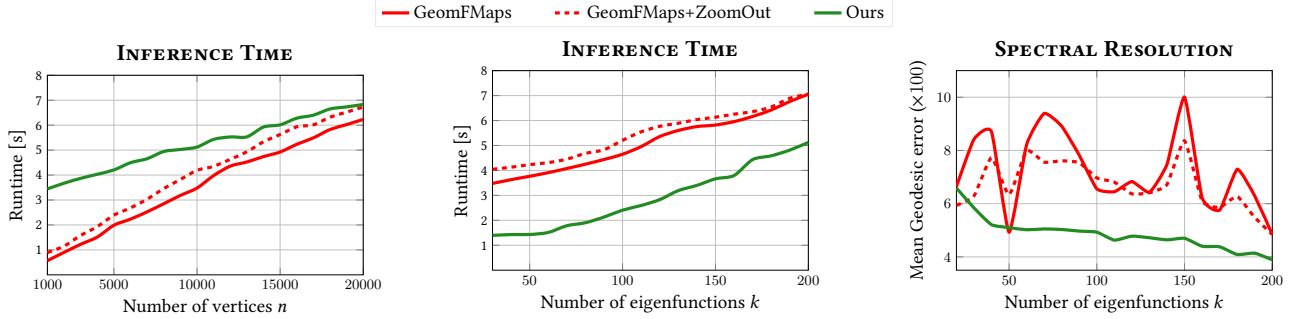


Fig. 13. Inference time and robustness to spectral resolution. We compare our method to the state-of-the-art supervised method GeomFMaps [Donati et al. 2020] (with and without ZoomOut [Melzi et al. 2019b]). **Left:** Runtime comparison with a different number of vertices. Compared to GeomFMaps, our method requires more computational time due to the choice of a larger number of eigenfunctions (200 versus 30). Nevertheless, for shapes with higher resolution, the runtime between GeomFMaps and our approach become comparable. **Middle:** Runtime comparison with a different number of eigenfunctions (while the number of vertices for shapes is fixed to 10k). Our method is faster than GeomFMaps when the number of eigenfunctions is the same for both methods. **Right:** We observe an unstable matching performance for GeomFMaps w.r.t. the spectral resolution. In contrast, our unsupervised method is more robust to the choice of the spectral resolution and consistently outperforms GeoFMaps.

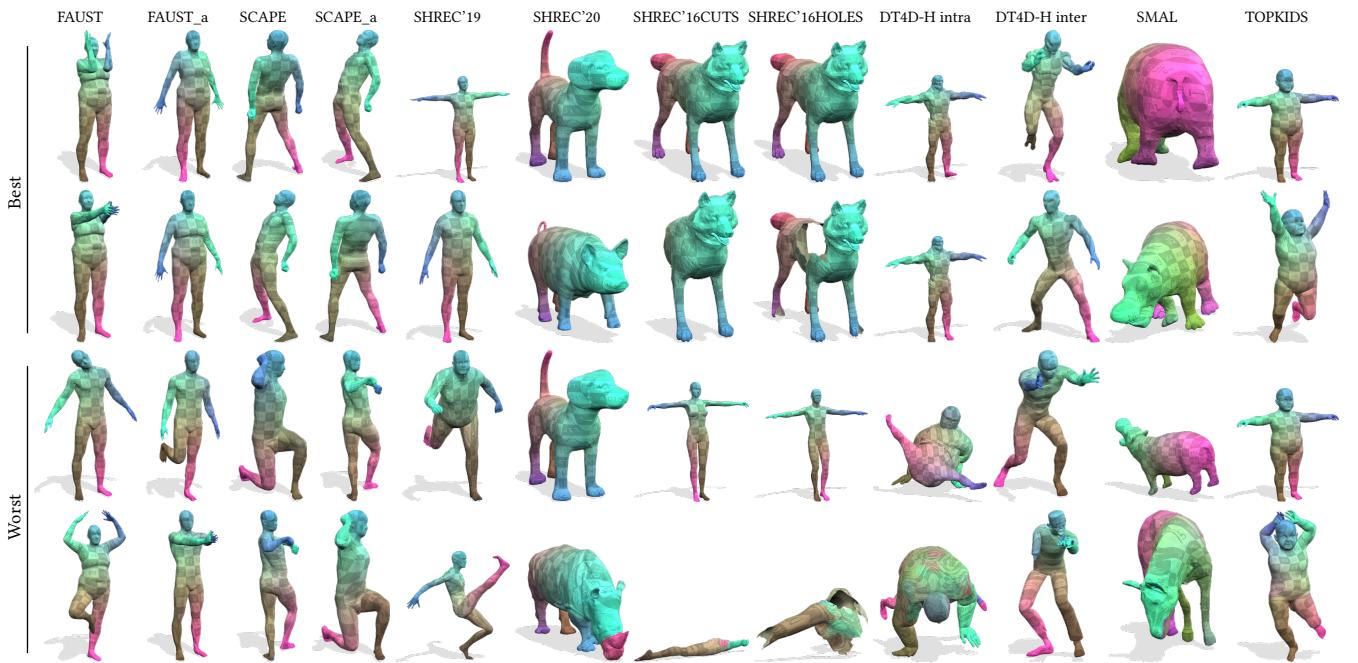


Fig. 14. Best and worst pair on each of the evaluated datasets w.r.t. geodesic error of the matchings computed with our approach. We can see that the worst matchings still form reasonable results in most cases, and that poor geodesic error scores originate from geometrically inconsistent matchings (e.g. nose of the rhino in the sixth column), or left-right flips (e.g. kid in the last column).

row and the last row, we observe that the point map based on deep feature similarity (last row) is more accurate than converting the functional map computed by the functional map solver to a point-wise map (second row). Comparing the third and the last row shows that using the test-time adaptation leads to better matching performance. Comparing the fourth and the last row confirms that using the smoothness term $L_{\text{dirichlet}}$ improves matching performance. By comparing the fifth row and the last row, we can conclude that

soft point-wise maps provides more flexibility for shape matching, which in turn leads to better matching performance.

6.2 Robustness Against Spectral Resolution

Previous functional map methods are sensitive to the choice of spectral resolution. In contrast, we demonstrate that our method is more robust to the choice of the spectral resolution. For this experiment we consider the challenging non-isometric SMAL dataset

and compare our method to the state-of-the-art supervised method GeomFMaps [Donati et al. 2020].

Results. Fig. 13 (right) shows the matching performance w.r.t. the choice of the spectral resolution. We observe that GeomFMaps is unstable during the increase of the spectral resolution. In contrast, our method is much more robust to the choice of the spectral resolution.

6.3 Our method as an axiomatic approach

Since we observe that the test-time adaptation leads to better matching results (see Tab. 8), we experimentally analyse whether our method is capable of being used as an axiomatic method by optimising the feature extractor on a single pair of shapes individually. To this end, we compare the individual optimisation strategy (randomly initialised network weights for each shape pair) to our unsupervised training strategy for challenging non-isometric shape matching on the SMAL dataset.

Table 9. Non-isometric matching on SMAL. We compare our unsupervised training strategy to the individual optimisation strategy. We observe a large performance drop when using our method as an axiomatic approach.

Unsupervised vs Axiomatic	SMAL
Axiomatic (individual optimisation)	43.1
Ours (unsupervised training)	3.9

Results. Tab. 9 summarises the quantitative results, which shows that using our pipeline as an axiomatic method does not work well. We believe that this is because it lacks the collective regularisation from large datasets, so that the severe non-convexity of the optimisation problem may lead to poor local optima.

7 DISCUSSION AND LIMITATIONS

Our proposed shape matching approach sets the new state of the art on a wide range of benchmark datasets. Yet, there are also some limitations that give rise to interesting future research questions. Fig. 14 summarises the best and worst pair on each of the evaluated datasets w.r.t. geodesic error of the matchings computed with our method. We observe that most failure cases stem from geometrically inconsistent matchings and left-right flips. Fig. 15 showcases some hand-picked failure modes of our method. While such cases occur rarely (see our complete results in the supplementary material), these examples confirm the difficulty of matching under extreme conditions, such as large non-isometries between an elephant and a giraffe.

Our method is robust to the choice of spectral resolution, see Fig. 13. Opposed to previous works for which this choice is critical, our robustness allows us to choose a relatively large number of eigenfunctions. With that, high-frequency features are adequately retained, so that in turn symmetric flips that commonly occur in competing methods (e.g. of humans) can properly be resolved via the preserved geometric details (e.g. of the face). Yet, the number of used eigenfunctions still needs to be chosen carefully in order to balance matching performance and runtime. Further improving the runtime of our method, as well as finding a reasonable trade-off



Fig. 15. Examples of failure modes of our method. **Left:** Training our method exclusively on complete shapes (FAUST and SCAPE) and then testing on partial data leads to matching failures (shape 40 of the SHREC’19 dataset, i.e. the shape on the top, has a large missing part in the upper leg region). **Middle:** Extreme non-isometries, such as between an elephant and a giraffe, may lead to erroneous matchings (SHREC’20). **Right:** Cases with severe topological noise may also lead to local mismatches (TOPKIDS).

between matching performance and runtime, are both interesting future directions.

In this work we only consider the case of pairwise shape matching and we leave the investigation of respective multi-shape matching formulations, e.g. based on cycle-consistency constraints [Bernard et al. 2019; Cao and Bernard 2022; Gao et al. 2021; Huang and Guibas 2013; Huang et al. 2020], for future work. Furthermore, follow-up works might integrate the vertex-to-triangle matching formalism [Ezuz and Ben-Chen 2017] into our learning framework. Another interesting direction for future work is a generalisation of our approach to other types of shape representation, such as point clouds or neural shape representations.

8 CONCLUSION

In this work we propose the first unsupervised learning approach that provides a universal framework for shape matching under different challenging scenarios, including shapes with different discretisation, shapes with topological noise, highly non-isometric shapes, and partial shapes. The favourable matching quality, robustness and cross-dataset generalisation ability of our method was enabled mainly due to the combination of three aspects: differentiable point-wise map computation, coupling point-wise maps with functional maps, and test-time adaptation. We have demonstrated that by combining these aspects, many previous limitations of functional maps can be addressed, including the sensitivity to the spectral resolution, the need for post-processing to obtain point-wise maps, and the unsatisfactory matching quality in the case of non-isometric shape deformations. Overall, by bridging the gap between the theoretical advances of deep functional maps and their practical application in real-world settings, we believe that our method will be a valuable contribution to the field of computer graphics and beyond.

ACKNOWLEDGMENTS

Funded by the TRA Modellling (University of Bonn) as part of the Excellence Strategy of the federal and state governments and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 458610525.

REFERENCES

- Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. 2005. Scape: shape completion and animation of people. In *ACM SIGGRAPH*.
- Souhaib Attaike and Maks Ovsjanikov. 2022. NCP: Neural Correspondence Prior for Effective Unsupervised Shape Matching. In *NeurIPS*.
- Souhaib Attaike, Gautam Pai, and Maks Ovsjanikov. 2021. Dpfm: Deep partial functional maps. In *International Conference on 3D Vision (3DV)*.
- Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. 2011. The wave kernel signature: A quantum mechanical approach to shape analysis. In *ICCV*.
- Omri Azenctot and Rongjie Lai. 2021. A Data-Driven Approach to Functional Map Construction and Bases Pursuit. In *Computer Graphics Forum*, Vol. 40. Wiley Online Library, 97–110.
- Florian Bernard, Zeeshan Khan Suri, and Christian Theobalt. 2020. Mina: Convex mixed-integer programming for non-rigid shape alignment. In *CVPR*.
- Florian Bernard, Johan Thunberg, Paul Swoboda, and Christian Theobalt. 2019. Hippi: Higher-order projected power iterations for scalable multi-matching. In *ICCV*.
- Federica Bogo, Javier Romero, Matthew Loper, and Michael J Black. 2014. FAUST: Dataset and evaluation for 3D mesh registration. In *CVPR*.
- Michael M Bronstein and Iasonas Kokkinos. 2010. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *CVPR*.
- Dongliang Cao and Florian Bernard. 2022. Unsupervised Deep Multi-shape Matching. In *ECCV*.
- Aharon Cohen and Mirela Ben-Chen. 2020. Robust shape collection matching and correspondence from shape differences. In *Computer Graphics Forum*. Wiley Online Library.
- Luca Cosmo, Emanuele Rodola, Michael M Bronstein, Andrea Torsello, Daniel Cremers, and Y Sahillioglu. 2016. SHREC’16: Partial matching of deformable shapes. *Proc. 3DOR* 2, 9 (2016), 12.
- Huong Quynh Dinh, Anthony Yezzi, and Greg Turk. 2005. Texture transfer during shape transformation. *ACM Transactions on Graphics (ToG)* 24, 2 (2005), 289–310.
- Nicolas Donati, Etienne Corman, Simone Melzi, and Maks Ovsjanikov. 2022b. Complex functional maps: A conformal link between tangent bundles. In *Computer Graphics Forum*. Wiley Online Library.
- Nicolas Donati, Etienne Corman, and Maks Ovsjanikov. 2022a. Deep Orientation-Aware Functional Maps: Tackling Symmetry Issues in Shape Matching. In *CVPR*.
- Nicolas Donati, Abhishek Sharma, and Maks Ovsjanikov. 2020. Deep geometric functional maps: Robust feature learning for shape correspondence. In *CVPR*.
- Roberto M. Dyke, Yu-Kun Lai, Paul L. Rosin, Stefano Zappalà, Seana Dykes, Daoliang Guo, Kun Li, Riccardo Marin, Simone Melzi, and Jingyu Yang. 2020. SHREC’20: Shape correspondence with non-isometric deformations. *Computers & Graphics* 92 (2020), 28–43.
- Bernhard Egger, William AP Smith, Ayush Tewari, Stefanie Wuhrer, Michael Zollhoefer, Thabo Beeler, Florian Bernard, Timo Bolkart, Adam Kortylewski, Sami Romdhani, et al. 2020. 3d morphable face models—past, present, and future. *ACM Transactions on Graphics (ToG)* 39, 5 (2020), 1–38.
- Marvin Eisenberger, Zorah Lähner, and Daniel Cremers. 2019. Divergence-Free Shape Correspondence by Deformation. In *Computer Graphics Forum*. Wiley Online Library.
- Marvin Eisenberger, Zorah Lähner, and Daniel Cremers. 2020a. Smooth shells: Multi-scale shape registration with functional maps. In *CVPR*.
- Marvin Eisenberger, David Novotny, Gael Kerchenbaum, Patrick Labatut, Natalia Neverova, Daniel Cremers, and Andrea Vedaldi. 2021. Neuromorph: Unsupervised shape interpolation and correspondence in one go. In *CVPR*.
- Marvin Eisenberger, Aysim Toker, Laura Leal-Taixé, and Daniel Cremers. 2020b. Deep shells: Unsupervised shape correspondence with optimal transport. *NIPS* (2020).
- Davide Eynard, Emanuele Rodola, Klaus Glashoff, and Michael M Bronstein. 2016. Coupled functional maps. In *2016 Fourth International Conference on 3D Vision (3DV)*.
- Danielle Ezuz and Mirela Ben-Chen. 2017. Deblurring and denoising of maps between shapes. In *Computer Graphics Forum*. Wiley Online Library.
- Danielle Ezuz, Behrend Heeren, Omri Azenctot, Martin Rumpf, and Mirela Ben-Chen. 2019a. Elastic correspondence between triangle meshes. In *Computer Graphics Forum*, Vol. 38. Wiley Online Library, 121–134.
- Danielle Ezuz, Justin Solomon, and Mirela Ben-Chen. 2019b. Reversible harmonic maps between discrete surfaces. *ACM Transactions on Graphics (ToG)* 38, 2 (2019), 1–12.
- Maolin Gao, Zorah Lähner, Johan Thunberg, Daniel Cremers, and Florian Bernard. 2021. Isometric multi-shape matching. In *CVPR*.
- Thibault Groueix, Matthew Fisher, Vladimir G Kim, Bryan C Russell, and Mathieu Aubry. 2018. 3d-coded: 3d correspondences by deep deformation. In *ECCV*.
- Oshri Halimi, Or Litany, Emanuele Rodola, Alex M Bronstein, and Ron Kimmel. 2019. Unsupervised learning of dense shape correspondence. In *CVPR*.
- Benjamin Holzschuh, Zorah Lähner, and Daniel Cremers. 2020. Simulated annealing for 3d shape correspondence. In *2020 International Conference on 3D Vision (3DV)*.
- Qixing Huang, Fan Wang, and Leonidas Guibas. 2014. Functional map networks for analyzing and exploring large shape collections. *ACM Transactions on Graphics (ToG)* 33, 4 (2014), 1–11.
- Qi-Xing Huang, Bart Adams, Martin Wicke, and Leonidas J Guibas. 2008. Non-rigid registration under isometric deformations. In *Computer Graphics Forum*. Wiley Online Library.
- Qi-Xing Huang and Leonidas Guibas. 2013. Consistent shape maps via semidefinite programming. In *Computer Graphics Forum*. Wiley Online Library.
- Ruqi Huang, Jing Ren, Peter Wonka, and Maks Ovsjanikov. 2020. Consistent zoomout: Efficient spectral map synchronization. In *Computer Graphics Forum*. Wiley Online Library.
- Eric Jang, Shixiang Gu, and Ben Poole. 2016. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144* (2016).
- Vladimir G Kim, Yaron Lipman, and Thomas Funkhouser. 2011. Blended intrinsic maps. *ACM Transactions on Graphics (ToG)* 30, 4 (2011), 1–12.
- Diederik P Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *ICLR*.
- Artiom Kovnatsky, Michael M Bronstein, Alexander M Bronstein, Klaus Glashoff, and Ron Kimmel. 2013. Coupled quasi-harmonic bases. In *Computer Graphics Forum*. Wiley Online Library.
- Zorah Lähner, Emanuele Rodola, Michael M Bronstein, Daniel Cremers, Oliver Burghard, Luca Cosmo, Andreas Dieckmann, Reinhard Klein, and Yusuf Sahillioglu. 2016. SHREC’16: Matching of deformable shapes with topological noise. *Proc. 3DOR* 2, 10.2312 (2016).
- Lei Li, Nicolas Donati, and Maks Ovsjanikov. 2022. Learning Multi-resolution Functional Maps with Spectral Attention for Robust Shape Matching. *NIPS* (2022).
- Qinsong Li, Shengjun Liu, Ling Hu, and Xinru Liu. 2020. Shape correspondence using anisotropic Chebyshev spectral CNNs. In *CVPR*.
- Yang Li, Hikari Takehara, Takafumi Taketomi, Bo Zheng, and Matthias Nießner. 2021. 4dccomplete: Non-rigid motion estimation beyond the observable surface. In *ICCV*.
- Or Litany, Tal Remez, Emanuele Rodola, Alex Bronstein, and Michael Bronstein. 2017a. Deep functional maps: Structured prediction for dense shape correspondence. In *ICCV*.
- Or Litany, Emanuele Rodola, Alexander M Bronstein, and Michael M Bronstein. 2017b. Fully spectral partial shape matching. In *Computer Graphics Forum*. Wiley Online Library.
- Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. 2015. SMPL: A skinned multi-person linear model. *ACM Transactions on Graphics (ToG)* 34, 6 (2015), 1–16.
- Robin Magnet, Jing Ren, Olga Sorkine-Hornung, and Maks Ovsjanikov. 2022. Smooth non-rigid shape matching via effective Dirichlet energy optimization. In *International Conference on 3D Vision (3DV)*.
- Riccardo Marin, Marie-Julie Rakotosaona, Simone Melzi, and Maks Ovsjanikov. 2020. Correspondence learning via linearly-invariant embedding. In *NeurIPS*.
- Simone Melzi, Riccardo Marin, Emanuele Rodola, Umberto Castellani, Jing Ren, Adrien Poulenard, Peter Wonka, and Maks Ovsjanikov. 2019a. Shrec 2019: Matching humans with different connectivity. In *Eurographics Workshop on 3D Object Retrieval*.
- Simone Melzi, Jing Ren, Emanuele Rodola, Abhishek Sharma, Peter Wonka, and Maks Ovsjanikov. 2019b. ZoomOut: spectral upsampling for efficient shape correspondence. *ACM Transactions on Graphics (ToG)* 38, 6 (2019), 1–14.
- Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. 2012. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (ToG)* 31, 4 (2012), 1–11.
- Maks Ovsjanikov, Quentin Mérigot, Facundo Mémoli, and Leonidas Guibas. 2010. One point isometric matching with the heat kernel. In *Computer Graphics Forum*. Wiley Online Library.
- Gautam Pai, Jing Ren, Simone Melzi, Peter Wonka, and Maks Ovsjanikov. 2021. Fast sinkhorn filters: Using matrix scaling for non-rigid shape correspondence with functional maps. In *CVPR*.
- Ulrich Pinkall and Konrad Polthier. 1993. Computing discrete minimal surfaces and their conjugates. *Experimental mathematics* 2, 1 (1993), 15–36.
- Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *NIPS* (2017).
- Jing Ren, Simone Melzi, Maks Ovsjanikov, and Peter Wonka. 2020. MapTree: recovering multiple solutions in the space of maps. *ACM Transactions on Graphics (ToG)* 39, 6 (2020), 1–17.
- Jing Ren, Simone Melzi, Peter Wonka, and Maks Ovsjanikov. 2021. Discrete optimization for shape matching. In *Computer Graphics Forum*. Wiley Online Library.

- Jing Ren, Mikhail Panine, Peter Wonka, and Maks Ovsjanikov. 2019. Structured regularization of functional map computations. In *Computer Graphics Forum*. Wiley Online Library.
- Jing Ren, Adrien Poulenard, Peter Wonka, and Maks Ovsjanikov. 2018. Continuous and orientation-preserving correspondences via functional maps. *ACM Transactions on Graphics (ToG)* 37 (2018), 1–16.
- Emanuele Rodolà, Luca Cosmo, Michael M Bronstein, Andrea Torsello, and Daniel Cremers. 2017. Partial functional correspondence. In *Computer Graphics Forum*. Wiley Online Library.
- Emanuele Rodolà, Michael Moeller, and Daniel Cremers. 2015. Point-wise map recovery and refinement from functional correspondence. *arXiv preprint arXiv:1506.05603* (2015).
- Emanuele Rodolà, Michael Moeller, and Daniel Cremers. 2017. Regularized pointwise map recovery from functional correspondence. In *Computer Graphics Forum*. Wiley Online Library.
- Paul Roetzer, Paul Swoboda, Daniel Cremers, and Florian Bernard. 2022. A Scalable Combinatorial Solver for Elastic Geometrically Consistent 3D Shape Matching. In *CVPR*.
- Jean-Michel Roufosse, Abhishek Sharma, and Maks Ovsjanikov. 2019. Unsupervised deep learning for structured shape matching. In *ICCV*.
- Yusuf Sahillioglu. 2020. Recent advances in shape correspondence. *The Visual Computer* 36, 8 (2020), 1705–1721.
- Samuele Salti, Federico Tombari, and Luigi Di Stefano. 2014. SHOT: Unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding* 125 (2014), 251–264.
- Abhishek Sharma and Maks Ovsjanikov. 2020. Weakly supervised deep functional maps for shape matching. *NIPS* (2020).
- Nicholas Sharp, Souhaib Attaiki, Keenan Crane, and Maks Ovsjanikov. 2020. Diffusion-net: Discretization agnostic learning on surfaces. *arXiv preprint arXiv:2012.00888* (2020).
- Robert W Sumner and Jovan Popović. 2004. Deformation transfer for triangle meshes. *ACM Transactions on Graphics (ToG)* 23, 3 (2004), 399–405.
- Gary KL Tam, Zhi-Quan Cheng, Yu-Kun Lai, Frank C Langbein, Yonghuai Liu, David Marshall, Ralph R Martin, Xian-Fang Sun, and Paul L Rosin. 2012. Registration of 3D point clouds and meshes: A survey from rigid to nonrigid. *IEEE transactions on visualization and computer graphics* 19, 7 (2012), 1199–1217.
- Hugues Thomas, Charles R Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J Guibas. 2019. Kpconv: Flexible and deformable convolution for point clouds. In *ICCV*.
- Giovanni Trappolini, Luca Cosmo, Luca Moschella, Riccardo Marin, Simone Melzi, and Emanuele Rodolà. 2021. Shape registration in the time of transformers. *NIPS* (2021).
- Oliver Van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. 2011. A survey on shape correspondence. In *Computer Graphics Forum*. Wiley Online Library.
- Gul Varol, Javier Romero, Xavier Martin, Naureen Mahmood, Michael J Black, Ivan Laptev, and Cordelia Schmid. 2017. Learning from synthetic humans. In *CVPR*.
- Matthias Vestner, Roee Litman, Emanuele Rodolà, Alex Bronstein, and Daniel Cremers. 2017. Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space. In *CVPR*.
- Fan Wang, Qixing Huang, and Leonidas J Guibas. 2013. Image co-segmentation via consistent functional maps. In *ICCV*.
- Ruben Wiersma, Elmar Eisemann, and Klaus Hildebrandt. 2020. Cnns on surfaces using rotation-equivariant features. *ACM Transactions on Graphics (ToG)* 39, 4 (2020), 92–1.
- Thomas Windheuser, Ulrich Schlickewei, Frank R Schmidt, and Daniel Cremers. 2011. Geometrically consistent elastic matching of 3d shapes: A linear programming solution. In *ICCV*.
- Silvia Zuffi, Angjoo Kanazawa, David W Jacobs, and Michael J Black. 2017. 3D menagerie: Modeling the 3D shape and pose of animals. In *CVPR*.