IUST
Iran University of Science and Technology

Iranian Journal of Electrical and Electronic Engineering

Journal Homepage: ijeee.iust.ac.ir

Efficient 3D Shape Matching: Dense Correspondence for non-isometric Deformation

Amirreza Amirfathiyan*, Hossein Ebrahimnezhad* (C.A.).

Abstract: This paper presents an application of deep learning in computer graphics, utilizing learn-based networks for 3D shape matching. We propose an efficient method for shape matching between 3D models with non-isometric deformation. Our method organizes intrinsic and directional attributes in a structured manner. For this purpose, we use a hybrid feature derived from Diffusion-Net and spectral features. In fact, we combine learned-based intrinsic properties with orientation-preserving features and demonstrate the effectiveness of our method. We achieve this by first extracting features from Diffusion-Net. Then, we compute two maps based on the functional map networks to obtain intrinsic and directional features. Finally, we combine them to achieve a desired map that can resolve symmetry ambiguities on models with high deformation. Quantitative results on the TOSCA dataset indicate that the proposed method achieves lowest average geodetic error of 0.0023, outperforming state-of-the-art methods and reducing the error by 70.66%. We demonstrate that our method outperforms similar approaches by leveraging an accurate feature extractor and effective geometric regularizers, allowing for better handling of non-isometric shapes and resulting in reduced matching errors.

Keywords: 3D Shape Correspondence, 3D Shape Matching, Deep learning, Orientation Preserving.

1 Introduction

MACHINE learning for shapes matching is a critical issue in 3D shape analysis with applications in statistical shape analysis [1] and deformation transfer [2]. Early methods either concentrated on learning informative features to ensure that corresponding points have similar descriptors [3] or approached shape correspondence as a semantic segmentation problem. Approaches such as [4] and [5] aim to predict the vertex index on a ground truth template shape for each point on the surface. However, these methods offer minimal consistency in point correspondence predictions and may be sensitive to the underlying shape discretization [6]. Recent techniques have concentrated on predicting and applying a training loss across the complete mapping between shape pairs. This progress has been significantly aided by spectral methods, particularly the functional map representation [7], which encodes a mapping as a compact matrix utilizing the Laplacian Eigen-basis. Various approaches utilizing both supervised [8] and unsupervised losses [9] have been proposed using functional map representation. Central to these methods is the learning of feature functions that predict the overall functional map. Recent studies have demonstrated that this approach decreases the required training data, offers strong regularization for smoother maps, enhances the robustness of learned features to discretization changes, and removes the need for a fixed template shape. While using the compact functional map representation introduces a strong bias towards smooth approximately isometric correspondences, however, it still allows for both orientation-preserving and orientation-reversing correspondences. That means that the approach being discussed allows for the possibility

Iranian Journal of Electrical & Electronic Engineering, 2024.
Paper first received 19 October 2024 and accepted 15 December 2024.
* The authors are with the Department of Electrical Engineering, Sahand University of Technology, Tabriz, Iran.
E-mail: ebrahimnezhad@sut.ac.ir.
Corresponding Author: Hossein Ebrahimnezhad.

of matching points in a way that either maintains the original orientation (orientation preserving) or reverses it (orientation reversing). In other words, it can accommodate both types of correspondences between shapes or objects. This work permits a nuanced exploration of shape relations, enabling the mapping of geometric features with diverse properties. The compact functional map formulation leverages spectral information intrinsic to the manifold, facilitating the discovery of correspondences that respect both local and global geometric structures. Consequently, it becomes a powerful tool for applications such as shape matching, deformation analysis, and even shape retrieval, where the preservation of intrinsic features remains paramount. Although this property of functional maps is beneficial for tasks like symmetry detection, most practical applications require preserving orientation. However, restricting maps to orientation-preserving ones using this framework is not straightforward, often resulting in both local and global symmetry flips, leading to left/right ambiguities in organic shapes. Consequently, current state-of-the-art learning networks necessitate either a supervised loss [10] and [6] or a rigid pre-alignment [11] to resolve these symmetries.

In this paper, we use the recently proposed functional map representation [12] for overcoming ambiguities of symmetry that are based on the alignment of tangent vector fields. We propose a strategy that utilizes functional maps to learn features that align tangent bundles on surfaces. This approach aids in creating orientation-preserving maps and regularizes the learning process. We present loss functions tailored for functional maps, showing that our network can be trained entirely in an unsupervised manner, independent of rigid pre alignments or ground truth correspondences. The orientation-preserving features learned from our approach offer valuable signals for 3D shape matching tasks.

2 Related Works

3D shape matching is a well-established research area. Here, we review the most relevant works, emphasizing learning-based and unsupervised techniques. For a more comprehensive overview, interested readers can refer to surveys [13]. Our method builds on the functional map representation introduced in [7] and extended in various subsequent works, such as [14-18]. This framework's key advantage is its ability to represent and optimize maps as compact matrices, facilitate robust linear-algebraic regularization, and adapt to partial settings [19, 20]. In using this representation, the “descriptor” functions are important. They are used to estimate the underlying functional maps and must be computed. Early methods utilized descriptors like HKS [21] or WKS [22], or through optimization techniques [23]. Learning-based shape matching has been approached as a dense semantic segmentation problem (e.g., [4, 5]) or through template alignment [24]. However, these methods often demand substantial training data, rely on mapping to a template, and may struggle to generalize when connectivity changes occur [6].

Our approach aligns closely with methods that integrate learning with the functional map representation, enabling holistic evaluation of the map and facilitating direct training and testing on arbitrary shape pairs. Descriptor functions like SHOT [25] can be refined using a deep neural network to minimize supervised loss based on ground truth correspondences, as demonstrated in FM-Net [26]. This approach was later enhanced [10] to directly extract descriptor functions for functional map estimation from shape geometry using point-based feature extractors, incorporating a regularized map estimation layer. Spectral approaches use unsupervised learning while exploiting the functional map representation. They utilize desirable structural properties in the spectral domain [27] or geodesic distance preservation [9] instead of the supervised loss used in FM-Net. Other attributes, such as unsupervised alignment of heat kernels [28] or cycle consistency [29], have also been used to improve efficiency and accuracy.

These methods are intrinsic, enabling them to generalize well across pose changes and intrinsic symmetries like human shapes. In contrast, previous functional map methods typically refine descriptors such as SHOT, incorporating extrinsic information [28, 29], but they can be unstable with changes in connectivity. Recently, weak-supervision through extrinsic alignment [30] has been introduced to address symmetry ambiguity. Additionally, Deep Shells [31] refines SHOT features while utilizing 3D coordinate information to guide correspondence and mitigate symmetry issues. The symmetry ambiguity is a persistent challenge in unsupervised learning for non-isometric shape matching. This issue is significant because spectral methods often outperform extrinsic methods, like those in [24], in generalizing to unseen poses. The symmetry ambiguity is inherent in functional maps since their losses cannot be orientation-aware. The mixed functional maps presented in [12] focus on aligning tangent vector fields, extending the functional map framework to conformal maps between these fields on surfaces. A key feature of this work is its orientation awareness, achieved through angle-preserving correspondence across shapes without relying on axiomatic descriptors or additional regularization.

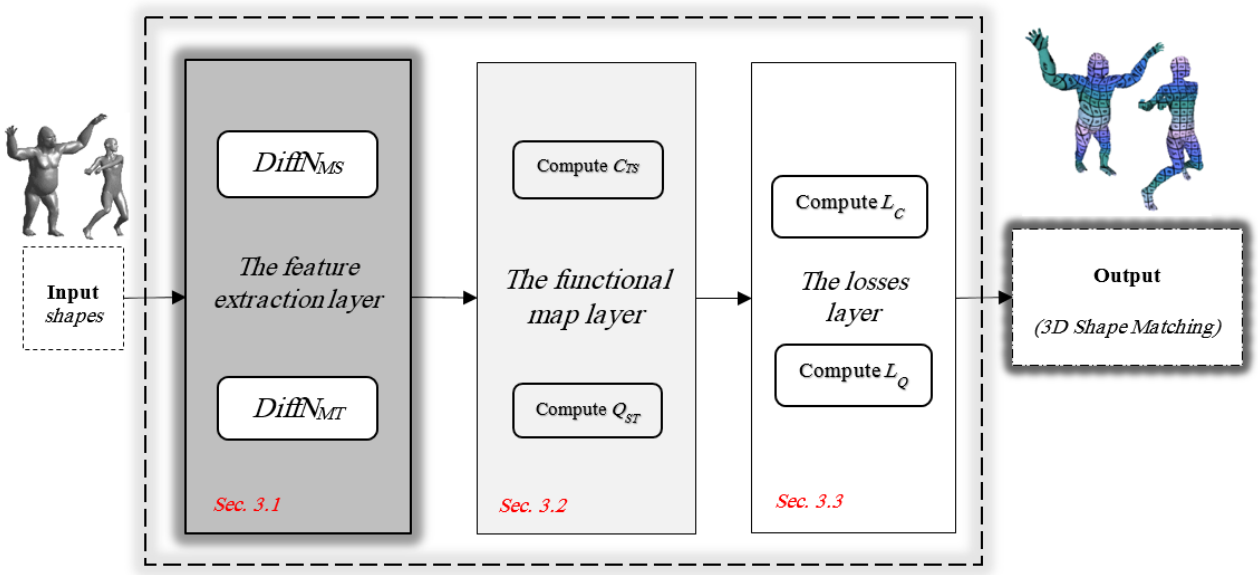

Fig. 1 Outline of the proposed method.

The proposed method introduces a novel angle-preserving unsupervised loss by leveraging the properties of tangent vector fields. Our method resolves symmetry ambiguities by calculating mixed functional maps from the gradients of features obtained through a machine-learning model, and then adding an extra loss function to these maps. This approach is effective when different shapes or structures appear similar due to their symmetrical properties, leading to inaccuracies. The additional loss helps to improve the reliability and consistency of the computations.

3 Proposed Method

In this section, our method is described in more details. The proposed method leverages deep learning by integrating diffusion techniques and spatial gradient features to reveal directional information. It consists of three main components: the feature extraction layer, the functional map layer, and the loss layer. Each component is detailed in the following subsections (3.1 to 3.3). An outline of the proposed method is illustrated in Figure 1.

3.1 The feature extraction layer

The proposed method utilizes the 3D shape feature extractor from [6] to extract features from both source (M_S) and target (M_T) shapes, producing feature vectors for the models. The feature extractor layer is composed of several *DiffN* blocks, as shown in Figure 2. Each *DiffN* block comprises three main components: (1) MLPs for modeling point-wise scalar functions of feature channels: For a mesh with n vertices, each possessing d scalar features, the primary element is a pointwise function f: $R^d \rightarrow R^d$, which is applied independently at each vertex to transform the features. A multilayer perceptron (MLP) with shared weights is utilized for this purpose, allowing effective fitting of arbitrary functions at each vertex. However, these MLPs are inadequate for capturing the spatial structure of the surface and do not facilitate communication between vertices. Therefore, a more sophisticated structure is necessary to overcome these limitations and harness the spatial relationships present in the data. (2) A learned diffusion operation for propagating information across the domain: In this work, a learned diffusion layer inspired by the heat equation h_t: $R^n \rightarrow R^n$ is defined, which diffuses a feature channel u over a learned time $t \in R_{\geq 0}$. In these networks, $h_t(u)$ is applied independently to each feature channel, with a unique learned time t for each. Learning the diffusion parameter allows the network to continuously optimize for spatial support ranging from local to global and selecting different receptive fields for each feature. To evaluate the diffusion layer h_t (u), spectral expansion is utilized, which relies solely on efficient dense arithmetic during evaluation. (3) Local spatial gradient features that broaden the network's filter space beyond radially symmetric filters to model directional filters: The learned diffusion layer enables information propagation across points on a shape but only supports radially symmetric filters centered at a point. In contrast, the spatial gradient features block expands filter options by calculating additional features at the vertices, allowing for directional filters. These components are combined to construct the *DiffN* block.

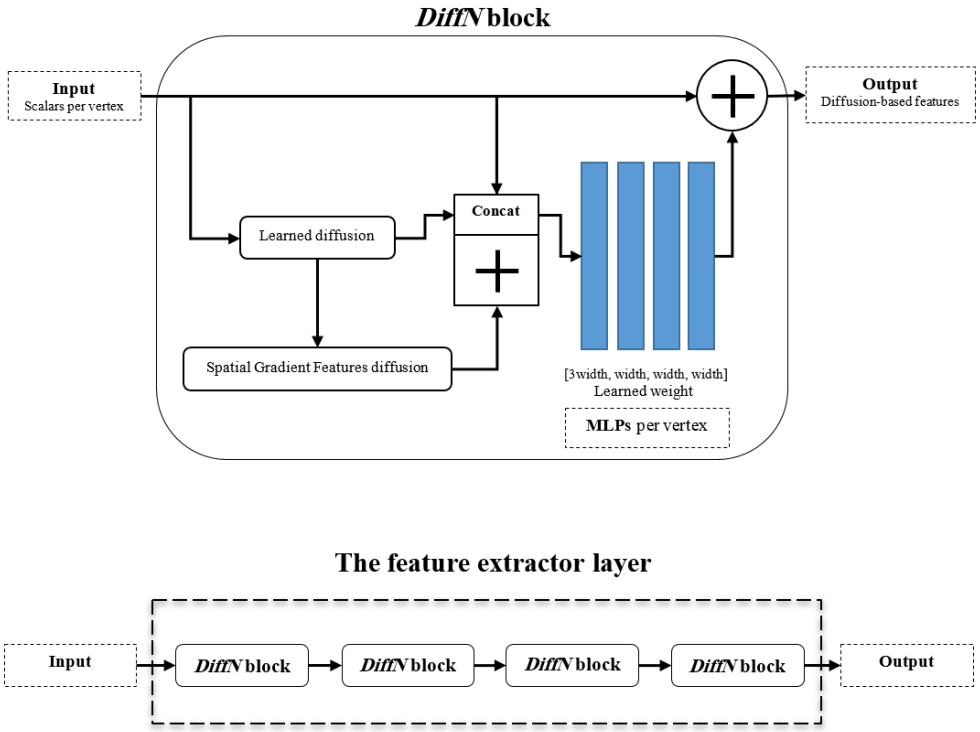

Fig. 2 The feature extractor layer consisting of several consecutive identical *DiffN* blocks.

For shapes M_S and M_T, with n_S and n_T vertices respectively, the learned source and target features are represented as $DiffN_{MS} \in R^{n_S \times d}$ and $DiffN_{MT} \in R^{n_T \times d}$, where d is the dimension of the feature vectors. In fact, the feature extractor layer is a network that operates on a fixed channel width d of scalar values throughout, with each *DiffN* block diffusing the features, constructing spatial gradient features, and feeding the result to an MLP. This layer presents a simple and effective architecture for learning on surfaces. In the feature extractor layer, the 4-block *DiffN* architecture is employed for shape matching, configured with a width of 32. The shape of the first and last linear layers is adapted to the input and output dimension. MLPs utilize ReLU activations and may include dropout after intermediate linear layers. The *DiffN* uses a vector of scalar values per vertex as input features, specifically the heat kernel signatures (HKS) sampled at 16 logarithmically spaced on $[10^{-2}, 1]$. The *DiffN* network is trained using the ADAM optimizer with an initial learning rate of 10^{-3}, a batch size of 1, over 20 epochs, and a learning rate decay of 0.5 every 5 epochs. Cross-entropy loss is applied for labeling tasks, and spectral acceleration is used to assess diffusion, except where truncated to a k = 100 Eigen basis. The features extracted through diffusion are robust and independent of shape triangulation. The outputs from this network, or learned features, are intrinsic and capture information about shape directionality through gradient blocks, as described in [6]. Consequently, the output of this layer generates orientation-aware features used later to estimate orientation-preserving maps. Using this strategy as a feature extractor outperforms recent methods, achieving state-of-the-art correspondence results.

3.2 The functional map layer

This layer is a non-learnable functional map layer based on deep structured prediction, which is introduced in [26]. Indeed, it estimates maps using source and target features obtained from the previous step. The main goal is to learn a mapping F_m: $M_S \rightarrow M_T$ for a pair of triangle meshes, M_S and M_T. The spectrum or the set of Laplace-Beltrami operator eigenvalues serve as a shape signature capable of effectively discriminating shapes. Furthermore, the eigenfunctions of the Laplace-Beltrami operator on a manifold M (M_S or M_T) form a basis that generalizes the Fourier basis to surfaces, which makes them widely used for representing functions on manifolds. The LBO eigenfunctions form the optimal basis for a truncated representation of all gradient-bounded functions on M in a min-max sense [32]. These eigenfunctions have been utilized for truncated representations of descriptors and permutation functions [7]. However, using an optimized basis for a subset of these functions may yield an even better representation. Here, some theoretical background is necessary to understand the Laplace operator on manifolds and its computation (See [33, 34] for detailed information). Let

$f \in \mathbb{R}^2$ be a real-valued function defined on a Riemannian manifold *M*. The Laplace-Beltrami operator is defined as Eq. (1):

$$\Delta f = div(\nabla f) = trace(Hessian(f)) \quad (1)$$

With ∇f the gradient of *f* and *div* the divergence on the manifold. If *M* is a domain in the Euclidean plane, the Laplace-Beltrami operator simplifies to the familiar expression as Eq. (2):

$$\Delta f = \frac{\partial^2 f}{(\partial x)^2} + \frac{\partial^2 f}{(\partial y)^2} \quad (2)$$

Thus, from the Helmholtz equation, Eq. (3) is derived:

$$\Delta f = -\lambda f \quad (3)$$

Where λ is a real scalar. This equation is significant because the family of eigenvalues $0 \leq \lambda_0 \leq \lambda_1 \leq \cdots \leq +\infty$ constitutes the shape spectrum, which is isometric invariant, depending solely on the gradient and divergence of the Riemannian structure of the manifold. Additionally, the Laplace Beltrami operator is Hermitian, ensuring that the eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ corresponding to distinct eigenvalues λ_i and λ_j are orthogonal. As a result, Eq. (4) can be expressed as:

$$\mathbf{v}_i . \mathbf{v}_j = \int_{\mathcal{M}} v_i v_j = 0 \;\; ; \;\; i \neq j \quad (4)$$

In addition, the i^{th} eigenvector is defined a coefficient over the function f as Eq. (5):

$$c_i = f . \mathbf{v}_i = \int_{\mathcal{M}} f \mathbf{v}_i \quad (5)$$

Therefore, the function f can be expanded as Eq. (6):

$$f = c_1 . \mathbf{v}_1 + c_2 . \mathbf{v}_2 + c_3 . \mathbf{v}_3 + \cdots \quad (6)$$

The discrete Laplace-Beltrami operator can be calculated at each vertex of a 3D mesh using the cotangent scheme proposed in [33] as shown in Eq. (7):

$$L(\mathbf{p}_i) = \frac{1}{2A_i} \sum_{\mathbf{p}_j \in \mathbb{N}_1(\mathbf{p}_i)} \left(\cot \alpha_{ij} + \cot \beta_{ij}\right)\left(\mathbf{p}_i - \mathbf{p}_j\right) \quad (7)$$

A_i represents the Voronoi region area around $\mathbf{p}_i$, α_{ij} and β_{ij} are the angles opposite to the arc $\overline{\mathbf{p}_i \mathbf{p}_j}$ and N_1 ($\mathbf{p}_i$) is the set of $\mathbf{p}_i$'s adjacent vertices. Refer to Figure 3 for further details. To numerically compute the Laplace-Beltrami operator, a matrix $L = \{L_{ij}\}$ can be calculated as shown in Eq. (8):

$$L_{ij} = \begin{cases} -\frac{1}{2A_i}\left(\cot \alpha_{ij} + \cot \beta_{ij}\right) & if\ \mathbf{p}_i\ is\ adjacent\ to\ \mathbf{p}_j \\ \sum_k \frac{1}{2A_i}\left(\cot \alpha_{ik} + \cot \beta_{kj}\right) & if\ \mathbf{p}_i = \mathbf{p}_j \\ 0 & otherwise \end{cases} \quad (8)$$

Where $\mathbf{p}_k$ are the adjacent vertices of $\mathbf{p}_i$. The eigenvalues and eigenvectors of this matrix are significant, leading to the problem framed in Eq. (9):

$$l\mathbf{v} = -\lambda \mathbf{v} \quad (9)$$

It is evident that *l* may not be symmetric, meaning $L_{ij} \neq L_{ji}$ when $A_i \neq Aj$, which is often the case. Nevertheless, it can be represented as a generalized eigenvalue problem where $l = S^{-1}M$.

Consequently, Eq. (10) is a modification of the latter equation.

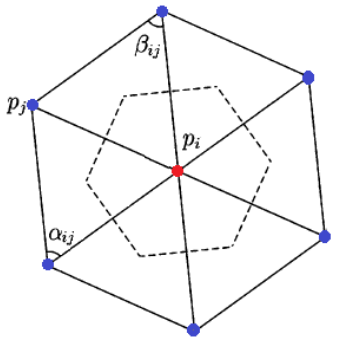

Fig. 3 Neighborhood configuration around $\mathbf{p}_i$. The dashed lines indicate the Voronoi region utilized for calculating the Laplace-Beltrami operator.

$$M\mathbf{v} = -\lambda S\mathbf{v} \quad (10)$$

Where, the stiffness matrix $M = \{\mathbf{M}_{ij}\}$ contains the cotangent weights as shown in Eq. (11):

$$M_{ij} = \begin{cases} \frac{1}{2}\left(\cot \alpha_{ij} + \cot \beta_{ij}\right) & if\ \mathbf{p}_i\ is\ adjacent\ to\ \mathbf{p}_j \\ \sum_k \frac{1}{2}\left(\cot \alpha_{ik} + \cot \beta_{kj}\right) & if\ \mathbf{p}_i = \mathbf{p}_j \\ 0 & otherwise \end{cases} \quad (11)$$

Moreover, the mass matrix $S = \{S_{ij}\}$ is a diagonal matrix of vertex area elements as shown in Eq. (12):

$$S_{ij} = \begin{cases} A_i & if\ i = j \\ 0 & 0therwise \end{cases} \quad (12)$$

This solution ensures that the eigenvalues and eigenvectors are real. Furthermore, two eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ corresponding to different eigenvalues λ_i and λ_j are orthogonal under the dot product S, as shown in Eq. (13):

$$\mathbf{v}_i . \mathbf{v}_j = \mathbf{v}_i^T S\, \mathbf{v}_j \quad ; \; i \neq j \quad (13)$$

The spectrum of the shape is the set of eigenvalues $\{\lambda_0, \lambda_1, \lambda_2 \ldots \lambda_{n-1}\}$. If the shape is closed, $\lambda_0 = 0$.

The proposed method uses the discrete Laplace-Beltrami operator (LBO), which preserves many properties of its continuous counterpart and accurately represents the geometric and topological properties of the underlying surface. To minimize dependence on the mesh effectively, the discrete Laplacian is utilized.

This study employs the functional map framework, which allows correspondences to be effectively represented using a compact matrix. Specifically, the map F_m: $M_S \rightarrow M_T$ is represented by the binary matrix Π_{MSMT}, where Π_{MSMT} $(i, j) = 1$ if and only if F_m $(i) = j$ for vertices *i* and *j* in M_S and M_T, respectively. The functional map C_{TS} is defined as $C_{TS} = l_{MS}^{\dagger}\, \Pi_{MSMT}\, l_{MT}$, with l_{MS} and l_{MT} being matrices containing the first *k* eigenfunctions of the Laplace-Beltrami operators of shapes M_S and M_T, respectively. The sign † denotes the pseudoinverse of a matrix. Note that C_{TS} is of size $k \times k$ with *k* typically between 20−100, and is typically orders of magnitude smaller than Π_{MSMT}.

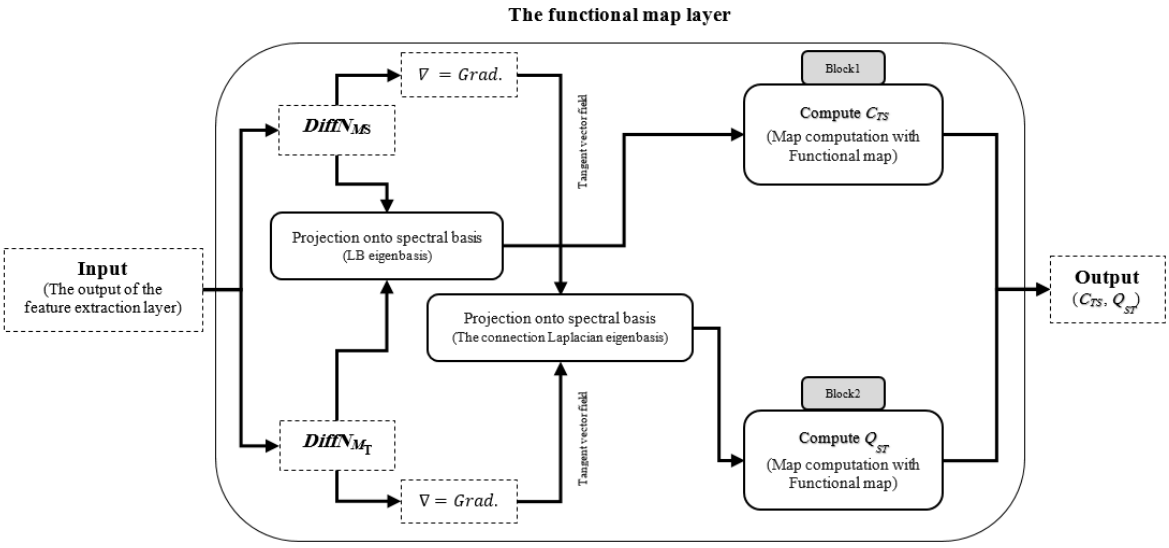

Fig. 4 The functional map layer architecture.

The basic pipeline for map recovery, introduced in [7], assumes the existence of descriptor functions that remain consistent under the unknown mapping. Thus, the optimal functional map $(C_{TS})_{opt}$ is computed as shown in Eq. (14):

$$(C_{TS})_{opt} = \underset{C_{TS}}{\operatorname{argmin}}(\|C_{TS}A_T - A_S\|_F^2) + \mu E_{reg}(C_{TS}) \quad (14)$$

Where A_S and A_T are the coefficients of descriptors in the spectral basis l_{MS} and l_{MT} and μ is a scalar regularization parameter. The first term promotes the preservation of descriptor functions, whereas the second serves as a regularizer that promotes structural properties, for example, $E_{reg}(C_{TS}) = \|C_{TS}\Delta_T - \Delta_S C_{TS}\|$, where Δ_S, Δ_T are diagonal $k \times k$ matrices of Laplacian eigenvalues (see more in [10]). The value of k, chosen as 100 in the proposed method, equals the first k Eigen-functions of the Laplace-Beltrami operators for shapes M_S and M_T. The final point-to-point map F_m: $M_S \rightarrow M_T$ can be obtained via nearest neighbor search between the rows of $l_{MS}C_{TS}$ and those of l_{MT} [35].

The proposed method uses the output of the feature extractor layer to estimate functional maps. This layer consists of two blocks, as illustrated in Figure 4.

In the first block, the features of the input shapes (*DiffN*) are projected onto the Laplace-Beltrami eigenbasis of the shapes to obtain spectral features $A_S = l_{MS}^{\dagger}DiffN_{MS}$ and $A_T = l_{MT}^{\dagger}DiffN_{MT}$. The functional map C_{TS} is then estimated as the solution to the following least-squares problem as shown in Eq. (15):

$$C_{TS} = \underset{C}{\operatorname{argmin}}(\|CA_T - A_S\|_F^2) \quad (15)$$

Where A_S and A_T are $l_x^{\dagger}.x.DiffN_x$ and l_x is the spectral basis with $x \in \{M_S, M_T\}$. It leads to Eq. (16):

$$C_{TS} = A_S {A_T}^{\dagger} \quad (16)$$

In the second block, the shape features (*DiffN*) are transformed into a tangent vector field using the discrete gradient operator, *Grad*, as has been described in [12]. These vector-valued descriptors are projected onto the connection Laplacian basis. The connection Laplacian (L_Q) behaves like a second derivative on vector fields with many of the same basic properties as the ordinary Laplacian. It is negative semidefinite, self-adjoint, and elliptic [36]. The connection Laplacian is associated defined by Eq. (17):

$$\Delta^{\nabla}\mathbf{v_f} = trace(\nabla^2\mathbf{v_f}) = -\nabla * \nabla\mathbf{v_f} \quad (17)$$

This leads to spectral feature vectors B_x with $x \in \{M_S, M_T\}$, and *Grad* (M_x) the gradient operator on shape M_x. In this case, the functional map Q_{ST} is estimated as shown in Eq. (18):

$$Q_{ST} = \underset{Q}{\operatorname{argmin}}(\|QB_S - B_T\|_F^2) \quad (18)$$

Where B_S and B_T are $L_{Q\_x}{}^{\dagger}.Grad(x).DiffN_x$ and L_{Q_x} is the connection Laplacian with $x \in \{M_S, M_T\}$, its closed-form solution is provided in Eq. (19):

$$Q_{ST} = B_S {B_T}^{\dagger} \quad (19)$$

Q_{ST} includes local orientation information based on the normal orientation derived from feature gradients. As a mixed linear representation, Q_{ST} can only depict orientation-preserving maps, as demonstrated in [12].

3.3 The Losses layer

In proposed method losses for the functional map layer in two case is estimated. The loss is unnecessary when estimating C_{TS} and Q_{ST} with the Laplacian regularizer from [10], as it naturally produces maps with low isometric loss. In contrast, FM-Net [27] enforces orthogonality on the estimated functional map C_{TS}, so only L_{ort} is implemented. It is calculated using Eq. (20):

$$L_{ort}(C_{TS}) = \left\|{C_{TS}}^T C_{TS} - I\right\|_F^2 \quad (20)$$

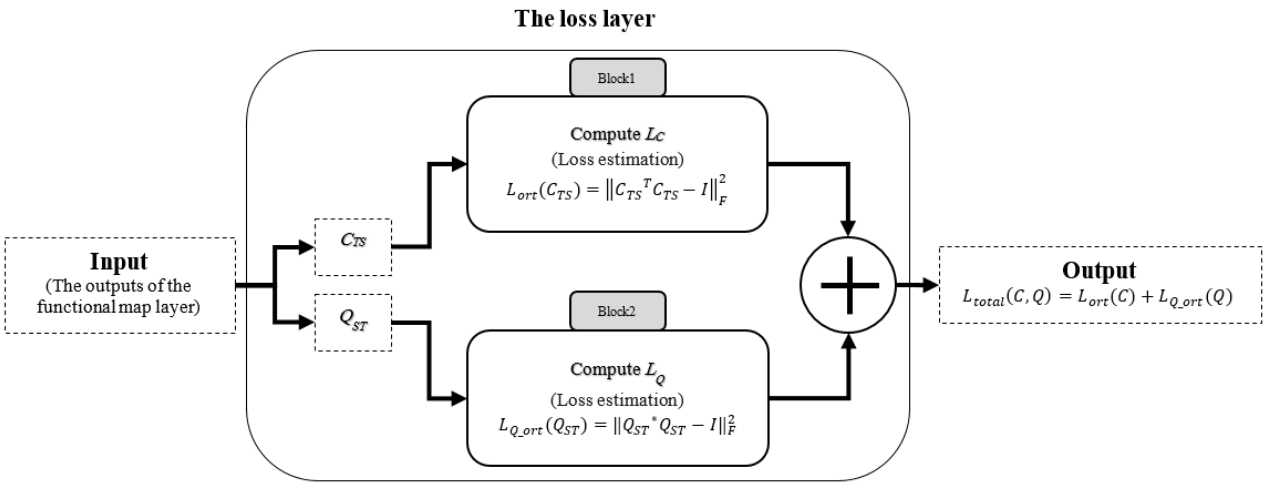

Fig. 5 The loss layer architecture.

The loss is referred to do not provide enough information or constraints to eliminate or prevent the presence of inherent symmetries in the shape matching. In other words, just having this loss does not guarantee that symmetry ambiguity will be excluded from the results. Also, as shown in [13], enforces orthogonality on the estimated functional map Q_{ST}, so only L_{Q_ort} is implemented. It is computed using Eq. (21):

$$L_{Q\_ort}(Q_{ST}) = \|Q_{ST}{}^* Q_{ST} - I\|_F^2 \quad (21)$$

The proposed method uses a weighted combination of two losses in the loss layer, ensuring that the learned descriptors create an isometric map while preserving orientation. This approach addresses symmetry ambiguities in 3D shape matching, and the entire computation is unsupervised. This layer consists of two blocks, as illustrated in Figure 5.

4 Implementation

The proposed method is implemented in PyTorch. The feature extraction layer follows the standard network structure from [6], utilizing 100 eigenfunctions of the Laplace-Beltrami operators for shapes M_S and M_T. Each block converts these 100 features into learned features of the same dimension, resulting in 100-dimensional vector field descriptors for the input shapes. Additionally, μ is set to 10^{-3} in the functional map layer for regularization, as noted in [10]. The network is trained with a batch size 1 for a number of epochs between 5 and 25. A learning rate of 10^{-3} is used with the ADAM optimizer.

5 Results

In this section, we show that our network can outperform state-of-the-art deep shape matching architectures on non-isometric datasets like TOSCA [37]. It contains human and animal models in different classes. The dataset can be split into two types of shape pairs: isometric pairs, which consist of two shapes from the same class, and non-isometric pairs, which consist of two shapes from different classes. To remove the bias present due to identical mesh connectivity within a dataset, the LRVD algorithm [38] is used to re-mesh the datasets. The proposed method is compared quantitatively and qualitatively with existing methods such as Zoom-Out [39], FM-Net [26], GeoFmaps [10], UFM-Net [9]. The proposed method is tested on non-isometric TOSCA shapes. The dataset, initially consisting of 100 shapes, is divided into 65 training shapes and 35 test shapes. Indeed, these include animal shapes from various species and poses, which often present considerable challenges for existing methods.

The Averaged Geodesic Errors of the experiment are reported in Table 1. Our approach effectively identifies correspondences between models undergoing non-isometric deformation. It outperforms unsupervised methods and achieves similar or superior performance compared to supervised ones.

Table 1 Quantitative results comparing the different methods with the proposed method on the TOSCA dataset.

Method	The type of Feature used	The Averaged Geodesic Errors
Zoom-Out [39]	-	0.0618
Geo-FMaps [10]	WKS	0.0042
FM-Net [26]	SHOT	0.0321
UFM-Net [9]	SHOT	0.0058
Proposed method	The feature extraction layer	**0.0023**

(a) (b) (c) (d) (e)

Fig. 6 Comparison of the qualitative results of (a) The proposed method with (b) Geo-Fmaps [10], (c) Zoom-Out [39], (d) UFM-Net [9], (e) FM-Net [26] on TOSCA dataset.

The proposed method demonstrates the lowest Averaged Geodesic Distance compared to other approaches in predicting the corresponding points between the source and target models. In other words, the corresponding points identified by our method are accurately mapped onto the target model. Figure 6 displays a qualitative result. The proposed method outperforms others, even in challenging case with strong non-isometric distortions (Figure 6). Meanwhile, both axiomatic method [39] and SHOT-based methods [9, 26] fail to predict accurate correspondences. The results show that the proposed method is robust to deformations that often cause methods in [9, 26, 39] to fail, typically due to their misinterpretation of anisotropy as meaningful geometric information. The method in [10] is a more robust learning-based approach, utilizing spectral filters in its feature extractor to reduce high-frequency overfitting. However, it struggles with non-isometric datasets, as it fails to learn accurate descriptors from intrinsically symmetric signals (like WKS as input), leading to poorer correspondence than the proposed method. In contrast, the proposed method excels by leveraging deep learning to focus on geometric features tied to surface orientation and deformation, producing orientation-aware maps. The qualitative results are showed that the proposed method is the only method that produces reasonable correspondence than the others, emphasizing the advantages of our accurate feature extractor combined with effective geometric regularizers, which enable our approach to handle non-isometric shapes effectively.

6 Conclusion

Our method utilizes deep learning and incorporates learned diffusion and spatial gradient features to provide directional information, consisting of three main components: the feature extraction layer, the functional map layer, and the losses layer. Results demonstrate that our network outperforms state-of-the-art shape matching architectures on non-isometric datasets like TOSCA. However, limitations include an assumption of relatively regular input models and a requirement for similar features in both shapes. Additionally, only full 3D models are considered. Future work aims to expand the framework to accommodate a wider variety of input models and address limitations related to partial or damaged 3D shapes and models with high deformation.

Limitations& Future Work: The proposed method assumes that the input models should be relatively regular, a condition that can be addressed through re-meshing if necessary. Additionally, our method assumes the existence of similar features in both input shapes. In fact, one would expect to see at least some similar large-scale structures between the shapes that are supposed to be in correspondence. All shapes are full 3D models and partial shapes or damaged models are not considered.

In the future, we aim to expand the framework of the proposed method to encompass a wider variety of input models, utilizing it to identify correspondences between partial or damaged 3D shapes and highly deformed models.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Amirreza Amirfathiyan: Conceptualization, Software, Methodology, Formal analysis, Writing - Original draft.

Hossein Ebrahimnezhad: Supervision, Revise and editing, Investigation.

Funding

No funding was received for this work.

Informed Consent Statement

Not applicable.

References

[1] L. Pishchulin, S. Wuhrer, T. Helten, C. Theobalt, and B. Schiele, "Building statistical shape spaces for 3d human modeling," *Pattern Recognition*, vol. 67, pp. 276-286, 2017.

[2] I. Baran, D. Vlasic, E. Grinspun, and J. Popović, "Semantic deformation transfer," in *ACM SIGGRApH 2009 papers*, pp. 1-6, 2009.

[3] R. Litman and A. M. Bronstein, "Learning spectral descriptors for deformable shape correspondence," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 1, pp.171-180, 2013.

[4] A. Poulenard and M. Ovsjanikov, "Multi-directional geodesic neural networks via equivariant convolution," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, pp. 1-14, 2018.

[5] R. Wiersma, E. Eisemann, and K. Hildebrandt, "CNNs on surfaces using rotation-equivariant features," *ACM Transactions on Graphics (ToG)*, vol. 39, no. 4, pp. 92: 1-92: 12, 2020.

[6] N. Sharp, S. Attaiki, K. Crane, and M. Ovsjanikov, "Diffusionnet: Discretization agnostic learning on surfaces," *ACM Transactions on Graphics (TOG)*, vol. 41, no. 3, pp. 1-16, 2022.

[7] M. Ovsjanikov, M. Ben-Chen, J. Solomon, A. Butscher, and L. Guibas, "Functional maps: a flexible representation of maps between shapes," *ACM Transactions on Graphics (ToG)*, vol. 31, no. 4, pp. 1-11, 2012.

[8] D. Cao and F. Bernard, "Self-Supervised Learning for Multimodal Non-Rigid 3D Shape Matching," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 17735-17744, 2023

[9] O. Halimi, O. Litany, E. Rodola, A. M. Bronstein, and R. Kimmel, "Unsupervised learning of dense shape correspondence," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4370-4379, 2019.

[10] N. Donati, A. Sharma, and M. Ovsjanikov, "Deep geometric functional maps: Robust feature learning for shape correspondence," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8592-8601, 2020.

[11] A. Sharma and M. Ovsjanikov, "Weakly supervised deep functional maps for shape matching," *Advances in Neural Information Processing Systems*, vol. 33, pp. 19264-19275, 2020.

[12] N. Donati, E. Corman, S. Melzi, and M. Ovsjanikov, "Complex functional maps: A conformal link between tangent bundles," in *Computer Graphics Forum*, vol. 41, no. 1: Wiley Online Library, pp. 317-334, 2022.

[13] Y. Sahillioğlu, "Recent advances in shape correspondence," *The Visual Computer*, vol. 36, no. 8, pp. 1705-1721, 2020.

[14] A. Gehre, M. Bronstein, L. Kobbelt, and J. Solomon, "Interactive curve constrained functional maps," in *Computer Graphics Forum*, vol. 37, no. 5: Wiley Online Library, pp. 1-12, 2018.

[15] D. Nogneng and M. Ovsjanikov, "Informative descriptor preservation via commutativity for shape matching," in *Computer Graphics Forum*, vol. 36, no. 2: Wiley Online Library, pp. 259-267, 2017.

[16] J. Ren, A. Poulenard, P. Wonka, and M. Ovsjanikov, "Continuous and orientation-preserving correspondences via functional maps," *ACM Transactions on Graphics (ToG)*, vol. 37, no. 6, pp. 1-16, 2018.

[17] M. Shoham, A. Vaxman, and M. Ben-Chen, "Hierarchical functional maps between subdivision surfaces," in *Computer Graphics Forum*, vol. 38, no. 5: Wiley Online Library, pp. 55-73, 2019.

[18] Y. Wang, B. Liu, K. Zhou, and Y. Tong, "Vector field map representation for near conformal surface correspondence," in *Computer Graphics Forum*, vol. 37, no. 6: Wiley Online Library, pp. 72-83, 2018.

[19] O. Litany, E. Rodolà, A. M. Bronstein, and M. M. Bronstein, "Fully spectral partial shape matching," in *Computer Graphics Forum*, vol. 36, no. 2: Wiley Online Library, pp. 247-258, 2017.

[20] E. Rodolà, L. Cosmo, M. M. Bronstein, A. Torsello, and D. Cremers, "Partial functional correspondence," in *Computer graphics forum*, vol. 36, no. 1: Wiley Online Library, pp. 222-236, 2017.

[21] J. Sun, M. Ovsjanikov, and L. Guibas, "A concise and provably informative multi-scale signature based on heat diffusion," in *Computer graphics forum*, vol. 28, no. 5: Wiley Online Library, pp. 1383-1392, 2009.

[22] M. Aubry, U. Schlickewei, and D. Cremers, "The wave kernel signature: A quantum mechanical approach to shape analysis," in *2011 IEEE international conference on computer vision workshops (ICCV workshops)*: IEEE, pp. 1626-1633, 2011.

[23] E. Corman, M. Ovsjanikov, and A. Chambolle, "Supervised descriptor learning for non-rigid shape matching," in *Computer Vision-ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part IV 13*: Springer, pp. 283-298, 2015.

[24] T. Groueix, M. Fisher, V. G. Kim, B. C. Russell, and M. Aubry, "3d-coded: 3d correspondences by deep deformation," in *Proceedings of the european conference on computer vision (ECCV*, pp. 230-246*)*, 2018.

[25] F. Tombari, S. Salti, and L. Di Stefano, "Unique signatures of histograms for local surface description," in *Computer Vision–ECCV 2010:*

11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part III 11: Springer, pp. 356-369, 2010.

[26] O. Litany, T. Remez, E. Rodola, A. Bronstein, and M. Bronstein, "Deep functional maps: Structured prediction for dense shape correspondence," in *Proceedings of the IEEE international conference on computer vision*, pp. 5659-5667, 2017.

[27] J.-M. Roufosse, A. Sharma, and M. Ovsjanikov, "Unsupervised deep learning for structured shape matching," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1617-1627, 2019.

[28] M. Aygün, Z. Lähner, and D. Cremers, "Unsupervised dense shape correspondence using heat kernels," in *2020 International Conference on 3D Vision (3DV*: IEEE, pp. 573-582*)*, 2020.

[29] D. Ginzburg and D. Raviv, "Cyclic functional mapping: Self-supervised correspondence between non-isometric deformable shapes," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16*: Springer, pp. 36-52, 2020.

[30] M. Eisenberger *et al.*, "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.

[31] M. Eisenberger, A. Toker, L. Leal-Taixé, and D. Cremers, "Deep shells: Unsupervised shape correspondence with optimal transport," *Advances in Neural information processing systems*, vol. 33, pp. 10491-10502, 2020.

[32] Y. Aflalo, H. Brezis, and R. Kimmel, "On the optimality of shape and data representation in the spectral domain," SIAM Journal on Imaging Sciences, vol. 8, no. 2, pp. 1141-1160, 2015.

[33] M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr, "Discrete differential-geometry operators for triangulated 2-manifolds," in Visualization and mathematics III, 2003: Springer, pp. 35-57, 2003.

[34] M. Reuter, S. Biasotti, D. Giorgi, G. Patanè, and M. Spagnuolo, "Discrete Laplace–Beltrami operators for shape analysis and segmentation," Computers & Graphics, vol. 33, no. 3, pp. 381-390, 2009.

[35] G. Pai, J. Ren, S. Melzi, P. Wonka, and M. Ovsjanikov, "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.

[36] N. Sharp, Y. Soliman, and K. Crane, "The vector heat method," ACM Transactions on Graphics (TOG), vol. 38, no. 3, pp. 1-19, 2019.

[37] A. M. Bronstein, M. M. Bronstein, R. Kimmel, A. Bronstein, M. Bronstein, and R. Kimmel, "Non rigid Correspondence and Calculus of Shapes," *Numerical Geometry of Non-Rigid Shapes*, pp. 239-259, 2009.

[38] D.-M. Yan, G. Bao, X. Zhang, and P. Wonka, "Low-resolution remeshing using the localized restricted Voronoi diagram," *IEEE transactions on visualization and computer graphics*, vol. 20, no. 10, pp. 1418-1427, 2014.

[39] S. Melzi, J. Ren, E. Rodola, A. Sharma, P. Wonka, and M. Ovsjanikov, "Zoomout: Spectral upsampling for efficient shape correspondence," *arXiv preprint arXiv:1904.07865*, 2019.

Amirreza Amirfathiyan was born in Tabriz/ Iran in 1978. He received his B.Sc. and M.Sc. degrees in Control Engineering and Telecommunication system from Sahand University of Technology in 2001 and 2017, respectively. His research interests include in Computer Graphics, Machine Vision, 3D Shape Correspondence, Image processing and statistical pattern recognition. He is now a Ph.D. student at Sahand University of Technology in telecommunication system.

Dr. Hossein Ebrahimnezhad was born in Iran 1971. He received his B.Sc. and M.Sc. degrees in Electronic and Communication Engineering from Tabriz University and K.N. Toosi University of Technology in 1994 and 1996, respectively. In 2007, he received his Ph.D. degree from Tarbiat Modares University. His research interests include image processing, computer Vision, 3D model processing and soft computing. Currently, he is a professor at Sahand University of Technology.