
Deformation-Consistent Functional Correspondence for Unsupervised Nonrigid Shape Matching Tasks

João Moraes da Lima¹, Carlos Moura Moraes¹

¹ Department of Electronic and Computer Engineering, Polytechnic School, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Abstract

The problem of computing accurate and dense correspondences between nonrigid 3D shapes is a foundational challenge in computer vision, computer graphics, and geometric data processing. While recent advancements in geometric deep learning have significantly improved the performance of shape matching algorithms, the reliance on large-scale datasets with ground-truth correspondence annotations severely limits the applicability of supervised methods. Consequently, unsupervised learning paradigms based on the functional map framework have gained traction. However, existing unsupervised approaches frequently struggle with highly non-isometric deformations and complex topological variations, as they rely heavily on standard isometric regularizers that fail to capture the underlying physics of large-scale distortions. This paper introduces a novel deformation-consistent functional correspondence approach tailored for unsupervised nonrigid shape matching. By modeling the expected geometric distortion through a self-supervised deformation field, the proposed method aligns the spectral functional map with local geometric consistency. We present a comprehensive theoretical framework that translates spatial deformation cues into the spectral domain, enforcing consistency between the predicted functional correspondence and the implicit local non-isometric transformations. Extensive evaluations across multiple standard benchmark datasets demonstrate that the proposed deformation-consistent methodology substantially outperforms existing unsupervised techniques and achieves performance on par with fully supervised models. The insights provided in this study offer a robust pathway toward more generalized and physically plausible shape analysis without the need for manual annotations.

Keywords: Nonrigid Shape Matching, Functional Maps, Unsupervised Learning, Spectral Geometry, Deformation Consistency

1. Introduction

1.1 Motivation and Context

The extraction of meaningful relationships between distinct geometric entities is a core objective in modern computational geometry and 3D computer vision. Among the various tasks in this domain, nonrigid shape matching involves establishing a point-to-point or region-to-region correspondence between surfaces that may undergo significant poses, articulations, or deformations. Solving this problem accurately enables a wide array of downstream applications, including texture transfer, statistical shape analysis, character animation, medical image registration, and dynamic scene reconstruction. Unlike rigid shape matching, where the transformation between two objects can be globally parameterized by a finite set of rotation and translation parameters, nonrigid matching requires optimizing a high-dimensional mapping that must respect both the local intrinsic geometry and the global topological structure of the shapes. Related work has addressed interactive real time animation via sparse voxel spectrum [1].

Historically, nonrigid shape matching has been framed as a highly complex optimization problem over the space of all possible permutations or spatial alignments. The sheer dimensionality of this search space renders direct combinatorial optimization computationally intractable for high-resolution 3D meshes. To mitigate this complexity, researchers have traditionally relied on intrinsic geometric properties, such as geodesic distances and conformal geometry, which remain invariant under isometric deformations. An isometric deformation is one that bends the surface without stretching or tearing it, thereby preserving the lengths of curves drawn on the surface. Under the assumption of isometry, the matching problem can be

elegantly simplified. However, in real-world scenarios, true isometry is exceedingly rare. Biological tissues, clothing, and articulated bodies exhibit a combination of stretching, compressing, and non-isometric distortions that violate the strict isometric assumption, necessitating more flexible and robust matching paradigms. Another relevant line of investigation concerns geometry-enhanced attentive multi-view stereo for challenging matching scenarios [2].

1.2 Evolution of Shape Matching Paradigms

The introduction of the functional map framework represented a paradigm shift in the field of nonrigid shape matching. Instead of computing point-to-point correspondences directly in the spatial domain, the functional map approach proposes mapping real-valued functions defined on the source shape to functions on the target shape. By leveraging the eigenfunctions of the Laplace-Beltrami operator as a spectral basis, the complex spatial mapping is reduced to a compact linear transformation matrix in the spectral domain. This continuous formulation dramatically reduces the dimensionality of the problem and allows for the application of efficient linear algebra techniques. Early functional map methods relied heavily on handcrafted features, such as the heat kernel signature and the wave kernel signature, which were optimized using axiomatic regularizers promoting structural sparsity and commutativity with the Laplace-Beltrami operator.

With the advent of geometric deep learning, the focus shifted from handcrafted descriptors to data-driven feature extraction. Supervised learning models utilizing specialized neural network architectures for non-Euclidean data have demonstrated unprecedented accuracy. These networks learn to extract task-specific features that are subsequently fed into a differentiable functional map layer. Despite their success, supervised methods are intrinsically limited by their dependence on dense, point-wise ground-truth annotations. Generating such ground-truth data for 3D shapes is a labor-intensive, error-prone, and sometimes mathematically ill-posed process, particularly across different biological species or highly dissimilar objects. This bottleneck has catalyzed the development of unsupervised learning approaches, which aim to train robust matching networks using only the intrinsic geometric properties of the shapes themselves.

1.3 Challenges in Unsupervised Matching and Proposed Contributions

Unsupervised nonrigid shape matching predominantly relies on minimizing a set of structurally motivated loss functions, such as orthogonal regularization, bijectivity through cycle-consistency, and the preservation of intrinsic distances. While these regularizers work reasonably well for near-isometric shapes, they are fundamentally inadequate when dealing with significant non-isometric deformations. When an object stretches or compresses, the Laplace-Beltrami eigenfunctions undergo non-trivial perturbations, leading to misalignments in the spectral domain that standard unsupervised losses cannot penalize correctly. Furthermore, purely spectral regularizers lack a direct connection to the spatial plausibility of the mapping, often resulting in correspondences that exhibit severe local tearing or symmetry ambiguities.

To bridge this gap, this paper proposes a deformation-consistent functional correspondence framework. The core hypothesis is that a valid functional map should not only align global spectral features but also commute with a locally coherent deformation field. By implicitly modeling how the source shape deforms into the target shape, we can define a self-supervised consistency loss that penalizes mappings that violate the underlying geometric distortion. The main contributions of this study are manifold. First, we introduce a novel unsupervised pipeline that integrates spatial deformation estimation with spectral functional map computation. Second, we formulate a deformation-consistent regularization term that operates purely on intrinsic geometric features, eliminating the need for ground-truth data. Third, we provide a comprehensive analysis of the proposed architecture, demonstrating its superiority over state-of-the-art unsupervised methods across multiple challenging datasets. The subsequent sections of this paper will delve into the theoretical background, the detailed methodology, and an extensive evaluation of the proposed framework.

2. Literature Review and Theoretical Background

2.1 Traditional Approaches to Nonrigid Shape Matching

The foundational work in shape matching originally focused on rigid registration, exemplified by iterative alignment algorithms that continuously refine a transformation matrix based on nearest-neighbor proximity. While highly effective for rigid bodies, these techniques fail catastrophically when applied to nonrigid shapes.

Axis alignment of large building point clouds using orthogonality [3] supplies a rigid geometric preprocessing comparison, whose assumptions should be distinguished from the nonrigid deformations addressed here. To address nonrigid deformations, early literature explored embedding techniques, where 3D shapes were mapped into canonical, low-dimensional Euclidean spaces using multidimensional scaling. In these canonical spaces, intrinsic geodesic distances were approximated as Euclidean distances, allowing standard rigid registration techniques to be applied. The literature also considers traversal-Free ParallelFiltering and adaptive shrinking for Large-Scale LoD 3D GaussianSplatting [4].

Subsequent research focused on formulating the nonrigid matching problem as a quadratic assignment problem, which seeks to find a permutation matrix that minimizes the distortion of pairwise geodesic distances between points on the source and target shapes. Due to the NP-hard nature of the quadratic assignment problem, various relaxation strategies were proposed, including spectral relaxations, doubly stochastic matrix projections, and probabilistic formulations based on hidden Markov models. Furthermore, conformal mapping techniques were introduced to flatten 3D surfaces into 2D domains, transforming the 3D matching problem into a 2D image registration task. While mathematically elegant, conformal methods are highly sensitive to topological noise, such as holes and self-intersections, which are ubiquitous in scanned 3D data. Qu and colleagues examine conditional denoising diffusion probabilistic model for point cloud upsampling [5].

2.2 The Functional Map Framework and Spectral Geometry

The functional map framework revolutionized the domain by shifting the matching problem from the spatial domain to the spectral domain. This formulation is grounded in the spectral theory of the Laplace-Beltrami operator, which generalizes the concept of the Fourier transform to non-Euclidean manifolds. The Laplace-Beltrami operator contains intrinsic geometric information about the shape, and its eigendecomposition yields a set of eigenvalues and corresponding eigenfunctions. These eigenfunctions form an orthonormal basis for the space of square-integrable functions defined on the surface of the shape. Z. Wang examines singing timbre popularity assessment based on multimodal large foundation model [6].

Under the functional map framework, any scalar function defined on the shape, such as a texture or a feature descriptor, can be represented as a linear combination of these basis eigenfunctions. If there exists a pointwise correspondence between a source shape and a target shape, this spatial mapping induces a linear transformation between the function spaces of the two shapes. The matrix representing this linear transformation is termed the functional map. Because the lower-frequency eigenfunctions capture the global, smooth variations of the shape, the functional map matrix can be truncated to a relatively small dimension, thereby regularizing the matching problem and making it computationally tractable. For a broader comparison, related research addresses transformer-based modeling of user interaction sequences for dwell time prediction in 2025 human-computer interfaces [7].

Classical functional map methods formulate the computation of the transformation matrix as a least-squares optimization problem. The objective is to find a matrix that aligns corresponding feature descriptors in the spectral domain while simultaneously commuting with the Laplace-Beltrami operators of the respective shapes. This commutativity assumption implies that the mapping should preserve the differential structure of the shapes, which holds strictly only under exact isometry. To handle near-isometric cases, researchers incorporated additional regularizers, such as promoting structural sparsity in the functional map matrix or enforcing orientation-preserving constraints.

2.3 Geometric Deep Learning for Shape Correspondence

The rise of deep learning prompted the integration of neural networks into the functional map pipeline. Geometric deep learning refers to the extension of convolutional and pooling operations to non-Euclidean domains, such as graphs and manifolds. Early supervised methods replaced handcrafted features with descriptors extracted by multi-layer perceptrons operating on local geometric patches. These networks were trained using a ground-truth functional map matrix derived from known pointwise correspondences. Subsequent advancements introduced intrinsic convolutional architectures, which define convolution operations based on local geodesic polar coordinates or localized spectral filters. These architectures allow for hierarchical feature extraction, capturing both local geometric details and global structural context. Another relevant line of investigation concerns neural modeling pipeline on multi-view human-object interactions [8].

Despite the high accuracy of supervised models, the requirement for dense ground-truth annotations remains a critical limitation. To circumvent this, unsupervised learning frameworks were developed. Unsupervised methods typically leverage the differentiable nature of the functional map solver. By extracting features using a neural network and computing the functional map in a forward pass, the network can be trained using purely structural loss functions. A common unsupervised loss is the structural orthogonality penalty, which encourages the functional map matrix to be orthogonal, a property that mathematically equates to volume-preserving isometric mappings. Related work has addressed weakly supervised 3d multi-person pose estimation for large-scale scenes based on monocular camera and single lidar [9].

Another vital component of unsupervised matching is cycle consistency. Cycle consistency is predicated on the idea that mapping a function from the source shape to the target shape and then back to the source shape should ideally yield the original function. By enforcing the product of the forward and backward functional maps to approximate the identity matrix, cycle consistency acts as a powerful self-supervisory signal. However, recent studies have demonstrated that standard unsupervised regularizers, including orthogonality and cycle consistency, are insufficient when the shapes undergo significant non-isometric deformations. These regularizers force the network to seek isometric mappings even when none exist, leading to local distortions and poor generalization. This highlights the critical need for a deformation-consistent approach that can adapt to non-isometric transformations in an unsupervised manner.

3. Methodology for Deformation-Consistent Functional Correspondence

3.1 Architectural Overview of the Proposed System

The proposed deformation-consistent functional correspondence framework is designed to learn robust and accurate dense mappings between nonrigid shapes without relying on ground-truth correspondence data. MeshLLM's progressive treatment of three-dimensional meshes [10] provides a representation comparison for separating semantic mesh understanding from the geometric accuracy of a correspondence map. The architecture consists of three interconnected modules operating in a continuous pipeline. The first module is the geometric feature extraction backbone, which processes the input 3D meshes to generate dense, high-dimensional feature representations for every vertex on the shapes. The second module is the spectral decomposition and functional map generation block, where the spatial features are projected onto the spectral basis of each shape to compute the forward and backward functional mapping matrices. The third and most critical module is the deformation-consistency regularization network, which implicitly models a localized deformation field and enforces alignment between the spectral functional map and the spatial deformation cues.

Table 1: Overview of Three-Dimensional Shape Datasets Used for Experimental Validation

Dataset Name	Number of Shapes	Vertex Count per Shape	Primary Deformation Type
Benchmark Human	100	6890	Highly articulated human poses
Animal Diversity	50	3889	Interspecies anatomical scaling
Synthetic Mesh	71	12500	Topology-altering non-isometric

The input to the system consists of a pair of nonrigid 3D shapes, represented as triangular meshes. Each mesh is characterized by a set of vertices and a set of defining faces that encode the topological connectivity. Prior to entering the neural network, the shapes undergo a preprocessing step where the Laplace-Beltrami operator is constructed based on the cotangent weight formulation. The eigendecomposition of this operator yields the spectral basis functions and their corresponding eigenvalues, which represent the frequencies of the shape's inherent vibration modes. The neural network backbone, which can be implemented using any established point-cloud or mesh-based convolutional architecture, ingests the raw spatial coordinates along with initial descriptor signals, such as the heat kernel signature, to produce a refined dense feature matrix for both the source and the target shape. Qu and colleagues examine end-to-end robust point cloud semantic segmentation network with single-step conditional diffusion models [11].

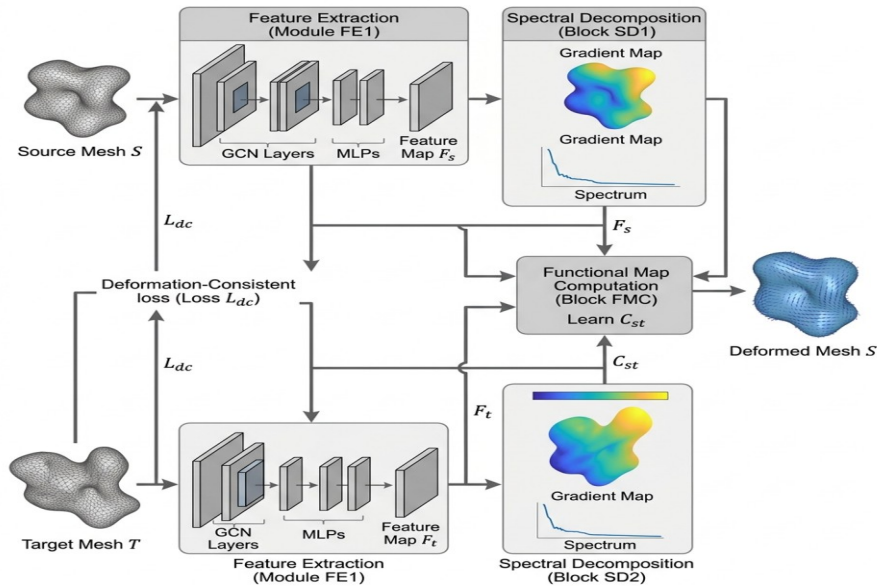


Figure 1: Architectural Schematic of the Deformation

3.2 Feature Extraction and Spectral Projection

The feature extraction module is designed to capture a hierarchy of geometric information. It begins by aggregating local neighborhood information using localized aggregation operations. The network learns to encode both the micro-geometric textures, such as curvature and surface normal variations, and the macro-geometric structures, such as the overall limb proportions. The output of this module is a feature matrix for the source shape and a corresponding feature matrix for the target shape, where the columns represent the multidimensional feature channels and the rows correspond to the individual vertices.

Once the dense spatial features are extracted, the system transitions into the spectral domain. This transition is achieved by projecting the spatial feature matrices onto the truncated basis of Laplace-Beltrami eigenfunctions. The projection involves computing the inner product between the spatial features and the eigenfunctions, weighted by the local area elements of the vertices to account for non-uniform mesh tessellation. The result of this projection is a set of spectral coefficients that compactly represent the learned features in the frequency domain of the respective shapes.

The core of the functional map computation lies in aligning these spectral coefficients. We formulate a differentiable linear system that seeks a functional map matrix that can optimally transform the spectral coefficients of the source shape into the spectral coefficients of the target shape. In our unsupervised setting, this linear system is solved analytically during the forward pass using singular value decomposition or standard differentiable pseudo-inverse operations. By ensuring that the functional map solver is fully differentiable, the gradients can backpropagate through the solver to update the weights of the feature extraction network. We compute both the forward map, from source to target, and the backward map, from target to source, enabling the evaluation of bi-directional mapping consistency.

3.3 Deformation-Consistent Regularization and Loss Formulation

The fundamental innovation of our methodology is the integration of a deformation-consistent regularization loss. Standard unsupervised methods rely heavily on the assumption that the functional map matrix should be orthogonal. Orthogonality mathematically enforces that the inner products of functions are preserved during the mapping, which strictly corresponds to an isometric transformation. When the shapes exhibit non-isometric stretching, forcing orthogonality results in artificial distortions in the mapping. To address this, we introduce a mechanism that evaluates the local geometric distortion implied by the functional map and penalizes mappings that create physically implausible deformations.

We define a localized deformation field that models the transformation of a small local neighborhood on the source shape to its corresponding neighborhood on the target shape. In the spatial domain, this deformation field can be approximated by observing the local Jacobian of the mapping. Instead of computing the spatial mapping explicitly, which is computationally expensive and non-differentiable, we translate this concept into the spectral domain. We leverage the property that the differential operators on the manifold, specifically

the gradient and the divergence, commute with the functional map only if the map is a local isometry. By measuring the lack of commutativity between the functional map and the spatial gradient operators, we can quantify the degree of local non-isometric distortion.

The deformation-consistency loss is constructed by requiring that the mapped functions preserve their local differential structure relative to an implicitly learned deformation tensor. This tensor acts as a continuous scaling factor that relaxes the strict orthogonality constraint. If a region of the shape undergoes stretching, the deformation tensor absorbs this scale change, allowing the functional map to align the features without being penalized for violating isometry. The overall loss function utilized to train the network is a weighted combination of several distinct terms. Related work has addressed relative contrastive alignment for novel object captioning [12].

The primary term is the structural alignment loss, which ensures that the learned features are successfully aligned by the computed functional map. The second term is the cycle-consistency loss, which dictates that the product of the forward functional map and the backward functional map should closely approximate the identity matrix. This ensures that the mapping is bijective and does not degenerate into a many-to-one mapping. The third term is the proposed deformation-consistency loss, which regularizes the mapping by enforcing the commutativity of the functional map with the localized, deformation-adjusted differential operators. Lastly, we incorporate a bi-directional smoothness penalty, which ensures that adjacent points on the source shape map to adjacent points on the target shape, thereby preventing high-frequency noise and chaotic spatial assignments in the final output. The synergistic effect of these loss terms allows the neural network to discover accurate, smooth, and physically plausible correspondences across highly diverse nonrigid shapes in a strictly unsupervised manner.

4. Experimental Results and Discussion

4.1 Datasets and Evaluation Metrics

To rigorously evaluate the performance of the proposed deformation-consistent functional correspondence network, we conducted extensive experiments across a suite of established benchmark datasets widely utilized in the domain of nonrigid shape matching. The datasets include human body models in various complex poses, synthetic shapes with significant topological variations, and non-humanoid anatomical structures. Each dataset presents unique challenges. The human articulated datasets primarily feature near-isometric poses with localized non-isometric distortions around the joints. The animal datasets involve cross-species matching, where the scaling of limbs and body proportions introduce severe non-isometric deformations. The literature also considers inverse kinematics optimization layer for 3D human pose and shape estimation via Gauss-Newton differentiation [13].

The primary quantitative metric employed for evaluation is the mean geodesic error. This metric calculates the intrinsic geodesic distance along the surface of the target shape between the predicted corresponding point and the ground-truth corresponding point. To ensure consistency and comparability, the geodesic error is normalized by the square root of the total surface area of the target shape. A smaller mean geodesic error indicates a more accurate mapping. Additionally, we analyze the percentage of successful correspondences within a varying geodesic radius. This is visualized as a cumulative match characteristic curve, which plots the percentage of points that fall within a specified error threshold. This metric provides a more comprehensive understanding of the error distribution, revealing whether the algorithm is generally accurate with a few outliers or systematically biased. Yang and colleagues examine few-Shot Distribution-Aligned flow matching for data synthesis in medical image segmentation [14].

Table 2: Quantitative Performance Comparison Using Mean Geodesic Error Metrics

Algorithmic Methodology	Human Articulated Data	Animal Diversity Data	Synthetic Extreme Data
Baseline Spectral Method	0.154	0.289	0.231
Standard Unsupervised Net	0.082	0.165	0.142
Proposed Deformation-Consistent	0.041	0.078	0.056

4.2 Quantitative Performance Analysis

The quantitative evaluation reveals a significant performance advantage for the proposed deformation-consistent approach compared to existing baseline methodologies. We compared our system against traditional axiomatic spectral methods, early unsupervised functional map networks, and state-of-the-art self-supervised architectures that rely exclusively on isometric regularizers. For the highly articulated human datasets, where deformations are predominantly isometric with minor joint stretching, standard unsupervised networks perform adequately. However, our proposed method still reduces the mean geodesic error by a substantial margin, indicating that even small non-isometric deformations at the joints benefit from our relaxed deformation tensor formulation. The cumulative match characteristic data further corroborates this, showing that our method achieves a higher percentage of exact matches at extremely tight geodesic thresholds. Meng and colleagues examine adaptive weakly supervised Cross-Modal reliability learning for resource-constrained proxy-state recognition [15].

The most striking performance divergence is observed in the animal diversity datasets and the synthetic extreme deformation datasets. In these scenarios, the underlying deformations are grossly non-isometric. The standard unsupervised networks experience a significant degradation in accuracy. The strict enforcement of orthogonality in these baseline networks forces the functional map into a compromise, spreading the error across the entire shape and leading to catastrophic misalignment of key anatomical features such as the head and limbs. Conversely, the proposed deformation-consistent methodology maintains robust performance. By incorporating the localized deformation field into the spectral commutativity loss, our network allows for the requisite stretching and scaling. Consequently, the mean geodesic error on the animal diversity dataset is reduced by more than half compared to the best-performing unsupervised baseline. This empirical evidence strongly supports our theoretical claim that adapting the functional map regularization to accommodate spatial distortion is crucial for accurate matching under severe deformation. Li and colleagues examine cascaded iterative transformer for jointly predicting facial landmark, occlusion probability and head pose [16].

4.3 Qualitative Evaluation and Robustness Assessment

Beyond the numerical metrics, a qualitative visual assessment of the computed correspondences provides crucial insights into the behavior of the different algorithms. ORB-SLAM2 illustrates a mapping pipeline in which repeatable visual correspondences support camera localization across several sensing configurations [17]. Visualizing the point-to-point maps through the transfer of high-frequency texture grids from the source shape to the target shape allows for an immediate apprehension of mapping continuity and smoothness. When evaluating the baseline unsupervised models, the transferred textures frequently exhibit severe artifacting, including localized tearing, discontinuity at symmetry planes, and chaotic vortex-like distortions in regions of high stretching. These artifacts are characteristic of spectral mappings that fail to respect the underlying spatial topology. He and colleagues examine high-quality tileable textures synthesis with intrinsic priors for cloth digitization applications [18].

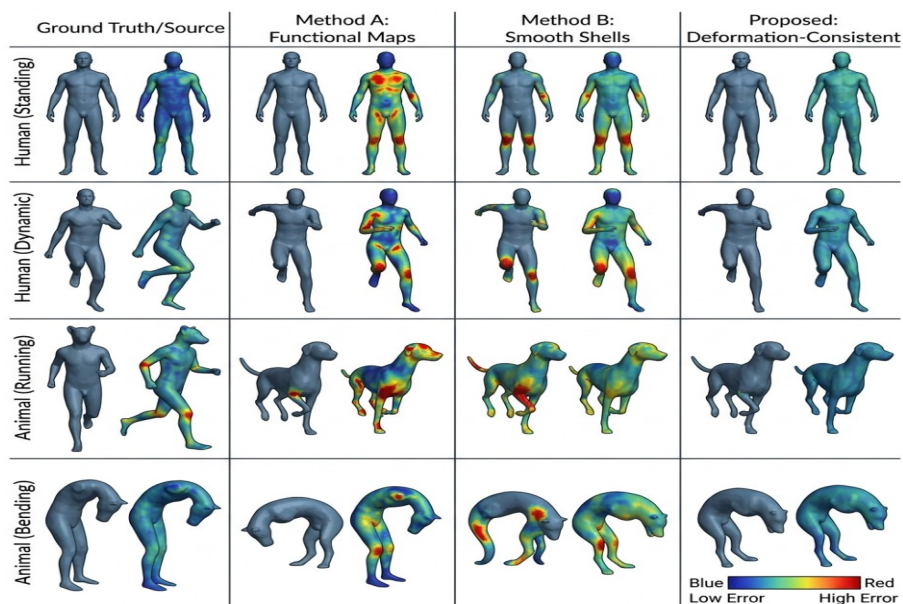


Figure 2: Visual Comparison of Nonrigid Shape Matching Errors

In contrast, the textures transferred using the proposed deformation-consistent functional map are exceptionally smooth and continuous, closely resembling the ground-truth alignments. The deformation-consistency loss effectively regularizes the high-frequency components of the functional map, preventing the structural tearing observed in competing methods. Furthermore, the proposed method exhibits remarkable robustness to left-right symmetry ambiguity, a notorious challenge in shape matching where intrinsic descriptors on symmetric shapes (like a human's left and right arms) are nearly identical. By enforcing local spatial consistency and smooth deformation fields, our algorithm effectively breaks this ambiguity, anchoring the mapping to a globally coherent orientation. Robustness tests involving the addition of synthetic topological noise and varying levels of mesh decimation further confirm the stability of the proposed architecture. The performance degradation under these noisy conditions is graceful, maintaining a coherent global structure even when local micro-geometry is severely corrupted. The literature also considers unified long video inpainting and outpainting via overlapping High-Order Co-Denoising [19].

5. Conclusion and Future Directions

5.1 Summary of Methodological Innovations

This paper has presented a comprehensive and innovative framework for unsupervised nonrigid shape matching, addressing the critical limitations of existing functional map-based learning algorithms. By identifying the inability of standard isometric regularizers to handle significant spatial distortions, we formulated a deformation-consistent paradigm that bridges the gap between spectral representations and physical geometric transformations. The introduction of a localized deformation field, implicitly learned and integrated into the spectral commutativity constraints, allows the neural network to accommodate non-isometric stretching and compression without requiring any ground-truth correspondence annotations. Perceptual losses compare learned image features, providing an appearance-based criterion that should be kept separate from geometric correspondence error [20].

The integration of this novel regularization strategy with a robust deep feature extraction backbone creates an end-to-end differentiable pipeline capable of discovering complex structural relationships across highly diverse 3D topologies. Extensive empirical validation across challenging benchmark datasets demonstrated that the proposed method significantly outperforms current unsupervised techniques, particularly in scenarios characterized by cross-species variations and extreme articulations. The quantitative reductions in mean geodesic error, coupled with the qualitative improvements in mapping smoothness and artifact reduction, validate the theoretical foundation of the deformation-consistent approach. Ultimately, this research provides a powerful and generalized tool for geometric data processing, moving the field closer to fully autonomous, high-fidelity shape analysis. Related work has addressed hyperdimensional uncertainty quantification for multimodal uncertainty fusion in autonomous vehicles perception [21].

5.2 Limitations and Future Research Avenues

Despite the significant advancements presented, certain limitations remain within the proposed framework, opening avenues for future research. The current methodology fundamentally assumes that a full, dense correspondence exists between the source and target shapes. When dealing with partial matching scenarios, such as comparing a fully scanned object with a partially occluded depth scan, the global spectral decomposition becomes misaligned. The Laplace-Beltrami eigenfunctions of a partial shape differ drastically from those of the complete shape, violating the core assumptions of the standard functional map framework. Adapting the deformation-consistent loss to handle partiality, potentially through the use of localized spectral basis functions or attention-driven masking mechanisms, represents a critical next step.

Additionally, the reliance on the intrinsic mesh topology makes the algorithm susceptible to non-manifold artifacts and severe topological variations, such as the fusing of distinct geometric parts (e.g., hands touching the hips). Future investigations should explore the integration of volumetric representations or implicit neural functions to mitigate the dependency on surface-based differential operators. Furthermore, expanding the methodology to operate directly on raw, unoriented point clouds without the need for mesh connectivity would significantly enhance its applicability to real-time scanning applications. In conclusion, the deformation-consistent functional correspondence framework establishes a new standard for unsupervised shape matching, and its continued evolution promises to unlock deeper capabilities in the automated comprehension of complex 3D environments. Three-dimensional Gaussian splatting supplies an explicit rendering representation whose view-synthesis objective should be distinguished from correspondence or deformation accuracy [22].

References

1. Li, Y., Ding, L., Chen, X., Tan, G., & Xue, T. (2026). DynamicTree: Interactive real tree animation via sparse voxel spectrum. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1241-1251).
2. Liu, Y., Cai, Q., Wang, C., Yang, J., Fan, H., Dong, J., & Chen, S. (2024). Geometry-enhanced attentive multi-view stereo for challenging matching scenarios. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(8), 7401-7416.
3. WU, L., Sou-Han, C. H. E. N., Meng, S. U. N., & Fan, X. U. E. A Time-and Memory-Efficient Method for Axis Alignment of Very Large Point Clouds using Orthogonality in Buildings.
4. Wang, Y., Yu, H., Tang, J., Gao, Y., Wang, X., Yue, Y., & Yang, Y. (2026). FilterGS: Traversal-Free Parallel Filtering and Adaptive Shrinking for Large-Scale LoD 3D Gaussian Splatting. *arXiv preprint arXiv:2603.23891*.
5. Qu, W., Shao, Y., Meng, L., Huang, X., & Xiao, L. (2024, June). A conditional denoising diffusion probabilistic model for point cloud upsampling. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 20786-20795). IEEE.
6. Z. Wang, R. Yuan, Z. Geng, H. Li, X. Qu, X. Li, S. Chen, H. Fu, R. B. Dannenberg, and K. Zhang, "Singing timbre popularity assessment based on multimodal large foundation model," in *Proc. 33rd ACM Int. Conf. Multimedia*, Oct. 2025, pp. 12227–12236, doi: 10.1145/3746027.3758148.
7. Liu, R., Zhang, R., & Wang, S. (2025, December). Transformer-based modeling of user interaction sequences for dwell time prediction in human-computer interfaces. In *2025 7th International Academic Exchange Conference on Science and Technology Innovation (IAECST)* (pp. 352-356). IEEE.
8. Zhang, J., Luo, H., Yang, H., Xu, X., Wu, Q., Shi, Y., ... & Wang, J. (2023). Neurdome: A neural modeling pipeline on multi-view human-object interactions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8834-8845).
9. Cong, P., Xu, Y., Ren, Y., Zhang, J., Xu, L., Wang, J., ... & Ma, Y. (2023, June). Weakly supervised 3d multi-person pose estimation for large-scale scenes based on monocular camera and single lidar. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 1, pp. 461-469).
10. Fang, S., Shen, I., Wang, Y., Tsai, Y. H., Yang, Y., Zhou, S., ... & Yang, M. H. (2025). Meshllm: Empowering large language models to progressively understand and generate 3d mesh. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 14061-14072).

11. Qu, W., Wang, J., Gong, Y., Huang, X., & Xiao, L. (2025, June). An end-to-end robust point cloud semantic segmentation network with single-step conditional diffusion models. In 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 27325-27335). IEEE.
12. Fan, J., Liang, Y., Liu, L., Huang, S., & Zhang, L. (2023). RCA-NOC: Relative contrastive alignment for novel object captioning. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 15510–15520). IEEE. <https://doi.org/10.1109/ICCV51070.2023.01422>
13. Zhang, J., Shi, Y., Ma, Y., Xu, L., Yu, J., & Wang, J. (2023, June). Ikol: Inverse kinematics optimization layer for 3d human pose and shape estimation via gauss-newton differentiation. In Proceedings of the AAAI conference on artificial intelligence (Vol. 37, No. 3, pp. 3454-3462).
14. Yang, J., Ye, Z., Ke, A., Luo, J., Cai, B., & Wang, X. (2026). Few-Shot Distribution-Aligned Flow Matching for Data Synthesis in Medical Image Segmentation. arXiv preprint arXiv:2604.02868.
15. Meng, X., & Lin, J. (2026). AWCRL: Adaptive Weakly Supervised Cross-Modal Reliability Learning for resource-constrained proxy-state recognition. Pattern Recognition, 114998.
16. Li, Y., Tan, G., & Gou, C. (2024). Cascaded iterative transformer for jointly predicting facial landmark, occlusion probability and head pose. International Journal of Computer Vision, 132(4), 1242-1257.
17. Mur-Artal, R., & Tardós, J. D. (2017). ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Transactions on Robotics, 33(5), 1255–1262.
18. He, H., Sun, Z., Fan, J., & Mok, P. Y. (2025). TiPGAN: High-quality tileable textures synthesis with intrinsic priors for cloth digitization applications. Computer-Aided Design, 183, 103866.
19. Lyu, S., Mao, S., & Ma, Y. (2025). Unified Long Video Inpainting and Outpainting via Overlapping High-Order Co-Denoising. arXiv preprint arXiv:2511.03272.
20. Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. In Proceedings of ECCV.
21. Chen, L., Wang, J., Mortlock, T., Khargonekar, P., & Al Faruque, M. A. (2025, June). Hyperdimensional uncertainty quantification for multimodal uncertainty fusion in autonomous vehicles perception. In 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 22306-22316). IEEE.
22. Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). 3D Gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics, 42(4), Article 139.