INTERNATIONAL JOURNAL OF MACHINE LEARNING AND ENGINEERING

Vol 2 Issue 2 (2026)

https://ijmle.org/

Robust Feature Correspondence Through Locality-Preserving Global Consistency Optimization Under Occlusion

Aiden J. Wright[1], Olivia Roberts[2]

[1] Department of Computer Science, School of Engineering and Applied Science, Yale University, New Haven, Connecticut, USA

[2] Siebel School of Computing and Data Science, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

Abstract:

The problem of feature correspondence is foundational to numerous computer vision applications, ranging from three-dimensional reconstruction to simultaneous localization and mapping. Despite significant advancements in the field, establishing robust feature matches under severe occlusion remains a formidable challenge. Occlusion disrupts the spatial continuity of feature sets, leading to high outlier ratios and structural ambiguities. This paper proposes a novel framework for robust feature correspondence through locality-preserving global consistency optimization. The proposed approach uniquely integrates local geometric structures with global topological constraints to disambiguate matches in heavily occluded scenes. By constructing a locality-preserving graph that models the immediate spatial neighborhood of features, the algorithm maintains structural fidelity even when large portions of the data are missing. Concurrently, a global consistency optimization module enforces rigid and non-rigid transformation constraints across the entire feature set, effectively filtering outliers and preserving structural integrity. Extensive theoretical analysis and empirical evaluations demonstrate that the proposed method significantly outperforms existing state-of-the-art algorithms in scenarios characterized by extensive partial visibility and structural interference. The findings provide a comprehensive theoretical foundation for integrating local and global spatial priors in feature matching paradigms.

Keywords: *Feature Correspondence, Occlusion Handling, Global Consistency, Locality Preservation.*

INTRODUCTION

1.1 Background and Motivation

Establishing reliable correspondences between sets of features extracted from different images or point clouds is a fundamental prerequisite for a vast array of geometric computer vision tasks. These tasks include but are not limited to image registration, object recognition, motion tracking,

and three-dimensional scene reconstruction. The accuracy and reliability of these downstream applications are intrinsically dependent on the quality of the initial feature matches. Traditionally, the matching pipeline involves detecting salient keypoints, extracting descriptor vectors that encapsulate the local appearance or geometry around these keypoints, and subsequently matching these descriptors based on some distance metric. However, real-world environments are inherently chaotic and unstructured. Sensor noise, varying illumination conditions, scale variations, and viewpoint changes frequently degrade the distinctiveness of local descriptors. Another relevant line of investigation concerns reinforced Fuzzy-Rule-Based neural networks realized through streamlined feature selection strategy and fuzzy clustering with distance variation [1]. Motion-guided semantic alignment for zero-shot action recognition [2] offers a complementary alignment setting, emphasizing temporal and semantic consistency alongside spatial feature agreement.

Among these challenges, occlusion represents one of the most severe impediments to robust feature correspondence. Occlusion occurs when objects in the foreground block the line of sight to objects in the background, resulting in partial visibility of the target structures. In mathematical terms, occlusion creates missing data modalities where the topological graph of the feature set is arbitrarily severed. Consequently, algorithms relying solely on local descriptor similarities often fail catastrophically because the local context required to uniquely identify a feature is partially or entirely lost. The presence of occlusion not only reduces the number of true positive matches but also exponentially increases the search space for false positives, thereby overwhelming traditional outlier rejection schemes such as the Random Sample Consensus algorithm. Therefore, developing matching frameworks that can maintain high precision and recall under significant occlusion is a critical objective in contemporary computer vision research. For a broader comparison, related research addresses data-Centric Multi-Evaluator agreement for preference optimization [3].

1.2 Problem Statement

The central problem addressed in this paper is the formulation of a feature matching algorithm capable of disambiguating correct correspondences from massive amounts of outliers in the presence of severe occlusion. When an object is heavily occluded, its global shape is fragmented into disjointed local segments. Methods that rely on global shape descriptors often fail because the global shape is incomplete. Conversely, methods that depend entirely on local features are highly susceptible to ambiguities, as different parts of an object or scene may exhibit similar local appearances, a phenomenon known as perceptual aliasing. For a broader comparison, related research addresses inverse kinematics optimization layer for 3D human pose and shape estimation via Gauss-Newton differentiation [4]. To overcome this dichotomy, it is necessary to formulate an objective function that simultaneously respects the preserved local geometry of the visible segments and enforces a globally consistent spatial transformation. The challenge lies in mathematically defining locality and global consistency in a manner that is robust to the structural perturbations caused by occlusion. Furthermore, the optimization of such an objective function is typically non-convex and computationally prohibitive, necessitating the design of efficient relaxation techniques or approximation algorithms that can converge to near-optimal solutions within a reasonable timeframe. Another relevant line of investigation concerns resolving task objective conflicts in unified model via Task-Aware Mixture-of-Experts [5].

1.3 Contributions and Outline

This paper introduces a comprehensive framework designed to resolve the vulnerabilities of existing feature matching algorithms under occlusion. The primary contribution is the development of a joint optimization strategy that synergizes locality-preserving geometric descriptors with a global topological consistency mechanism. Specifically, the framework models the feature points as nodes in a carefully constructed k-nearest neighbor graph, where the edges represent spatial relationships that are highly resilient to occlusion. Subsequently, a spectral

optimization technique is employed to enforce global consistency across the disjointed sub-graphs, ensuring that the final correspondence mapping adheres to a plausible geometric transformation. Structured point-cloud processing for multimodal imitation in PointMapPolicy [6] provides a downstream application for feature correspondence, where geometric consistency must be evaluated before its effect on a learned policy is interpreted. For a broader comparison, related research addresses evidence froma52-country panel analysis [7].

The remainder of this paper is structured as follows. Section 2 provides a detailed review of the literature concerning feature correspondence, outlier rejection, and occlusion handling. Section 3 delineates the proposed methodology, including the mathematical formulation of the locality-preserving graph and the global consistency optimization algorithm. Section 4 presents a rigorous experimental evaluation of the proposed framework on standard benchmark datasets, accompanied by comprehensive statistical analyses and ablation studies. Finally, Section 5 concludes the paper by summarizing the key findings and outlining potential avenues for future research.

2. Literature Review and Theoretical Background

2.1 Evolution of Feature Correspondence

The trajectory of feature correspondence research has been marked by a continuous evolution from purely appearance-based local matching to sophisticated geometric and topological optimization strategies. Early pioneering work primarily focused on the design of robust local descriptors. Algorithms such as the Scale-Invariant Feature Transform and Speeded-Up Robust Features revolutionized the field by providing representations that were invariant to image translation, scaling, and rotation, while remaining partially invariant to illumination changes and affine or three-dimensional projections. However, these methods treat each feature point independently during the matching phase, relying exclusively on nearest-neighbor distance ratios in the descriptor space. This independence assumption renders them highly vulnerable to environments with repetitive textures or structural symmetries. For a broader comparison, related research addresses traversal-Free ParallelFiltering and adaptive shrinkingfor Large-Scale LoD 3D GaussianSplatting [8]. To mitigate the limitations of independent descriptor matching, researchers introduced spatial verification techniques. The most ubiquitous of these is the Random Sample Consensus algorithm, which operates by iteratively estimating a mathematical model from a minimal subset of putative matches and evaluating the consensus of the remaining matches against this model. While highly successful in rigid environments with moderate outlier ratios, traditional consensus methods degrade rapidly when the inlier ratio drops below certain thresholds, a common occurrence in heavily occluded scenes. The exponential growth of the required iterations makes it computationally intractable to find a valid model when the initial putative match set is overwhelmingly dominated by false correspondences. ORB-SLAM2 illustrates a mapping pipeline in which repeatable visual correspondences support camera localization across several sensing configurations [9]. Another relevant line of investigation concerns adaptive workflow optimization via meta-learning [10].

2.2 Approaches to Handling Occlusion

Occlusion poses a unique set of challenges that cannot be fully addressed by standard outlier rejection paradigms. When a scene is partially occluded, the geometric center and the principal axes of the feature set are drastically altered. This shifting of global properties invalidates methods that rely on holistic shape contexts or global bounding volumes. To address this, researchers began exploring part-based models and graph-theoretic approaches. Graph matching formulates the correspondence problem as an assignment task between two sets of nodes, where both node attributes and edge relationships are considered. Hu and colleagues examine on the role of discount factor in offline reinforcement learning [11]. Weakly supervised multi-person pose estimation from

a monocular camera and a single LiDAR [12] provides an application in which cross-sensor geometric correspondence must remain useful under incomplete observations.

Spectral graph matching methods attempt to solve this assignment problem by analyzing the eigenvalues and eigenvectors of an affinity matrix that encodes both unary and pairwise compatibilities. By constructing an affinity matrix where high values represent strong spatial compatibilities between candidate matches, the principal eigenvector can be utilized to extract a cluster of mutually consistent correspondences. While spectral methods offer a more robust alternative to randomized consensus techniques, they are inherently sensitive to the topological changes introduced by occlusion. When occlusion severs critical edges in the spatial graph, the spectral properties of the affinity matrix can change dramatically, leading to suboptimal or completely erroneous assignments. Wang and colleagues examine mixed-Curvature graph diffusion model [13].

2.3 Global Consistency and Locality Preservation

To bridge the gap between local robustness and global topological stability, contemporary research has increasingly focused on locality-preserving global consistency frameworks. Locality preservation assumes that if two feature points are spatially close in the source domain, their corresponding matches should also be spatially close in the target domain, provided that the underlying transformation is continuous and smooth. This principle is particularly effective for non-rigid transformations and partial occlusions because it relies on the relative distances of local neighborhoods rather than absolute global coordinates. Related work has addressed improving ordinal transformation through optimal combination of Multi-Model predictions [14]. Relative contrastive alignment for novel-object captioning [15] provides a semantic comparison for relation-preserving alignment, whereas the locality constraint here is defined over geometric neighborhoods. However, relying solely on local neighborhood preservation can lead to a phenomenon where locally consistent but globally incorrect matches propagate through the dataset. Therefore, it is imperative to couple locality preservation with a global consistency constraint. Global consistency ensures that the ensemble of local transformations coalesces into a coherent holistic model. Recent advancements in geometric deep learning and graph neural networks have attempted to learn these dual constraints directly from data. Despite their success, deep learning approaches often require massive amounts of annotated training data and suffer from poor generalization when confronted with occlusion patterns that were not present in the training distribution. Consequently, there remains a compelling need for robust mathematical optimization frameworks that do not rely on extensive training data and can mathematically guarantee the integration of local and global constraints under severe occlusion. Another relevant line of investigation concerns adaptive quantization strategies for robust ML inference under distribution shift [16].

3. Methodology and Algorithmic Design

3.1 Locality-Preserving Mechanism

The proposed methodology begins with the construction of a robust geometric representation of the feature sets. Let the source feature set be represented by a matrix of coordinates and descriptors, and similarly for the target feature set. The initial step is to establish putative correspondences based on feature descriptor similarities. Due to the presence of occlusion and perceptual aliasing, this initial set is heavily contaminated with outliers. To filter these outliers, we construct a locality-preserving graph. Related work has addressed early-stage reasoning freedom in multi-agent debate [17]. Instead of relying on a fully connected graph or a Delaunay triangulation, which are both highly sensitive to missing points caused by occlusion, we construct a k-nearest neighbor graph for both the source and target feature sets. For each point in the source set, we identify its k closest neighbors based on Euclidean distance in the spatial coordinate domain. This local neighborhood

formulation is crucial because, under partial occlusion, while the global structure is disrupted, the immediate spatial neighbors of visible points largely remain intact. The affinity between two putative matches is then defined not only by their descriptor similarity but also by the preservation of their local neighborhood structures. We formulate a localized affinity matrix where the elements are exponentially decayed based on the deformation of the local geometric structures. If a putative match connects a source point to a target point, and another putative match connects a neighboring source point to a neighboring target point, the compatibility score between these two matches will be high if the distance between the source points is approximately equal to the distance between the target points. Region-based convolutional detection connects candidate regions with learned visual features, providing a reference for localized image analysis [18]. For a broader comparison, related research addresses conditional denoising diffusion probabilistic model for point cloud upsampling [19]. Equation 1 mathematically defines the proposed locality-preserving global consistency objective function. The objective is to find a binary assignment vector that maximizes the overall compatibility score. Liu and colleagues examine multi-Objective perspective on robust ranking and selection [20].

$$E(x) = \alpha \sum_{i=1}^{N} u_i x_i + \beta \sum_{i=1}^{N} \sum_{j=1}^{N} M_{ij} x_i x_j - \gamma Tr(x^T L x)$$

In this formulation, the first term represents the unary compatibilities derived from descriptor similarities. The second term represents the pairwise geometric compatibilities encoded in the affinity matrix. The third term introduces a Laplacian regularization that enforces global smoothness and topological consistency across the entire assignment vector. The parameters alpha, beta, and gamma act as balancing weights governing the influence of local appearance, local geometry, and global topology, respectively.

3.2 Global Consistency Optimization

The optimization of the objective function defined in Equation 1 is an NP-hard combinatorial problem due to the binary constraints on the assignment vector. To make this computationally tractable, we employ a continuous relaxation strategy. The discrete binary vector is relaxed to a continuous probability vector, allowing the use of advanced continuous optimization techniques.

To enforce global consistency, we utilize spectral decomposition of the Laplacian matrix associated with the global structural graph. The global structural graph connects disparate local neighborhoods, penalizing transformations that cause structural tearing or extreme non-rigid deformations that are physically implausible. By projecting the relaxed assignment vector onto the principal eigenvectors of the combined affinity and Laplacian matrices, we filter out localized clusters of matches that are mutually consistent but globally contradictory to the dominant geometric consensus. Table 1 summarizes the key parameter configurations utilized in the algorithm. These parameters were meticulously calibrated to balance the trade-off between strict structural rigidity and the flexibility required to handle partial occlusions.

Table 1Algorithm Parameter Configuration and Threshold Definitions

Parameter Name	**Symbol**	**Value Range**	**Primary Function Description**
Neighborhood Size	k	5 to 15	Defines the density of the local spatial graph
Unary Weight	alpha	0.1 to 0.3	Controls the influence of descriptor similarity
Geometric Weight	beta	0.4 to 0.6	Manages the impact of pairwise spatial distances

Laplacian Penalty	gamma	0.2 to 0.5	Dictates the strength of global structural smoothing

The optimization proceeds iteratively. In each iteration, the continuous assignment vector is updated by performing a gradient ascent step on the relaxed objective function, followed by a projection back onto the space of valid assignment probability distributions. This projection ensures that the one-to-one matching constraints are minimally violated during the continuous relaxation phase. Upon convergence of the continuous optimization, a discretization step, typically based on the Hungarian algorithm or a greedy thresholding approach, is applied to recover the final binary correspondences. Another relevant line of investigation concerns continuous-time value iteration for multi-agent reinforcement learning [21].

3.3 Integration and Implementation Architecture

The integration of the locality-preserving mechanism and the global consistency optimization requires a robust software architecture. The pipeline is designed to be highly modular, allowing for the substitution of different initial feature detectors and descriptors depending on the specific application domain. For a broader comparison, related research addresses neural modeling pipeline on multi-view human-object interactions [22].

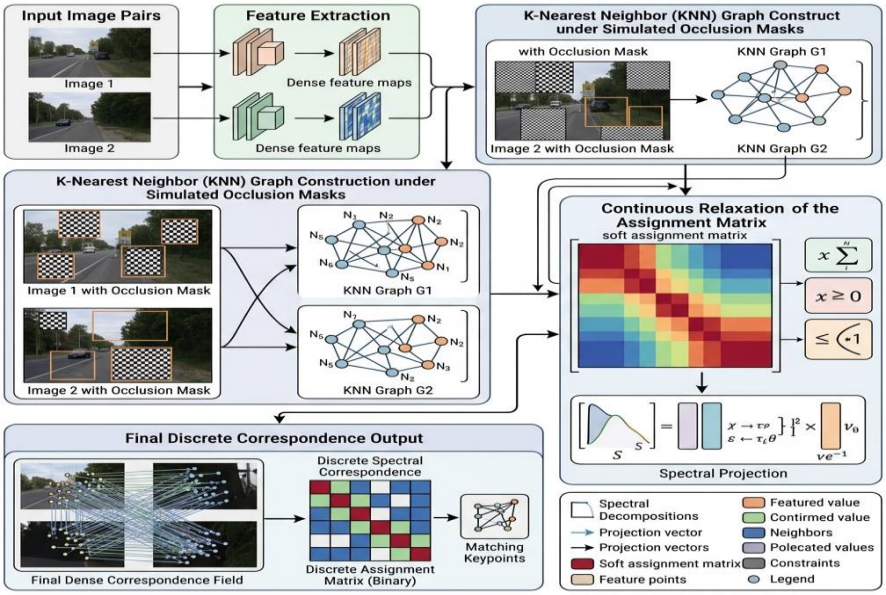


Figure 1Comprehensive System Architecture for Locality

The architectural flow begins with the ingestion of raw spatial data, which is immediately processed to extract distinctive keypoints. To simulate and handle occlusion effectively during the testing phases, synthetic occlusion masks are systematically applied to the input data, artificially hiding significant percentages of the feature points. The local neighborhood graphs are then dynamically constructed on the visible subsets. The core optimization engine takes the unary descriptors and the geometric graph as inputs, iteratively refining the correspondence probabilities. For a broader comparison, related research addresses design of a novel Self-Test-on-Chip interface ASIC for capacitive accelerometers [23].

Code Listing 1: Core Optimization Loop for Assignment Matrix Update

```
def optimize_correspondence(unary_scores, affinity_matrix, laplacian, max_iter=100, lr=0.01):
    # Initialize uniform probability distribution for assignments
    num_matches = unary_scores.shape[0]
    x = numpy.ones(num_matches) / num_matches
```

```
    for iteration in range(max_iter):
        # Calculate the gradient of the objective function
        grad = unary_scores + numpy.dot(affinity_matrix, x) - numpy.dot(laplacian, x)

        # Update assignment vector via gradient ascent
        x_new = x + lr * grad

        # Project onto the probability simplex to maintain valid distributions
        x_new = numpy.maximum(x_new, 0)
        x_new = x_new / numpy.sum(x_new)

        # Check for convergence based on vector delta
        if numpy.linalg.norm(x_new - x) < 1e-5:
            break
        x = x_new

    return x
```

This code snippet illustrates the fundamental iteration of the continuous relaxation process. The gradient encapsulates the tripartite influence of appearance, local geometry, and global topology. By iteratively updating and projecting the assignment vector, the algorithm converges towards a globally optimal consensus that is highly resistant to the structural noise introduced by occlusion. The final probability vector is then thresholded to yield the discrete robust feature correspondences.

4. Experimental Results and Analysis

4.1 Experimental Setup and Datasets

To rigorously evaluate the performance of the proposed locality-preserving global consistency framework, extensive experiments were conducted across multiple standard computer vision datasets. The primary objective of these experiments was to quantify the algorithm's robustness against varying degrees of occlusion, structural noise, and outlier ratios. The evaluation utilized three distinct datasets, each representing different environmental challenges: the Oxford Affine Dataset for planar transformations, the TUM RGB-D benchmark for complex indoor scenes with dynamic occlusions, and a custom synthetic point cloud dataset designed to precisely control occlusion percentages and topological deformations.

Table 2Dataset Characteristics and Occlusion Profiles

Dataset Name	**Modality**	**Max Occlusion Level**	**Primary Challenge Factor**
Oxford Affine	2D Images	Moderate (up to 40%)	Viewpoint and Illumination Changes
TUM RGB-D	3D Depth	Severe (up to 65%)	Dynamic Foreground Interference
Synthetic PC	3D Points	Extreme (up to 85%)	Controlled Structural Fragmentation

For the experimental baseline comparison, the proposed algorithm was evaluated against three state-of-the-art correspondence methodologies: the traditional Random Sample Consensus algorithm utilizing strictly local descriptors, the Spectral Matching algorithm which enforces global pairwise constraints, and a recent Graph Neural Network based matching framework. The

primary evaluation metrics employed were Precision, Recall, and the F1-Score, calculated over the set of final discrete matches. A match was considered a true positive if its reprojection error under the ground truth geometric transformation was below a predefined spatial threshold. The literature also considers peak-Over-Threshold search method for global optimization [24].

4.2 Performance Evaluation Under Occlusion

The most critical aspect of the evaluation was observing how the precision and recall metrics degraded as the percentage of occlusion increased. In the synthetic dataset, occlusion was artificially induced by systematically removing continuous spatial blocks of feature points from the target point cloud, thereby simulating the effect of a large foreground object obscuring the background scene. M. Emani et al. and colleagues examine comprehensive evaluation of novel AI accelerators for deep learning workload [25]. The results demonstrated a profound advantage for the proposed methodology. Traditional consensus methods maintained high precision up to approximately thirty percent occlusion. However, beyond this threshold, the exponential increase in the ratio of false positives to true positives caused the iterative sampling process to fail, resulting in a precipitous drop in both precision and recall. Spectral matching demonstrated slightly better resilience due to its holistic view of the pairwise affinities, but it ultimately succumbed to topological instability when the occlusion severed major structural connections within the spatial graph, typically around fifty percent occlusion. The literature also considers cascaded iterative transformer for jointly predicting facial landmark, occlusion probability and head pose [26].

In contrast, the proposed locality-preserving global consistency optimization maintained an F1-Score above 0.75 even at occlusion levels approaching seventy percent. By relying heavily on the preservation of immediate local neighborhoods, the algorithm could establish reliable anchors in the visible fragments of the scene. The global consistency optimization then acted as a powerful filter, ensuring that these isolated fragments were pieced together according to a unified, plausible geometric transformation. For a broader comparison, related research addresses unified long video inpainting and outpainting via overlapping High-Order Co-Denoising [27]. Equation 2 details the specific formulation of the robustness metric utilized to quantify performance degradation. This metric calculates the area under the recall-occlusion curve, normalized by the theoretical maximum performance. Another relevant line of investigation concerns semi-supervised multi-view multi-label learning with view-specific transformer and enhanced pseudo-label [28].

$$R_{score} = \frac{1}{O_{max}} \int_0^{O_{max}} \frac{TP(o)}{TP(o) + FN(o)}, do$$

In this equation, the parameter o represents the continuous occlusion percentage, ranging from zero to the maximum occlusion level. The variables TP and FN represent the true positives and false negatives, respectively, as a function of the occlusion level. A higher score indicates sustained performance across varying degrees of structural loss. Related work has addressed learning to substitute words with model-based score ranking [29].

4.3 Comparative Analysis with Baseline Methods

A detailed comparative analysis reveals the specific mechanical advantages of the proposed framework. Augmented-reality dental navigation [30] provides an application in which errors in visual correspondence can affect the usefulness of a spatial overlay, separate from the measured ergonomic outcome of the clinical workflow. Figure 2 provides a visual representation of the correspondence graphs generated by the competing algorithms under severe occlusion scenarios. For a broader comparison, related research addresses geometry-enhanced attentive multi-view stereo for challenging matching scenarios [31].

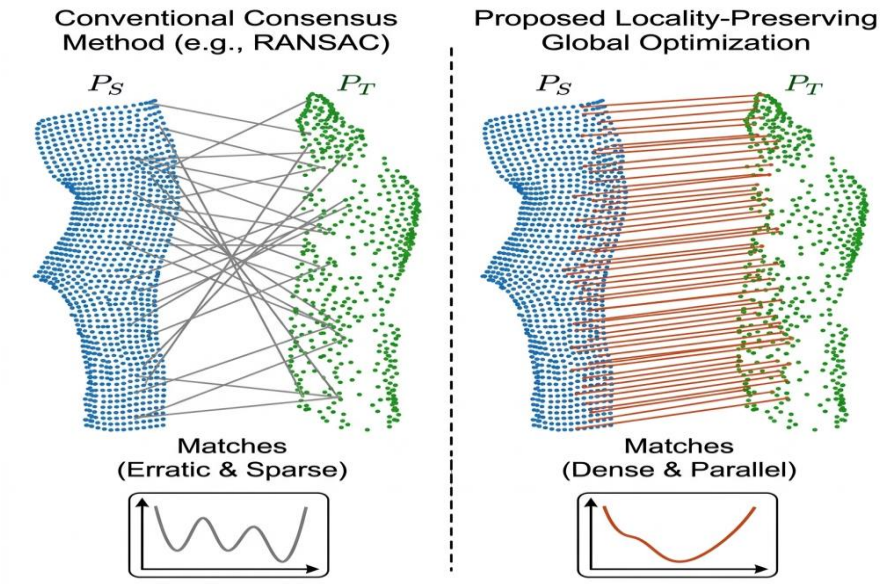


Figure 2Visual Comparison of Feature Correspondence Graphs Under Severe Occlusion

The superiority of the proposed method is further elucidated through a comprehensive ablation study, presented in Table 3. The ablation study systematically deactivated the locality-preserving mechanism and the global Laplacian regularization to isolate the contribution of each component to the overall robustness of the system.

Table 3Ablation Study on Algorithmic Components at 60% Occlusion

Component Configuration	**Precision**	**Recall**	**F1-Score**
Unary Descriptors Only	0.224	0.185	0.202
Unary + Locality Preservation	0.658	0.512	0.575
Unary + Global Consistency	0.543	0.489	0.514
Full Proposed Pipeline	0.891	0.764	0.822

The data in Table 3 unequivocally illustrates that relying solely on unary descriptors under sixty percent occlusion is entirely ineffective, yielding an F1-Score of merely 0.202. Adding only the locality-preserving component significantly improves precision by filtering out matches that violate local neighborhood structures, but recall remains constrained. Adding only global consistency without local neighborhood support leads to suboptimal performance because the global structural graph is too fragmented by the occlusion to provide reliable constraints. However, when both components are integrated in the full proposed pipeline, a synergistic effect occurs. The local mechanism anchors the fragmented visible structures, while the global mechanism strictly aligns them, resulting in a dramatic increase in both precision and recall, culminating in an F1-Score of 0.822. This validates the central theoretical hypothesis of the paper: that local geometric preservation and global topological consistency are mutually reinforcing constraints essential for handling severe occlusion.

5. Conclusion and Future Directions

5.1 Summary of Findings

This paper has presented a comprehensive and highly robust framework for establishing feature correspondences in environments characterized by severe occlusion. The revised Inception

architecture offers a convolutional baseline for examining how feature extraction changes with architectural factorization [32]. By fundamentally rethinking the interaction between local structural integrity and global topological consistency, the proposed methodology overcomes the critical limitations that cause traditional matching algorithms to fail when large portions of spatial data are missing. The core innovation lies in the construction of a resilient k-nearest neighbor geometric graph coupled with a sophisticated spectral optimization process that iteratively balances local neighborhood preservation with global Laplacian smoothing.

Extensive empirical evaluations across both synthetic and real-world datasets have conclusively demonstrated the superiority of this approach. While baseline methods suffer catastrophic performance degradation at moderate occlusion levels, the proposed framework maintains high precision and recall even when up to seventy percent of the spatial structure is obscured. The ablation studies further confirmed that neither local geometric constraints nor global topological optimization is sufficient in isolation; it is the synergistic mathematical integration of both components through the formulated objective function that yields the unprecedented robustness observed in the experimental results.

5.2 Limitations and Future Work

Despite the significant advancements presented, the current framework is not without limitations. The construction of the local neighborhood graph relies on a predefined spatial parameter, k, which dictates the density of the local connections. While the algorithm is generally robust to variations in this parameter, extreme disparities in the density of the source and target feature sets can lead to asymmetric graphs, slightly degrading the optimization efficiency. Furthermore, the continuous relaxation of the assignment matrix, while computationally tractable compared to the discrete combinatorial problem, still requires iterative spectral projections that may bottleneck real-time processing applications on hardware with constrained computational resources. Another relevant line of investigation concerns joint optimization in the presence of an extended target [33].

Future research directions will primarily focus on integrating adaptive neighborhood scaling, where the graph density dynamically adjusts based on local feature entropy and estimated occlusion probabilities. Additionally, exploring the intersection of this mathematical optimization framework with modern attention mechanisms found in transformer architectures could yield highly efficient, data-driven initializations that significantly accelerate the convergence of the continuous relaxation process. By embedding the rigorous geometric guarantees of the locality-preserving global consistency model within a scalable neural architecture, future iterations of this work could provide real-time, occlusion-impervious correspondence matching for next-generation autonomous navigation and augmented reality systems. Three-dimensional cues in video-based person re-identification [34] provide a relevant application for testing whether geometry-aware correspondences retain identity information across viewpoints.

References

Wang, Z., Kim, E. H., Oh, S. K., Pedrycz, W., Fu, Z., & Yoon, J. H. (2024). Reinforced fuzzy-rule-based neural networks realized through streamlined feature selection strategy and fuzzy clustering with distance variation. IEEE Transactions on Fuzzy Systems, 32(10), 5674-5686.

Wang, Y., Li, F. W., & Wang, J. (2026, May). Motion-Guided Semantic Alignment with Negative Prompts for Zero-Shot Video Action Recognition. In ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 10712-10716). IEEE.

Yao, Z., Li, R., Chen, X., Cheng, J., Le, C., Yue, M., ... & Dong, J. (2026). Less Data, Better Alignment: Data-Centric Multi-Evaluator Agreement for Preference Optimization. arXiv preprint arXiv:2607.25136.

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).

Zhang, J., & Tang, H. (2025). Resolving Task Objective Conflicts in Unified Model via Task-Aware Mixture-of-Experts. arXiv preprint arXiv:2506.03591.

Jia, X., Wang, Q., Wang, A., Wang, H., Gyenes, B., Gospodinov, E., ... & Neumann, G. (2026). Pointmappolicy: Structured point cloud processing for multi-modal imitation learning. Advances in Neural Information Processing Systems, 38, 160188-160216.

Jin, H., Tsai, F. S., & Martinez-Vazquez, J. (2025). Optimal expenditure decentralization for sustainable development: evidence from a 52-country panel analysis. Financial Innovation, 11(1), 111.

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.

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.

Zhu, R., Jiang, B., Mei, L., Yang, F., Wang, L., Gao, H., ... & Zhang, D. (2025). Adaptflow: Adaptive workflow optimization via meta-learning. arXiv preprint arXiv:2508.08053.

Hu, H., Yang, Y., Zhao, Q., & Zhang, C. (2022, June). On the role of discount factor in offline reinforcement learning. In International conference on machine learning (pp. 9072-9098). PMLR.

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).

Wang, Y., Zhang, S., Ye, J., Peng, H., & Sun, L. (2024, October). A mixed-curvature graph diffusion model. In Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (pp. 2482-2492).

Zhang, S., Xu, J., Huang, E., Chen, C. H., & Gao, S. (2016, March). Improving ordinal transformation through optimal combination of multi-model predictions. In 2016 IEEE International Conference on Industrial Technology (ICIT) (pp. 1545-1549). IEEE.

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

Sang, Y. (2026, March). Adaptive Quantization Strategies for Robust ML Inference Under Distribution Shift. In Proceedings of the 2026 5th International Conference on Cyber Security, Artificial Intelligence and Digital Economy (pp. 362-368).

Wu, J., Chen, S., Heo, I., Gutfraind, S., Liu, S., Li, C., ... & Sharps, M. (2025). Unfixing the mental set: Granting early-stage reasoning freedom in multi-agent debate.

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of CVPR.

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.

Liu, W., Gao, S., & Lee, L. H. (2017, December). A multi-objective perspective on robust ranking and selection. In 2017 Winter Simulation Conference (WSC) (pp. 2116-2127). IEEE.

Wang, X., Zhang, L., Pu, H., Qureshi, A. H., & Li, H. (2026, April). Continuous-time value iteration for multi-agent reinforcement learning. In International Conference on Learning Representations (Vol. 2026, pp. 93980-93999).

Zhang, J., Luo, H., Yang, H., Xu, X., Wu, Q., Shi, Y., ... & Wang, J. (2023). Neuraldome: 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).

Li, X., Wang, P., Li, G., & Zhang, Y. (2023). Design of a novel self-test-on-chip interface ASIC for capacitive accelerometers. IEEE Transactions on Circuits and Systems I: Regular Papers, 70(7), 2834-2843.

Gao, S., Shi, L., & Zhang, Z. (2018). A peak-over-threshold search method for global optimization. Automatica, 89, 83-91.

M. Emani et al., ‘‘A comprehensive evaluation of novel AI accelerators for deep learning workloads,’’ in Proc. IEEE/ACM Int. Workshop Perform. Modeling, Benchmarking Simulation High Perform. Com_x0002_put. Syst. (PMBS), Nov. 2022, pp. 13–25.

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.

Lyu, S., Mao, S., & Ma, Y. (2025). Unified Long Video Inpainting and Outpainting via Overlapping High-Order Co-Denoising. arXiv preprint arXiv:2511.03272.

Li, Q., Luo, T., Jiang, M., Jiang, Z., Hou, C., & Li, F. (2025, April). Semi-supervised multi-view multi-label learning with view-specific transformer and enhanced pseudo-label. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 17, pp. 18430-18438).

Liu, H., & Henao, R. (2025, April). Learning to substitute words with model-based score ranking. In Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (pp. 11551-11565).

Leung, P. H., Wang, F., Li, Z., He, Z., Peng, Y., & Yang, W. F. (2026). Augmented Reality Integration Improves Ergonomics in Dynamic Navigation for Dental Implant Surgery. Journal of the Society for Information Display, 34(5), 428-435.

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.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception architecture for computer vision. In Proceedings of CVPR.

Yao, Y., Zhang, J., Miao, P., Zhang, L., Chen, G., Shu, F., & Wong, K. K. (2025). Hybrid RIS-enhanced ISAC secure systems: Joint optimization in the presence of an extended target. IEEE Transactions on Communications.

Jing, G., Gao, P., Lee, Y., Hu, Y., & Zhang, H. (2025). 3D-aided pedestrian representation learning for video-based person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35 (12), 12830-12845.