HyperDSC: Reject Outliers with High-Order Deep Spatial Compatibility Learning for 3D Registration

Kezheng Xiong*[1], Zhehan Xie†[1], Wenkai Lin†[2], Fangyu Zhou†[1]
31520241154530, 23020241154459, 36920241153237, 23020241154478
[1]School of Informatics, [2]Institution of Artificial Intelligence
{xiongkezheng, xiezhehan, linwenkai, zhoufangyu}@stu.xmu.edu.cn

Abstract

Point cloud registration, which aims to align two point clouds, is a fundamental task in computer vision and robotics. However, the presence of outliers (mismatched correspondences) in the point clouds makes the registration task challenging. Recently, a surge of interest has been observed in outlier rejection methods, formulating the outlier removal as an inlier/outlier classification problem. Despite their success, they are either memory-intensive, or time-consuming. Moreover, they fail to consider high-level learned representations by solely relying on pairwise relationships, which further restricts their performance. Hypergraph, a generalization of the graph, have proven effective in capturing high-order relationships among objects. In this work, we propose a novel outlier rejection method for 3D registration termed HyperDSC, a novel deep neural network that explicitly models high-order spatial compatibility relationships with hypergraphs. Firstly, we proposed Compatibility-Aware Hypergraph Convolution(CAHConv) to learn high-order relationships among correspondences with geometric priors embedded. Built upon CAHConv, we further introduce a novel Hierarchical Bi-directional Aggregation and Fusion Block(HBAF) to learn and propagate learned representations with clean-but-sparse 2nd-order compatibility hypergraphs and nosiy-but-dense 1st-order compatibility hypergraphs. With superior ability on capturing high-order relationships, our method achieves state-of-the-art performance. Extensive experiments on both outdoor and indoor datasets demonstrate the effectiveness of our method.

Introduction

Point cloud registration, which aims to align two point clouds, is a fundamental task in computer vision and robotics. With the advent of powerful point cloud descriptors(Qin et al. 2022) and deep learning techniques, the performance of existing efforts has advanced significantly. However, the presence of outliers (mismatched correspondences) introduced by feature-matching still makes the task challenging.

Recently, a surge of interest has been observed in outlier rejection methods, which formulate outlier removal as an inlier/outlier classification problem. Despite their success, they suffer from expensive computational costs. PointDSC(Bai et al. 2021) devises memory-intensive non-local blocks, which has a quadratic complexity w.r.t the number of correspondences. MAC(Zhang et al. 2023) rely on maximal clique search, which has been proven to be NP-hard and thus time-consuming, while FastMAC(Zhang et al. 2024) approximates maximal cliques with stochastic sampling, trading off accuracy for efficiency. Moreover, existing efforts fail to consider high-level representations, leveraging only purely-geometric heuristics and pairwise relationships. This partiality further restricts their performance.

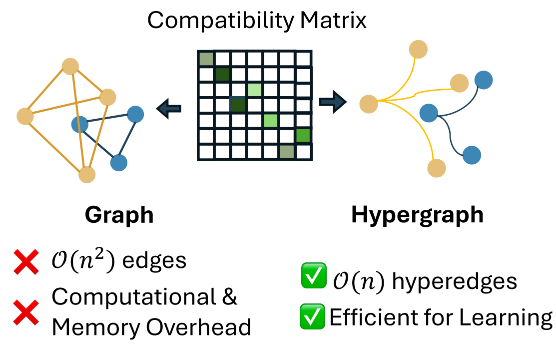

Figure 1: **Motivation.** Representing high-order consistency among correspondences with hypergraphs introducing fewer relationships, improving the efficiency of feature learning, compared to pairwise relationships of graphs.

As shown in Fig. 1, learning high-order consistency among correspondences is the key to improving the performance of outlier rejection(Chen et al. 2022b). Hypergraphs, which is a generalization of graphs, have proven effective in capturing high-order relationships among objects, whereas spatial compatibility metrics are efficient in capturing geometric priors. Based on these insights, we propose a novel outlier rejection method for 3D registration termed HyperDSC, a novel deep neural network that explicitly models high-order spatial compatibility relationships with hypergraphs.

Specifically, we represent spatial-compatibility relationships among correspondences with hypergraphs, avoiding

*Correspondence Author
†These authors contributed equally.

an overabundance of pairwise relationships. Hunter(Yao et al. 2023) directly apply HGNN(Gao et al. 2022) convolutions for correspondence hypergraphs. However, they neglect the geometric priors embedded in the correspondences, which are crucial for point cloud registration. Therefore, we proposed Compatibility-Aware Hypergraph Convolution(CAHConv) to learn high-order relationships among correspondences with geometric priors embedded. We emprically observe that hypergraphs constructed with spatial compatibility (SC) metrics are noisy but enables feature exchanges among abundant correspondences, while hypergraphs constructed with second-order spatial compatibility (SC^2) metrics are clean, but sometimes rejects potential inliers when too many outliers are present. To leverage the advantages of both, we further introduce a novel Hierarchical Bi-directional Aggregation and Fusion Block(HBAF) built upon CAHConvto learn and propagate learned representations with clean-but-sparse 2^{nd}-order compatibility hypergraphs and nosiy-but-dense 1^{st}-order compatibility hypergraphs. Finally, we train the model with node-wise contrastive and classification learning to build a robust feature space that facilitates outlier rejection.

Overall, our contributions are as follows:

- We propose a novel outlier rejection method for 3D registration termed HyperDSC, a novel deep neural network that explicitly models high-order spatial compatibility relationships with hypergraphs.
- CAHConv and HBAF are introduced for effective learning on hypergraphs. Node-wise contrastive and classification learning are introduced for robust feature learning.
- Extensive experiments on 3DMatch(Zeng et al. 2017) and KITTI Odometry demonstrate the effectiveness of our method.

Related Works

Learning-based Registration Methods. Learning-based registration methods fall into two categories: direct registration methods and correspondence-based methods. Direct registration methods (Fu et al. 2021; Wang and Solomon 2019; Xu et al. 2021; Aoki et al. 2019; Huang, Mei, and Zhang 2020) directly estimate the transformation between two point clouds in an end-to-end way, either with soft correspondences, or by regressing the transformation directly from a global feature vector. However, such methods could potentially fail in large-scale scenes. Correspondence-based methods (Choy, Park, and Koltun 2019; Deng, Birdal, and Ilic 2018) first extract correspondences between two point clouds, and then estimate the transformation with robust pose estimators. However, traditional robust estimators suffer from slow convergence and are sensitive to outliers. To address this, deep robust estimators (Choy, Dong, and Koltun 2020; Bai et al. 2021; Lee et al. 2021) utilize deep neural networks to reject outliers and compute the transformation. Although these methods require a training procedure, they improve accuracy and speed. Xiong et al. (2024b) is the first to introduce skeletal priors as geometric cues to facilitate feature learning. Based on a similar insight of combining low-level and high-level information, Xiong et al. (2024a) further proposed a reliable unsupervised registration method by combining both low-level geometric cues and high-level learned features.

Hypergraphs. The hypergraph is a generalization of graph structures. Zhou, Huang, and Scholkopf (2006) first proposed the concept of hypergraph learning in 2007, extending the spectral clustering algorithm of undirected graphs to hypergraph structures, and further introducing spectral hypergraph embedding and transductive inference on hypergraphs. In recent years, deep learning methods based on hypergraphs have been widely studied. HGNN (Yang et al. 2023) was the first to propose convolution operations on hypergraphs from the spectral domain to learn complex data associations. Bai, Zhang, and Torr (2021) introduced attention mechanisms to hypergraph convolutional networks to further enhance performance. HGNN+ (Gao et al. 2022) is an extension of HGNN, improving spectral domain convolution to spatial domain convolution for more stable performance.

Method

Priliminaries

Problem Formulation. Given two point clouds $\mathcal{P} = \{\mathbf{p}_i \in \mathbb{R}^3 | i = 1, \dots, N\}$ and $\mathcal{Q} = \{\mathbf{q}_i \in \mathbb{R}^3 | i = 1, \dots, M\}$, our goal is to align the two point clouds by estimating a rigid transformation $\mathbf{T} = \{\mathbf{R}, \mathbf{t}\}$, where $\mathbf{R} \in SO(3)$ is a 3D rotation matrix and $\mathbf{t} \in \mathbb{R}^3$ is a 3D translation vector. The transformation can be solved by:

$$\min_{\mathbf{R},\mathbf{t}} \sum\nolimits_{(\mathbf{p}_{x_i}, \mathbf{q}_{y_i}) \in \mathcal{C}^\star} \|\mathbf{R}\mathbf{p}_{x_i} + \mathbf{t} - \mathbf{q}_{y_i}\|^2, \tag{1}$$

where $\mathcal{C}^\star$ denotes the set of correspondences between two point clouds $\mathcal{P}$ and $\mathcal{Q}$. In reality, $\mathcal{C}^\star$ is usually unknown. Hence, we need to establish accurate correspondences $\mathcal{C}$ between two point clouds for a good transformation. Typically, we retreive a noisy set of correspondences $\mathcal{C}^{'}$, and then remove outliers to obtain its clean subset $\mathcal{C} \subset \mathcal{C}^{'}$.

Hypergraph. A hypergraph is defined as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $\mathcal{V}$ represents the set of vertices, $\mathcal{E}$ denotes the set of hyperedges, and $\mathbf{W}$ assigns weights to these hyperedges. In analogy to simple graphs, the structure of a hypergraph is often captured by its incidence matrix $\mathbf{H} \in 0, 1^{|\mathcal{V}| \times |\mathcal{E}|}$, where $\mathbf{H}(v_i, e_j)$ equals 1 if the vertex v_i is connected to the hyperedge e_j, and 0 otherwise. The degree of a hyperedge e_j and a vertex v_i are then given by $D(e_j) = \sum_{i=1}^{|V|} \mathbf{H}(v_i, e_j)$ and $D(v_i) = \sum_{j=1}^{|\mathcal{E}|} \mathbf{W}(e_j)\mathbf{H}(v_i, e_j)$, respectively. Unlike edges in simple graphs, which only connect two vertices at a time, hyperedges can link more than two vertices, allowing them to compactly represent a hypothesis by connecting the vertices that correspond to data points in a sampled minimal subset. When the vertices represent data points, hyperedges provide an elegant way to capture correspondences. Therefore, different from existing efforts using graphs(Zhang et al. 2023, 2024) to represent spatial compatibility, we resort to hypergraphs, thereby capturing high-order relationships and avoiding the introduction of too many pairwise relationships.

Pipeline

The overall pipeline of HyperDSC is illustrated in Fig. 2. Given input correspondences, we first construct hyperedges by leveraging both first and second-order spatial compatibility metrics. Then, the proposed novel module, HBAF, is interleaved N times to learn high-order relationships among correspondences with geometric priors embedded. The HBAF module aggregates the input features with two stand-alone CAHConv blocks, one for 2^{nd}-order compatibility hypergraphs and the other for 1^{st}-order compatibility hypergraphs. Then, we design a selective fusion mechanism to fuse the learned features from the two CAHConv blocks, leveraging the advantages of both types of hypergraphs. Finally, we employ an MLP to predict the initial correspondence confidence, and then post-process the correspondences to obtain the final transformation. The network is end-to-end trained with the proposed node-wise contrastive and classification learning.

Hypergraph Construction

We construct hyperedges by leveraging length consistency. Given a correspondence set $\mathcal{C} = \{(x_i, y_j)\}_N$, we first compute the BLE matrix. For each correspondence C_i, we identify the correspondences that exhibit stronger spatial consistency with $\mathcal{C}$ and treat them as its partners. The set of these partners is defined as $P(\mathcal{C}_i) = \{C_j \mid d_{ij} < \theta\}$. From this partner set, we randomly select k data points to form a hyperedge $\mathcal{E}(\mathcal{C}_i)$. This process is repeated for each correspondence, resulting in a hypergraph where the vertices are $\mathcal{V} = \mathcal{C}$ and the hyperedges are $\mathcal{E} = \{\mathcal{E}(C_i) \mid i = 1, 2, 3, \ldots, N\}$. We assign the initial correspondence feature ${}^{(i,j)}\mathbf{F}_{\mathcal{C}} = \mathrm{Cat}\left(x_i, y_j, \mathbf{F}_i, \mathbf{F}_j\right)$ to a correspondence ${}^{(i,j)}\mathcal{C} = (x_i, y_j)$.

Hypergraph Feature Embedding

Compatibility-Aware Hypergraph Convolution. Existing effort, such as Hunter(Yao et al. 2023), directly apply HGNN(Gao et al. 2022) for learning on correspondence hypergraphs. However, they neglect the geometric priors embedded in the correspondences, which are crucial for learning high-order relationships. To address this, we propose a novel Compatibility-Aware Hypergraph Convolution(CAHConv) to learn high-order relationships among correspondences with geometric priors embedded.

Given a compatibility matrix $\mathcal{C}_{ij} = 1 - (d_{ij}/\sigma_d)^2$ where $d = |\|\mathbf{p}_i - \mathbf{q}_i\|_2 - \|\mathbf{p}_j - \mathbf{q}_j\|_2|$ is the length consistency between two correspondences and σ_d is a hyperparameter, We compute hyperedge-wise geometric embedding $\mathbf{E}_k$ for hyperedge $\mathcal{E}_k$ by aggregating the compatibility of its vertices:

$$\mathbf{E}_k = \sum_{i=1}^{|\mathcal{V}|} \sum_{j=i}^{|\mathcal{V}|} \frac{\mathbf{H}_{ij} \cdot \mathbf{H}_{jk} \cdot \mathcal{C}_{ij}}{D(\mathcal{E}_k)} \tag{2}$$

Then, the input feature is updated as $\mathbf{F}_i = \sum_{j=1}^{|\mathcal{V}|} \alpha_{ij} \mathbf{F}_j \mathbf{W}$, where $\mathbf{W}$ is a learnable weight matrix, and α_{ij} is computed as:

$$\alpha_{ij} = \frac{1}{\sqrt{D(\mathcal{V}_i) D(\mathcal{V}_j)}} \sum_{k=1}^{|\mathcal{E}|} \frac{\mathbf{H}_{ik} \mathbf{H}_{jk} \mathbf{E}_k}{D(\mathcal{E}_k)} \tag{3}$$

Hierarchical Bi-directional Aggregation and Fusion Block. Built-upon CAHConv, we further introduce a novel HBAF to learn and propagate learned representations with both 2^{nd}-order compatibility hypergraphs and 1^{st}-order compatibility hypergraphs.

To aggregates features from both types of hypergraphs, it is crucial to decide how features from different hypergraphs are fused. Drawing inspiration from SKNet(Li et al. 2019), we propose a selective fusion mechanism, as is depicted in Fig. 2. Specifically, we first compute the aggregated features $\mathbf{F}_1$ and $\mathbf{F}_2$ from both hypergraphs with the proposed CAHConv. Then, the attention weights $\mathcal{A} \in \mathbb{R}^2$ for respective aggregated features are computed by applying a channel-wise MLP $\mathcal{F}_A$ to their summation followed by a softmax operation.

$$\mathcal{A} = \mathrm{softmax}(\mathcal{F}_A(\mathbf{F}_1 + \mathbf{F}_2)) \tag{4}$$

Finally, we aggregate the resultant features as the weighted summation of the input:

$$\mathbf{F} = \mathcal{A}_1 \mathbf{F}_1 + \mathcal{A}_2 \mathbf{F}_2 \tag{5}$$

The proposed HBAF is interleaved N times to form the hypergraph feature embedding module. After that, an MLP is employed to predict the initial correspondence confidence $\{v_i\}$ using the learned features.

Post-Processing

Based on the learned features and predicted confidences, we employed Seed Selection, Neural Spectral Matching and Hypothesis Selection from PointDSC(Bai et al. 2021) to obtain the final correspondences. Then, we estimate the transformation $\mathbf{T}$ using the final correspondences based on weighted SVD(Choy, Dong, and Koltun 2020).

Loss

Node-wise Contrastive Learning. Contrastive learning is widely used (Qin et al. 2022; Huang et al. 2021) to train registration models. However, existing learned pose estimators(Bai et al. 2021) mainly formulate outlier rejection as a classification problem and adopt a cross-entropy loss. Without direct feature-space supervision, the learned representations may not be optimal for outlier rejection, and the training may be unstable. To address this, we opt to facilitate the learning of robust feature representations in a metric learning manner. We thus propose a novel node-wise contrastive learning strategy. To the best of our knowledge, we are the first to introduce node-wise contrastive learning for outlier rejection in 3D registration.

Specifically, we encourage the learned representations of inliers to be close to each other and far from outliers, and vice versa. For input correspondences $\mathcal{C} = \{\mathbf{c}_i\}$ with predicted labels $\mathbf{b}$ and associated features $\mathbf{F}$, we define the pairwise cluster label matrix $\mathbf{M} = \{M_{ij}\}$ as:

$$M_{ij} = \begin{cases} 1, & \text{if } b_i = b_j, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Then, we adopt the Circle Loss(Sun et al. 2020) to learn the

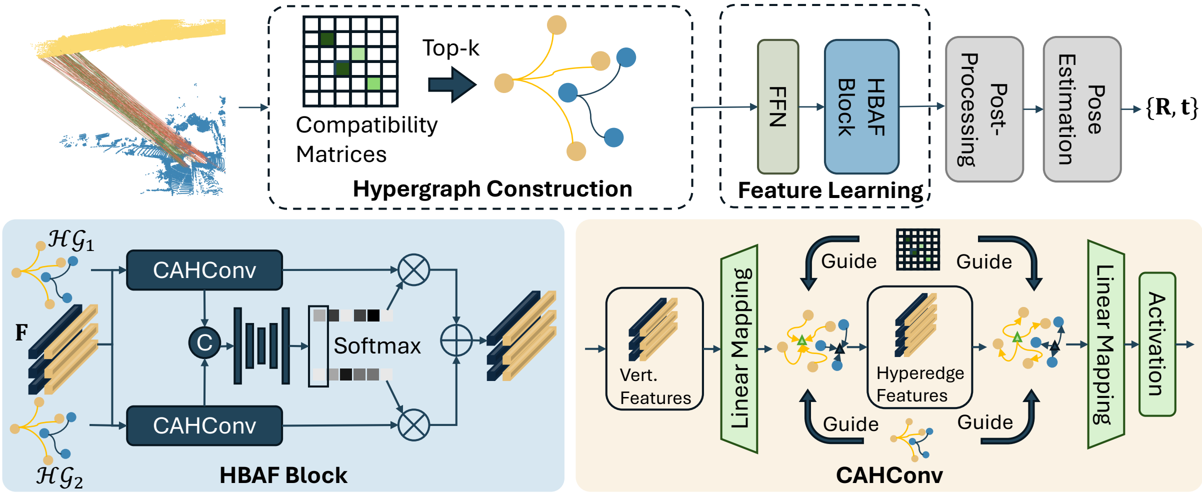

Figure 2: **The Overall Pipeline of HyperDSC.** We first construct 1st- and 2nd-order hyperedges. Then, the proposed HBAF-module learns high-order relationships among correspondences with geometric priors embedded, which aggregates the features with two stand-alone CAHConv for each hypergraph and fuse the features with a selective fusion mechanism. Finally, we post-process the correspondences using learned features and estimate the resultant transformation.

feature representations:

$$\mathcal{L}_c = \frac{1}{|\mathcal{C}|}\sum_{\mathcal{C}_i \in \mathcal{C}} \log[1 + \sum_{M_{ij}=1} e^{\beta_p^{i,j}(d_i^j - \Delta_p)} \cdot \sum_{M_{ik}=0} e^{\beta_n^{i,k}(\Delta_n - d_i^k)}] \tag{7}$$

where $d_i^j = \|\mathbf{F}_i - \mathbf{F}_j\|_2$ is the distance in the feature space, and weights $\beta_p^{i,j} = \gamma(d_i^j - \Delta_p)$ and $\beta_n^{i,k} = \gamma(\delta_n - d_i^k)$ are used to highlight the patches with high overlap ratio. We empirically set hyperparameters $\Delta_p = 0.1$ and $\Delta_n = 1.4$. Circle Loss provides a more effective mechanism for selecting and prioritizing informative pairs, improving the generalization of learned embeddings.

Node-wise Classification Learning. To facilitate the classification prediction of inlier/outlier labels, we also train the model with a node-wise classification loss. Cross-Entropy Loss, despite its simplicity and wide adoption(Bai et al. 2021), may not be optimal for outlier rejection. In inlier/outlier classification, a large portion of correspondences are easy to classify, while the remaining are hard to distinguish. Due to the imbalanced nature of the task in regard of hard/easy samples, the model may be biased towards easy samples, thereby hindering effective training.

Drawing insight existing efforts, we resort to Focal Loss(Ross and Dollár 2017) to address the issue, which highlights hard samples in a self-adaptive manner:

$$\mathcal{L}_f = -\frac{1}{|\mathcal{C}|}\sum_{\mathcal{C}_i \in \mathcal{C}} \alpha_t (1 - p_t)^\gamma \log(p_t), , \tag{8}$$

where α and γ are hyperparameters, and p_t is the model's predicted probability for the correct class. Focal loss dynamically scales down the loss for well-classified examples, focusing on hard examples that the model struggles with.

Loss Aggregation. The final loss is a weighted sum of the node-wise contrastive loss and the node-wise classification loss:

$$\mathcal{L} = \lambda \mathcal{L}_c + (1 - \lambda)\mathcal{L}_f, \tag{9}$$

where λ is a hyperparameter that balances the two losses.

Experiments

Experiment Settings

Datasets We mainly evaluate the proposed method on two representative and challenging public datasets, namely 3DMatch and KITTI Odometry. Both datasets adhere to official splits for training and testing. The evaluation protocol follows the standard settings of PointDSC(Bai et al. 2021). Following Predator(Huang et al. 2021), we also evaluate our method on the 3DLoMatch benchmark dataset, which is a low-overlap challenging subset of the 3DMatch dataset.

Baselines For traditional methods, we mainly compare our method with FGR(Zhou, Park, and Koltun 2016), SM(Leordeanu and Hebert 2005), RANSAC(Fischler and Bolles 1981), TEASER++(Yang, Shi, and Carlone 2020), SC2-PCR(Chen et al. 2022b). For learning-based methods, we compare with DGR(Choy, Dong, and Koltun 2020), PointDSC(Bai et al. 2021), VBReg(Jiang et al. 2023), and DHVR(Lee et al. 2021). To retreive the initial input correspondences, we use FCGF(Zeng et al. 2017) for 3DMatch descriptors and FPFH(Rusu, Blodow, and Beetz 2009) descriptors for KITTI Odometry.

Implementation Details We implement our method with PyTorch and train the model on a single NVIDIA RTX 3090 GPU. We use the Adam optimizer with an initial learning rate of $1e-4$ and a batch size of 16. The model is trained for 100 epochs with a learning rate decay of 0.1.

Table 1: **Comparison with State-of-the-Arts.** Quantitative results on the KITTI Odometry and 3DMatch datasets.

Method	KITTI(FPFH)			3DMatch(FCGF)		
	RR%	RRE(○)	RTE(m)	RR%	RRE(○)	RTE (m)
FGR	1.26	1.69	47.18	79.17	2.93	8.56
SM	75.50	0.66	15.01	86.57	2.29	7.07
RANSAC	89.37	1.22	25.88	91.50	2.49	7.54
TEASER++	64.14	1.04	14.85	85.77	2.73	8.66
DGR	73.69	1.67	34.74	91.30	2.40	7.48
DHVR	–	–	–	89.40	2.19	6.95
SC2-PCR	97.84	0.39	9.09	93.10	2.04	6.53
PointDSC	98.20	0.58	7.27	92.42	2.05	6.49
VBReg	98.92	0.32	7.17	93.53	2.04	6.49
HyperDSC	**99.46**	**0.32**	**7.17**	**93.71**	**2.04**	**6.49**

Table 2: RR% with different numbers of correspondences on 3DLoMatch benchmark dataset

Model	5000	1000	250
FGR	18.6	16.9	12.4
SM	32.4	31.4	23.5
RANSAC	37.6	35.9	25.9
TEASER++	42.8	39.5	25.7
DHVR	50.4	46.4	34.6
SC2_PCR	57.4	51.8	36.2
TR_DE	49.5	48.4	34.3
PointDSC	55.8	46.8	26.7
VBReg	58.3	52.9	34.5
HyperDSC	**58.7**	**54.86**	**36.2**

Comparison With State-of-the-Art Methods

Evaluation on KITTI Odometry and 3DMatch. We first compare our method with state-of-the-arts on two representative datasets, KITTI Odometry and 3DMatch. Following existing efforts(Fu et al. 2021; Bai et al. 2021), for KITTI Odometry, we use FPFH descriptors, while for 3DMatch, we use FCGF descriptors. As shown in Table 1, our method achieves state-of-the-art performance on both datasets, demonstrating superior effectiveness in both indoor and outdoor scenarios.

Evaluation on 3DLoMatch 3DLoMatch is a low-overlap (10% 30%) challenging subset of the 3DMatch dataset. We additionally compare with TR_DE(Chen et al. 2022a), which is a traditional method. The registration recall (RR) with different numbers of correspondences is shown in Table 2. Our method consistently outperforms existing efforts, demonstrating the effectiveness and robustness of our method in challenging scenarios.

Analysis

Ablation Study of Modules. We first validate the effectiveness of proposed CAHConv and HBAF. We take SC-NonLocal from PointDSC(Bai et al. 2021) and VB-NonLocal from VBReg(Jiang et al. 2023) as baselines to highlight the effectiveness of our method. Our method consistently outperforms the baselines on KITTI dataset, as shown in Table 3. Using HBAF outperforms other fusion strategies such as concatenation and summation, demonstrating the superiority of our selective mechanism.

Ablation Study of Losses. We also validate the loss choices by comparing with alternative configurations using standard cross-entropy loss and Spectral Matching Loss(Bai et al. 2021). We report RR% on 3DLoMatch benchmark in Table 4. With the proposed node-wise contrastive and classification learning, our method learns robust and discriminative feature representations, outperforming the alternatives. We find that the Spectral Matching Loss can be regarded as a special (and weakened) case of contrastive loss, which will be further investigated in future work.

Table 3: **Ablation Study of Modules.** HBAF(Sum) and HBAF(Cat) denote the alternative configurations to replace the selective fusion strategies of HBAF with **Sum**mation and Con**cat**enation.

Configuration	RR%	F1	Time (s)
SC-NonLocal	98.20	92.71	6.20
VB-NonLocal	98.92	92.69	8.20
HBAF(Sum)	98.75	92.21	3.48
HBAF(Cat)	99.10	93.26	3.48
HBAF	**99.46**	**93.43**	3.84

Table 4: **Ablation Study of Losses.** CE denotes cross-entropy loss, FO denotes focal loss, SPM denotes spectral matching loss, and CL denotes circle loss.

CE	FO	SPM	CL	RR%
✓		✓		54.74
	✓	✓		54.80
✓			✓	54.80
	✓		✓	**54.86**

Conclusion

In this work, we propose a novel outlier rejection method for 3D registration termed HyperDSCto explicitly model high-order spatial compatibility with hypergraphs. We proposed Compatibility-Aware Hypergraph Convolution(CAHConv) to learn high-order relationships among correspondences with geometric priors embedded. Built upon that, we introduce a novel Hierarchical Bi-directional Aggregation and Fusion Block(HBAF) to learn and propagate learned representations with different compatibility hypergraphs. Extensive experiments on both outdoor and indoor datasets demonstrate the effectiveness of our method.

References

Aoki, Y.; Goforth, H.; Srivatsan, R. A.; and Lucey, S. 2019. Pointnetlk: Robust & efficient point cloud registration using pointnet. In *Proceedings of the CVPR*, 7163–7172.

Bai, S.; Zhang, F.; and Torr, P. H. 2021. Hypergraph convolution and hypergraph attention. *Pattern Recognition*, 110: 107637.

Bai, X.; Luo, Z.; Zhou, L.; Chen, H.; Li, L.; Hu, Z.; Fu, H.; and Tai, C.-L. 2021. Pointdsc: Robust point cloud registration using deep spatial consistency. In *Proceedings of the CVPR*, 15859–15869.

Chen, W.; Li, H.; Nie, Q.; and Liu, Y.-H. 2022a. Deterministic point cloud registration via novel transformation decomposition. In *Proceedings of the CVPR*, 6348–6356.

Chen, Z.; Sun, K.; Yang, F.; and Tao, W. 2022b. Sc2-pcr: A second order spatial compatibility for efficient and robust point cloud registration. In *Proceedings of the CVPR*, 13221–13231.

Choy, C.; Dong, W.; and Koltun, V. 2020. Deep global registration. In *Proceedings of the CVPR*, 2514–2523.

Choy, C.; Park, J.; and Koltun, V. 2019. Fully convolutional geometric features. In *Proceedings of the IEEE/CVF ICCV*, 8958–8966.

Deng, H.; Birdal, T.; and Ilic, S. 2018. Ppf-foldnet: Unsupervised learning of rotation invariant 3d local descriptors. In *Proceedings of the ECCV*, 602–618.

Fischler, M. A.; and Bolles, R. C. 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6): 381–395.

Fu, K.; Liu, S.; Luo, X.; and Wang, M. 2021. Robust point cloud registration framework based on deep graph matching. In *Proceedings of the CVPR*, 8893–8902.

Gao, Y.; Feng, Y.; Ji, S.; and Ji, R. 2022. HGNN+: General hypergraph neural networks. *TPAMI*, 45(3): 3181–3199.

Huang, S.; Gojcic, Z.; Usvyatsov, M.; Wieser, A.; and Schindler, K. 2021. Predator: Registration of 3d point clouds with low overlap. In *Proceedings of the CVPR*, 4267–4276.

Huang, X.; Mei, G.; and Zhang, J. 2020. Feature-metric registration: A fast semi-supervised approach for robust point cloud registration without correspondences. In *Proceedings of the CVPR*, 11366–11374.

Jiang, H.; Dang, Z.; Wei, Z.; Xie, J.; Yang, J.; and Salzmann, M. 2023. Robust outlier rejection for 3d registration with variational bayes. In *Proceedings of the CVPR*, 1148–1157.

Lee, J.; Kim, S.; Cho, M.; and Park, J. 2021. Deep hough voting for robust global registration. In *Proceedings of the IEEE/CVF ICCV*, 15994–16003.

Leordeanu, M.; and Hebert, M. 2005. A spectral technique for correspondence problems using pairwise constraints. In *Tenth IEEE ICCV (ICCV'05) Volume 1*, volume 2, 1482–1489. IEEE.

Li, X.; Wang, W.; Hu, X.; and Yang, J. 2019. Selective kernel networks. In *Proceedings of the CVPR*, 510–519.

Qin, Z.; Yu, H.; Wang, C.; Guo, Y.; Peng, Y.; and Xu, K. 2022. Geometric transformer for fast and robust point cloud registration. In *Proceedings of the CVPR*, 11143–11152.

Ross, T.-Y.; and Dollár, G. 2017. Focal loss for dense object detection. In *proceedings of the CVPR*, 2980–2988.

Rusu, R. B.; Blodow, N.; and Beetz, M. 2009. Fast point feature histograms (FPFH) for 3D registration. In *2009 ICRA*, 3212–3217. IEEE.

Sun, Y.; Cheng, C.; Zhang, Y.; Zhang, C.; Zheng, L.; Wang, Z.; and Wei, Y. 2020. Circle loss: A unified perspective of pair similarity optimization. In *Proceedings of the CVPR*, 6398–6407.

Wang, Y.; and Solomon, J. M. 2019. Deep closest point: Learning representations for point cloud registration. In *Proceedings of the IEEE/CVF ICCV*, 3523–3532.

Xiong, K.; Xiang, H.; Xu, Q.; Wen, C.; Shen, S.; and Wang, C. 2024a. Mining and Transferring Feature-Geometry Coherence for Unsupervised Point Cloud Registration. *NeurIPS*.

Xiong, K.; Zheng, M.; Xu, Q.; Wen, C.; Shen, S.; and Wang, C. 2024b. SPEAL: Skeletal Prior Embedded Attention Learning for Cross-Source Point Cloud Registration. In *Proceedings of the AAAI*, volume 38, 6279–6287.

Xu, H.; Liu, S.; Wang, G.; Liu, G.; and Zeng, B. 2021. Omnet: Learning overlapping mask for partial-to-partial point cloud registration. In *Proceedings of the IEEE/CVF ICCV*, 3132–3141.

Yang, H.; Shi, J.; and Carlone, L. 2020. Teaser: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2): 314–333.

Yang, Y.; Wu, Z.; Wu, L.; Zhang, K.; Hong, R.; Zhang, Z.; Zhou, J.; and Wang, M. 2023. Generative-Contrastive Graph Learning for Recommendation. *Proceedings of the 46th SIGIR*.

Yao, R.; Du, S.; Cui, W.; Ye, A.; Wen, F.; Zhang, H.; Tian, Z.; and Gao, Y. 2023. Hunter: Exploring high-order consistency for point cloud registration with severe outliers. *TPAMI*.

Zeng, A.; Song, S.; Nießner, M.; Fisher, M.; Xiao, J.; and Funkhouser, T. 2017. 3dmatch: Learning local geometric descriptors from rgb-d reconstructions. In *Proceedings of the CVPR*, 1802–1811.

Zhang, X.; Yang, J.; Zhang, S.; and Zhang, Y. 2023. 3D registration with maximal cliques. In *Proceedings of the CVPR*, 17745–17754.

Zhang, Y.; Zhao, H.; Li, H.; and Chen, S. 2024. FastMAC: Stochastic Spectral Sampling of Correspondence Graph. *arXiv preprint arXiv:2403.08770*.

Zhou, D.; Huang, J.; and Scholkopf, B. 2006. Learning with Hypergraphs: Clustering, Classification, and Embedding. In *NeurIPS*.

Zhou, Q.-Y.; Park, J.; and Koltun, V. 2016. Fast global registration. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, 766–782. Springer.