Pattern Recognition 136 (2023) 109267

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Hybrid optimization with unconstrained variables on partial point cloud registration

Yuanjie Yan[a,b], Junyi An[a,b], Jian Zhao[c,*], Furao Shen[a,d,**]

[a] State Key Laboratory for Novel Software Technology, Nanjing University, China
[b] Department of Computer Science and Technology, Nanjing University, China
[c] School of Electronic Science and Engineering, Nanjing University, China
[d] School of Artificial Intelligence, Nanjing University, Nanjing, China

ARTICLE INFO

Article history:
Received 3 June 2022
Revised 2 December 2022
Accepted 14 December 2022
Available online 16 December 2022

Keywords:
Point cloud registration
Optimization

ABSTRACT

3D point cloud registration is a fundamental problem in computer vision (CV) and computer graphics (CG). Recently, a series of learning-based algorithms have been proposed to show the advantages in registration accuracy and inference speed. However, those learning-based methods usually ignore transformations with constrained rotations and translations in registration. In this paper, we propose a novel hybrid optimization method to solve the constrained rotational and translational transformations. A mapping function is introduced to deal with the restrained variables in optimization. Our method achieves superior performance on the Multi-View Partial Point dataset, which won the first place on the registration challenge in ICCV 2021. The method is also validated on the synthetic datasets ModelNet, ICL-NUIM, and the realistic 3DMatch dataset. We demonstrate that the global optimization methods still have great potential research for point cloud registration. The code is available at https://github.com/Dizzy-cell/HOUV.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Point clouds play an important role in 3D applications, which are received by radars and depth cameras. Due to occlusion and other issues, only partial point clouds are acquired for objects in the realistic scene. To obtain the complete point clouds, it is crucial *t* integrate partial point clouds related to the same object. Point cloud registration aims to find a rigid transformation to align partial point clouds. It is widely leveraged in map reconstruction, 3D modelling [1] and autonomous driving [2] in reconstruction applications. Besides, as a key technique in SLAM [3], point cloud registration can match the current partial point cloud with the global map to locate the agent in the environment. Therefore, point cloud registration is a fundamental task in computer vision and robotics.

Point cloud registration is usually solved by minimizing the aligned geometric errors in the literature of optimization [4]. Traditional registration method aims to generate the rotation matrix and translation vector through optimization, like Iterative Closest Point (ICP) [5]. As ICP is limited to the rigid transformation with small rotations, FGR [6] extends it to the global registration with large rotations and translations. GO-ICP [7] adopts the branch-and-bound (BnB) scheme to perform the ICP on the entire motion space $SE(3)$. Deng et al. [8] introduce a novel optimization function related to the intersection points between pairwise partial point clouds and a random straight line. However, the speed and performance of these methods are unsatisfactory for the global registration. Recently, with the development of 3D deep learning techniques, more and more learning-based methods are proposed to challenge the optimization methods. Learning-based registration methods are typically much faster than iterative global registration methods. DCP [9] achieves the end-to-end point cloud registration, but it only shows good performance on the ModelNet [10]. DeepGMR [11] matches the points to hybrid probability distributions whose parameters are estimated by a neural network and achieves distribution-to-distribution correspondences on ICL-NUIM [12]. Deep Global Registration (DGR) [13] proposes a differentiable framework for pairwise registration, including the feature correspondence prediction, pose estimation and pose refinement. Horache et al. [14] improve the generalization of learning-based methods on unknown datasets by transfer learning.

The performances of those registration methods are validated on synthetic and realistic datasets. However, some synthetic datasets like ModelNet [10] are too simplistic to compare the registration methods. For example, DeepGMR and RGM methods

* Corresponding author at: School of Electronic Science and Engineering, Nanjing University, Nanjing, China.
** Corresponding author at: School of Artificial Intelligence, Nanjing University, Nanjing, China.
E-mail addresses: jianzhao@nju.edu.cn (J. Zhao), frshen@nju.edu.cn (F. Shen).

https://doi.org/10.1016/j.patcog.2022.109267
0031-3203/© 2022 Elsevier Ltd. All rights reserved.

achieve almost 100% registration recall on ModelNet. To foster the research of point cloud registration, the Multi-View Partial (MVP) point cloud dataset is proposed in the ICCV 2021 Workshop: Sensing, Understanding and Synthesizing Humans.[1] Different from the previous datasets, the MVP dataset aims to align partial-to-partial point clouds from different viewpoints. MVP needs to deal with three issues. (1) global registration, i.e., finding the rotation matrix over the entire SO(3) space. (2) uncompleted matching, i.e., solving with the outliers between a pair of partial-to-partial point clouds. (3) symmetric structure, i.e., point cloud objects have symmetrical and similar structures. Therefore, MVP dataset is a great challenge for previously proposed registration methods.

In this paper, we propose a novel Hybrid Optimization method with Unconstrained Variables for 3D registration, abbreviated as HOUV. HOUV directly defines a smooth transformation matrix to align a pair of partial point clouds through optimization. Considering the point matching in registration, the chamfer distance (CD) [15] is usually utilized to measure the distance for a pair of partial point clouds. Based on CD, we propose two variant losses. (1) **Local CD loss** solves the incomplete matching in pairwise partial point clouds. (2) **Projected CD loss** improves the transformation accuracy with global registration.

Then, to ensure the overlapping regions of pairwise partial point clouds, most registration tasks restrict the rotation angle and the translation distance. Exploiting the boundaries for fine-grained optimization can improve the performance of registration. However, learning-based methods pay little attention to this issue which is addressed by implicit learning from the training data. Optimization methods usually explicitly restrict these variables by truncation functions. In HOUV, we adopt the mapping function to replace the constrained variables with unconstrained ones for smooth optimization. Inspired by the branch-and-bound method [16], we divide the SE(3) space into several separate subspaces. Based on the range of transformation, HOUV handles subspaces by introducing boundaries. Furthermore, due to the symmetries or local similarities of objects, both the optimization and the learning-based methods are prone to get stuck in local minima. We propose a batch strategy by randomly initializing multiple groups of variables and a programmable optimization strategy to alleviate the local minimum problem.

To the best of our knowledge, HOUV is a novel method to directly optimize unconstrained variables in registration, surpassing the advanced optimization-based and learning-based methods. The advantages of HOUV includes (1) *Unsupervised*, i.e., our method can be directly applied to the registration of any point cloud without learning on the training dataset. (2) *Fine-grained programmability*, i.e., the optimization strategies are programmable to fit the sample or dataset. We can control the hyperparameters to align different degrees of overlap on pairwise point clouds. Our work demonstrates the advantages of optimization methods with fine-grained constraints. HOUV improves registration accuracy even without the aid of a deep learning model. Our contributions are summarized as below:

- We propose a novel hybrid optimization method with unconstrained variables on pairwise partial point cloud registration.
- To address the local and global point matching, we propose the local CD loss and projected CD loss based on Chamfer Distance. The proposed optimization strategies are programmable and flexible to fit the sample or dataset.
- We verify the performances of the proposed method on the synthetic and realistic datasets. Our method demonstrates great flexibility and generalization on fine-grained optimization.

[1] https://sense-human.github.io

2. Related work

2.1. Optimization-based registration methods

A large number of registration methods are associated with optimized methods, most of which align the partial cloud points by the correspondence searching and transformation estimation. Iterative Closest Points (ICP) [5] is a representative and classical method. ICP iteratively estimates the correspondences and computes the transformation via SVD. ICP variants [17,18] consider more complicated issues like noise, partiality, and sparsity to improve the robustness. However, most optimization-based methods are prone to find local minimum solutions, since the objective function is non-convex. To avoid the local minimum, Go-ICP [7] presents a branch-and-bound (BnB) method by searching the 3D motion space $SE(3)$. Without the acceleration of GPUs, Go-ICP is several orders of magnitude slower than other ICP variants. Since part-to-part point clouds are matched only at the corresponding points, Trimmed ICP (TrIPP) [19] processes the top-k points instead of all. Another branch of the optimization-based method involves probabilistic models, which is often related to GMMs and the EM algorithm. GMMReg [20] proposes a unified framework in the considering of noises and outliers, where point clouds are represented by Gaussian mixture models. Ma et al. [21] adopt a variational mixture model to align the partial point clouds by variational inference.

The performances of optimization-based methods depend on the matching accuracy between pairwise partial point clouds. However, those methods are sensitive to similar structures in partial-to-partial point cloud registration. As the symmetry structure of objects in partial point clouds, the local correspondences may be incorrectly matched. Furthermore, due to the difference degrees of overlap, the corresponding points are still a great challenge for optimization methods in global registration.

2.2. Learning-based registration methods

With the development of deep learning technologies on the point cloud, learning the latent representations of point clouds is leveraged for registration. Feature learning methods learn robust features, which are matched via optimization registration methods. The end-to-end learning methods directly estimate the transformation using the neural network models to align pair-wise partial point clouds.

The feature-learning registration method aims to combine the traditional registration method and feature extraction. The point cloud network extracts the corresponding features, and the optimization method is responsible for aligning those features. FCGF [22] adopts the 3D fully-convolutional network to capture the geometric features. Based on the feature descriptors by FCGF, Deep Global Registration (DGR) [13] proposes a differentiable Weighted Procrustes algorithm for closed-form pose estimation. Furthermore, Iterative Distance-Aware similarity Matrix convolution network (IDAM) [23] proposes a novel learnable pipeline by iteratively performing feature extraction and registration. Zhang et al. [24] propose a self-supervised equivalent transformation through Siamese network.

The end-to-end learning method estimates the transformation to solve the partial point cloud registration by the neural network. Inspired by ICP, DCP [9] constructs a transformation network consisting of a point cloud embedding, an attention module and a differentiable SVD layer. Similar to GMMReg, Deep Gaussian Mixture Registration (DeepGMR) [11] designs a neural network that extracts pose-invariant correspondences from raw point clouds with a differentiable Gaussian mixture model. Huang et al.[25] present a fast feature-metric registration framework, which minimizes feature-

metric projection errors instead of correspondences. RGM [26] develops a deep graph matching to calculate a soft correspondence matrix. Graph matching not only establishes the correspondence of the local geometric features for each point, but also considers the structure and topology of the global registration.

3. Hybrid optimization with unconstrained variables

First, we formally define the partial point clouds registration problem. Then, the hybrid optimization with unconstrained variables method (HOUV) is introduced to construct the rotation matrix and translation vector individually. Our proposed hybrid loss includes two variants of Chamfer Distance to deal with the partial and global point matching respectively. To avoid instability of optimization caused by random initialization, we employ a batching strategy to improve the stability of HOUV. Considering of the transformation subspace in registration, HOUV narrows the search space by introducing boundaries. We adjust the boundaries about registration by leveraging the prior knowledge of point clouds.

3.1. Formalization of the registration

In the 3D coordinate system, we have a rigid scene and a moving observation point. The point clouds $P = \{\mathbf{p}_i \in \mathbb{R}^3 \mid i = 1, \ldots, N\}$ and $Q = \{\mathbf{q}_j \in \mathbb{R}^3 \mid j = 1, \ldots, M\}$ are obtained from the viewpoints $\mathbf{v}_p$ and $\mathbf{v}_q$, respectively. N and M represent the number of points in each point cloud. Assume $\mathbf{v}_q$ is transformed from $\mathbf{v}_p$ by an unknown rigid rotation matrix $\mathbf{R}_{view}$ and translation vector $\mathbf{t}_{view}$, i.e.,

$$\mathbf{v}_q = \mathbf{R}_{view} \cdot \mathbf{v}_p + \mathbf{t}_{view}. \tag{1}$$

If $\{\mathbf{R}_{view}, \mathbf{t}_{view}\}$ are known, we can align partial point clouds of the same scene to obtain the fused point clouds. When $\mathbf{v}_p$ and $\mathbf{v}_q$ are unknown, we can directly align pairwise partial point clouds by $\{\mathbf{R}_{pq}, \mathbf{t}_{pq}\}$.

The registration aims to estimate a rigid transformation $\{\mathbf{R}_{pq}, \mathbf{t}_{pq}\}$ to align a source point cloud P and a target point cloud Q, where $\mathbf{R}_{pq} \in SO(3), \mathbf{t}_{pq} \in \mathbb{R}^3$. When pairwise partial point clouds have overlapping regions from viewpoints $\mathbf{v}_p$ and $\mathbf{v}_q$, P and Q have local matching points with noises. The transformation $\{\mathbf{R}_{pq}, \mathbf{t}_{pq}\}$ can be estimated from the matching points by

$$(\mathbf{R}_{pq}, \mathbf{t}_{pq}) = \arg\min_{\mathbf{R},\mathbf{t}} \sum_{(\mathbf{p}_k, \mathbf{q}_k) \in CP} (||\mathbf{R} \cdot \mathbf{p}_k + \mathbf{t} - \mathbf{q}_k||^2), \tag{2}$$

where CP is made up of the subsets from P and Q. The corresponding points are matched one by one in the CP set. The corresponding points represent the overlap region in pairwise particle point clouds. It is not a trivial issue to find suitable subsets with correctly corresponding points.

3.2. Transformation matrix with unconstrained variables

The transformation consists of a rotation matrix and a translation vector. Following the Rodrigues' formula [27], any rotation matrix in three dimensions can be defined by its axis $\mathbf{v}$ and angle θ. The rotation matrix $\mathbf{R}$ can be represented by the rotation axis and angle, i.e.,

$$\mathbf{R} = \cos(\theta)\mathbf{I} + (1 - \cos(\theta))\mathbf{v}\mathbf{v}^T + \sin(\theta)\mathbf{v}^\wedge, \tag{3}$$

where $||\mathbf{v}|| = 1$, $\theta \in [0, \pi]$ and

$$\mathbf{v}^\wedge = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}. \tag{4}$$

The vector $\mathbf{v} = (v_x, v_y, v_z)$ can be any unit vector. Point cloud registration aims to estimate $\{\mathbf{R}_{pq}, \mathbf{t}_{pq}\}$ from point cloud P to Q. We define the variables $\{\mathbf{v}_{pq}, \theta_{pq}\}$ to generate the rotation matrix $\mathbf{R}_{pq}$ by (3). Since the modulo of $\mathbf{v}$ is restricted to one in optimization, we normalize $\mathbf{v}_{pq}$ to solve the constraint as

$$\mathbf{v} = \frac{\mathbf{v}_{pq}}{||\mathbf{v}_{pq}||} \tag{5}$$

As $\cos(\cdot)$ is a symmetric and periodic function, we directly optimize θ_{pq} on $\mathbb{R}$. Therefore, we reconstruct the $\mathbf{R}_{pq}$ with unconstrained variables $\{\mathbf{v}_{pq}, \theta_{pq}\}$.

Then, we focus on the translation distance along the translation direction. The translation direction is generally regarded as a unit vector. The translation distance is usually limited. One reason is that the corresponding points are used for point cloud centroid alignment, the other is that the translation distance is bounded in realistic environments. The translation vector $\mathbf{t}$ can be represented by a translation direction $\mathbf{u}$ and a distance d as,

$$\mathbf{t} = d\mathbf{u}. \tag{6}$$

To ensure the legitimacy of the translation, we constrain the translation distance: $d <= d_{\max}$. Constant $d_{\max}$ represents the upper bound of translation distance. This constraint is known as "box constraint" in the optimization literature. There are three different methods to solve this problem [28], i.e.,

1) Perform standard gradient descent, then clip all variables to satisfy the constraints.
2) Incorporate the constraint functions into the objective function and solve the generalized Lagrangian function.
3) Introduce unconstrained variables instead of the original constrained variables to optimize the objective function.

Variables tend to be concentrated in the boundaries using the first method. The second method requires the objective function to be convex according to the Lagrangian duality. The third method is adopted to alleviate the problem of being trapped in boundary regions. We use the mapping function to replace the constrained d with an unconstrained d_{pq}. A mapping function extends the variable d_{pq} to the $\mathbb{R}$ space in terms of translational distance. We reconstruct (6) with the unconstrained variables $\{d_{pq}, \mathbf{u}_{pq}\}$ and mapping function sigmoid$(\cdot)$ as

$$\mathbf{t}_{pq} = d_{\max}\text{sigmoid}(d_{pq})\frac{\mathbf{u}_{pq}}{||\mathbf{u}_{pq}||}. \tag{7}$$

With a similar idea, we also use the $\sin(\cdot)$ as a mapping function, i.e.,

$$\mathbf{t}_{pq} = 1/2(d_{\max} + d_{\max}\sin(d_{pq}))\frac{\mathbf{u}_{pq}}{||\mathbf{u}_{pq}||}. \tag{8}$$

Different mapping functions have biases to generate translation distances. The mapping function is selected according to the distribution of the translation distances in the dataset.

3.3. Hybrid optimization objective

To deal with the registration problem in (2), we introduce the chamfer distance as an evaluation metric for two aligned point clouds. The original CD takes each point into account to find the nearest point in the other point cloud. For pairwise partial point clouds, the CD only works reasonably on the corresponding points related to the overlap regions. In this paper, we adopt two variants of the CD loss to align a pair of point clouds with the local and global registration.

Local CD loss Since a pair of partial point clouds have only related to the corresponding points, it is not reasonable to minimize the CD loss on the entire points. The local CD loss focuses on the local matching points, improving the tolerance for outliers on registration. The local CD loss is defined by

$$\begin{aligned}\mathcal{L}_{CD_{local}}(P, Q, \alpha) = &\frac{1}{|P_\alpha|}\sum_{\mathbf{p}\in P_\alpha} \min_{\mathbf{q}\in Q} \|\mathbf{p} - \mathbf{q}\|^2 \\ &+ \frac{1}{|Q_\alpha|}\sum_{\mathbf{q}\in Q_\alpha} \min_{\mathbf{p}\in P} \|\mathbf{q} - \mathbf{p}\|^2,\end{aligned} \tag{9}$$

where P_α and Q_α are subsets of P and Q, respectively. α is related to the proportion of overlapping. $|P_\alpha| = \alpha|P|$ and $|Q_\alpha| = \alpha|Q|$. P_α is selected from P using the following processes: (1). In each round, the point p in P is selected with the smallest value according to $\min_{\mathbf{q}\in Q} ||\mathbf{p}-\mathbf{q}||^2$. Then the point is removed from P. (2). Repeat process (1) and stop until $\alpha|P|$ points are selected. The subset Q_α is generated in the same way. The local CD loss adjusts α to control the number of matching points from the original point cloud to the target point cloud. With the hyperparameter α (only one hyperparameter needs to be adjusted), the local CD loss can be used in various registration tasks with different overlap regions.

Projected CD loss The local CD loss only pays attention to the local alignment but ignores the global point matching. When the corresponding points in the overlap are more than the chosen points with the fixed hyperparameters α, the local CD loss fails to work in this situation. The points in the three-dimensional space are projected onto a two-dimensional plane, which retains the contour feature for global registration. Based on this observation, we project the point cloud onto the $x-y$, $y-z$, $x-z$ planes and calculate the projected CD loss as

$$\begin{aligned}\mathcal{L}_{CD_{uv}}(P,Q) = &\frac{1}{|P|}\sum_{\mathbf{p}\in P}\min_{\mathbf{q}\in Q}\|\mathbf{p}_{uv}-\mathbf{q}_{uv}\|^2 \\ &+\frac{1}{|Q|}\sum_{\mathbf{q}\in Q}\min_{\mathbf{p}\in P}\|\mathbf{q}_{uv}-\mathbf{p}_{uv}\|^2,\end{aligned} \tag{10}$$

where uv denotes a projected plane, $\mathbf{p}_{uv}$ and $\mathbf{q}_{uv}$ are the projection vectors.

Hybrid loss The hybrid loss consists of the local CD and projected CD loss, which is defined as

$$\begin{aligned}\mathcal{L}_{hybrid}(P,Q) = &\mathcal{L}_{CD_{local}}(P,Q,\alpha) \\ &+\beta(\mathcal{L}_{CD_{xy}}(P,Q)+\mathcal{L}_{CD_{yz}}(P,Q) \\ &+\mathcal{L}_{CD_{xz}}(P,Q)),\end{aligned} \tag{11}$$

where β is a balance weight. Hyperparameter β can be estimated by the gradients of variables. With the hybrid loss, we iteratively update the unrestricted variables by gradient descent.

3.4. Strategies about optimization

Predicting rigid transformation by gradient descent is a non-convex optimization problem. We apply HOUV with multiple initialization instances to alleviate the local extrema problem. For each pairwise partial point cloud, we initialize the rotation axis $\mathbf{v}_{pq}$ and angle θ into 32 groups. With the parallel computing on GPUs, those variables of instances are optimized at the same time. At last, the transformation of the smallest local CD loss is chosen as the final result for registration. We assume that the optimal registration on pairwise partial point clouds satisfies the minimum local CD loss in those groups. Algorithm 1 shows our method in details.

HOUV can not only solve the registration in the entire SE(3) space, but also adjust the search subspace by introducing the boundaries. The boundaries come from the prior range of transformation by data analysis. When the entire $SE(3)$ space is divided into several subspaces, HOUV can perform the fine-grained optimization on each subspace. This strategy improves the inference speed and accuracy, partially alleviating the problem of symmetric point cloud objects. For symmetrical objects, the partial point clouds are still aligned when adding π or $\pi/2$ degrees to the rotation angle. By specifying the search subspace, HOUV can solve the above problem. Similar to (7), we restrict the rotation angle θ_{pq} to $[l, r]$ as follows

$$\theta_{pri} = l + (r-l)\sin(\theta_{pq}), \tag{12}$$

where l and r are the lower bound and upper bound for the rotational angle, respectively. HOUV with optimization strategies can be flexibly adapted to the distribution of the dataset. We summarize HOUV with the prior of the dataset in Algorithm 2.

Algorithm 1 Hybrid optimization method with unconstrained variables ($HOUV_{ori}$).

Input: Pairwise partial point clouds P and Q.
Parameter: The group size g, learning rate l, epochs $iter$, α, β.
Output: Rotation $\mathbf{R}_{pq}$ and transformation $\mathbf{t}_{pq}$ with unconstrained variables.

```
Initialize the variables {v_pq, θ, u_pq, d_pq} with g groups randomly.
for each set about variables in group g do
    while not reach iter do
        Calculate rotation R_pq and translation t_pq, following (3) and (8).
        Calculate the hybrid loss following (11).
        Perform backpropagation to compute the gradient of the variables with respect to the hybrid loss.
        Follow the gradient descent method and update the variables.
    end while
end for
Choose the result with the smallest local CD loss.
return {v_pq, θ, u_pq, d_pq, R_pq, t_pq}.
```

Algorithm 2 Hybrid optimization method with unconstrained variables based on prior data ($HOUV_{pri}$).

Input: Pairwise partial point clouds P and Q.
Parameter: The group size g, learning rate l, epochs $iter$, α, β.
Output: Rotation $\mathbf{R}_{pq}$ and transformation $\mathbf{t}_{pq}$ with unconstrained variables.

```
Initialize the variables {v_pq, θ, u_pq, d_pq} with g groups randomly.
for each set about variables in group g do
    while not reach iter do
        Set the max rotational angle and translation distance following the prior distribution of data following (12).
        Calculate rotation R_pq and translation t_pq, following (3) and (8).
        Calculate the hybrid loss following (11).
        Perform backpropagation to compute the gradient of the variables with respect to the hybrid loss.
        Follow the gradient descent method and update the variables.
    end while
end for
Choose the result with the smallest local CD loss.
return {v_pq, θ, u_pq, d_pq, R_pq, t_pq}.
```

4. Experiment

In this section, we provide the detailed experimental settings, then evaluate the proposed methods on three synthetic datasets, including MVP, ModelNet and ICL-NUIM. We also verify the performance of registration methods on the real-world 3DMatch dataset. Ablation studies investigate the effectiveness of the proposed methods and losses. We also analyse the hyperparameters of HOUV methods. In the end, we discuss some related topics about HOUV.

4.1. Dataset and metric

Multi-view partial point cloud (MVP) Pan et al. [29] is a high-quality multi-view partial point cloud dataset, which contains over 100,000 high-quality scans. For each CAD model, the partial point clouds are acquired from 26 uniformly distributed camera poses in MVP. The rotation angles are mostly within $[0, \pi/4]$, and the rest

are unrestricted in $[0, \pi]$. The ratio is roughly 4 : 1. The translation distance is limited to [0,0.25].

ModelNet Wu et al. [10] includes 12,311 meshed CAD models from 40 categories. For each object in the dataset, we randomly sample 1024 points as the source point cloud P, and then apply a random transformation on P to obtain the target point cloud Q by shuffling those points. Gaussian noise from $\mathcal{N}(0, 0.01)$ is added to simulate real-world noise. Actually, ModelNet is not strictly a pairwise partial point clouds dataset. The rotation angle is uniformly selected from $[0, 2\pi]$ and the translation distance is limited to [0,0.5]. For fairness, we directly use the three datasets generated by Yuan et al. [11], i.e., datasets with clean, noisy and unseen on ModelNet.

ICL-NUIM Handa et al. [12] denotes the Imperial College London and National University of Ireland Maynooth dataset for the evaluation of visual odometry, 3D reconstruction, and SLAM algorithms. Following [11], they use the point clouds derived from RGB-D scans in the Augmented ICL-NUIM dataset [30] and then augment it with the realistic sensor noise and arbitrary transformation between input pairs. ICL-NUIM constrains the rotation angle in $[0, 2\pi]$ and the translation distance in [0,0.05].

3DMatch Zeng et al. [31] consists of 3D point cloud pairs from various real-world scenes with ground truth transformations esimated from RGB-D reconstruction pipelines. We compare our unsupervised methods and other methods on test dataset where the pair partial point clouds is generated with at least 30% overlap [32].

Metrics Following the MVP challenge, we choose the isotropic rotation error Rot_{err}, isotropic translation error $Trans_{err}$ and mean square error (MSE) as the metrics. We denote the result transformation $\mathbf{T}_{pq} = \{\mathbf{R}_{pq}, \mathbf{t}_{pq}\}$ and the ground truth $\mathbf{T}_{gr} = \{\mathbf{R}_{gr}, \mathbf{t}_{gr}\}$. Rot_{err} calculates the angular distance between $\mathbf{R}_{pq}$ and $\mathbf{R}_{gr}$ rotations as

$$Rot_{err} = \frac{180}{\pi} \arccos\left(\frac{\left(\text{Trace}\left(\mathbf{R}_{pq}\mathbf{R}_{gr}{}^{T}\right) - 1\right)}{2}\right). \tag{13}$$

For the convenience of analysis, the angles are expressed in degrees. $Trans_{err}$ and MSE are evaluated using the L_2 distance of the translation vector and the corresponding points in pairwise partial point clouds. The MSE is defined by

$$MSE = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{T}_{pq}(\mathbf{p}_i) - \mathbf{T}_{gr}(\mathbf{p}_i)\|^2, \tag{14}$$

where $\mathbf{T}_{pq}(\mathbf{p}_i) = \mathbf{R}_{pq}\mathbf{p}_i + \mathbf{t}_{pq}$ and $\mathbf{T}_{gr}(\mathbf{p}_i) = \mathbf{R}_{gr}\mathbf{p}_i + \mathbf{t}_{gr}$. Following [11], we evaluate the RMSE and recall metrics on ModelNet and ICL-NUM.

For the 3DMatch dataset, the evaluation metrics are slightly different from other datasets. Average Rot_{err} and $Trans_{err}$ are calculated only on these successfully registered pairs. The main reason is that the results of failed estimation make the metrics unreliable compared with the ground truth. Following [13], the successful registration meets the $Tran_{err}$ threshold 30cm and the Rot_{err} threshold 15 degrees. The recall is the ratio of the successful matching points in pairwise registrations.

4.2. Comparing methods and implementations

We test the ICP [5], fast global registration (FGR) [33] as the basic iterative optimization methods. The ICP and FGR are implemented by Open3D.[2] Euler is a constrained optimization method which derived from $HOUV_{pri}$. We also evaluate three learning-based methods, DCP [9], IDAM [23] and DeepGMR [11] as benchmarks in the MVP registration competition. RGM [26] is the current state-of-the-art method on ModelNet. HOUV denotes the default $HOUV_{ori}$ for fair comparison. We set the hyperparameters $\alpha = 0.5$ and $\beta = 0.1$ for HOUV. As for the ModelNet and ICL-NUM datasets, the results of other methods are referred to Yuan et al. [11], including HGMR [34], PointNetLK [35] and ICP with RANSAC [36].

For iterative optimization methods, we set the number of iterations to 500. We randomly initialize the 32 group variables for one pairwise partial point cloud. Adam [37] is chosen as the default optimizer and the learning rate is set to 0.001. For MVP dataset, we also use $HOUV_{pri}$ to improve the performance in the competition. We divide 180 degrees into four intervals with 45 degrees in terms of rotational angle. For ICP and FGR methods, we also randomly initialize 32 groups for each method. We select the final transformation matrix with the highest fitness. For learning-based registration methods, DCP, IDAM and DeepGMR are pretrained with data augmentation.

4.3. Performances on datasets

HOUV method shows superior performances on synthetic datasets. As shown in Table 1, HOUV is superior to other methods on MVP dataset. Fig. 1 visualizes the aligned results with HOUV. While the learning-based methods surpass the traditional optimization methods like ICP and FGR, their performances are weaker than our proposed methods. We investigate the reason for the large rotation error on traditional methods. For MVP dataset, most point cloud objects have a symmetrical structure as shown in Fig. 1. When greedily applying the chamfer distance on pairwise partial point clouds, the registration results will overlap as much as possible but ignore the existence of similar structures. From the 32 groups of ICP, we select the highest fitness as the final registration result. This strategy fails when the proportion of overlap is smaller than the hyperparameter α in pairwise partial point clouds. Methods like ICP may cause an extra 180 or 90 rotation error. As for HOUV, we can adjust α to control the overlapping area by the local CD loss. However, each pairwise partial point cloud has a different proportion of overlapping area in the dataset. An adaptive adjustment of α is not used in HOUV. Instead of the adaptive strategy, we set the fixed $\alpha = 0.5$ to control the local registration. The fixed α is derived from the average overlap on the dataset.

For ModelNet and ICL-NUIM, Table 2 shows the results of HOUV and other methods whose results are referred to Yuan et al. [11]. Since ModelNet and ICL-NUIM can be effectively solved with many registering methods, HOUV does not demonstrate significant improvements compared with other methods like RGM. Although the experiment results are limited, HOUV demonstrates generalization on multiple synthetic datasets.

HOUV has also been verified with various scenes in realistic 3D Match dataset. Table 3 shows the registration results on 3DMatch. The registration is implemented on downsampling point clouds with 5 cm voxels. We only evaluate the HOUV method on the 3DMatch, the other results come from Choy et al. [13]. As shown in Table 3, HOUV outperforms the classical registration methods. The registration accuracy of HOUV is also validated on real-world datasets. We investigate the reason about the slightly inferior performance of HOUV to DGR. Firstly, HOUV is a unsupervised method compared to DGR. DGR learns the prior object features of similar scenes in the training dataset. Secondly, there is a wide distribution of transformations and overlapping regions in the 3DMatch dataset. HOUV cannot fit all pairwise point clouds with the fixed hyperparameter α. For failed samples, HOUV can adjust the hyperparameter α to align those partial point clouds. Fig. 2 shows the registration results on different scenes of 3DMatch.

[2] http://www.open3d.org/

Table 1
The results of our method and other registration methods on MVP dataset.

Method	ICP	FGR	GO-ICP	DCP	IDAM	DeepGMR	RGM	Euler	DGR	HOUV
Rot_{err}	35.87	30.47	56.76	26.27	22.87	49.95	17.4	5.63	20.89	**3.05**
$Trans_{err}$	0.14	0.15	0.24	0.23	0.23	0.38	0.10	0.03	0.21	**0.02**
MSE	0.68	0.60	0.79	0.64	0.62	0.70	0.42	0.13	0.57	**0.07**

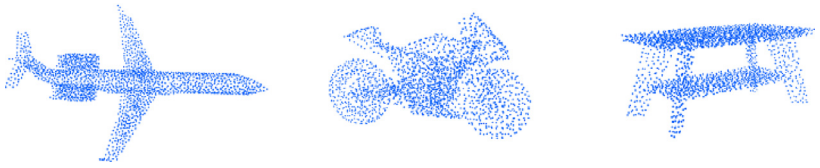

Fig. 1. The aligned results with HOUV method on MVP dataset. Each color represents a partial point cloud. Each color means a partial point cloud. The red sample represents the point clouds P. The green sample represents the point clouds Q. In the aligned row, we demonstrate the P after performing the predicted transformation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Average RMSE and recall with threshold 0.2 on ModelNet and ICL-NUIM.

	ModelNet clean		ModelNet noisy		ModelNet unseen		ICL-NUIM	
	RMSE	Re@0.2	RMSE	Re@0.2	RMSE	Re@0.2	RMSE	Re@0.2
ICP	0.53	0.41	0.53	0.41	0.59	0.32	1.16	0.27
FGR	0.19	0.79	0.19	0.79	0.23	0.75	0.15	0.87
HGMR	0.52	0.44	0.52	0.45	0.54	0.43	0.72	0.50
ICP + RANSAC	0.08	0.91	0.42	0.49	0.30	0.67	0.17	0.84
PointNetLK	0.51	0.44	0.56	0.38	0.68	0.13	1.29	0.08
DCP	0.02	0.99	0.08	0.94	0.34	0.54	0.64	0.16
DeepGMR	**<0.01**	0.99	**<0.01**	0.99	0.01	0.99	0.07	0.99
RGM	**<0.01**	1.0	**<0.01**	0.99	**<0.01**	0.99	0.04	0.99
DGR	**<0.01**	0.99	**<0.01**	0.99	0.02	0.99	0.05	0.99
HOUV	**<0.01**	0.99	**<0.01**	0.99	**<0.01**	0.99	**0.01**	0.99

Table 3
The results of our method and other registration methods on 3DMatch dataset.

Method	ICP	FGR	GO-ICP	DCP	DGR	HOUV
Rot_{err}	8.25	4.08	5.38	8.42	2.43	3.42
$Trans_{err}$(cm)	18.1	10.6	14.7	21.4	7.34	7.77
Recall	6.04%	42.7%	22.9%	3.22%	91.3%	60%

4.4. Ablation studies

In Table 4, we investigate the different components about HOUV. As shown in the first and fourth rows of Table 4, the projected CD loss improves the rotational accuracy in pairwise partial point clouds with symmetric structures. Compared with the results in the second and fourth rows, HOUV decreases the translation error by replacing the variable d with unconstrained variable d_{pq}. Using the mapping function, unconstrained variable d_{pq} greatly reduces the translation error along the translation direction. We do not divide the optimization intervals in the third row. The performance of HOUV is improved by partitioning more fine-grained intervals for rotational angles. Due to the multi GPUs parallelization, the inference time does not increase linearly with the number of intervals when those strategies are employed.

Fig. 3 visualizes the results of large rotational errors without projected CD loss. In this figure, we analyse the HOUV with the local CD loss and the projected loss on the same samples. The partial point cloud P and Q are received from different viewpoints. We perform the predicted transformation to align P with Q. In the second column, HOUV fits well in the local region, but fails on the global registration when only the local CD loss is used. The reason is mainly that we set the hyperparameter $\alpha = 0.5$ for the local CD loss. The projected CD loss can solve the sample that the proportion of overlap is bigger than the hyperparameter α. As shown in the third column, the projected CD loss can repair the flaws of local CD loss to align those examples.

We also analyse the mapping functions on the registration. Table 5 demonstrates the different mapping functions for the translation error. The distributions of translation distance are shown in Fig. 4. Since the interval of each bin is 0.01, the translation distance between [0,0.25] is equally divided into 25 parts. The distribution of translation distance is a uniform distribution between $[0, 0.25]$ in MVP dataset. With the saturation of the sigmoid function, the Adam optimizer tends to push variables to boundaries. The distribution of translation distances about the sine mapping function is roughly consistent with the uniform distribution of the prior data.

4.5. Hyperparameter analysis

As an optimization-based registration method, the performance of HOUV is related to hyperparameter for each sample. In previous experiments, we evaluate the performance and generalization of HOUV with the hypeparameters $\alpha = 0.5$ and $\beta = 0.1$ on the synthetic and realistic datasets. For the failed samples with the $Rot_{err} > 10$ in the dataset, we can interactively tune the hyperparameters of HOUV to align pairwise partial point clouds by visualizations. Fig. 5 illustrates the effect of hyperparameter α on registration. The pairwise point clouds fail to align with the default hyperparameter $alpha = 0.5$ in Fig. 5. The overlapping of the pairwise point clouds was less than 50% by visualizing the failed sample. As shown in the right sub-graph of the Fig. 5, we adjust $\alpha = 0.3$ to make a successful match.

We set the hyperparameter $\beta = 0.1$ so that the local CD loss dominates in the optimization. The projected CD loss achieves coarse alignment for the global registration. Since the projected CD

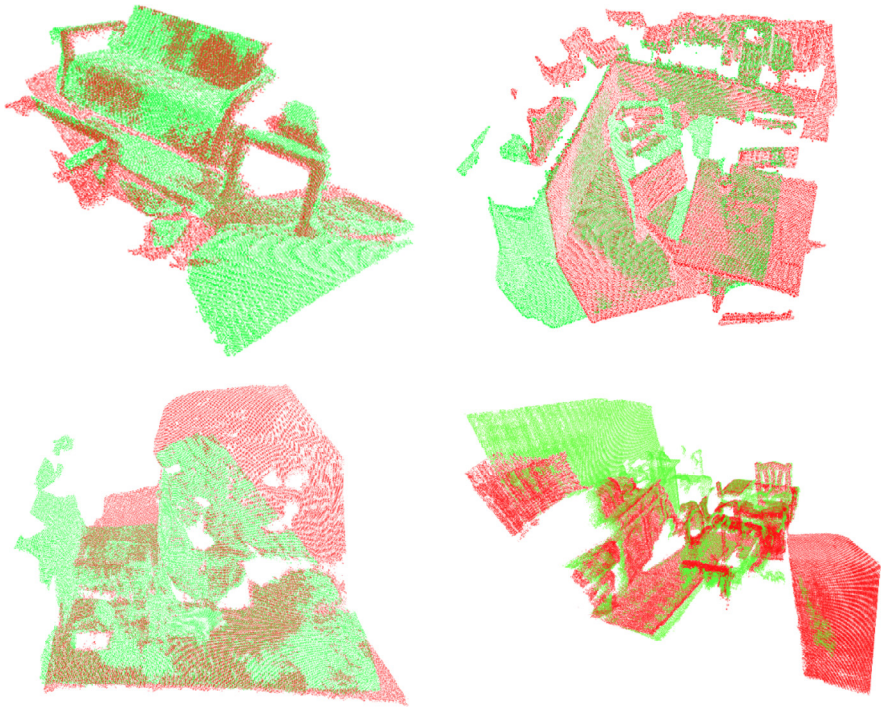

Fig. 2. The registration results of HOUV about pairwise partial point clouds on test scenes in 3DMatch. The red sample represents the point clouds P. The green sample represents the point clouds Q. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
The ablation experiments of the HOUV on MVP.

Local CD	Projected CD	Unconstrained T	Strategies	Rot error	Trans error	MSE
✓		✓	✓	3.4654	0.0229	0.0799
✓	✓		✓	3.2349	0.0291	0.0795
✓	✓	✓		3.8390	0.0238	0.0734
✓	✓	✓	✓	2.9987	0.0213	0.0719

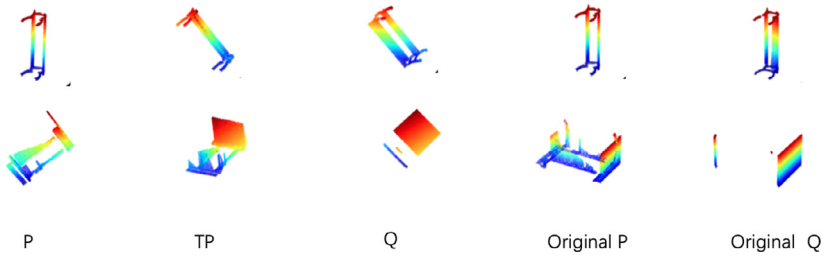

Fig. 3. The registration examples are shown about HOUV without the projected loss on MVP dataset. The red sample represents the point clouds P. The green sample represents the point clouds Q. The blue sample represents the aligned P with the projected CD loss. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

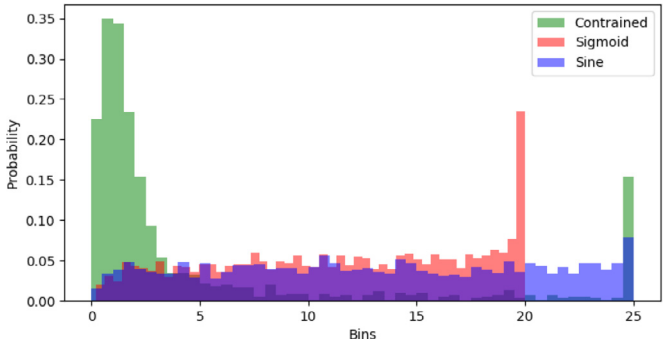

Fig. 4. The distributions of translation distance with mapping functions on MVP. The constrained refers to the method without the mapping function.

Table 5
The performances about mapping functions on MVP dataset.

Function	Constrained	Sigmoid	Sine
Rot_{err}	0.08	0.04	0.02
$Trans_{err}$	0.03	0.03	0.02
MSE	0.41	0.25	0.08

loss converges faster than the local CD loss, the hyperparameter β is robust for the registration.

The group size g can solve the multiple local minimums in the hybrid loss. A single initialization of the unconstrained variables $\{\mathbf{v}_{pq}, \theta, \mathbf{u}_{pq}, d_{pq}\}$ makes the registration results unstable. Table 6 shows the results of different initialization groups. To some extent,

Fig. 5. The registration results with different hyperparameter α in same pairwise point clouds.

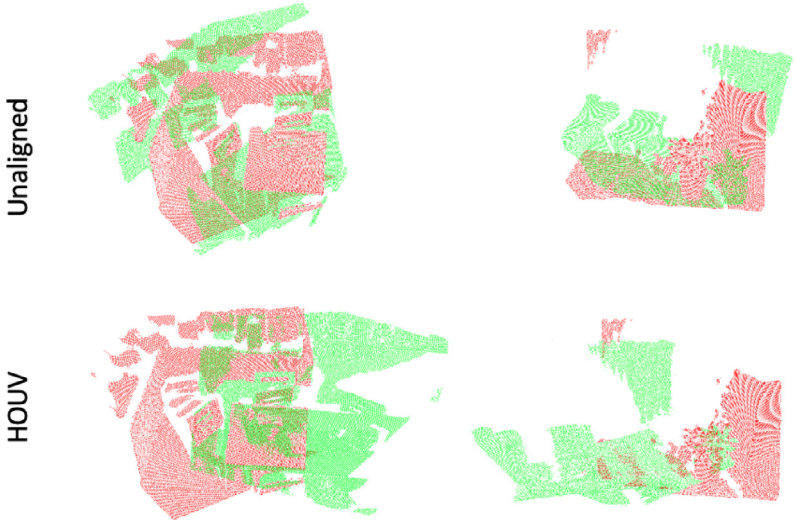

Fig. 6. The registration results are aligned with the large rotational angle. The red sample represents the point clouds *P*. The green sample represents the point clouds *Q*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
The performances abut groups on MVP validation dataset.

Group size	1	4	8	16	32	64
Rot_{err}	14.81	10.87	7.08	4.97	4.07	4.53
$Trans_{err}$	0.08	0.06	0.04	0.02	0.02	0.02
MSE	0.28	0.13	0.09	0.07	0.06	0.07

the bigger group size g improves the stability and the accuracy of alignment. However, we choose the transformation result with the smallest local CD loss from the group, which cannot guarantee optimal solutions due to hyperparameters setting. The registration accuracy of group size $g = 64$ is lower than that of group size $g = 32$. In addition, the inference time also increases linearly with the group size when computing resources are limited. Based on the computing capacity of the GPUs, we set the group size to be 32 about HOUV in terms of performance.

In the 3Dmtach dataset, the partial point clouds are obtained from the motion trajectory of the same scene. The rotation angle in the adjacent segments of trajectory is limited to $[0, \pi/4]$. HOUV is applied to align two non-adjacent partial point clouds whose rotation angle is larger than $\pi/4$. Following the strategy in (12), we limit the rotational angle in $[l, r]$, where l and r are the lower bound and upper bound for the rotational angle. As shown in the Fig. 6, we constraint the rotational angle in $[\pi/4, \pi/2]$ to estimate the transformation with the large rotation angle.

4.6. Discussions

The failed examples Fig. 7 shows some failed examples in MVP. Considering the symmetry of the CAD model, HOUV often fails to obtain the correct transformation when aligning these samples. The main reason is that selecting the result with smallest local CD loss cannot obtain the correct transformation from the group results. To solve this issue, we adopt to specify an optimization subspace for HOUV. In Fig. 7, the result *TP* is obtained in the first row when the rotation angle is calculated within $[0, \pi]$. Comparing the original P and Q at the same coordinates, the original point cloud pairs are locally aligned due to the viewpoint deviation. When the rotation angle is assumed to be between $[0, \pi]$, $HOUV_{ori}$ makes approximately π angle error due to object symmetry. If the rotation angle is assumed to be within $[0, \pi/4]$, $HOUV_{pri}$ can predict roughly correct transformations to align those samples. $HOUV_{pri}$ can be programmed to interactively align the special sample through visual analysis.

The speed of inference Table 7 shows the inference speed of iterative registration methods. We test the average time to align a pair of cloud points by the $HOUV_{ori}$, $HOUV_{pri}$ and $HOUV_{not}$ methods. $HOUV_{not}$ means $HOUV_{ori}$ without the projected CD loss. We perform $HOUV_{mul}$ to accelerate inference using data parallelism on multiple GPUs. The above methods use the same default hyperparameters, i.e., 32 initialization groups and 500 iterations. The experiments are tested on a 1080Ti GPU. From Table 7, the inference speed is acceptable for HOUV, which is slower than ICP but faster

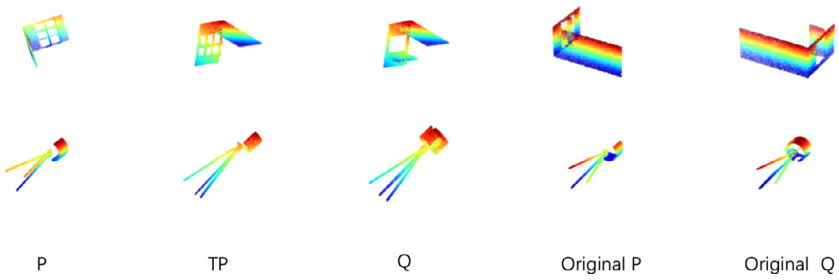

Fig. 7. The failed examples are shown with our method on validation dataset of MVP. The red sample represents the point clouds *P*. The green sample represents the point clouds *Q*. The blue sample represents the ground truth about aligned *P*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7
The average time of inference (s) for one sample with optimization-based methods.

Method	ICP	ICP + RANSAC	GO-ICP	$HOUV_{ori}$	$HOUV_{pri}$	$HOUV_{not}$	$HOUV_{mul}$
Cost (s)	0.07	0.91	10.32	7.36	29.83	3.78	1.67

Table 8
The performances about the iterations on validation dataset of MVP.

Iterations	50	100	200	400	500
Rot_{err}	11.39	7.87	4.63	4.73	4.53
$Trans_{err}$	0.08	0.04	0.02	0.03	0.02
MSE_{err}	0.17	0.10	0.07	0.08	0.07
Cost (s)	0.98	1.65	3.07	6.3	7.8

than GO-ICP. Besides, FGR is faster than other methods. The reason is that there is no iterative point matching in the pipeline of FGR, which differs from that of ICP and our methods.

We investigate the performance of HOUV with different iterations in Table 8. Reducing the number of iterations can linearly decrease the inference time. In the previous experiments, we mainly focus on the registration accuracy when setting the default iteration to be 500. For most pairwise point clouds, the default epoch 500 is redundant. Furthermore, implementing a faster nearest neighbour method on GPUs is beneficial for accelerating HOUV [38]. When the performance of HOUV drops slightly without the projected CD loss, it can speed up the registration.

Advantages and disadvantages Considering the advantages, HOUV as an optimization algorithm has superior generalization without supervision information. HOUV has a programmable optimization strategy that can flexibly select the mapping function to fit the data. Furthermore, our method has better interpretability than learning-based methods. In the process of optimization, we can visualize the alignment process and make interactive analysis. As the nearest neighbour selection is the bottleneck, HOUV requires the acceleration of GPUs, which limits the scope of its applications. For HOUV, the choice of hyperparameters is critical. Hyperparameter α controls the proportion of overlap in pairwise point clouds. The bounds l and r are introduced to specify the optimization subspace. The number of iterations can balance the speed and accuracy of registration. Those hyperparameters improve the flexibility of HOUV in different registration cases.

5. Conclusion

In this paper, we propose a hybrid optimization method with unconstrained variables on point cloud registration. We propose the local CD loss and projected CD loss. The local CD loss deals with the local point matching and the projected CD loss solves the global registration. We also implement tunable hyperparameters to constrain the optimized variables. Those strategies not only improve the accuracy in registration but also decrease the inference time. Our studies demonstrate that the optimization-based methods still have great research potential for point cloud registration compared to the learning-based methods. In the future, we can improve HOUV from the following aspects. (1): Adaptive hyperparameter selection. In HOUV, we choose the static hyperparameters by the prior of data. Studying the dynamic mechanism based on overlapping regions can improve the stability of HOUV. (2): Accelerating HOUV with learning-based models. The speed bottleneck of HOUV is the iterative process. We may reduce the number of iterations with the help of the learned model.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant 62276127.

References

[1] J.-E. Deschaud, IMLS-SLAM: scan-to-model matching based on 3D data, in: IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 2480–2485.
[2] Y. Li, L. Ma, Z. Zhong, F. Liu, M.A. Chapman, D. Cao, J. Li, Deep learning for Lidar point clouds in autonomous driving: a review, IEEE Trans. Neural Netw. Learn. Syst. 32 (8) (2020) 3412–3432.
[3] T. Bailey, H. Durrant-Whyte, Simultaneous localization and mapping (slam): part ii, IEEE Robot. Autom. Mag. 13 (3) (2006) 108–117.
[4] F. Pomerleau, F. Colas, R. Siegwart, A review of point cloud registration algorithms for mobile robotics, Found. Trends Robot. 4 (1) (2015) 1–104.
[5] P.J. Besl, H.D. Mckay, A method for registration of 3-D shapes, IEEE Trans. Pattern Anal. Mach. Intell. 14 (2) (1992) 239–256.
[6] Q.-Y. Zhou, J. Park, V. Koltun, Fast global registration, in: Proceedings of the European Conference on Computer Vision, Springer, 2016, pp. 766–782.
[7] J. Yang, H. Li, D. Campbell, Y. Jia, Go-ICP: a globally optimal solution to 3D ICP point-set registration, IEEE Trans. Pattern Anal. Mach. Intell. 38 (11) (2015) 2241–2254.
[8] Z. Deng, Y. Yao, B. Deng, J. Zhang, A robust loss for point cloud registration, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 6138–6147.
[9] Y. Wang, J.M. Solomon, Deep closest point: learning representations for point cloud registration, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 3523–3532.

[10] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, J. Xiao, 3D ShapeNets: a deep representation for volumetric shapes, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1912–1920.
[11] W. Yuan, B. Eckart, K. Kim, V. Jampani, D. Fox, J. Kautz, DeepGMR: learning latent Gaussian mixture models for registration, in: European Conference on Computer Vision, Springer, 2020, pp. 733–750.
[12] A. Handa, T. Whelan, J. McDonald, A.J. Davison, A benchmark for RGB-D visual odometry, 3Dreconstruction and slam, in: IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2014, pp. 1524–1531.
[13] C. Choy, W. Dong, V. Koltun, Deep global registration, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2514–2523.
[14] S. Horache, J.-E. Deschaud, F. Goulette, 3D point cloud registration with multi-scale architecture and unsupervised transfer learning, in: International Conference on 3D Vision (3DV), 2021, pp. 1351–1361.
[15] H. Fan, H. Su, L.J. Guibas, A point set generation network for 3D object reconstruction from a single image, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 605–613.
[16] R.I. Hartley, F. Kahl, Global optimization through searching rotation space and optimal estimation of the essential matrix, in: 2007 IEEE 11th International Conference on Computer Vision, IEEE, 2007, pp. 1–8.
[17] F. Pomerleau, F. Colas, R. Siegwart, S. Magnenat, Comparing ICP variants on real-world data sets, Auton. Robots 34 (3) (2013) 133–148.
[18] R.B. Rusu, N. Blodow, M. Beetz, Fast point feature histograms (FPFH) for 3Dregistration, in: IEEE International Conference on Robotics and Automation, IEEE, 2009, pp. 3212–3217.
[19] D. Chetverikov, D. Svirko, D. Stepanov, P. Krsek, The trimmed iterative closest point algorithm, in: International Conference on Pattern Recognition, vol. 3, 2002, pp. 545–548.
[20] B. Jian, B.C. Vemuri, Robust point set registration using Gaussian mixture models, IEEE Trans. Pattern Anal. Mach. Intell. 33 (8) (2010) 1633–1645.
[21] X. Ma, S. Xu, J. Zhou, Q. Yang, Y. Yang, K. Yang, S.H. Ong, Point set registration with mixture framework and variational inference, Pattern Recognit. 104 (2020) 107345.
[22] C. Choy, J. Park, V. Koltun, Fully convolutional geometric features, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 8958–8966.
[23] J. Li, C. Zhang, Z. Xu, H. Zhou, C. Zhang, Iterative distance-aware similarity matrix convolution with mutual-supervised point elimination for efficient point cloud registration, in: European Conference on Computer Vision (ECCV), 2020.
[24] Z. Zhang, J. Sun, Y. Dai, D. Zhou, X. Song, M. He, Self-supervised rigid transformation equivariance for accurate 3D point cloud registration, Pattern Recognit. 104 (2022) 108784.
[25] X. Huang, G. Mei, J. Zhang, Feature-metric registration: a fast semi-supervised approach for robust point cloud registration without correspondences, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 11366–11374.
[26] K. Fu, S. Liu, X. Luo, M. Wang, Robust point cloud registration framework based on deep graph matching, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 8893–8902.
[27] J.S. Dai, Euler–Rodrigues formula variations, quaternion conjugation and intrinsic connections, Mech. Mach. Theory 92 (2015) 144–152.
[28] N. Carlini, D. Wagner, Towards evaluating the robustness of neural networks, in: IEEE Symposium on Security and Privacy (SP), IEEE, 2017, pp. 39–57.
[29] L. Pan, X. Chen, Z. Cai, J. Zhang, H. Zhao, S. Yi, Z. Liu, Variational relational point completion network, arXiv preprint arXiv:2104.10154(2021).
[30] S. Choi, Q.-Y. Zhou, V. Koltun, Robust reconstruction of indoor scenes, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 5556–5565.
[31] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, T. Funkhouser, 3DMatch: learning local geometric descriptors from RGB-D reconstructions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1802–1811.
[32] H. Deng, T. Birdal, S. Ilic, PPFNet: global context aware local features for robust 3D point matching, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 195–205.
[33] Q.-Y. Zhou, J. Park, V. Koltun, Fast global registration, in: European Conference on Computer Vision, Springer, 2016, pp. 766–782.
[34] B. Eckart, K. Kim, J. Kautz, HGMR: hierarchical Gaussian mixtures for adaptive 3Dregistration, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 705–721.
[35] Y. Aoki, H. Goforth, R.A. Srivatsan, S. Lucey, Pointnetlk: robust & efficient point cloud registration using pointnet, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 7163–7172.
[36] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395.
[37] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980(2014).
[38] D. Qiu, S. May, A. Nüchter, GPU-accelerated nearest neighbor search for 3d registration, in: International Conference on Computer Vision Systems, Springer, 2009, pp. 194–203.

Yuanjie Yan received the B.S. degree in the department of Computer Science and Technology from Jilin University in 2018. He is pursuing the Ph.D. degree in Computer Science and Technology at Nanjing University under the supervision of Prof. Furao Shen. His research interests include computer vision, neural networks, point clouds, etc.

Junyi An received the B.S. degree in the school of physics from Nanjing University in 2018. He is pursuing the Ph.D. degree in Computer Science and Technology at Nanjing University under the supervision of Prof. Furao Shen. His research interests include computer vision, the structure of neural networks.

Jian Zhao received the B.S. degree from Nanjing University, Nanjing, China, the M.S. degree from Hamburg University of Technology, Hamburg, Germany, and the Ph.D. degree in electrical engineering from Swiss Federal Institute of Technology (ETH) Zurich, Switzerland. He is an Associate Professor with the School of Electronic Science and Engineering in Nanjing University. His research interests include deep neural networks, mathematical optimization, and wireless communications.

Furao Shen received the Ph.D. degree in intelligent system science from Tokyo Institute of Technology, Japan, in 2006. He is currently a professor in the Department of Computer Science and Technology. His research interests include neural network, data mining, robot intelligence.