



A novel partial point cloud registration method based on graph attention network

Yanan Song^{1,2} · Weiming Shen³ · Kunkun Peng⁴

Accepted: 22 December 2021 / Published online: 8 February 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Point cloud registration is a challenging task due to sparsity and unknown initial correspondence information. The traditional registration methods tend to converge to local optimal solutions and rely on good initial correspondence information. Deep learning-based methods show good adaptability to initial information and noises, but they cannot effectively cope with partial-to-partial registration scenes. This paper proposes a partial point cloud registration method based on graph attention network. The context information of the point cloud is obtained by a message passing mechanism. The attention features of the key registration points are extracted by an attention network. The key matching points are chosen by a key point selection module. Virtual correspondences are generated based on these key points and their features. A rigid transformation is obtained based on the virtual registration by a singular value decomposition layer. The performance of the proposed method is evaluated in three scenarios based on the ModelNet40 dataset. Experimental results show that the proposed method is robust to arbitrary initial positions and noises. It obtains higher registration accuracy than traditional methods while maintaining low network complexity.

Keywords Point cloud registration · Partial correspondence · Graph attention network · Deep learning

1 Introduction

Point cloud [1] is a very important data format that represents the geometric shape of an object. It is widely used in object segmentation [2, 3] and point cloud registration [4, 5]. Point cloud registration is a key technology in applications such as robotics and scene construction [6–8]. It aims to estimate the rigid transformation between point clouds. The key task is to determine the correspondence between points. However, point cloud is sparse and unstructured, which brings great challenges to accurately finding the correspondence. The

point cloud registration task becomes more difficult when the point cloud is partially visible [9] and low overlap [10].

Iterative closest point (ICP) [11] is a traditional point cloud registration method that is widely used to fine-tune the registration result. It iteratively explores the correspondence and estimates the rigid transformation. Unfortunately, it is easy for the ICP method to fall into a local optimum during the iterative update process. The global optimal solution is explored according to a branch-and-bound scheme in the globally optimal registration method Go-ICP [12]. But it consumes a lot of computing resources and reduces the registration efficiency. The global registration speed is further improved by the Fast Global Registration (FGR) method [13]. But FGR cannot effectively deal with noises and outliers in the point cloud.

Recently, more attention has been paid to point cloud registration methods based on deep learning. PointNetLK [14] is an earlier registration method based on deep learning. PointNet framework [15] is used to extract the point cloud features. The feature distance between point clouds is minimized by Lucas & Kanade (LK) algorithm. Deep Closest Point (DCP) [16] builds a soft correspondence between point clouds based on rigid-invariant features extracted by a transformer network [17]. Singular value decomposition (SVD)

✉ Weiming Shen
shenwm@hust.edu.cn

¹ College of Computer Science and Technology, Zhejiang University, Hangzhou, China
² Institute of Computing Innovation, Zhejiang University, Hangzhou, China
³ State Key Lab. of Digital Manufacturing Equipment & Technology, Huazhong University of Science & Technology, Wuhan, China
⁴ School of Management, Wuhan University of Science and Technology, Wuhan, China

is used by DCP to estimate the rigid transformation based on the soft correspondence. These learning-based methods are insensitive to noises and initial registration positions. Their registration performance exceeds that of traditional point cloud registration methods. But their deep learning networks are designed to handle one-to-one correspondence scenes, which cannot handle partial-to-partial registration scenes. PRNet [18] selects key matching points based on a key point detection network layer. It performs partial point cloud registration task based on these selected key points. VCRNet [19] generates virtual corresponding points by a soft pointer based method. The best corresponding points are selected according to the feature distance between partial point clouds. Although these methods can be used in partial registration scenes, the registration performance is still achieved by the iterative registration process. The network structure becomes complicated, and it requires more computing resources.

This paper proposes a one-shot registration method for partial point cloud registration scenes. The features related to the registration task are extracted by a graph attention network. The graph attention network is designed by introducing a message passing mechanism into an attention module. A key point selection module is proposed to select key matching points. Virtual corresponding points are generated based on these selected key points and their features. The SVD is used to solve the rigid transformation according to the virtual matching points.

In the rest of this paper, Sect. 2 summarizes previous research efforts related to the point cloud registration. Section 3 introduces the proposed partial registration method in detail. Section 4 provides the experimental studies based on ModelNet40 dataset [20]. Section 5 gives the conclusions and comments for future work.

2 Related work

The point cloud registration methods are usually composed of optimization-based methods and deep learning-based methods.

ICP [11] is a well-known optimization-based registration method, which alternately finds nearest corresponding points and calculates a rigid transformation in the current corresponding state. However, ICP only performs well in fine matching processes, and it cannot handle scenes with poor initial positions. ICP-based variants [21, 22] were proposed to improve the ability to deal with noises and sparsity. But these methods still rely on accurate initial registration positions. Moreover, it is easy for these methods to converge to local optima. Some methods were proposed to explore the global optimal solution of the point cloud registration. Go-ICP [12] is an earlier global optimization algorithm for point

cloud registration. The method integrates the ICP method into a branch-and-bound scheme, and it searches matching results from the entire 3D space. Riemannian optimization [23] and convex relaxation [24] were also proposed to identify the global optimal solution in point cloud registration task. But they consume more computing resources than ICP, which leads to a low solution efficiency. Yang et al. [25] proposed a fast and certifiable point cloud registration method, which can handle scenarios with a large number of outlier correspondences. Fast Global Registration (FGR) [13] discards the iterative sampling and local refinement, which further accelerates the registration process. However, FGR needs to calculate the distance between points during the process of constructing the correspondences. Its ability is limited in dealing with noises and outliers.

Recently, deep learning-based methods are widely used to extract point cloud features. These deep learning methods learn key features from a large number of point clouds [15, 26]. PCNN [27] designs an extension operator and a restriction operator, and it applies the convolutional neural network directly to the point clouds. DensePoint [28] extends the CNN framework to irregular point configuration, and it learns densely contextual representation of point clouds. Edge-Conv [29] extracts topological features of point cloud by integrating local neighborhood points.

The methods based on deep learning are also used to handle point cloud registration tasks, and they show better registration performance than traditional methods. PointNetLK [14] provides a new path for solving point cloud registration based on deep learning. It treats PointNet as a learnable image function and integrates the classical image alignment algorithm LK into the PointNet. PCRNet [30] extracts the shape features of two registered point clouds by using the PointNet. A Siamese architecture is used to directly output rigid transformation based on these extracted features. DCP [16] maps the point clouds to permutation invariant features by the PointNet and transformer network. A pointer network [31] is used to predict the soft correspondences between two point clouds. 3DRegNet [32] consists of classification block and regression block. It classifies the point correspondences into inliers/outliers, which significantly improves the registration efficiency. Although these methods show strong robustness to complex environments such as noises and initial positions, they are designed based on the assumption of one-to-one correspondence registration. It is difficult for them to handle partial point cloud registration tasks.

PRNet [18] was proposed to use a key point detection module to identify key matching points. The module only detects the key points shared by two point clouds, which removes interference from those non-corresponding points. Li et al. [33] designed an iterative distance-aware similarity matrix convolution module that integrates feature information and

spatial information into the point registration process. The proposed module attempts to match points based on the Euclidean offset and entire geometric features. Yew and Lee [34] developed RPM-Net that builds soft assignments between two point clouds by using an annealing layer and Sinkhorn layer from RPM [35]. The sensitivity to the initialization is reduced by using the feature distance instead of spatial distance. Fu et al. [36] also handled the partial registration through the construction of soft correspondences. The proposed network extracts the deep features of each point based on a graph matching strategy, which reduces the sensitivity to outliers. Wu et al. [37] proposed a feature interactive representation learning network, which designs a multi-level feature interaction mechanism. However, these deep learning-based methods still rely on iterative strategies to improve the registration accuracy. They require significant computing resources and have low registration efficiency.

Different from the methods mentioned above, this paper introduces an effective key point selection module to deal with the partial registration problem. A graph attention network is proposed to extract the attention registration features. These features are insensitive to noises and initial positions. In addition, the proposed method does not rely on the iterative optimization strategy, which significantly improves the registration efficiency.

3 Proposed method for partial point cloud registration

Given two point clouds $\mathbf{X} = \{x_j \in \mathbb{R}^3 \mid j = 1, \dots, J\}$ and $\mathbf{Y} = \{y_k \in \mathbb{R}^3 \mid k = 1, \dots, K\}$. x_j and y_k represent the coordinates of the points. J and K are the numbers of points in \mathbf{X} and \mathbf{Y} . The main task is to find a rigid transformation $[\mathbf{R}, \mathbf{t}]$ that aligns the two point clouds. $\mathbf{R} \in SO(3)$ is a rotation matrix, and $\mathbf{t} \in \mathbb{R}^3$ is a translation matrix.

The points in the two point clouds may not correspond one-to-one because of the sparsity and perspective variations. This work considers the point cloud registration problem with only partial correspondences. It means that the point x_j in \mathbf{X} does not necessarily find the exact corresponding point y_k in \mathbf{Y} .

Point cloud is a collection that contains many points. The structural information of the point cloud is essential for the point cloud registration task. Moreover, key points play an important role in the registration task. The message passing mechanism from Graph Neural Network [38] can effectively capture the structural information in a graph. The attention mechanism from Transformer [17] can improve the attention to key information. Therefore, this paper proposes an effective partial point cloud registration method based on graph attention network. The designed registration network integrates the message passing mechanism into the attention

architecture. The overall framework of the proposed registration network is shown in Fig. 1. It mainly contains a feature extraction module, a key point selection module, and a virtual corresponding point generation module. Each module is described in detail in the following subsections.

The features \mathbf{F}_X and \mathbf{F}_Y are first extracted by a PointNet style network from the target point cloud \mathbf{X} and the source point cloud \mathbf{Y} . The graph attention network (GAttention) is then used to extract the attention features \mathbf{F}_{AX} and \mathbf{F}_{AY} based on the features \mathbf{F}_X and \mathbf{F}_Y . The key points and their corresponding features are obtained through the key point selection module. Finally, the virtual registration points are obtained by the virtual corresponding point generation module according to the selected key points and their features. The SVD method is used to solve the rigid transformation $[\mathbf{R}, \mathbf{t}]$ between the two point clouds.

3.1 Feature extraction

The feature extraction network consists of a multilayer perceptron (MLP) layer and a graph attention network. The structure of the multilayer perceptron is shown in Fig. 1. The graph attention network (GAttention) is given in Fig. 2. It is mainly composed of a multi-head attention network and a multilayer perceptron layer. Residual connections are also added to these two network layers. It is worth noting that the designed graph neural network is a modular structure. This module can be stacked multiple times to increase the network depth.

The multi-head attention layer is the core component of the graph neural network as shown in Fig. 3. The network takes as input the point cloud features $\mathbf{F}_X/\mathbf{F}_Y$ and their coordinates X/Y . The spatial structure information is obtained by the message passing layer. The attention parameters Q and V are generated by the multilayer perceptron layer. The parameter K is returned by adding the structure information to the parameter Q , which enhances the ability of the attention mechanism to perceive the point cloud structure information. The parameters Q , K and V are input to the attention layer followed by a multilayer perceptron. The attention features $\mathbf{F}'_{AX}/\mathbf{F}'_{AY}$ are obtained by the multilayer perceptron layer.

The attention layer is a self-attention network. The attention is expressed as follows [17].

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

where d_k is the dimension of the parameter K .

The message passing layer is a key operation in the graph neural networks. The operation is also called neighborhood aggregation, and it is mainly responsible for message passing between nodes in a graph. The message passing is implemented through a hierarchical structure. The operation

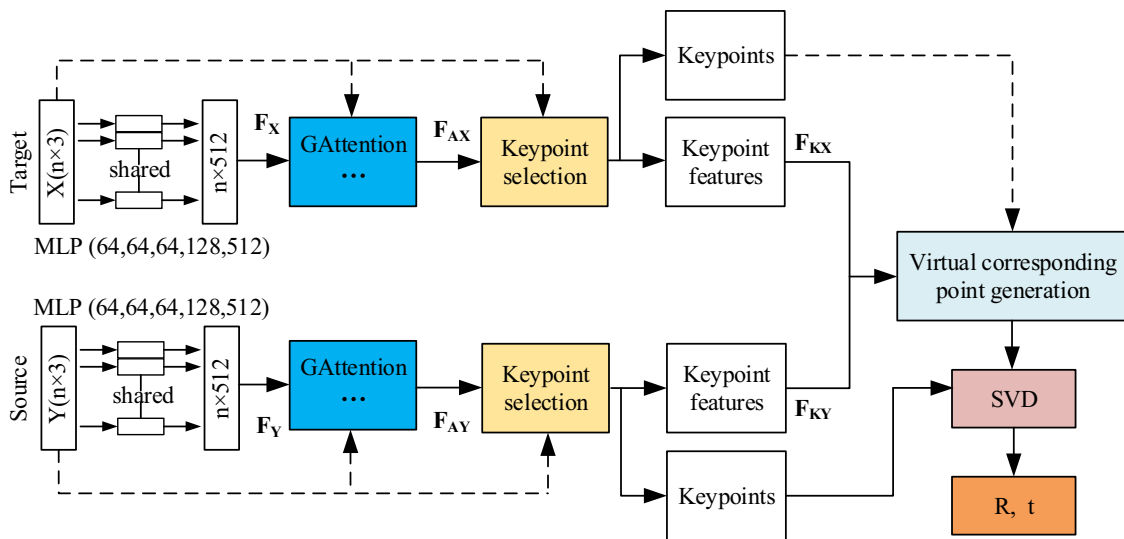


Fig. 1 The overall framework of the proposed partial registration method

Fig. 2 The graph attention network

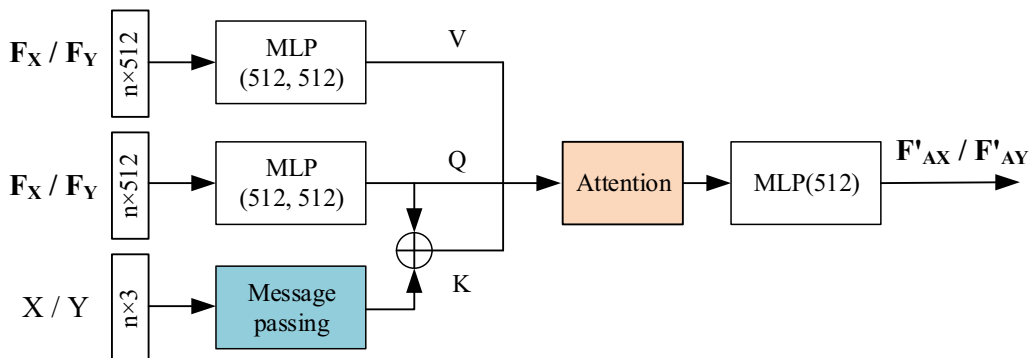
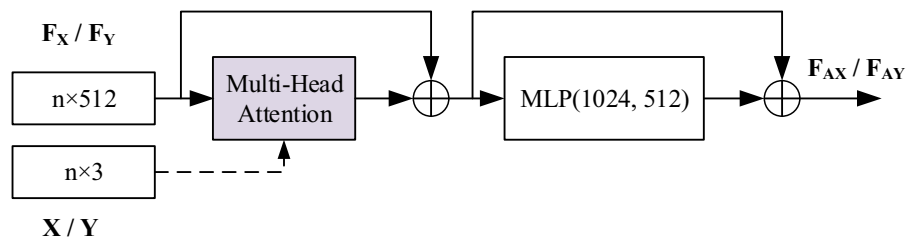


Fig. 3 The multi-head attention network

aggregates the state of the neighborhood nodes and iteratively updates the hidden state of a node. The structure of the message passing network is shown in Fig. 4. The network takes as input the point cloud coordinates. It contains a neighborhood aggregation layer, a multilayer perceptron layer and a maxpooling layer. The neighborhood aggregation layer can be used iteratively to capture multiscale spatial structure information. The parameters k_1 and k_2 represent the number of neighborhood points aggregated in each iteration.

To reduce the computational complexity, the neighborhood aggregation layer searches for neighborhood points in the coordinate space in each iteration, which is represented by a dashed line.

The neighborhood aggregation layer builds a message passing process for each point in the point cloud. It first searches the k neighborhood points of each point by using the ball query method. These k points are then converted to a local coordinate system centered on the searched point.

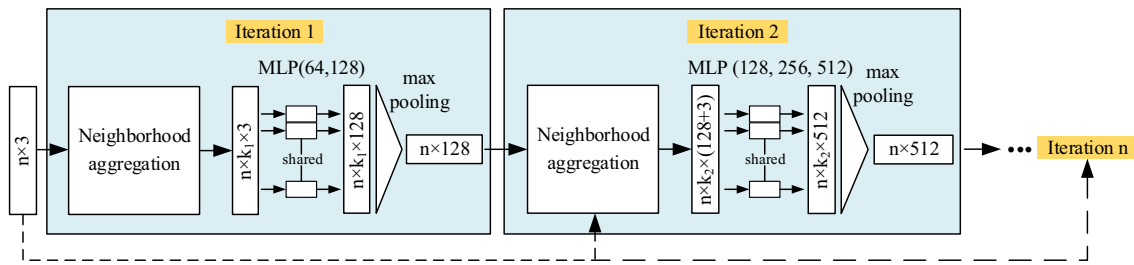


Fig. 4 The message passing network

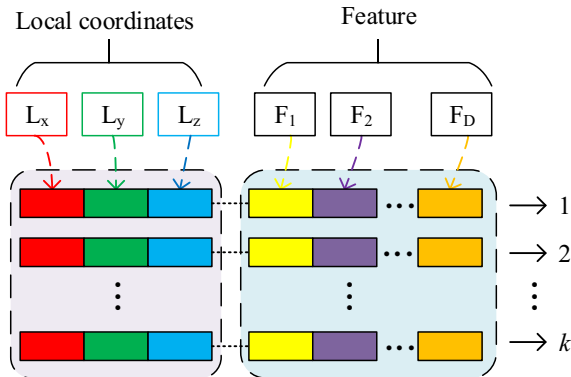


Fig. 5 The structure of the aggregation feature

Finally, the aggregation features are obtained by combining the features of the searched point and the local coordinates of these k points. Assuming that the feature dimension of a searched point is $1 \times D$, the aggregation feature of the searched point is shown in Fig. 5. $L_x, L_y,$ and L_z represent the local coordinates of the k neighborhood points. F_D represents the D -dimensional element of the feature for the searched point.

3.2 Key point selection

Only some points are matched in partial point cloud registration scenario. The irrelevant points are removed by selecting the key matching points. The key points are selected based on the mean value of the point cloud features. The key point selection module takes as input the attention features F_{AX} / F_{AY} . These attention features are used to calculate the feature mean of each point. The feature mean value represents the importance of each point to the registration task.

The key point selection module selects the key matching points based on the feature mean of each point. Using X_k and Y_k to represent the k key points of the target point cloud \mathbf{X} and the source point cloud \mathbf{Y} .

$$\begin{cases} X_k = X(\text{topk}(\text{mean}(F_{Ax_1}), \dots, \text{mean}(F_{Ax_i}), \dots, \text{mean}(F_{Ax_n}))), x_i \in X \\ Y_k = Y(\text{topk}(\text{mean}(F_{Ay_1}), \dots, \text{mean}(F_{Ay_i}), \dots, \text{mean}(F_{Ay_n}))), y_i \in Y \end{cases} \quad (2)$$

where $\text{topk}(\cdot)$ represents the index of the top k largest elements, $\text{mean}(\cdot)$ represents the mean value of the feature, F_{Ax_i} and F_{Ay_i} are the attention features of points x_i and y_i , respectively.

3.3 Virtual corresponding point generation

Since the point clouds \mathbf{X} and \mathbf{Y} are not one-to-one correspondence, the selected key points cannot fully represent the matching relationship between the two point clouds. The virtual corresponding points are used to represent the virtual mapping relationship between point clouds. The virtual matching relationship is expressed by the selected key point features. The virtual mapping from each point x_i into the elements of Y_k is expressed as follows [16].

$$m(x_i, Y_k) = \text{softmax}(F_{kY}, F_{kxi}^T) \quad (3)$$

where F_{kX} and F_{kY} represent the attention features corresponding to the selected key points X_k and Y_k , x_i is the element from X_k , F_{kxi} denotes the i -th row of the matrix F_{kX} . It should be noted that each row in F_{kX} represents the attention feature of a point.

The virtual matching point y_i corresponding to the element x_i is calculated based on the following representation.

$$y_i = Y_k^T m(x_i, Y_k) \quad (4)$$

Finally, the rigid transformation $[\mathbf{R}, \mathbf{t}]$ is estimated by the SVD method based on the pairing $x_i \rightarrow y_i$ over all i .

3.4 Loss function

The proposed registration network employs a direct loss strategy to train the network parameters. The loss function is constructed based on the deviation between the ground truth and the predicted rigid transformation.

$$L = \|\mathbf{R}^T \mathbf{R}^g - \mathbf{I}\|^2 + \|\mathbf{t} - \mathbf{t}^g\|^2 \quad (5)$$

where \mathbf{R} and \mathbf{R}^g are the predicted rotation transformation and the ground truth, \mathbf{t} and \mathbf{t}^g represent the translation transformation. \mathbf{I} stands for the identity matrix.

4 Experimental studies and discussions

The performance of the proposed method is evaluated based on the ModelNet40 dataset [20]. The dataset contains 12,311 CAD models with 40 object categories. Each object consists of 1,024 points that are randomly sampled in the CAD model. The rigid transformation is randomly generated along the x , y , z coordinate axis. The rotation transformation is randomly generated in the range of $[0, 45^\circ]$. The translation transformation is generated in $[-0.5, 0.5]$ along each axis. The partial scans of target point cloud \mathbf{X} are randomly sampled from each object point cloud. The source point cloud \mathbf{Y} is generated by applying the random transformation on the target point cloud \mathbf{X} . Only 768 points are randomly retained in point clouds \mathbf{X} and \mathbf{Y} .

The mean absolute error (MAE), mean squared error (MSE), and root-mean-squared error (RMSE) are measured to evaluate the network performance. The symbols R and t are used to represent the rotation indicator and the translation indicator. These indicators tend to zero when the point clouds are perfectly registered. It is worth noting that the angular measurement is in units of degrees.

The proposed method is implemented based on the Pytorch framework. All experiments use only one Nvidia TITAN V GPU on Red Hat 4.8.5–28.

4.1 Experimental setup

In this section, the performance of the proposed method is evaluated in three partial registration scenarios.

The registration ability for unseen point clouds is evaluated in point clouds with unseen objects. In this scenario, the ModelNet40 dataset is split into a training set with 9,843 models and a test set with 2,468 models. The dataset is the original point cloud that is used directly without any processing.

The robustness to noises is evaluated in point clouds with noises. The scenario is obtained by adding Gaussian noises to the point clouds with unseen objects. The Gaussian noises are generated by sampling from $N(0, 0.01)$ that is clipped to $[-0.05, 0.05]$. The separation of the ModelNet40 dataset is similar to the unseen object scenario.

The generalization performance is evaluated in point clouds with unseen categories. The training set and test set are split from the ModelNet40 dataset based on categories. The training set contains the first 20 categories and the test set contains the remaining 20 categories.

In the experiments, the number of the selected key points in the key point selection module is set to 512. The neighborhood aggregation layer is used iteratively once in the message passing network. The radius in the ball query strategy is 0.3, and the number of the neighborhood points is 8. The designed graph neural network is stacked twice. The multi-head attention network employs 8 heads. The initial learning rate is 0.0001. The learning rate will be multiplied by 0.1 when the minimum test loss does not drop for 50 consecutive epochs. The batch size is set to 16 for unseen objects and noise scenarios, and it is 8 for unseen categories scenario. The maximum epoch is 500. The network uses the ADAM optimizer.

4.2 Experiments on point clouds with unseen objects

In these experiments, all categories are observed during training and testing phases, but unseen objects appear at the test time. The corresponding experimental results are given in Table 1.

As it can be seen from Table 1, the classical ICP method performs the worst among the listed methods. ICP is invalid for partial point cloud registration tasks. Although the ICP variants have achieved important performance improvements, their registration performances are still lower than the learning-based method. The learning-based method can directly learn high-level point cloud features from a large number of point clouds. Compared with traditional manual features, the learned features can more effectively express the shape features of the partial point clouds. The learned features help improve the performance of the partial point cloud registration. It is worth noting that although PointNetLK is also a registration method based on deep learning, its performance is lower than Go-ICP and FGR. PointNetLK is designed based on PointNet whose ability is weak in extracting the partial point cloud features. DCP-v2 is a very effective partial point cloud registration method based on deep learning. It is designed based on classical Transformer structure that can enhance attention to the point cloud features. But its performance is still lower than that of the proposed method. The results show that the proposed graph attention network is very important for extracting the key registration features of point cloud pairs. The key registration feature is an important factor to improve the performance of the partial point cloud registration task.

Although the proposed method achieves better rotation transformation performance, it performs weaker on the translation transformation indicator. The proposed method pays more attention to the rotation transformation which is more important for the partial registration task. Nevertheless, the proposed method still provides a small translation error. The results show that the proposed graph attention network can extract the shape features of objects. It can effectively esti-

Table 1 The registration results on point clouds with unseen objects

Method	MSE(R)	RMSE(R)	MAE(R)	MSE(t)	RMSE(t)	MAE(t)
ICP [11]	1134.552	33.683	25.045	0.0856	0.293	0.250
Go-ICP [12]	195.985	13.999	3.165	0.0011	0.033	0.012
FGR [13]	126.288	11.238	2.832	0.0009	0.030	0.008
PointNetLK [14]	280.044	16.735	7.550	0.0020	0.045	0.025
DCP-v2 [16]	45.005	6.709	4.448	0.0007	0.027	0.020
The proposed method	41.109	6.412	4.513	0.0022	0.047	0.035

Bold indicates the best value among the listed methods

mate the rigid transformation between unseen objects even in incompletely corresponding scenes.

These comparison methods optimize the registration results by gradually changing the relative position between two point clouds. On the contrary, the proposed method abandons the strategy that iteratively optimizes the registration result. It directly predicts the rigid transformation between point clouds, but the registration results are significantly improved. The proposed graph attention network can make full use of the local spatial structure information of the point cloud, and it is more sensitive to the key matching point features. The proposed method is more adaptable to any initial registration positions, and it does not need to adjust the registration position step by step. In addition, the designed network consumes less computing resources due to the one-shot registration strategy. The registration efficiency is significantly improved compared to the iterative registration methods.

4.3 Experiments on point clouds with noises

It is difficult to avoid noises in the process of generating point clouds. Noises change the shape information of the original object, which destroys the corresponding relationship between point clouds. Noises will cause serious interference in the task of predicting the point cloud registration relationship.

To evaluate the performance of the proposed method in point clouds with noises, the registration experiments with Gaussian noises are completed in this section. The corresponding experimental results are provided in Table 2. The results show that the proposed method still achieves significant performance improvements, and it shows strong robustness to noises. On one hand, the proposed graph attention network combines the attention mechanism and the message passing mechanism. The attention mechanism pays more attention to the features related to the registration task and ignores irrelevant information. The message passing mechanism can capture the spatial structure information of the point cloud in a cascaded manner. Point cloud features are constructed based on the context information instead of the coordinate information of a single point. These features allow

point cloud coordinates to fluctuate within a certain range. The message passing mechanism can reduce the influence of the noises on the registration features. On the other hand, the rigid transformation is estimated based on key points and their features. These key points can alleviate the impact of the noises on the registration task. In addition, these key points are selected based on the feature mean of each point, which can significantly weaken the noise information in the key point features. Therefore, the proposed method can effectively deal with noises in the partial registration task.

Compared with the noise-free results provided in Table 1, the performances of most comparison methods are reduced in the noisy point cloud data. It shows that these methods are less robust to noises. Traditional registration methods are designed based on manual features, and they cannot adapt to complex registration scenarios. It is interesting that the registration performance of the proposed method is further improved in the noisy data. That is because the proposed method is a data-driven registration method. Noises play an important role in enriching the diversity of the registration point cloud data, and they have the effect of data enhancement. Rich data are beneficial to improve the registration performance of the proposed method. These results further show that the proposed graph attention network can fully mine point cloud registration features. The proposed learning-based registration method can effectively adapt to complex point cloud registration scenarios.

To show the robustness of the proposed method to noises from different aspects, the registration model trained on noisy data is also used to test the noise-free point cloud data in Sect. 4.2. The experimental results are provided in Table 3. The registration results from training model without noise are provided in Sect. 4.2. The results from training model with noise are obtained by applying the training model in Sect. 4.3 to the noise-free point cloud.

As can be seen from Table 3, the noisy training model significantly improves the registration performance of the proposed method on noise-free data. Noises improve the registration performance of the proposed method instead of weakening it. It shows that the proposed method has strong robustness to noises. In fact, the proposed method can make full use of the noises to enrich the diversity of training data,

Table 2 The registration results on unseen point clouds with noises

Method	MSE(R)	RMSE(R)	MAE(R)	MSE(t)	RMSE(t)	MAE(t)
ICP [11]	1229.670	35.067	25.564	0.0860	0.294	0.250
FGR [13]	764.671	27.653	13.794	0.0048	0.070	0.039
Go-ICP [12]	150.320	12.261	2.845	0.0008	0.028	0.029
PointNetLK [14]	397.575	19.939	9.076	0.0032	0.057	0.032
DCP-v2 [16]	47.378	6.883	4.534	0.0008	0.028	0.021
The proposed method	35.729	5.977	4.141	0.0021	0.046	0.035

Bold indicates the best value among the listed methods

Table 3 The registration results on noise-free point cloud for different training models

Training model	MSE(R)	RMSE(R)	MAE(R)	MSE(t)	RMSE(t)	MAE(t)
No noise	41.109	6.412	4.513	0.0022	0.047	0.035
With noise	36.052	6.004	4.145	0.0022	0.047	0.035

Bold indicates the best value among the listed methods

which is conducive to improving the registration performance.

4.4 Experiments on point clouds with unseen categories

The point cloud registration task may encounter different types of objects. It is impossible for the proposed method to encounter all categories in the training phase. The ability to register unseen categories determines the generalization performance and application scope of the proposed method.

To test the performance of the proposed method in the point clouds with unseen categories, the registration model is trained and tested on different object categories. The experimental results are given in Table 4. The proposed method still provides high performance even in the registration task with unseen categories. In fact, the effect of the object category on registration performance comes from the object shape. Different shapes may lead to different point cloud distributions. These unseen distributions will bring challenges to the point cloud registration task. However, the proposed method constructs the features of each point based on the local spatial structure of the point cloud, and it does not rely on the distribution of the overall point cloud. Moreover, the designed attention mechanism can prompt the network to capture the key registration features. This further alleviates the interference caused by the unseen categories to the registration task. Therefore, the proposed method can obtain better registration performance in the point clouds with unseen categories.

Comparing Tables 4 and 1, the registration performance of the proposed method for unseen categories is lower than the performance for unseen objects. On one hand, the ModelNet40 dataset is split based on categories, which reduces the size of the training data. The reduction in the amount of data leads to a decrease in the registration performance of the proposed method. On the other hand, the proposed network

is designed to handle partial point cloud registration tasks. Unseen point cloud shape distribution brings huge uncertainty to the partial point cloud correspondences. Although the new categories bring huge challenges to partial point cloud registration tasks, the proposed method still obtains a competitive registration results. It reflects that the generalization performance of the designed registration network is strong for new object categories and complex scenes.

4.5 Visualization of experimental results

To clearly show the registration performance of the proposed method, some partial registration results are visualized in Fig. 6. The initial point cloud state input to the network is presented in the first row. The true correspondence between the point cloud pairs is shown in the second row. The prediction result of the proposed method is displayed in the third row. The source point cloud is represented by yellow points, and the target point cloud is represented by blue points. If the points in the two point clouds are completely coincident, the position is represented by a yellow point. Therefore, most of the area is covered by yellow points in the second row.

As can be seen from the first line, there are random initial corresponding positions between the point cloud pairs. Even in the face of relatively different initial correspondences, the proposed method still gives accurate registration results. The results are highly consistent with the ground truth shown in the second line. Moreover, the proposed method is a one-shot method, and it directly predicts the registration results without any iterative optimization process. It reflects that the designed registration network is robust to any initial corresponding state. The ability is strong for the proposed method to adapt to complex registration environment.

As shown in Fig. 6, the point cloud is incomplete. The point cloud information can only express the local shape of an object. Only partial regions match each other in these

Table 4 The registration results on point clouds with unseen categories

Method	MSE(R)	RMSE(R)	MAE(R)	MSE(t)	RMSE(t)	MAE(t)
ICP [11]	1217.618	34.894	25.455	0.0860	0.293	0.251
Go-ICP [12]	157.072	12.533	2.940	0.0009	0.031	0.010
FGR [13]	98.635	9.932	1.952	0.0014	0.038	0.007
PointNetLK [14]	526.401	22.943	9.655	0.0037	0.061	0.033
DCP-v2 [16]	95.431	9.769	6.954	0.0010	0.034	0.025
The proposed method	50.970	7.139	5.158	0.0032	0.056	0.043

Bold indicates the best value among the listed methods

**Fig. 6** Qualitative registration results

point cloud pairs. The proposed method can still accurately predict the rigid transformation between point cloud pairs based only on these partially visible points. It shows that the designed graph attention network can accurately capture the key registration features of the point cloud. These key features are extracted based on the spatial matching information between the point cloud pairs instead of the overall shape information of a single object. In addition, the designed key point selection module greatly reduces the dependence on the point cloud information of the entire object. This module further improves the adaptability of the algorithm to the partial point cloud registration tasks.

The ModelNet40 dataset contains 40 categories. Five of these categories are shown in Fig. 6. There are different geometric shapes and symmetry relationships in these categories. But the proposed method performs well for object shapes with different topologies. The designed graph attention network extracts registration features based on local point pairs, and it pays more attention to task-related high-level features. These registration features weaken the adverse effects of dif-

ferent point cloud distributions on the registration results. The designed network can be used for point cloud registration tasks with different object categories. The results further demonstrate the strong generalization performance of the proposed method.

To show the performance of the key point selection module, some selected key matching points are provided in Fig. 7. The key matching points from the source point cloud are displayed in the first row. The key points from the target point cloud are shown in the second row. The yellow point represents the original point cloud, and the blue point represents the selected key matching point.

Comparing the source point cloud and the target point cloud, the location of the key matching points is consistent. Moreover, the selected key matching points are clustered together. Although the two registration point clouds are partially visible, the proposed key point selection module can still accurately identify the key matching points. It shows that the proposed method pays more attention to the common vis-

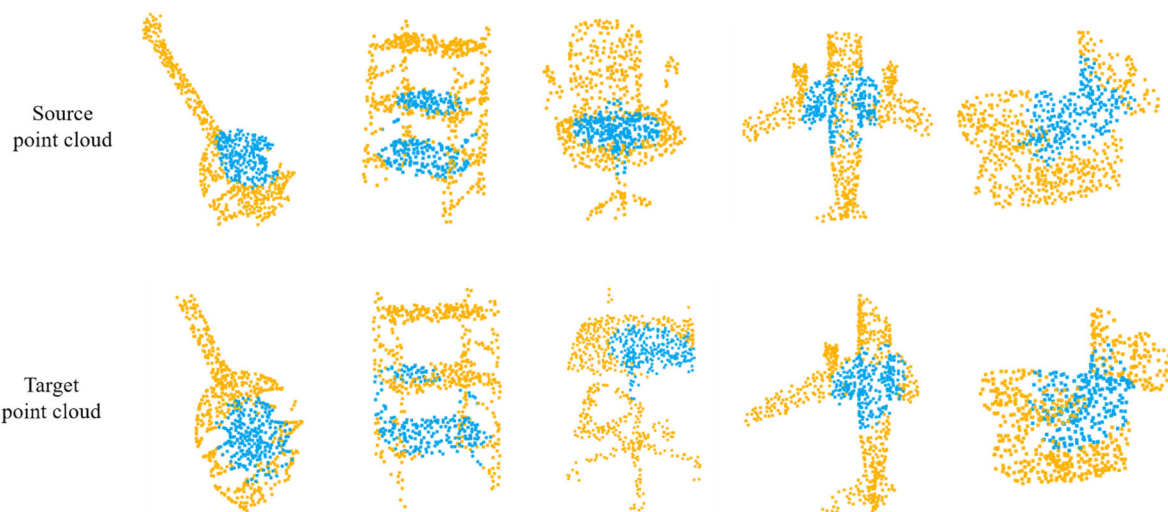


Fig. 7 Qualitative key matching points

ible area between two point clouds. The proposed method can effectively handle partial point cloud registration tasks.

To demonstrate the performance of the virtual corresponding point generation module, the virtual correspondences for some point clouds are shown in Fig. 8. The green point represents the source point cloud, and the red point represents the target point cloud. The virtual corresponding points between the two point clouds are connected by blue lines.

As shown in Fig. 8, the generated virtual corresponding points come from the same part of the object. Although the relative position between point clouds is arbitrary, the proposed method can still accurately generate the virtual corresponding positions. It further shows that the proposed method is insensitive to the initial position between point clouds. The proposed method can be adapted to the registration task with any initial positions.

5 Conclusions and future work

This paper proposes an effective partial point cloud registration method without iterative processes. The point cloud registration features are first extracted by the designed graph attention network. The key point selection module is then designed to select the key registration points and their corresponding features. Virtual matching points are constructed based on these key points and features. Finally, the rigid transformation between partially registered point clouds is estimated by the SVD method based on these virtual points. The experimental results on ModelNet40 dataset show that the proposed method obtains accurate rigid transformation. It suggests that the designed graph attention network can effectively extract key registration features even in the partial corresponding scenes. The proposed method obtains signifi-

cant performance improvements on point clouds with noises. It suggests that the designed network can effectively avoid adverse interferences and is robust to noises and outliers. The proposed method also gives a higher performance improvement on different unseen object categories. It suggests that the proposed method can adapt to a wider range of registration scenes, and its generalization ability is strong. The main contributions of this work are summarized as follows.

- An effective partial point cloud registration method is proposed based on deep learning without any iterative optimization process. The proposed method improves the accuracy of the partial registration task while reducing the network complexity.
- The graph attention network is designed by integrating the message passing mechanism into the attention module. The network enhances the attention to task-related features and makes full use of the local spatial structure information of the point cloud.
- A key point selection module is proposed to select key registration points and their corresponding features. This module judges the registration importance of each point based on the feature mean value of its neighborhood points. It can significantly improve the robustness of the proposed method to noises and outliers.

The proposed method extracts registration features from two point clouds separately, and it does not consider the interaction between the two point clouds in the feature extraction stage. This may not be conducive to the information exchange between point cloud pairs. In the future, the designed graph neural network will take point cloud pairs as input and further explore the registration relationship between point clouds. In addition, the designed key point selection module cannot be

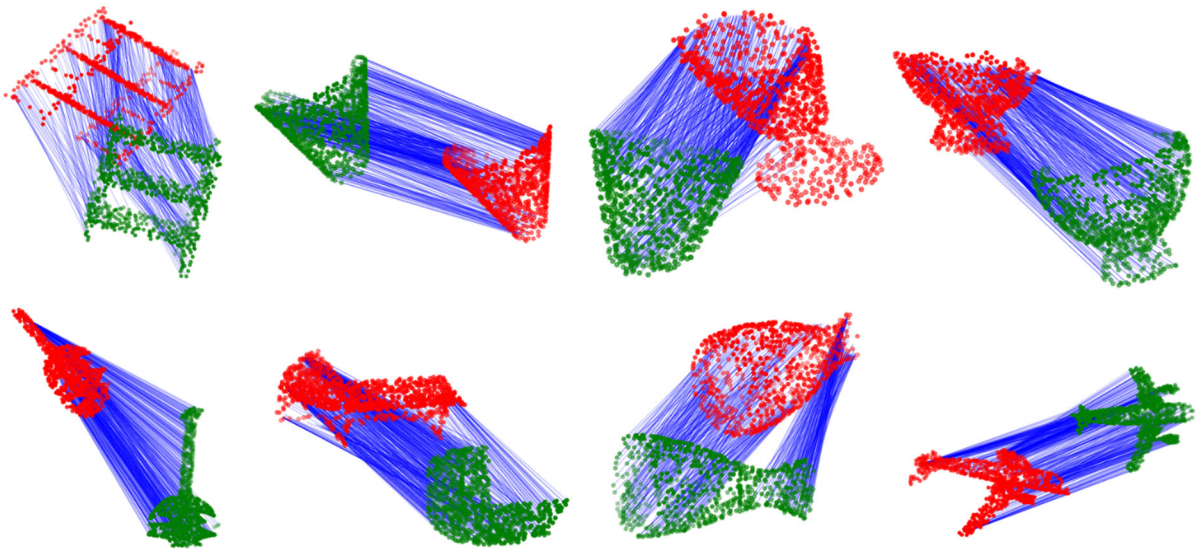


Fig. 8 Virtual correspondence between point clouds

iteratively optimized with the network. In our future work, the module will be designed as a network with optimizable parameters based on point cloud features.

Acknowledgements This work was supported by China Postdoctoral Science Foundation [Grant Number 2021M692778].

References

- Prieto, S.A., Adan, A., Quintana, B.: Preparation and enhancement of 3D laser scanner data for realistic coloured BIM models. *Vis. Comput.* **36**(1), 113–126 (2020)
- Liu, T.R., Cai, Y.Y., Zheng, J.M., Thalmann, N.M.: BEACon: a boundary embedded attentional convolution network for point cloud instance segmentation. *Vis. Comput.* (2021). <https://doi.org/10.1007/s00371-021-02112-7>
- Meng, H.-Y., Gao, L., Lai, Y.-K., Manocha, D., net Vv-: Voxel vae net with group convolutions for point cloud segmentation, In: Proceedings of IEEE/CVF International Conference on Computer Vision, pp. 8500–8508 (2019)
- Hu, L., Xiao, J., Wang, Y.: An automatic 3D registration method for rock mass point clouds based on plane detection and polygon matching. *Vis. Comput.* **36**(4), 669–681 (2020)
- Wang, C., Xu, Y.H., Wang, L., Li, C.M.: Fast structural global registration of indoor colored point cloud. *Vis. Comput.* (2021). <https://doi.org/10.1007/s00371-021-02295-z>
- Dong, K., Gao, S.S., Xin, S.Q., Zhou, Y.F.: Probability driven approach for point cloud registration of indoor scene. *Vis. Comput.* (2020). <https://doi.org/10.1007/s00371-020-01999-y>
- Gojcic, C., Zhou, J.D., Wegner, L.J., Guibas, T., Birdal: Learning multiview 3d point cloud registration, In: Proceedings of IEEE/CVF conference on computer vision and pattern recognition, pp. 1759–1769 (2020)
- Choy, C., Dong, W., Koltun, V.: Deep global registration, In: Proceedings of IEEE/CVF conference on computer vision and pattern recognition, pp. 2514–2523 (2020)
- Lee, D., Hamsici, O.C., Feng, S., Sharma, P., Gernoth, T., Deep-PRO: Deep partial point cloud registration of objects, In: Proceedings of IEEE/CVF International Conference on Computer Vision, pp. 5683–5692 (2021)
- Huang, S., Gojcic, Z., Usvyatsov, M., Wieser, A., Schindler, K., PREDATOR: Registration of 3D point clouds with low overlap, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 4267–4276 (2021)
- Besl, P.J., McKay, H.D.: A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992)
- Yang, J., Li, H., Campbell, D., Jia, Y.: Go-ICP: A globally optimal solution to 3D ICP point-set registration. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(11), 2241–2254 (2016)
- Zhou, Q.-Y., Park, J., Koltun, V.: Fast global registration, In: Proceedings of European Conference on Computer Vision, pp. 766–782 (2016)
- Aoki, Y., Goforth, H., Srivatsan, R.A., Lucey, S., Pointnetlk: Robust & efficient point cloud registration using pointnet, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 7163–7172 (2019)
- Charles, R.Q., Su, H., Mo, K., Guibas, L.J., PointNet: Deep learning on point sets for 3D classification and segmentation, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 77–85 (2017)
- Wang, Y., Solomon, J.M.: Deep closest point: learning representations for point cloud registration, In: Proceedings of IEEE International Conference on Computer Vision, pp. 3523–3532 (2019)
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need, In: Proceedings of Advances in Neural Information Processing Systems, pp. 5998–6008 (2017)
- Wang, Y., Solomon, J. PRNet: self-supervised learning for partial-to-partial registration, In: Proceedings of International Conference on Neural Information Processing Systems, pp. 8814–8826 (2019)
- Wei, H., Qiao, Z., Liu, Z., Suo, C., Yin, P., Shen, Y., Li, H., Wang, H.: End-to-End 3D Point cloud learning for registration task using virtual correspondences, In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2678–2683 (2020)
- Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., Xiao, J.: 3D ShapeNets: A deep representation for volumetric shapes, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1912–1920 (2015)

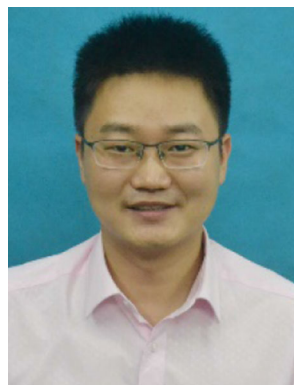
21. Rusinkiewicz, S.: A symmetric objective function for ICP. *ACM Trans. Graphics* **38**(4), 1–7 (2019)
22. Basdogan, C., Ozireli, A.C.: A new feature-based method for robust and efficient rigid-body registration of overlapping point clouds. *Vis. Comput.* **24**(7), 679–688 (2008)
23. Rosen, D.M., Carlone, L., Bandeira, A.S., Leonard, J.J.: SE-Sync: A certifiably correct algorithm for synchronization over the special euclidean group. *Int. J. Robot. Res.* **38**(2–3), 95–125 (2019)
24. Maron, H., Dym, N., Kezurer, I., Kovalsky, S., Lipman, Y.: Point registration via efficient convex relaxation. *ACM Trans. Graphics* **35**(4), 1–12 (2016)
25. Yang, H., Shi, J., Carlone, L.: Teaser: Fast and certifiable point cloud registration. *IEEE Trans. Rob.* **37**(2), 314–333 (2020)
26. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space, In: *Proceedings of Advances in Neural Information Processing Systems*, pp. 5099–5108 (2017)
27. Atzmon, M., Maron, H., Lipman, Y.: Point convolutional neural networks by extension operators. *ACM Trans. Graphics (TOG)* **37**(4), 1–12 (2018)
28. Liu, Y., Fan, B., Meng, G., Lu, J., Xiang, S., Pan, C.: Denspoint: Learning densely contextual representation for efficient point cloud processing, In: *Proceedings of IEEE/CVF International Conference on Computer Vision*, pp. 5239–5248 (2019)
29. Wang, Y., Sun, Y.B., Liu, Z.W., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph CNN for learning on point clouds. *ACM Trans. Graphics* **38**(5), 1–12 (2019)
30. Sarode, V., Li, X., Goforth, H., Aoki, Y., Srivatsan, R.A., Lucey, S., Choset, H.: PCRNNet: Point cloud registration network using PointNet encoding. *arXiv preprint arXiv:1908.0790* (2019)
31. Vinyals, O., Fortunato, M., Jaitly, N.: Pointer networks, In: *Proceedings of Advances in Neural Information Processing Systems*, pp. 2692–2700 (2015)
32. Pais, G.D., Ramalingam, S., Govindu, V.M., Nascimento, J.C., Chellappa, R., Miraldo, P.: 3dregnet: A deep neural network for 3d point registration, In: *Proceedings of IEEE/CVF conference on computer vision and pattern recognition*, pp. 7193–7203 (2020)
33. Li, J., Zhang, C., Xu, Z., Zhou, H., Zhang, C.: Iterative distance-aware similarity matrix convolution with mutual-supervised point elimination for efficient point cloud registration, In: *Proceedings of European Conference on Computer Vision*, pp. 378–394 (2020)
34. Yew, Z.J., Lee, G.H.: Rpm-net: Robust point matching using learned features, In: *Proceedings of IEEE conference on computer vision and pattern recognition*, pp. 11824–11833 (2020)
35. Gold, S., Rangarajan, A., Lu, C.-P., Pappu, S., Mjolsness, E.: New algorithms for 2D and 3D point matching: pose estimation and correspondence. *Pattern Recogn.* **31**(8), 1019–1031 (1998)
36. Fu, K., Liu, S., Luo, X., Wang, M.: Robust point cloud registration framework based on deep graph matching, In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8893–8902 (2021)
37. Wu, B., Ma, J., Chen, G., An, P.: Feature interactive representation for point cloud registration, In: *Proceedings of IEEE/CVF International Conference on Computer Vision*, pp. 5530–5539 (2021).
38. Ying, C., Cai, T., Luo, S., Zheng, S., Ke, G., He, D., Shen, Y., Y, T.: Do transformers really perform bad for graph representation?, *arXiv preprint arXiv:2106.05234* (2021)



Yanan Song received the M.S. degree in mechanical design from University of Shanghai for Science and Technology, Shanghai, China, in 2016, and received the Ph.D. degree in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2020. He is currently a Postdoc in the College of Computer Science and Technology, Zhejiang University, Hangzhou, China. His research interests include deep learning and computer vision.



Weiming Shen (M'98 SM'02 F'12) received his Bachelor and Master's degrees from Northern (Beijing) Jiaotong University, China (in 1983 and 1986 respectively) and his PhD degree from the University of Technology of Compiègne, France (in 1996). He is currently a Professor at Huazhong University of Science and Technology, China, and an Adjunct Professor at the University of Western Ontario, Canada. His research interest includes intelligent software agents, machine vision, IoT, Big Data, and their applications in industry. He is a Fellow of IEEE, a Fellow of Canadian Academy of Engineering, a Fellow of the Engineering Institute of Canada, and a licensed Professional Engineer in Ontario, Canada.



Kunkun Peng received the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2016. He is currently an Associate Professor with the School of Management, Wuhan University of Science and Technology, Wuhan, China. His current research interests include intelligent optimization and scheduling, machine learning.