Registration of Point Clouds: A Survey

Dongfang Xie
Academy for Eng. & Tech. Fudan University Shanghai, China
dfxie20@fudan.edu.cn

Wei Zhu
Academy for Eng. & Tech. Fudan University Shanghai, China
Wzhu20@fudan.edu.cn

Fengxiang Rong
Academy for Eng. & Tech. Fudan University Shanghai, China
fxrong20@fudan.edu.cn

Xu Xia
Mobile Communication Research Dept. China Telecom Research Institute Beijing, China
xiaxu@chinatelecom.cn

Huiliang Shang*
Academy for Eng. & Tech. Fudan University Shanghai, China
shanghl@fudan.edu.cn

2021 International Conference on Networking Systems of AI (INSAI) | 978-1-6654-0859-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/INSAI54028.2021.00034

***Abstract*—The registration of point cloud is essentially to obtain a relatively accurate coordinate transformation matrix through operation, and unify the point cloud data from multiview into the particular coordinate system through rigid transformations such as rotation and translation. Generally speaking, the registration is to discover the position conversion matrix of the overlap between clouds, which have an important effect in the domain of robot and computer vision. The purpose of this article is to comprehensively summarize the current progress of point cloud registration from two dimensions: algorithm optimization methods and deep learning methods. This paper first points out the possible application fields and development direction of point cloud registration in the future, then makes a comparison between different algorithms, and finally makes a proper analysis of the advantages and disadvantages of each algorithm.**

Keywords—registration, point clouds, survey, traditional methods, deep learning

I. INTRODUCTION

The point clouds from multiple perspectives are registered as a unified perspective to form a whole scene point cloud, which is the registration. The essence is to transform the data points measured in different coordinate systems to obtain the overall data model. The key to the problem is how to obtain the corresponding rotation and translation transformation in order that the distance of the three-dimensional data measured under the two viewing angles after coordinate transformation is the smallest. Since the concept of quaternions was proposed in 1987, quaternions have become one of the most common representations of rotation.

Previously, point cloud registration was only a committed segment in the fields related to robot [1] and computer vision [2]. Now its application scope has been greatly increased, especially in the application fields such as visualization and data processing. For example, 3D modeling [3], 3D mapping based on slam [4], thermal 3D mapping in dynamic scenes [5], acquisition of 3D objects [8], 3D scene reconstruction [8]–[10], 3D medical images[12], and other applications.

Due to the brilliance of point cloud registration in many fields, related research and articles are increasing year by year. There are lots of valuable ideas in the papers of registration in computer vision and related fields[2], [14]. Joaquim Salvi *et al.* [13] investigate the different techniques used for corresponding points and different view image registration before 2007 and supply a coarse to fine system framework. Gary K.L. Tam *et al.* [14] solved the registration problem of non-rigid objects by using the method of data fitting. François Pomerleau et al. [13] study different ICP based algorithms in the past two decades and their applications in the field of mobile robots.

However, these surveys do not start from the perspective of traditional methods and deep learning methods, so this paper decided to focus on the popular point cloud registration algorithms in recent years. It is not only a summary of past studies, but also to promote the progress of point cloud registration in academia and even industry. This paper considers point cloud registration from the following two dimensions: algorithm optimization methods (called traditional method in this paper) and deep learning methods from the point of view of algorithm, in which the traditional methods are subdivided into methods based on mathematical optimization, feature based and ICP[16] based. Classification of registration algorithms as shown in Fig. 1.

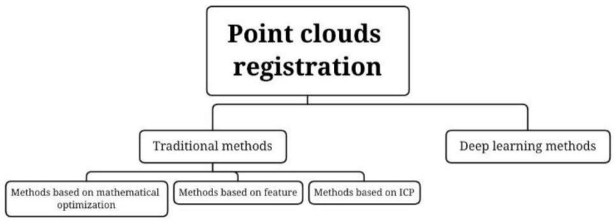


Fig. 1. Classification of registration algorithms

The following parts is arranged as shown below: Section Ⅱ introduces the traditional methods of registration in three sections, then analyzes the characteristics and problems of these algorithms, and tries to analyze the advantages and disadvantages of some commonly used algorithms. Section Ⅲ introduces the relevant algorithms and characteristics of deep learning methods. Section Ⅳ introduces several evaluation indexes to measure the registration accuracy of point clouds. Section Ⅴ summarizes the full text, reviews the past and looks forward to the future.

II. Traditional methods

A. Methods based on mathematical optimization

These methods mainly use mathematical tools to optimize the transformation matrix or propose new algorithms based on greedy algorithm to achieve better effect or/and lower time complexity. For example, the Random Sample Consensus (RANSAC) [17]， estimates the correlation function of the algorithm structure from the data sets containing "external points" by using an iterative method. The principle of RANSAC is to discover the most suitable function matrix H to maximize the quantity of data points in the data set that meet the matrix, such as

$$s\begin{bmatrix}x'\\y'\\1\end{bmatrix}=\begin{bmatrix}h_{11}&h_{12}&h_{13}\\h_{21}&h_{22}&h_{23}\\h_{31}&h_{32}&h_{33}\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix} \quad (1)$$

Where (x, y) represents the corner position of the target figure, (x', y') is the corner position of the scene figure, and s is the scale parameter. An example of the algorithm is shown in Fig. 2.

The advantage of RANSAC is that it has good robustness in estimating structure function. For instance, it can evaluate effective variables from a data set including numerous external points. However, the disadvantages of RANSAC are also obvious. First, function does not restrict the number of iterations for calculating parameters, If the algorithm limits the upper limit of iteration times, the final result may not be the optimal solution, or so far as to get the poor outcome; Second, The feasibility of RANSAC obtaining a believable architecture is directly proportional to the number of iterations. Moreover, RANSAC must also set relevant parameters such as threshold.

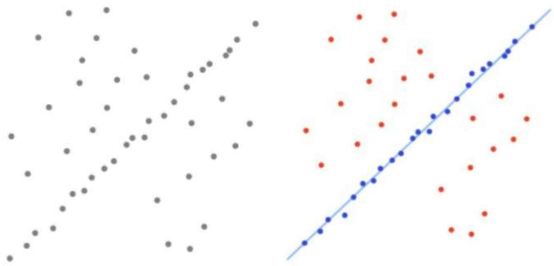


Fig. 2. The left picture: A dataset containing many external points; The right picture: The line found by RANSAC (the external point does not affect the result)

Florian Bernard *et al.* [18] studied the convex relaxation of quadratic optimization problems on permutation matrices by mathematical methods. Similarly, nadav Dym *et al.* [19] and Itay kezurer *et al.* [20] each proposed a new convex relaxation method to build the correspondence between points. Wolfgang Forstner *et al.* [21] raised and analyzed a method based on plane correspondence by dividing the point clouds into different planes and matching them in multiple point clouds by considering the uncertainty of a given plane. Someone usually solve the point cloud alignment problem with nonlinear rotation constraints by calculating Lagrangian Duality.

Most of the existing registration algorithms obtain the corresponding relationship by extracting invariant descriptors, but they are often troubled by noise. Somebody suggested that a point cloud registration algorithm without correspondence using the randomization method. The robustness of the algorithm is proved by experiments. Somebody else also raised a robust registration method for two groups of 3D points even if there are a large number of outliers. A robust matching algorithm considering metrics is raised by J. Minguezjie *et al.* [22]. This algorithm considers both the translation and rotation of the object when estimating the plane displacement.

The following introduces a very commonly used algorithm, Normal Distribution Transformation (NDT) algorithm [23], which uses a unique standard optimization algorithm to achieve the best matching between different point clouds, The reason for the low time complexity of the algorithm is that it does not calculate the characteristics and matching of corresponding points between point clouds. Therefore, we can consider using the optimization method to make this match best.

Another famous registration algorithm based on mathematical optimization is 4-Points Congruent Sets (4PCS) [24]. 4PCS uses the RANSAC algorithm framework, which reduces the spatial matching operation by constructing congruent four-point pairs with matching, and then speeds up the registration process. Specifically, it means to construct coplanar 4-point sets in the input points P and Q with arbitrary posture, use affine invariance constraints to manage the corresponding points in this particular set, use Largest Common Pointset method to discover the maximum overlap four-point pairs after registration, get a satisfactory answer, and complete the point cloud rough matching.

The spatial topology and matching schematic diagram of the 4PCS is shown in Fig. 3. For P and Q at any initial position, first make the P as the reference, and then find a long baseline in P that meets the requirements. As shown on the left part of Fig. 3, Then, the topological information of coplanar 4-point set $\mathrm{B} = \{a, b, c, d\}$ is extracted, and two scale factors r_1, r_2 between 4-point sets are calculated according to (2). The two scale factors have affine invariance when the point cloud rotates and translates.

$$\begin{cases} r_1 = \| a - e \| \| a - b \| \\ r_2 = \| c - e \| \| c - d \| \end{cases} \quad (2)$$

Calculate the four possible intersection positions of $Q\ (q_1, q_2 \in Q)$ according to (3), and then calculate the coordinates of the intersection points of all the medium and long baselines of Q, compare the intersection coordinates to determine the matching set, $e_i \approx e_j$ means to find the corresponding uniform four-points, and i, j represents the pair of wide baseline points in the ith and jth Q, respectively.

$$\begin{cases} e_1 = q_1 + r_1(q_2 - q_1) \\ e_2 = q_1 + r_2(q_2 - q_1) \end{cases} \quad (3)$$

The congruent four-point pair of $\mathrm{B} = \{a, b, c, d\}$ in Fig. 3 is $\mathrm{C} = \{q_1, q_3, q_4, q_5\}$, finding the coplanar 4-point sets of all P in the point cloud is denoted as $E = \{B_1, B_2, \dots, B_m\}$, here m represents all the 4-point sets in P. Repeat the above to gain $\mathrm{D} = \{C_1, C_2, \dots, C_n\}$, here n is the same as m.

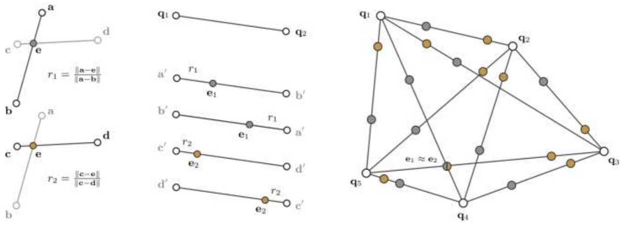


Fig. 3. Spatial topology and matching schematic diagram of 4PCS[24]

Most of the existing registration algorithms obtain the corresponding relationship by extracting invariant descriptors, but the 4PCS algorithm has a better performance in the face of registration tasks in complex scenes, so over the years, multifarious improved algorithms have been raised. The typical Super4PCS [25] algorithm adopts a unique index method to cut down the time complexity of 4PCS algorithm from $O(n^2)$ to $O(n)$, where n is the quantity of points. The Super4PCS is robust even in the face of large-scale scenic spot cloud registration. The k-4PCS [26] algorithm replaces the random sampling points of the 4PCS algorithm by detecting the key points, which solves the problem that the point cloud needs to be greatly downsampled in the 4PCS algorithm, and then realizes the point cloud matching with higher precision. Semantic-keypoint 4PCS [27] algorithm mainly solves the registration problem of urban building scene, first extracting the semantic key points of buildings at different levels, and then using semantic points to replace the random sampling points of the original algorithm for registration. 2PNS [28] algorithm is an improved algorithm of Super4PCS, in this algorithm, the matching rule is defined by the topological relationship from any two points in space and the normal vector, which makes the algorithm particularly ideal in less noise environment. Experiments show that the algorithm is 5.2 times faster in noise-free scenes, and the algorithm also has a good effect on scene matching with small overlap (Even if the overlap is only 5%, the registration can be successful). MSSF-4PCS [29] adds normal vectors to constrain the matching links of coplanar 4-point sets, and speeds up the running speed of the algorithm by reducing the matching number of 4-point sets. It further optimizes the matching results of point pairs and optimizes the precision of registration by calculating and matching the point features in the neighborhood of R radius of congruent 4-point sets. The main contribution of V4PCS [30] is to consider volume consistency and expand coplanar to non-coplanar, which cuts down the computational complexity and improves the computational efficiency. The Generalized4PCS [31] algorithm improved and extended the key core of the original algorithm, and no longer limited the points to a single plane. This improvement greatly optimized the efficiency of the algorithm. SuperGeneralized4PCS [32] algorithm is a combination of Super4PCS and Generalized4PCS, which combines intelligent indexing strategy with non-coplanar optimization, and also speeds up the algorithm by reducing the matching number of four point sets.

4PCS and its derivative algorithms are all under the framework of RANSAC registration, and they can quickly reconstruct and match similar spatial topologies in the scene. No matter the scene with small overlap, the scene with serious point noise or the scene with large range of noise can have a good registration effect.

B. Methods based on feature

Normal vector and curvature are the most basic geometric features of point cloud. Although the calculation is simple and fast, in most cases, the point cloud will contain many similar or the same eigenvalues, so that more information and properties cannot be obtained. Radu Bogdan Rusu *et al.* [33] proposed a commonly used point cloud geometric feature description operator Point Feature Histogram (PFH) in 2008, PFH is a quantification of the spatial differences in the neighborhood of the central point, and a histogram is obtained by the method of mathematical statistics to describe the relevant information of the adjacent central point. The purpose of PFH is to transform the geometric characteristics of the adjacent mean curvature of a point into a multi-dimensional histogram, which provides a very informative feature representation. The point cloud is not only invariant under the six-dimensional transformation (rotation and translation of each three-dimensional), but also well adapted to different degrees of sampling or different degrees of noise. However, computing FPH features is very resource-consuming, the computational efficiency is relatively low, and its algorithm complexity ($O(nk^2)$,Where k represents the quantity of neighborhoods of p_i) is difficult to meet the real-time requirements. Therefore, in the second year, the author proposed an improved version named FPFH[34], The input of FPFH is the same as before, and the output become a histogram that reflects the relevant characteristics of the neighborhood around each point. Different from PFH, FPFH mainly improves the calculation method of the neighborhood of the central point to speed up the calculation speed of the algorithm, and finally realizes the time complexity of $O(nk)$.

Another commonly used point based feature descriptor is Signature of Histogram of Orientation (SHOT) [35] proposed by Samuele salti *et al.* in 2014. Within the constructed local reference frame (rotation and translation invariance), the SHOT feature descriptor calculates the topological features of the neighborhood of feature points, preserves and normalizes the features in the histogram (robust to point cloud density and noise). In addition to point features, point cloud registration can also be based on line features. Iterative closest line (ICL) [36] provides a registration algorithm based on line features.

C. Methods based on ICP

There are two cases of registration, one is to find an approximate rotation translation matrix of the two point clouds without knowing the relative position, which we call the coarse registration. The other is to further calculate a more accurate rotation translation matrix when the initial value of a rotation translation matrix is known, which we call the fine registration. At present, fine registration has basically been fixed to use ICP algorithm and its derivative algorithms, so it is set as a separate module in this paper.

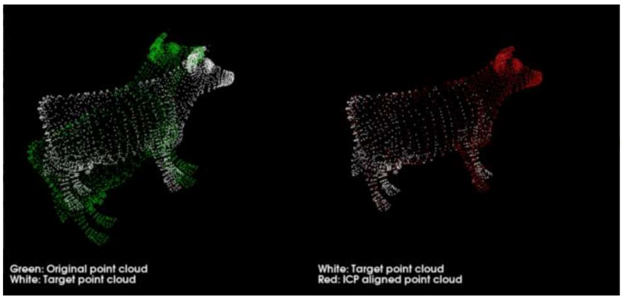
The core of ICP [16] employ Singular Value Decomposition (SVD) [37] algorithm. The basic idea is to give two sets of point clouds $X = \{x_1, x_2, \dots\dots, x_{N_x}\}$, $P = \{p_1, p_2, \dots\dots, p_{N_p}\}$, Where x_i and p_i represent point cloud coordinates, and N_x and N_p indicate the number of point clouds. Solve the rotation matrix R and the shift vector T that minimize the (4).

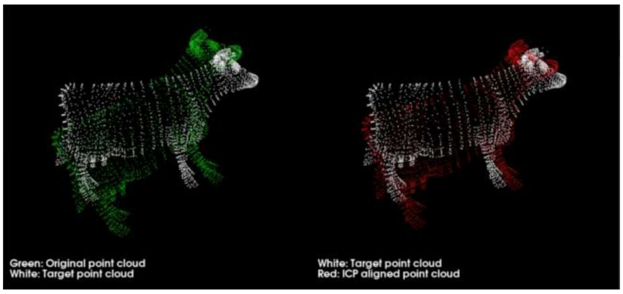
$$E(R,T) = \frac{1}{N_p}\sum_{i=1}^{N_p} \| x_i - Rp_i - T \|^2 \quad (4)$$

The approximate flow of the algorithm is as follows:

- Given X and P, let the cloud X be the source point cloud and the cloud P be the target.
- For every point in the source X, find the nearest corresponding point p_i in the target P, as the corresponding point in P, to form the initial corresponding point.
- Not all the corresponding relations in the initial corresponding point set are correct, and the wrong corresponding relationship will affect the final registration result, so the direction vector threshold is used to eliminate the wrong corresponding point pairs.
- The rotation matrix R and the shift vector T are calculated to minimize the mean square error between X and P.
- Set a certain threshold ε and the maximum iterations N_{max}, act the rigid body transformation obtained in the previous step on the X, get a new point cloud X', and calculate the distance error between X' and P. if the error of two iterations is less than the threshold ε or the current iterations is greater than N_{max}, the iteration ends, otherwise the initial registration point set is updated to X' and P, and the above steps are repeated until the convergence conditions are met.

However, according to [38], we know that the points in two point clouds cannot represent the same position in space. Therefore, using the distance from point to point as the error equation is bound to introduce random error. And If the difference between point clouds is large, the ICP algorithm has the possibility of falling into local optimization, so it cannot get a better matching result. The experimental effect of ICP is shown in Fig. 4 and Fig. 5.





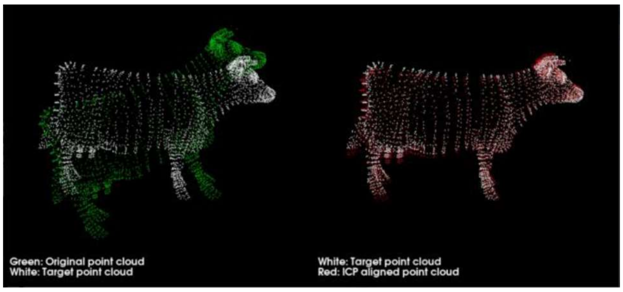


Fig. 4. Experimental effect diagram of ICP，The upper, middle and lower are ICP algorithm iterations 1, 8 and 16 times respectively

Fig. 5. ICP algorithm falls into local optimal solution

Andrew Wfitzgibbon *et al.* improved the ICP algorithm by using the Levenberg-MarQuardt (LM) [39] algorithm, and then proposed the LM-ICP [40] algorithm. The LM algorithm can provide a numerical solution of nonlinear minimization (local minimization), and it has the advantages of both the gradient descent and the Gauss-Newton iteration. When λ is very small, the step size is equal to the Newtonian step size, and when λ is very large. LM algorithm is not only fast in optimization, but also insensitive to overparameterization, and can efficaciously handle the trouble of redundant parameters, so that the chance of the cost function obtaining the local minimum is greatly reduced. The experimental effect of LM-ICP is seen in Fig. 6.

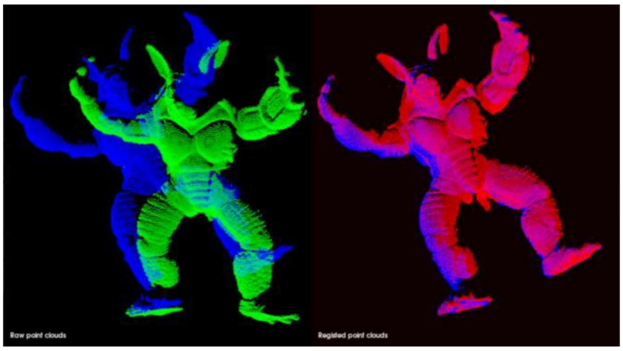


Fig. 6. Experimental effect diagram of LM-ICP

Point-to-plane ICP [41] improves the cost function on the basis of ICP, adds the surface normal vector of the target point, uses the nearest neighbor method to discover the nearest point, and then pulls the distance of the nearest point along the normal direction of the fitting plane at the target point. The biggest difference between point-to-line ICP [42] and the original ICP is that it improves the error equation. The original ICP takes the gap from the point to point as the error, while the PL-ICP uses the gap from the point to the connection between the two nearest points as the error, and PL-ICP can achieve second-order convergence in a short time. The improvement of ICP by GICP is that a Gaussian probability model is added to the cost function. The core idea is how to treat and derive the objective function of ICP algorithm from the point of view of probability. Although the main part of the algorithm is basically unchanged, the addition of Gaussian probability model not only ensures robustness but also reduces the complexity. SparseICP optimizes the algorithm by using sparse induced norms to achieve better registration results. The algorithm idea and main body flow of NICP [43] are basically the same as ICP. The main improvement point is to deal with outliers and error terms, in addition to constraining the distance between the corresponding points, and adding constraints on the normal vector and the curvature of the surface where the point cloud is located. Not only the distance from the point to the tangent plane of the corresponding point is considered, but also the angle difference of the normal vector of the corresponding point is considered. VGICP [44] improves ICP on real-time requirements. The reason why ICP is slower than NDT [23] is that ICP has a large number of kd-tree/octree searches, while NDT correlates domains in a voxel way, thus avoiding numerous neighborhood searches when ICP is under iteration. Although the real-time performance of VGICP is much better than that of traditional ICP, the voxel method will also bring the problem of voxel mesh size division. I think that how to balance accuracy and running efficiency in practical application will be a problem to be solved in the future. After years of development, the derivative algorithm of ICP has been able to solve different problems from different directions, but the framework of ICP is after all the product of the last century. We look forward to the emergence of a new precise registration algorithm without ICP in the future.

III. Deep learning methods

As mentioned earlier, local feature descriptors of 3D geometric information play an important role in many fields, although many authors have proposed a variety of 3D feature descriptors in the past decade. However, it is difficult to generate ideal reusable and distinguishing local descriptors for point cloud data from diverse devices and scenes. Because in many cases, 3D point clouds contain more noise, or incomplete.

The emergence of PPFNet is to make 3D point clouds generate ideal and robust 3D local feature descriptors. Referring to PointNet [45], this paper uses deep learning method to generate 3D local feature descriptors that are easy to distinguish and anti-rotation, and designs a new loss function, which can embed multiple samples of the same kind or different kinds into an Euclidean space at the same time, and the difference between samples is expressed by the Euclidean distance of its Eigenvectors. However, the problem with PPFNet is that it will take up memory twice, the effect will not be so ideal in some tasks, and the generalization ability of point cloud scenarios is not good enough. So in the same year, the author proposed PPF-FoldNet [46] to solve this problem. PPF-FoldNet made improvements on the basis of PPFNet, combining the encoder of PointNet and the decoder of FoldingNet [47]. The experimental results show that PPF-FoldNet has excellent rotation invariance.

In recent years, probably from 2017, The hot spot in the field of registration has gradually turned to solving problems in the way of DL, and many of them [48] [49]are not limited to finding 3D feature descriptors. For example, 3DFeat-Net is a framework that can not only detect key points, but also calculate key point descriptors. Under the condition of weak supervision, the algorithm obtains the 3D feature detector and descriptor through learning, and matches the point cloud. The biggest highlight of the algorithm is that there is no need for manual labeling, and the contribution of each descriptor is adjusted by the attention mechanism. SaraSabour *et al.* [50] raised a new architecture with a shape very much like a capsule, which uses iterative routing mechanism to achieve better experimental results than convolution networks. YueWang *et al.* [51] raised an end-to-end Deep Closest Point (DCP) algorithm, which uses the Transformer network structure to calculate a "hypothetical target point cloud". By using the loss function to constrain the corresponding relationship between known point clouds, the registration of point clouds is finally realized. RPM-Net [52] first realizes the initial value insensitivity by learning and merging a variety of features. Then we use the differentiable Sinkhorn algorithm [53] to effectively deal with the outer points of the matching process. At the same time, a quadratic network is used to learn to optimize the parameters of the annealing algorithm to get better execution results. The greatest contribution of RPM-Net is to solve the problem of sensitivity to initial rigid transformation and noise and outliers in point cloud rigid registration. PartNet is a very ingenious frame evolved from Recurrent Neural Network (RNN) [54]. PartNet divides 3D point clouds recursively from top to bottom to form a fine-grained hierarchy. It has achieved very good results in the registration at the component level. By optimizing the neural network (NN) model, someone raised a ordinary data-driven way to integrate multi-level local geometric features, and made a breakthrough in the lightweight and rotation invariance of the model.

IV. Evaluation criteria

At present, there is not a recognized evaluation standard of point cloud registration. Some methods calculate the distance after registration, some articles calculate the correct number of pairs before and after matching, and some calculate translation error and rotation error under different coincidence rates. Moreover, traditional methods and deep learning methods are two completely different fields, and evaluation methods may not necessarily communicate with each other. This paper briefly introduces several commonly used evaluation functions.

A. Root Mean Square Error

$$MSE = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \widehat{X_i}\right)^2}{n}} \tag{5}$$

The formula is defined as shown in (5), where n is the quantity of pairs of corresponding points, and $\hat{X_i}$ is the true value of Euclidean distance after registration. In the absolutely ideal state, after full registration, the gap between the corresponding points is 0, so the true value of the Euclidean distance also be 0.

B. Hausdorff Distance（HD）

Two point sets $A = \{a_1, a_2, \dots\}$ and $B = \{b_1, b_2, \dots\}$ in Euclidean space, HD is used to survey the distance between these points, as defined in (6).

$$H(A, B) = max[h(A, B), h(B, A)] \tag{6}$$

where

$$h(A,B) = \max_{a \in A} \min_{b \in B} \| a - b \|$$
$$h(B,A) = \max_{b \in B} \min_{a \in A} \| b - a \| \tag{7}$$

$H(A,B)$ is called bi-directional HD, $h(A,B)$ is known as the one-way HD from point set A to B. The corresponding $h(B,A)$ is known as the one-way HD from the point set B to A.

C. Benchmark

For deep learning, the dataset generally chooses the ModelNet40 dataset in PointNet [45]. Network architecture this article chooses the PCRNet [67] architecture, which is easy to reproduce, as shown in the following Fig. 7.

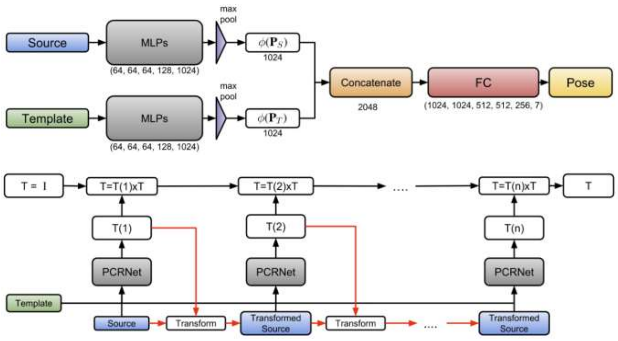


Fig. 7. The architecture of PCRNet[67]

There are more Loss options, such as MSE Loss, Loss about Euler Angle, Chamfer Distance (CD) Loss and Earth Mover's Distance (EMD) [55] Loss for point clouds. The main evaluation indexes are the isotropy of rotation angle and Euler angle and the running time of the algorithm.

V. Conclusion

In recent years, point cloud registration has become a critical technology in all fields of all walks of life, so it is a very valuable topic to study point cloud registration.

This paper makes a comprehensive research on registration from two point of view of traditional methods and deep learning methods. Among them, the focus of traditional methods is the extraction and feature description of key points, and the realization method is to discover the optimal solution of nonlinear problems. The advantage is that it can ensure convergence, no training and good generalization ability for unknown scenes. The disadvantage is that many complex mathematical methods are needed to inhibit the changes of noise from equipment acquisition, a few outliers irrelevant to the object, uneven density of different parts of the same object and local overlap of objects that are very likely to occur, all these will greatly increase the computational cost and waste the power of calculation.. The deep learning method designs different network architectures and loss functions, uses deep learning to extract features, and then estimates the corresponding relationship. its advantage is that the point features based on deep learning can provide robust and accurate correspondence search, and even get a very good registration effect in some specific scenes. The disadvantage is that the interpretability of the algorithm is poor, most DL algorithms rely heavily on training data and have very high requirements for them, and the registration performance in unknown scenarios may be greatly degraded.

Finally, I think there are four main aspects of Future developments in registration: (1) how to deal with the relationship between real-time and accessibility. (2) how to deal with the relationship between accuracy and time. (3) how to make the point cloud algorithm more generalized. (4) At present, the number of points in large-scale point clouds is even more than 10 million, and the existing algorithms are either unable to handle such large-scale data, or run very slowly, so how to balance the direct relationship between efficiency, speed and computing power will become very important. (5) If different input devices are used to scan, it is possible to obtain point clouds with different information even at the same location, so how to pursue more robust ways to solve non-homologous troubles. (6) Although DL and mathematical optimization methods are introduced separately in this paper, how to make an efficacious and perfect combination of the two is the light of the future.

Acknowledgment

This work was supported by Shanghai Key Research Laboratory of NSAI.

References

[1] M. Liu, "Robotic Online Path Planning on Point Cloud," *IEEE Trans. Cybern.*, vol. 46, no. 5, pp. 1217–1228, May 2016, doi: 10.1109/TCYB.2015.2430526.

[2] L. Cheng *et al.*, "Registration of laser scanning point clouds: A review," *Sensors*, vol. 18, no. 5, p. 1641, 2018.

[3] M. J. F. A. F. Joaquim and A. Jorge, "Towards 3D modeling using sketches and retrieval," in *Eurographics Workshop on Sketch-Based Interfaces and Modeling 2004*, 2004, p. 127.

[4] P. Kim, J. Chen, and Y. K. Cho, "SLAM-driven robotic mapping and registration of 3D point clouds," *Autom. Constr.*, vol. 89, pp. 38–48, 2018.

[5] M. Weinmann, J. Leitloff, L. Hoegner, B. Jutzi, U. Stilla, and S. Hinz, "THERMAL 3D MAPPING FOR OBJECT DETECTION IN DYNAMIC SCENES.," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 2, no. 1, 2014.

[6] N. Mellado, M. Dellepiane, and R. Scopigno, "Relative scale estimation and 3D registration of multi-modal geometry using Growing Least Squares," *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 9, pp. 2160–2173, 2015.

[7] F. Poux, R. Neuville, L. Van Wersch, G.-A. Nys, and R. Billen, "3D point clouds in archaeology: Advances in acquisition, processing and knowledge integration applied to quasi-planar objects," *Geosciences*, vol. 7, no. 4, p. 96, 2017.

[8] Z. Ouyang, Y. Liu, C. Zhang, and J. Niu, "A cgans-based scene reconstruction model using lidar point cloud," in *2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)*, 2017, pp. 1107–1114.

[9] P. Wu, W. Li, and M. Yan, "3D scene reconstruction based on improved ICP algorithm," *Microprocess. Microsyst.*, vol. 75, p. 103064, 2020.

[10] H. Liu, Y. Yan, X. Zhang, and Z. Wu, "3D scene reconstruction based on 3D laser point cloud combining UAV images," in *2nd ISPRS International Conference on Computer Vision in Remote Sensing (CVRS 2015)*, 2016, vol. 9901, p. 99011C.

[11] L. Yang and R. Chakraborty, "A GMM based algorithm to generate point-cloud and its application to neuroimaging," in *2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops)*, 2020, pp. 1–4.

[12] A. P. Placitelli and L. Gallo, "3D point cloud sensors for low-cost medical in-situ visualization," in *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, 2011, pp. 596–597.

[13] J. Salvi, C. Matabosch, D. Fofi, and J. Forest, "A review of recent range image registration methods with accuracy evaluation," *Image Vis. Comput.*, vol. 25, no. 5, pp. 578–596, 2007.

[14] G. K. Tam *et al.*, "Registration of 3D point clouds and meshes: A survey from rigid to nonrigid," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 7, pp. 1199–1217, 2012.

[15] F. Pomerleau, F. Colas, and R. Siegwart, “A review of point cloud registration algorithms for mobile robotics,” *Found. Trends Robot.*, vol. 4, no. 1, pp. 1–104, 2015.

[16] P. J. Besl and N. D. McKay, “Method for registration of 3-D shapes,” in *Sensor fusion IV: control paradigms and data structures*, 1992, vol. 1611, pp. 586–606.

[17] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[18] F. Bernard, C. Theobalt, and M. Moeller, “Ds*: Tighter lifting-free convex relaxations for quadratic matching problems,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4310–4319.

[19] N. Dym, H. Maron, and Y. Lipman, “Ds++: A flexible, scalable and provably tight relaxation for matching problems,” *ArXiv Prepr. ArXiv170506148*, 2017.

[20] I. Kezurer, S. Z. Kovalsky, R. Basri, and Y. Lipman, “Tight relaxation of quadratic matching,” in *Computer Graphics Forum*, 2015, vol. 34, no. 5, pp. 115–128.

[21] W. Forstner and K. Khoshelham, “Efficient and accurate registration of point clouds with plane to plane correspondences,” in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 2165–2173.

[22] J. Minguez, F. Lamiraux, and L. Montesano, “Metric-based scan matching algorithms for mobile robot displacement estimation,” in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 2005, pp. 3557–3563.

[23] P. Biber and W. Straßer, “The normal distributions transform: A new approach to laser scan matching,” in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, 2003, vol. 3, pp. 2743–2748.

[24] D. Aiger, N. J. Mitra, and D. Cohen-Or, “4-points congruent sets for robust pairwise surface registration,” in *ACM SIGGRAPH 2008 papers*, 2008, pp. 1–10.

[25] N. Mellado, D. Aiger, and N. J. Mitra, “Super 4pcs fast global pointcloud registration via smart indexing,” in *Computer graphics forum*, 2014, vol. 33, no. 5, pp. 205–215.

[26] P. W. Theiler, J. D. Wegner, and K. Schindler, “Markerless point cloud registration with keypoint-based 4-points congruent sets,” *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 1, no. 2, pp. 283–288, 2013.

[27] X. Ge, “Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets,” *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 344–357, 2017.

[28] C. Raposo and J. P. Barreto, “Using 2 point+ normal sets for fast registration of point clouds with small overlap,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 5652–5658.

[29] Z. Xu, E. Xu, Z. Zhang, and L. Wu, “Multiscale sparse features embedded 4-points congruent sets for global registration of TLS point clouds,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 286–290, 2018.

[30] J. Huang, T.-H. Kwok, and C. Zhou, “V4PCS: Volumetric 4PCS algorithm for global registration,” *J. Mech. Des.*, vol. 139, no. 11, 2017.

[31] M. Mohamad, D. Rappaport, and M. Greenspan, “Generalized 4-points congruent sets for 3d registration,” in *2014 2nd international conference on 3D vision*, 2014, vol. 1, pp. 83–90.

[32] M. Mohamad, M. T. Ahmed, D. Rappaport, and M. Greenspan, “Super generalized 4pcs for 3d registration,” in *2015 International Conference on 3D Vision*, 2015, pp. 598–606.

[33] R. B. Rusu, Z. C. Marton, N. Blodow, and M. Beetz, “Learning informative point classes for the acquisition of object model maps,” in *2008 10th International Conference on Control, Automation, Robotics and Vision*, 2008, pp. 643–650.

[34] R. B. Rusu, N. Blodow, and M. Beetz, “Fast point feature histograms (FPFH) for 3D registration,” in *2009 IEEE international conference on robotics and automation*, 2009, pp. 3212–3217.

[35] S. Salti, F. Tombari, and L. Di Stefano, “SHOT: Unique signatures of histograms for surface and texture description,” *Comput. Vis. Image Underst.*, vol. 125, pp. 251–264, 2014.

[36] M. Alshawa, “ICL: Iterative closest line A novel point cloud registration algorithm based on linear features,” *Ekscentar*, no. 10, pp. 53–59, 2007.

[37] G. H. Golub and C. Reinsch, “Singular value decomposition and least squares solutions,” in *Linear algebra*, Springer, 1971, pp. 134–151.

[38] K. S. Arun, T. S. Huang, and S. D. Blostein, “Least-squares fitting of two 3-D point sets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 5, pp. 698–700, 1987.

[39] J. J. Moré, “The Levenberg-Marquardt algorithm: implementation and theory,” in *Numerical analysis*, Springer, 1978, pp. 105–116.

[40] A. W. Fitzgibbon, “Robust registration of 2D and 3D point sets,” *Image Vis. Comput.*, vol. 21, no. 13–14, pp. 1145–1153, 2003.

[41] K.-L. Low, “Linear least-squares optimization for point-to-plane icp surface registration,” *Chap. Hill Univ. N. C.*, vol. 4, no. 10, pp. 1–3, 2004.

[42] A. Censi, “An ICP variant using a point-to-line metric,” in *2008 IEEE International Conference on Robotics and Automation*, 2008, pp. 19–25.

[43] J. Serafin and G. Grisetti, “NICP: Dense normal based point cloud registration,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 742–749.

[44] K. Koide, M. Yokozuka, S. Oishi, and A. Banno, “Voxelized gicp for fast and accurate 3d point cloud registration,” *EasyChair Prepr.*, no. 2703, 2020.

[45] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation,” 2017, pp. 652–660. Accessed: Sep. 15, 2021. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Qi_PointNet_Deep_Learning_CVPR_2017_paper.html

[46] H. Deng, T. Birdal, and S. Ilic, “Ppf-foldnet: Unsupervised learning of rotation invariant 3d local descriptors,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 602–618.

[47] Y. Yang, C. Feng, Y. Shen, and D. Tian, “Foldingnet: Point cloud auto-encoder via deep grid deformation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 206–215.

[48] J. Li, C. Zhang, Z. Xu, H. Zhou, and C. Zhang, “Iterative distance-aware similarity matrix convolution with mutual-supervised point elimination for efficient point cloud registration,” in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXIV 16*, 2020, pp. 378–394.

[49] V. Sarode *et al.*, “PCRNet: Point cloud registration network using PointNet encoding,” *ArXiv Prepr. ArXiv190807906*, 2019.

[50] S. Sabour, N. Frosst, and G. E. Hinton, “Dynamic routing between capsules,” *ArXiv Prepr. ArXiv171009829*, 2017.

[51] Y. Wang and 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.

[52] Z. J. Yew and G. H. Lee, “Rpm-net: Robust point matching using learned features,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11824–11833.

[53] M. Cuturi, “Sinkhorn distances: Lightspeed computation of optimal transport,” *Adv. Neural Inf. Process. Syst.*, vol. 26, pp. 2292–2300, 2013.

[54] W. Zaremba, I. Sutskever, and O. Vinyals, “Recurrent neural network regularization,” *ArXiv Prepr. ArXiv14092329*, 2014.

[55] Y. Rubner, C. Tomasi, and L. J. Guibas, “The earth mover’s distance as a metric for image retrieval,” *Int. J. Comput. Vis.*, vol. 40, no. 2, pp. 99–121, 2000.