A comprehensive survey on point cloud registration

UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

XuTaosi
xutaosi@umac.mo
2024.5.10

Introduction

arXiv.org
https://arxiv.org/abs/2103.02690

[2103.02690] A comprehensive survey on point cloud registration - arXiv.org

DOI: https://doi.org/10.48550/arXiv.2103.02690

Authors: Huang, Xiaoshui, Mei, Guofeng, Zhang, Jian, Abbas...

已发布: 2021/03/03

RelatedEntitiesURL: http://arxiv.org/pdf/2103.02690.pdf

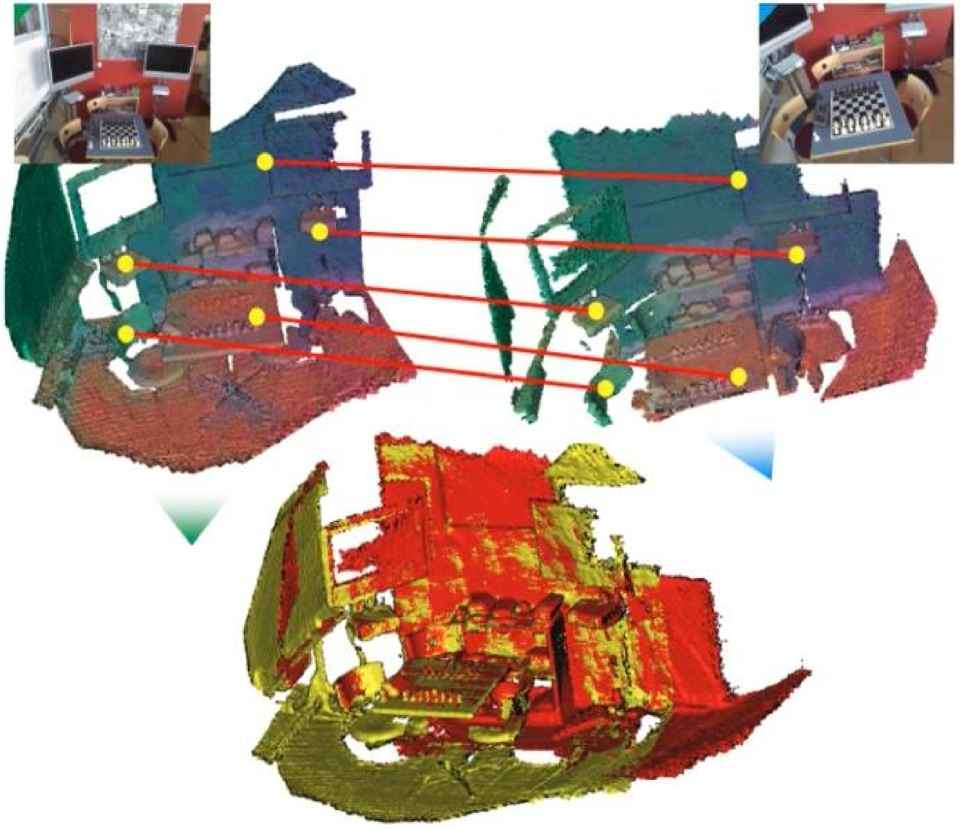
Subject: Computer Vision and Pattern Recognition (cs.CV)

Cite As: arXiv:2103.02690 [cs.CV]

Introduction

- Summarized the development of point cloud registration from 1992 to 2021, including same-source and cross-source; traditional optimization and current deep learning methods; also summarized the relationship between optimization strategies and deep learning.
- Established a new cross-source point cloud benchmark and evaluated the performance of the most advanced registration algorithms and made comparisons.

Problem Definition



Denote $\mathbf{x}_i{}^T (i \in [1, M])$ and $\mathbf{y}_i{}^T (j \in [1, N])$ as row vectors from matrices $X \in \mathbb{R}^{M \times 3}$ and $Y \in \mathbb{R}^{N \times 3}$ respectively. X and Y represent two point clouds, and $\mathbf{x}_i$ and $\mathbf{y}_j$ are the coordinates of the i_{th} and i_{th} points in the point clouds respectively. Suppose X and Y have K pairs of correspondences. The goal of registration is to find the rigid transformation parameters g (rotation matrix $R \in \mathcal{SO}(3)$ and translation vector $t \in \mathbb{R}^3$) which best aligns the point cloud X to Y as shown below:

$$\underset{R \in \mathcal{SO}(3), t \in \mathbb{R}^3}{\arg\min} \ \| d(X, g(Y)) \|_2^2 \tag{1}$$

where $d(X, g(Y)) = d(X, RY + t) = \sum_{k=1}^{K} \| \mathbf{x}_k - (R\mathbf{y}_k + t) \|_2$

Challenges

- **Same-source challenges:** Noise and outliers, Partial overlap.
- **Cross-source challenges:** Noise and outliers, Partial overlap, Density difference, Scale variation

Categories

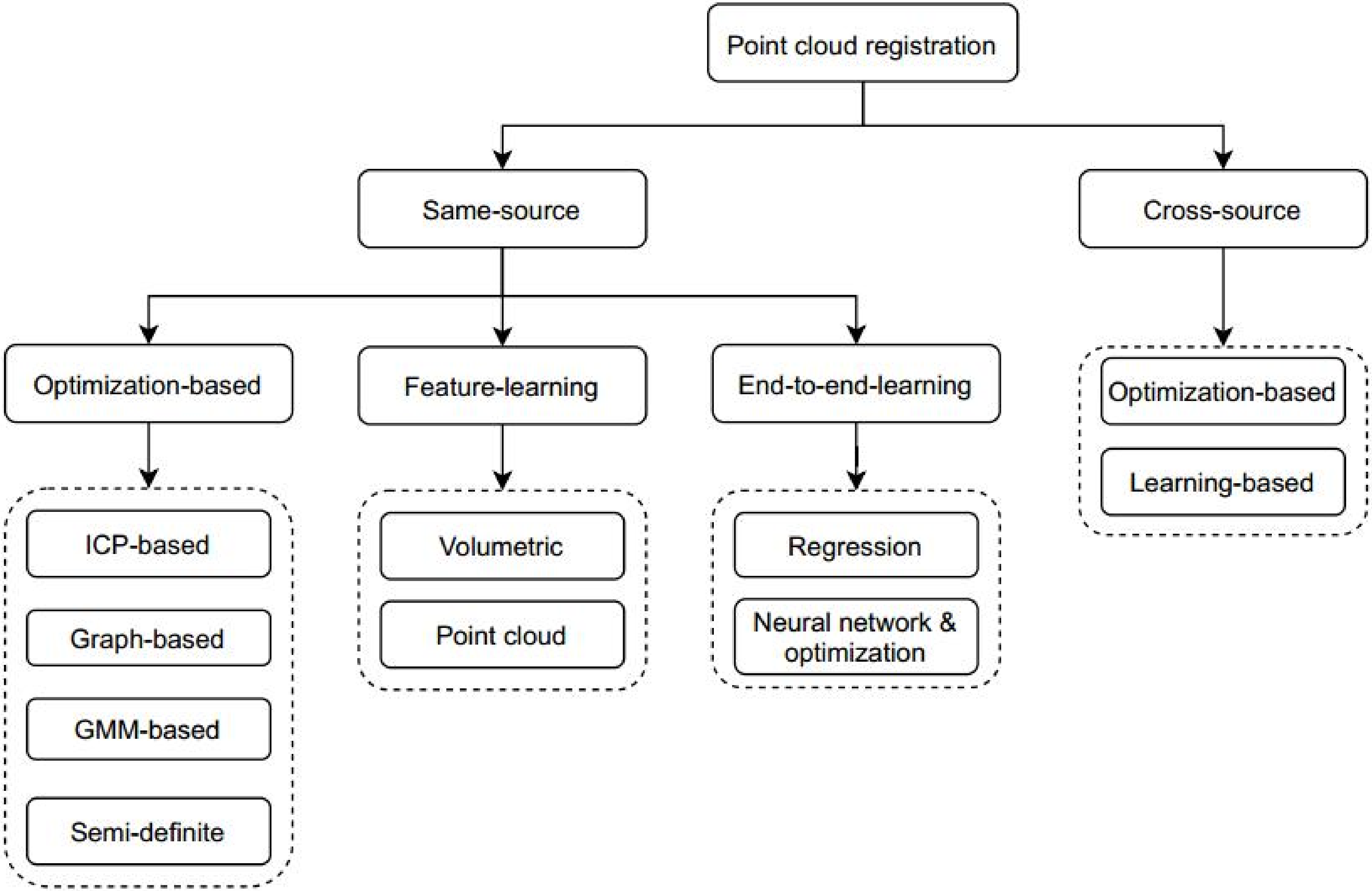
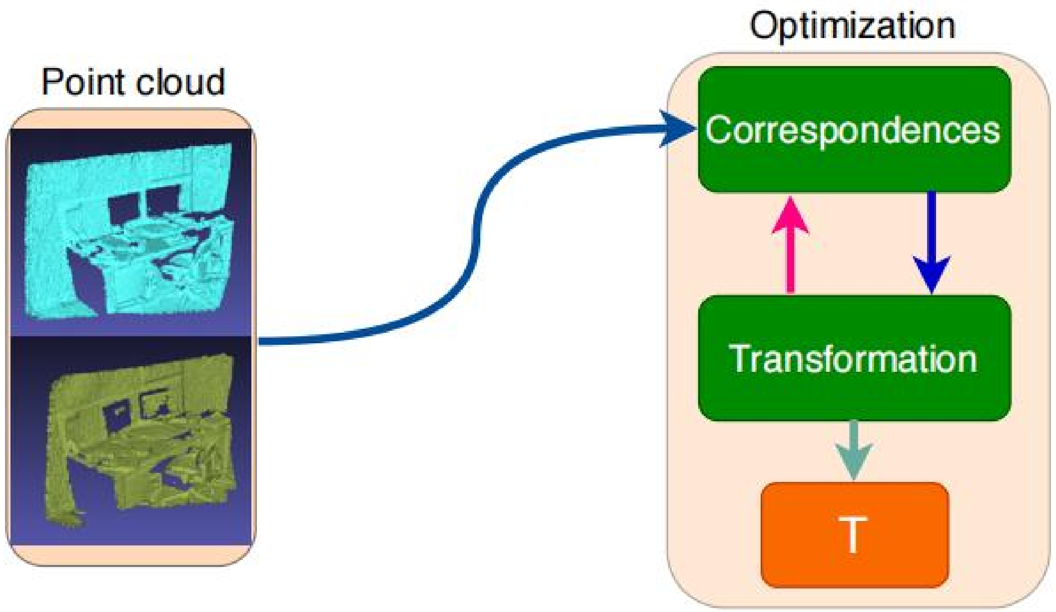


Fig. 1: Taxonomy of point cloud registration

Optimisation-based registration methods



(a) An optimization-based framework for point cloud registration. Given two input point clouds, the correspondences and transformation between these point clouds are iteratively estimated. The algorithm outputs the optimal transformation **T** as the final solution.

Advantages :

1) Rigorous mathematical theories could guarantee their convergence.
2) They require no training data and generalize well to unknown scenes.

Limitations :

Many sophisticated strategies are required to overcome the variations of noise, outliers, density variations and partial overlap, which will increase the computation cost.

Optimisation-based registration methods

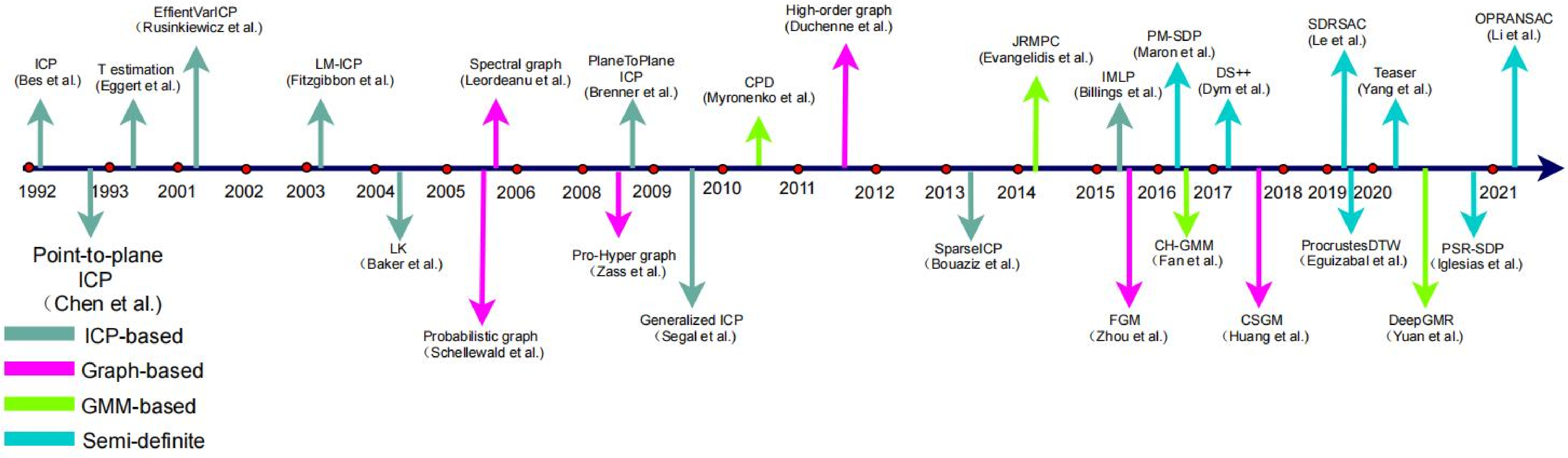


Fig. 3: Chronological overview of the most relevant optimization-based methods.

ICP-based registration

- **point-to-plane**

$$\underset{R\in\mathcal{SO}(3),t\in\mathbb{R}^3}{\arg\min}\ \{\sum_{k=1}^{K} w_k\|\mathbf{n_k} * (\mathbf{x}_k - (R\mathbf{y}_k + t)\|^2)\} \qquad (2)$$

where w_k is the weights of each correspondence, $\mathbf{n_k}$ is the surface normal at point $\mathbf{x}_k$, $\mathbf{x}_k$ and $\mathbf{y}_k$ are point-correspondence pairs on point cloud X and Y.

ICP-based registration

- **Generalized ICP(协方差矩阵)**

$$\arg\min_{T}\{\sum_{k=1}^{K}\|d^T(C_k^Y+\mathbf{T}C_k^X\mathbf{T}^T)^{-1}d\|^2\} \tag{3}$$

where $\{C_k^X\}$ and $\{C_k^Y\}$ are covariance matrices associated with the point cloud X and Y. $\mathbf{T}$ is the transformation parameters that consists of R and t, d is a distance metric.

ICP-based registration

- **Generalized ICP**

For standard point-to-point ICP, it is a special case by setting $C_k^Y = I$ and $C_k^X = 0$. Also, for point-to-plane ICP is a limiting case of this generalized ICP by setting $C_k^Y = P_k^{-1}$ and $C_k^X = 0$, where P_k^{-1} is the surface normal at x_k. The

point-to-point: $\arg\min_T \sum_{k=1}^{K} d_k^T (I + T \cdot 0 \cdot T^T)^{-1} d_k = \arg\min_T \sum_{k=1}^{K} d_k^T I^{-1} d_k = \arg\min_T \sum_{k=1}^{K} \|d_k\|^2$

point-to-plane: $d_k = (x_k - (Ry_k + t)) \cdot n_k$ $\quad$ $\arg\min_{R,t} \sum_{k=1}^{K} ((x_k - (Ry_k + t)) \cdot n_k)^2 \cdot \alpha_k$

Here, α_k is a scalar representation of P_k^{-1}, which increases the weighting in the direction of the normal, making the registration more focused on the orientation of the plane.

um 澳大

ICP-based registration

- **plane-to-plane**

In addition, plane-to-plane distance metric [10], [48], [33] is adopted to estimate the correspondences. The objective is similar to point-point distance metric, which is

$$\underset{R \in \mathcal{SO}(3), t \in \mathbb{R}^3}{\arg\min} \{ \sum_{k=1}^{K} \| \mathbf{nx}_k - (R\mathbf{ny}_k + t \|^2) \} \tag{4}$$

where $\mathbf{nx}$ and $\mathbf{ny}$ are surface normal of point cloud X and Y.

Graph-based registration

- **Principle:** Graph-based registration methods primarily rely on converting point clouds into graph representations, where points in the point cloud serve as vertices of the graph, and the geometric and other attribute relationships between points serve as edges. The advantage of this method lies in its ability to utilize algorithms from graph theory to find the optimal correspondence between two graphs, thereby deriving the best transformation between the point clouds.

- **Optimization strategies:** Quadratic Assignment Problem (QAP二次分配): Graph matching can be formalized as a QAP, which is a well-known NP-hard problem that requires heuristic or approximate algorithms for resolution.

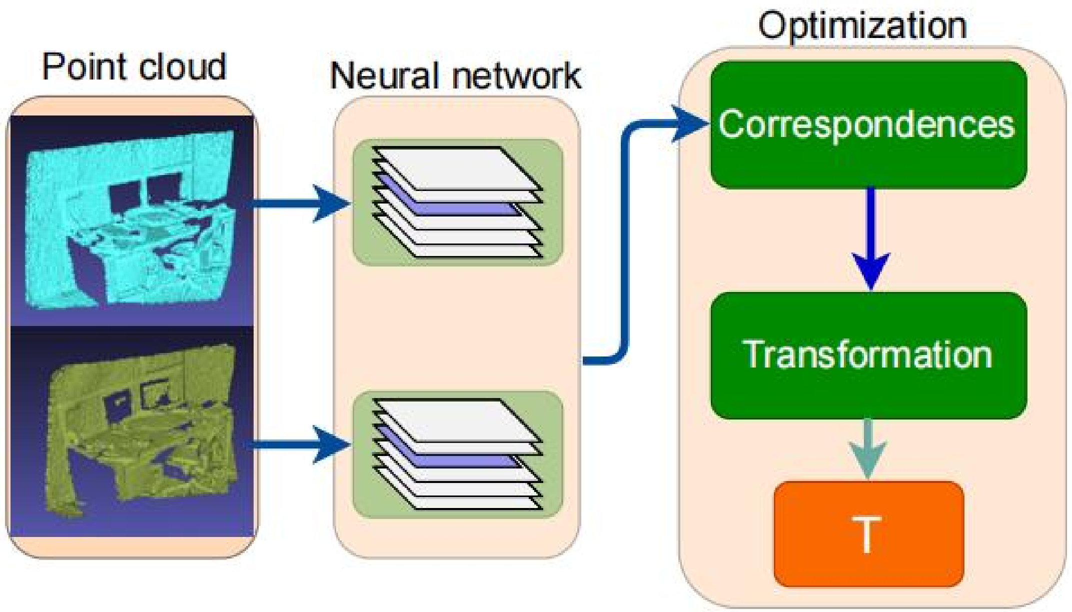
- **Solutions:** Relaxation Techniques: Common relaxation techniques include spectral relaxation and semidefinite programming relaxation. These techniques simplify the constraints of the original optimization problem, making it easier to solve.

GMM(高斯混合模型)-based registration

The critical idea of GMM-based methods is to formulate the registration problem into a likelihood maximization of input data.

Step1 Initialization : Set initial Gaussian Mixture Model parameters for the two point clouds, including means, covariances, and mixing weights.
Step2 Expectation-Maximization (EM) Algorithm Execution:
E-step: Calculate the probability of each point belonging to each Gaussian component.
M-step: Update the parameters of the Gaussian components and the transformation matrix (rotation and translation) to improve the alignment of the two point clouds.
Step3 Iterative Optimization: Repeat the E-step and M-step until the parameters converge or meet the stopping criteria.

Feature learning methods for registration



(b) A feature learning-based framework for point cloud registration. Given two input point clouds, the features are estimated using a deep neural network. Then, correspondence and transformation estimation run iteratively to estimate the final solution **T**.

Advantages :

1) Point features based on deep learning can provide robust and accurate correspondence searching.

2) Accurate correspondences can achieve precise registration results through a simple RANSAC method.

Limitations :

1) They require a large amount of training data.

2) If there is a significant difference in distribution between unknown scenes and the training data, the registration performance will significantly drop.

3) They use an independent training process to learn a standalone feature extraction network, and the learned feature network is used to determine point-to-point matching rather than registration.

Feature learning methods for registration

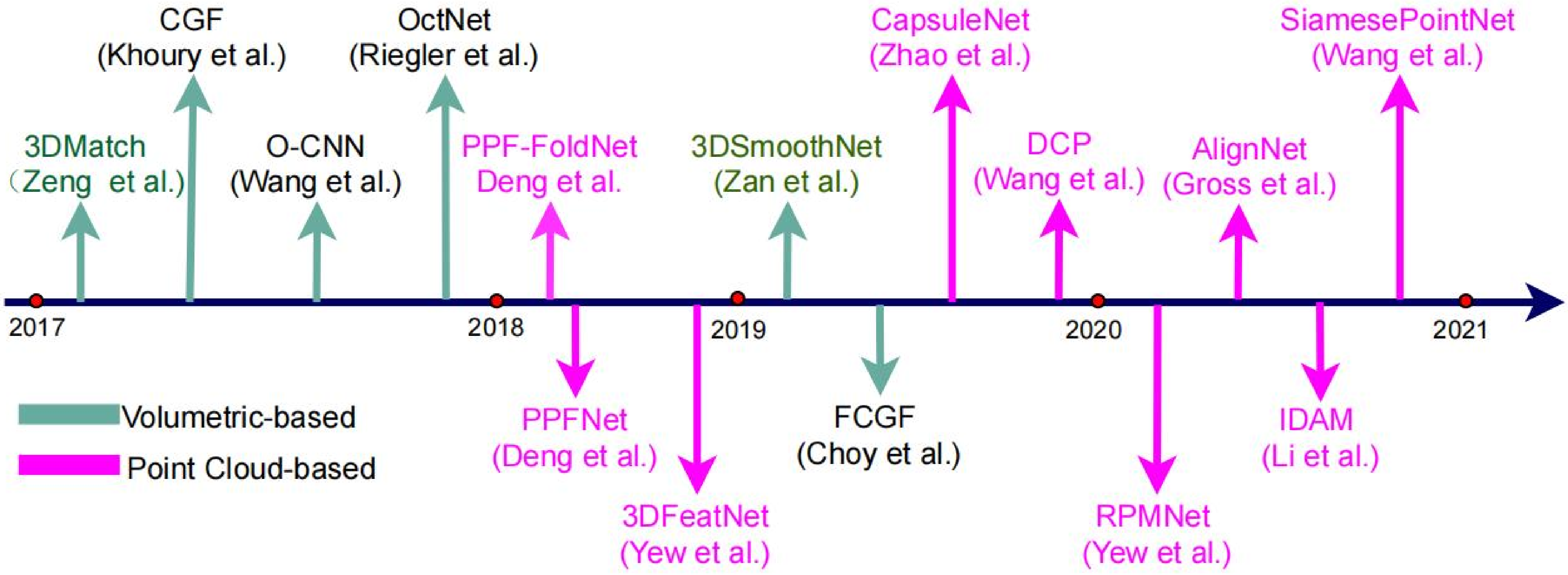


Fig. 4: Chronological overview of the most relevant feature-learning registration methods.

Learning on volumetric data

- **3DMatch:**

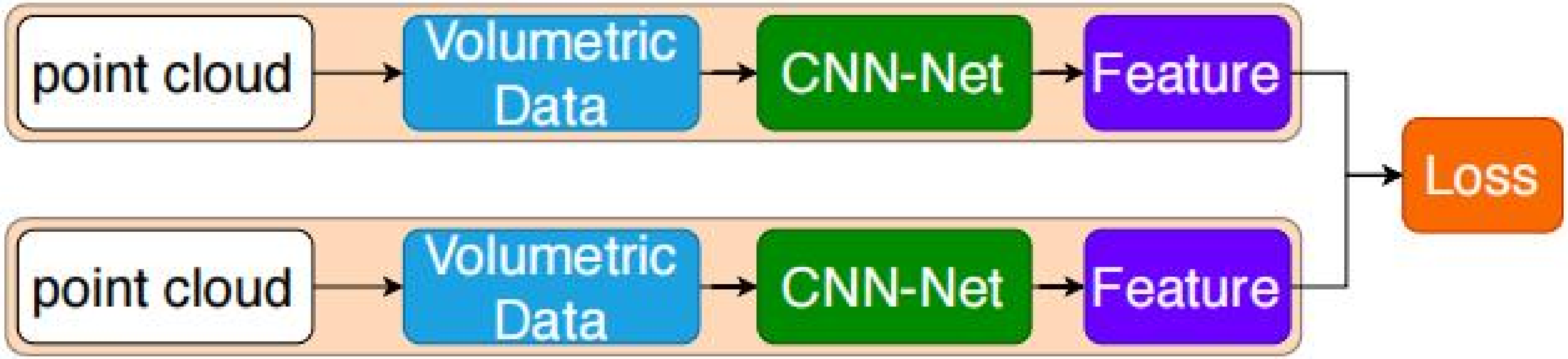


Fig. 5: The overall framework of 3DMatch, which is an example of neural networks in Figure 2b using volumetric data.

Learning on volumetric data

- **3DSmoothNet:**

具体而言，利用所有点的协方差矩阵的特征分解估计局部参考帧(LRF)。使用LRF对点云进行对齐后，对输入网格进行高斯平滑处理，得到平滑密度值(SDV)体素化。然后，将SDV输入网络进行特征提取。

- **FCGF:**

To improve the efficiency of volumetric-based descriptor, FCGF uses $1 \times 1 \times 1$ kernel to extract a fast and compact metric features for geometric correspondence.

Learning on point cloud

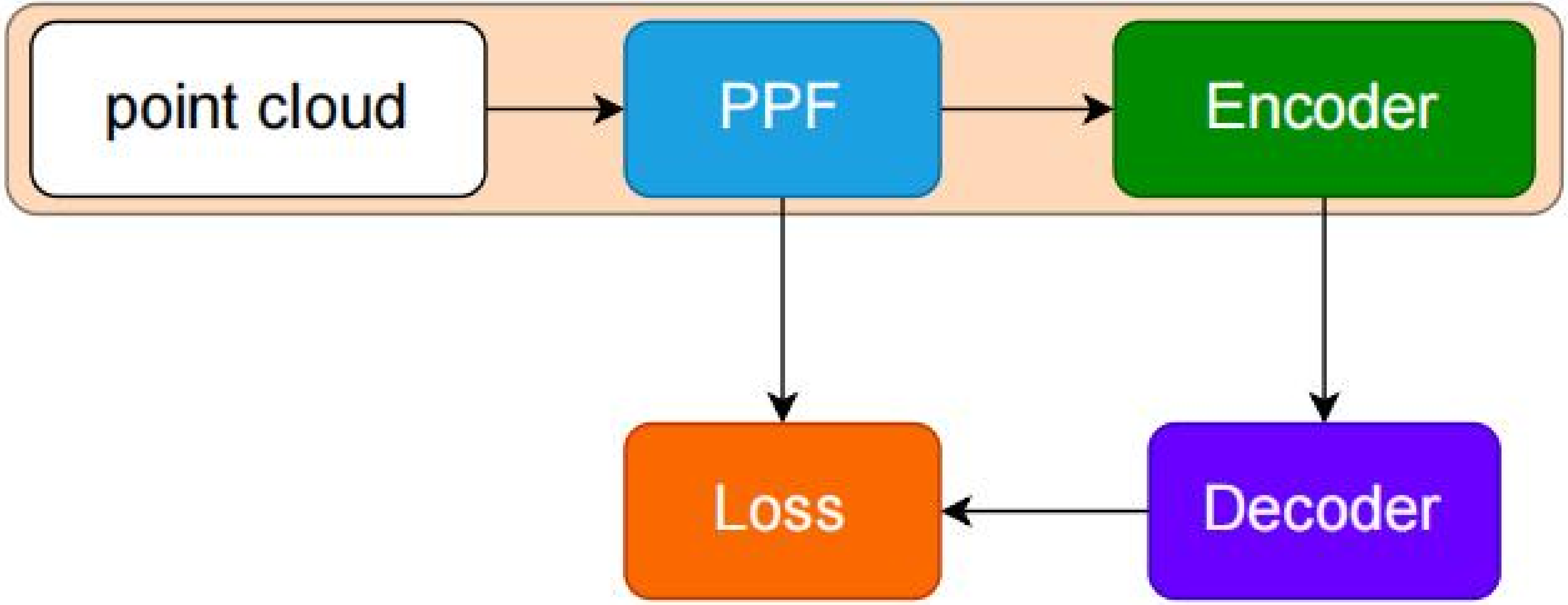


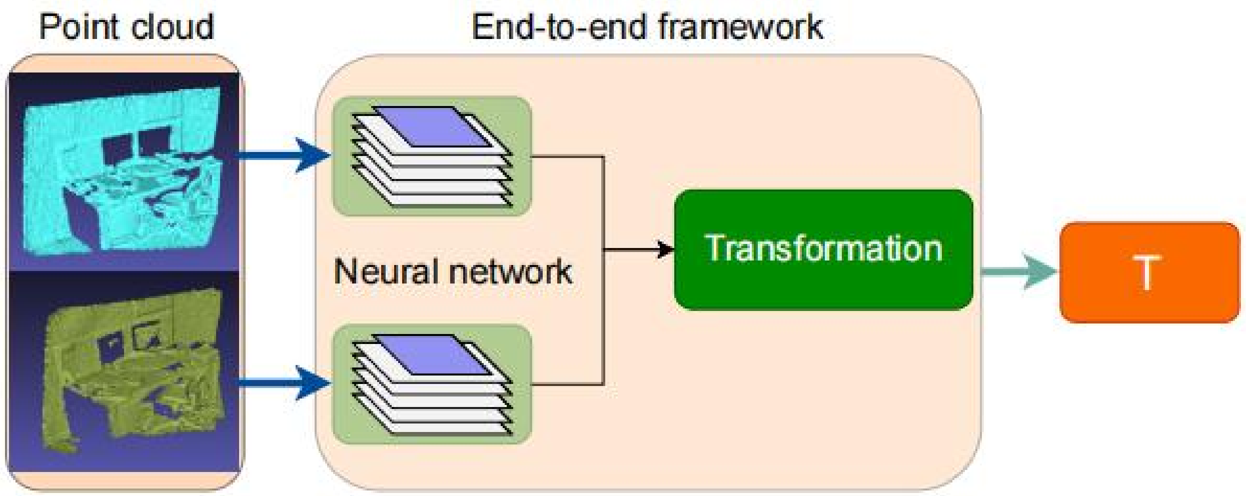
Fig. 6: The overall framework of PPFNet,, which is an example of neural networks in Figure $\boxed{2b}$ using point cloud.

Learning on point cloud

PPF-FoldNet：为了解决注释数据需求问题，PPF-FoldNet提出了一种无监督的方法。这种方法利用PointNet编码特征，并使用解码器将特征解码回与输入相同的数据。整个网络通过使用Chamfer损失优化输入和输出之间的差异来进行优化 。

SiamesePointNet：这种方法通过层次化的编解码器架构来生成感兴趣点的描述符，进一步处理点云数据以提取有用的特征 。

End-to-end learning-based registration methods



(c) An end-to-end learning-based framework for point cloud registration. Given two input point clouds, an end-to-end framework is used to estimate the final solution **T**.

优点：

1) 神经网络是专为配准任务设计和优化的;

2) 它能结合传统数学理论和深度神经网络的优点。

局限性：

1) 回归方法把变换参数估计视为黑箱，而且距离度量基于坐标的欧几里得空间，对噪声和密度差异敏感;

2) 特征度量配准方法未能充分考虑局部结构信息，这对于配准来说非常重要。

End-to-end learning-based registration methods

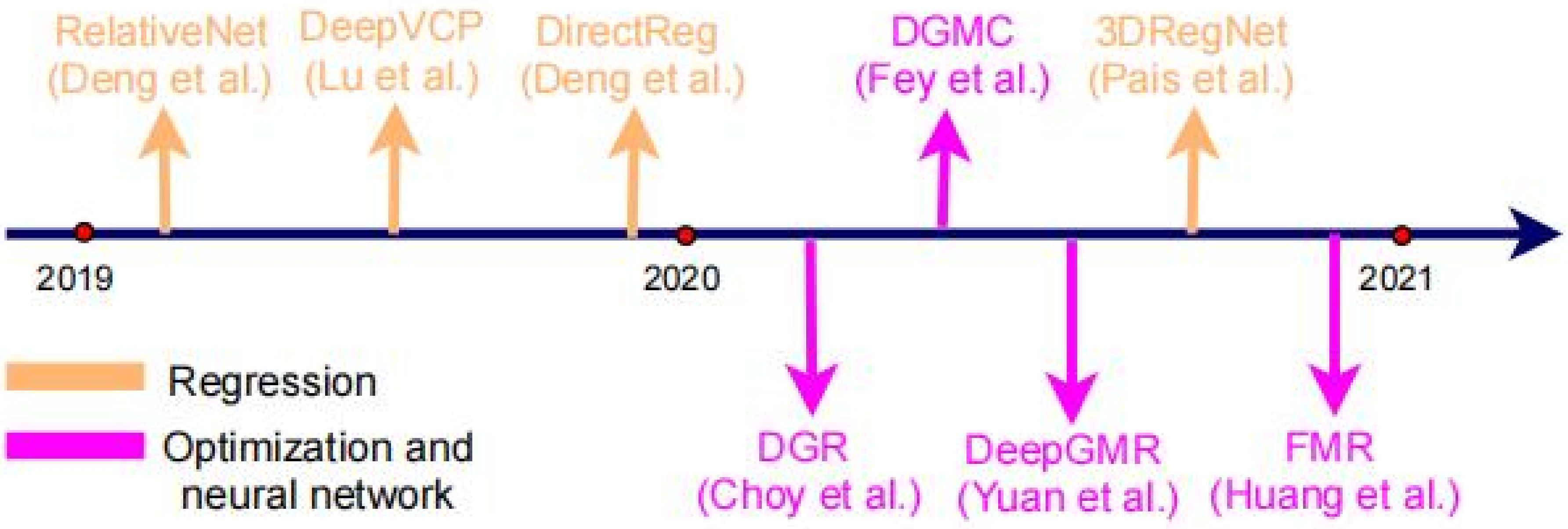


Fig. 7: Chronological overview of the most relevant end-to-end learning registration methods.

Registration by regression(回归)

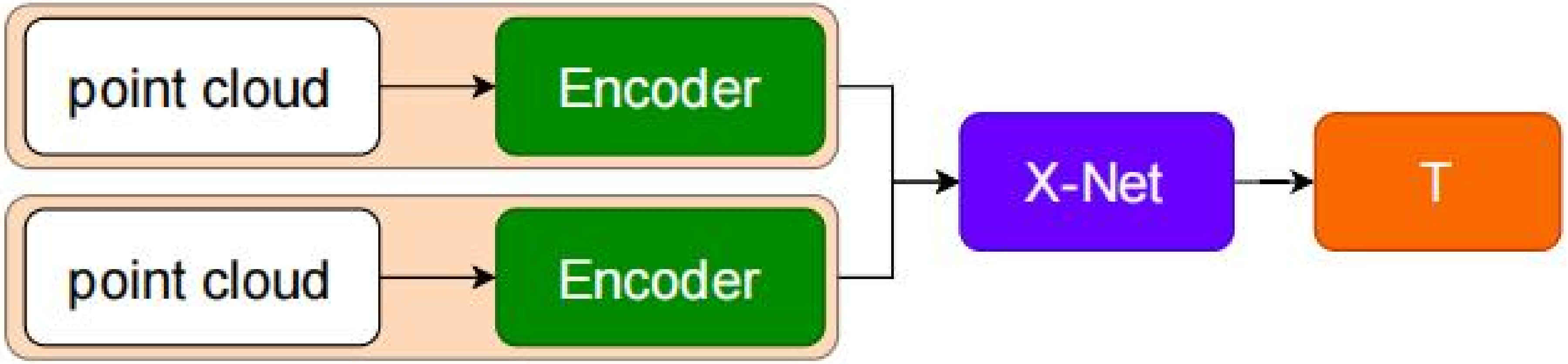


Fig. 8: The overall framework of end-to-end learning-based regression methods, which is an example of Figure 2c. Two global features are extracted firstly. Then, the two features fed into a X-Net to estimate a transformation matrix T.

Registration by optimization and neural network

Registration by Optimization and Neural Network 这一部分，主要探讨了如何将传统的配准相关优化理论与深度神经网络相结合来解决配准问题。以下是该方法的核心思想和关键技术：

- **结合理论和网络：**

该类别的主要思想是将传统的配准优化理论与深度神经网络技术结合起来，通过深度学习来提取点云的全局特征，然后使用优化算法来估计变换矩阵，以实现两个点云之间的精确对齐。

- **PointNetLK方法：**

具体实例如PointNetLK，使用PointNet来提取两个输入点云的全局特征，然后利用逆向组合（Inverse Compositional, IC）算法来估计变换矩阵。该方法的目标是最小化两个特征之间的差异，通过有限差分梯度计算来近似Jacobian估计，允许应用计算效率高的逆向组合Lucas-Kanade算法。

- **改进和应用：**

Huang等人对PointNetLK方法进行了改进，引入了自编码码器和点距离损失，减少了对标签的依赖 。此外，DeepGMR方法利用神经网络学习姿态不变的点到分布参数的对应关系，然后将这些对应关系输入到GMM优化模块中来估计变换矩阵 。

- **DGR方法：**

DGR提出了一个六维卷积网络架构来预测内点的可能性，并通过加权Procrustes模块来估计变换 。这些方法展示了传统优化方法与深度学习策略结合可以获得比以往方法更好的精确度。

这些技术表明，将深度学习策略与优化理论结合是解决复杂点云配准问题的有力途径，可以提供高精度和高效率的配准结果。

Cross-source registration

(d) An framework for cross-source point cloud registration. Given two input point clouds, a registration framework is designed to overcome cross-source challenges and estimate the final solution $\mathbf{T}$.

Advantages:
Integration of Advantages from Multiple Sensors: The cross-source point cloud registration utilizes the characteristics of multiple sensors, providing detailed 3D visual information for tasks such as augmented reality and construction.

Disadvantages:
Low Accuracy and High Time Complexity，Low Technological Maturity and Insufficient Research.

Cross-source registration

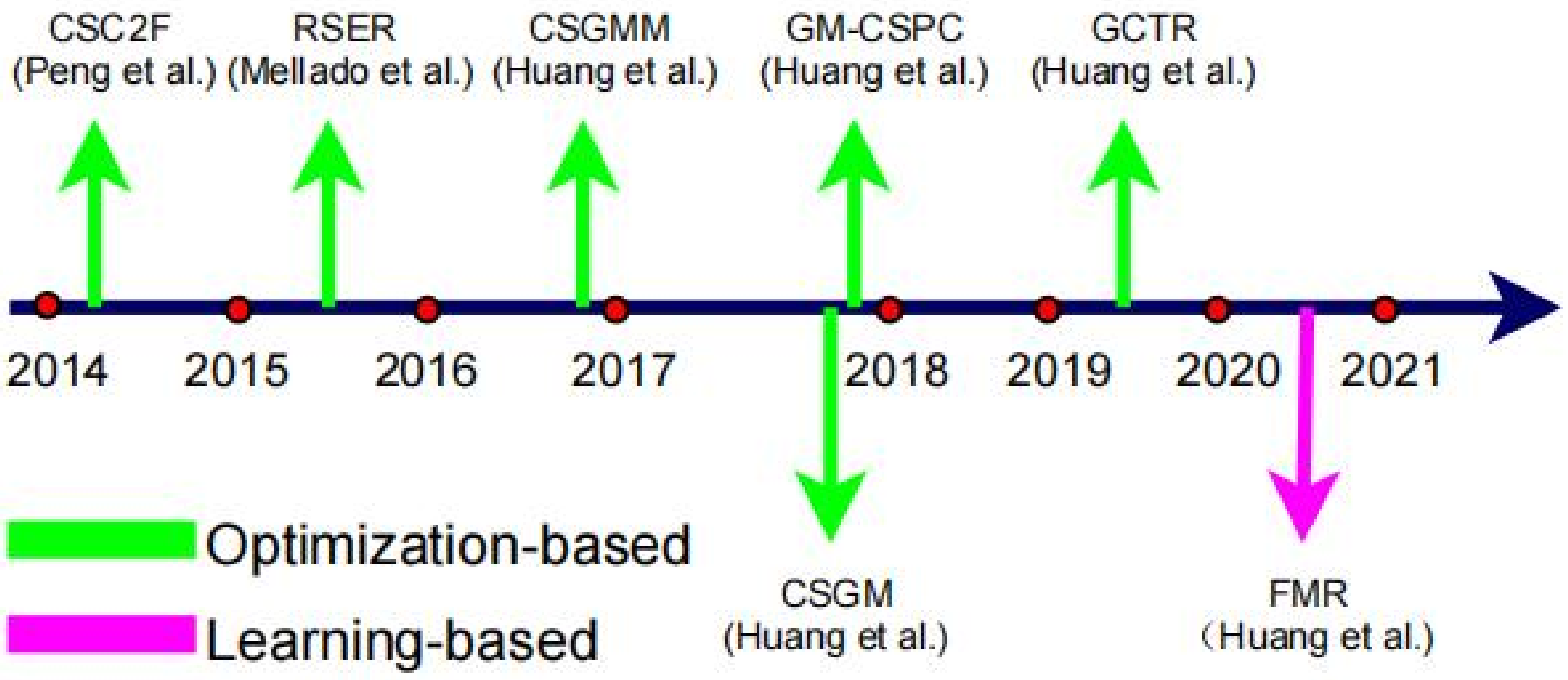


Fig. 9: Chronological overview of the most relevant cross-source point cloud registration methods.

Optimization-based methods

Optimization-based point cloud registration methods employ complex strategies to address cross-source registration challenges. While these methods guarantee convergence and generalization, they require high computational costs and exhibit varying performance across datasets.

Learning-based methods

FMR is the first learningbased method to solve the cross-source point cloud registration. This method combines the optimization and deep neural network and estimates the transformation by minimizing the global feature difference.

The connection between optimization-based methods and deep learning

The connections between deep learning and optimizationbased methods are: the deep learning technique could serve as a feature extraction tool to replace the original point coordinate. The conventional optimization could provide a theoretical guarantee for the convergence.

Evaluations

Dataset	sensor	sceneNum	indoor	outdoor	dense	sparse	ground-truth	xyz	corlor
3DMatch	depth	56	✓	×	✓	×	synthetic	✓	✓
KITTI	LiDAR	8	×	✓	×	✓	synthetic	✓	×
ETHdata	LiDAR	8	×	✓	×	✓	synthetic	✓	✓
3DCSR	Indoor	21	✓	✓	✓	✓	manual	✓	✓

Table I: Summary of existing same-source and cross-source dataset.

Same-source dataset

Methods	Average Recall	Thresholds
ICP(p2point)[120]	6.04	TE(0.3m),RE(15°)
ICP(p2plane)[120]	6.59	TE(0.3m),RE(15°)
Super4PCS[68]	21.6	TE(0.3m),RE(15°)
GO-ICP[106]	22.9	TE(0.3m),RE(15°)
FGR[119]	42.7	TE(0.3m),RE(15°)
RANSAC [86]	66.1	TE(0.3m),RE(15°)
SpinImage [46]	34	rmseP(0.2m)
SHOT [88]	27	rmseP(0.2m)
FPFH [86]	40	rmseP(0.2m
USC [93]	43	rmseP(0.2m)
PointNet [80]	48	rmseP(0.2m)
CGF [50]	56	rmseP(0.2m)
3DMatch [114]	67	rmseP(0.2m)
PPFNet [19]	71	rmseP(0.2m)
FCGF [17]	82	rmseP(0.2m)
DGR [16]	91.3	TE(0.3m),RE(15°)
PointNetLK [3]	1.61	TE(0.3m),RE(15°)
DCP [99]	3.22	TE(0.3m),RE(15°)

Table II: Comparison on 3DMatch datasets.

Methods	Average Recall	Thresholds
FGR [119]	0.2	TE(0.6m),RE(5°)
RANSAC [86]	34.2	TE(0.6m),RE(5°)
FCGF [17]	98.2	TE(0.6m),RE(5°)
DGR [16]	98.0	TE(0.6m),RE(5°)
FPFH [86]	58.95	TE(2m),RE(5°)
USC [93]	78.24	TE(2m),RE(5°)
CGF [50]	87.81	TE(2m),RE(5°)
3DMatch [114]	83.94	TE(2m),RE(5°)
3DFeatNet [110]	95.97	TE(2m),RE(5°)

Table III: Comparison on KITTI datasets.

Methods	Average Recall	Thresholds
FPFH [86]	67	TE(2m),RE(5°)
USC [93]	100	TE(2m),RE(5°)
CGF [50]	92.1	TE(2m),RE(5°)
3DMatch [114]	33.3	TE(2m),RE(5°)
3DFeatNet [110]	95.2	TE(2m),RE(5°)

Table IV: Comparison on ETHdata datasets.

New cross-source benchmark

In this paper, we introduce a benchmark dataset for crosssource point cloud registration to bridge this gap. Specifically, the dataset is captured using recent popular sensors: **LiDAR, Kinect and camera sensors**. In total, **202 pairs of point clouds**, where two scenes are captured using Kinect and RGB camera, 19 scenes are acquired from LiDAR and Kinect sensors.

Benchmark dataset: 3DCSR: We have two kinds of cross-source point cloud:
(1) Kinect and Lidar and (2) Kinect and 3D reconstruction.

New cross-source benchmark

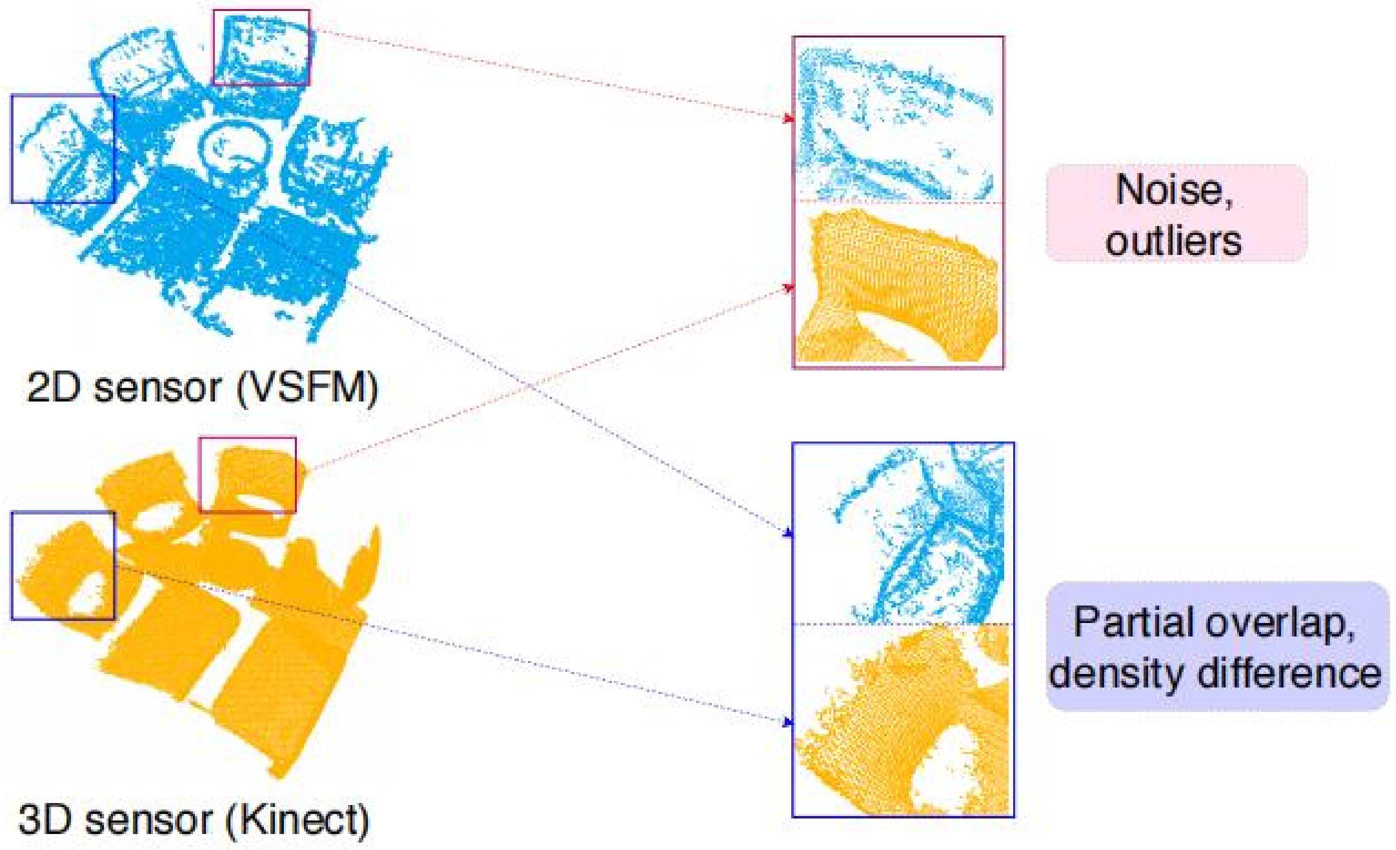


Fig. 10: An example shows the challenges of cross-source point clouds. Considerable noise, outlier, density difference and partial overlap universally exist in the cross-source pair.

New cross-source benchmark

Type	Method	Recall	TE	RE(deg)	Time(s)
Same source	FGR	1.49%	0.07	10.74	2.23
	PointnetLK	0.50%	0.09	12.54	2.25
	FMR	17.8%	0.10	4.66	0.28
	DGR	36.6%	0.04	4.26	0.87
Cross source	[77]	24.3%	0.38	5.71	0.19
	[42]	1.0%	0.71	8.57	18.1
	[69]	3.47%	0.13	8.30	0.03
	GCTR[39]	0.50%	0.17	7.46	15.8

Table V: Quantitative comparisons on the cross-source dataset.

Since [77] uses ICP, [42] uses Gaussian mixture model alignment, [69] uses RANSAC to solve the cross-source registration problem

Applications

- Construction eg: BIM (Building Information Modelling)
- Mining space
- Autonomous driving
- Robotics
- Geological and geotechnical data for geomechanical analysis

Open issues and future directions

- Robust and accurate registration
- Efficiency
- Partial overlap
- Fusion of deep learning and registration mathematical theories

Thank You!

Avenida da Universidade, Taipa, Macau, China
Tel : (853) 8822 8833 Fax : (853) 8822 8822
Email : info@um.edu.mo Website : www.um.edu.mo

澳大