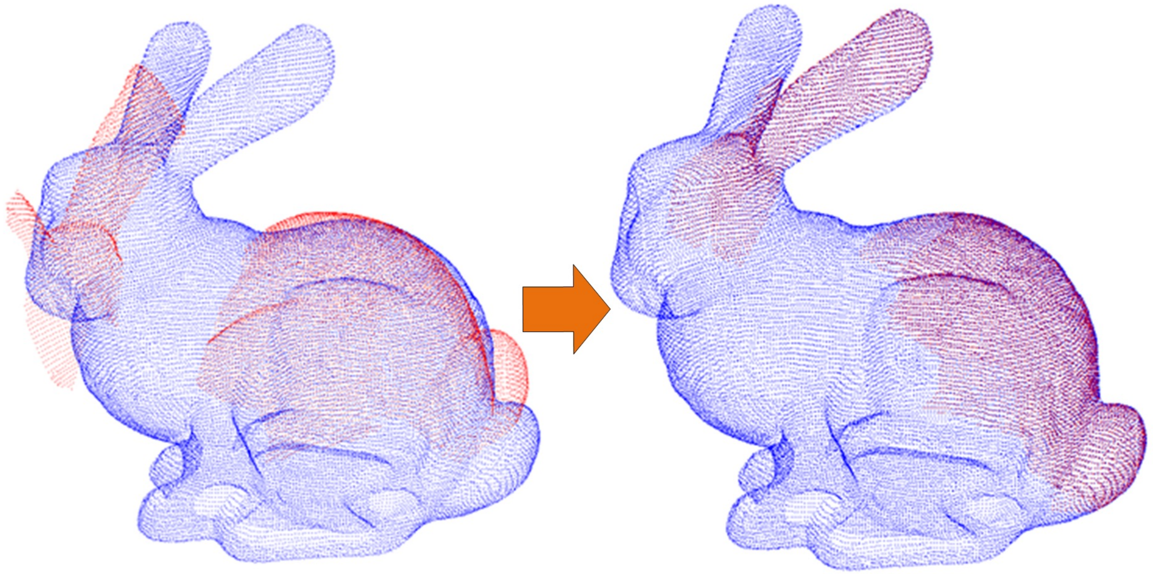
The Survey on Project：3D Reconstruction



姜一洲

jiang-yz19@mails.tsinghua.edu.cn

Contents

1 Problem Restatement & Understanding

Question：*Specifically, given a pair of 3D models, each 3D model has per-point coordinates, RGB color, semantic labels and instance labels. The algorithm will output:*

(1) *Whether the pair of models are the same place?*

(2) *If* (1) *is true, return the transformation matrix to align the two models.*

3D Loop Closure

How to combine point cloud and semantic labels to judge the same place

(*loop closure in SLAM*)

Point Cloud Registration

How to use all the information to get the correct transformation matrix

(*pose estimation in SLAM*)

2 The Survey on Point Cloud Registration

Introduction： Registration is a transformation estimation problem between two point clouds, which has a unique and critical role in numerous computer vision applications

This part will be：

——> 2.1 Optimization-based Registration

——> 2.2 Feature-learning Registration

——> 2.3 End-to-end learning Registration

——> 2.4 Coarse Registration and Fine Registration

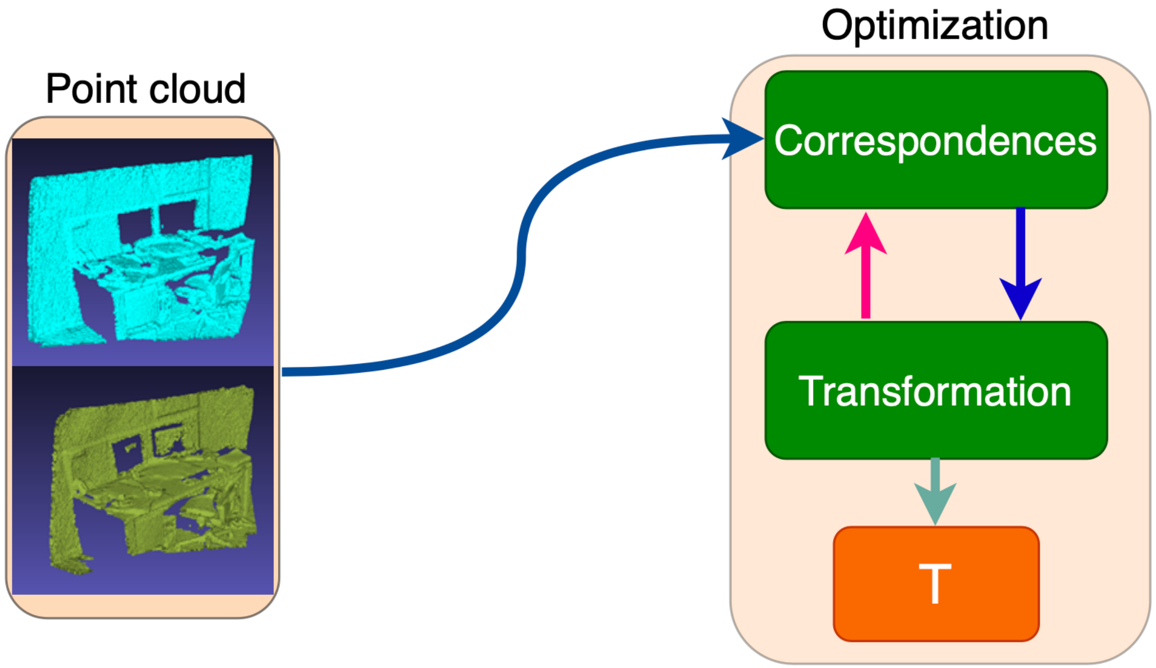
——> 2.5 Cross-Source Point Cloud Registration

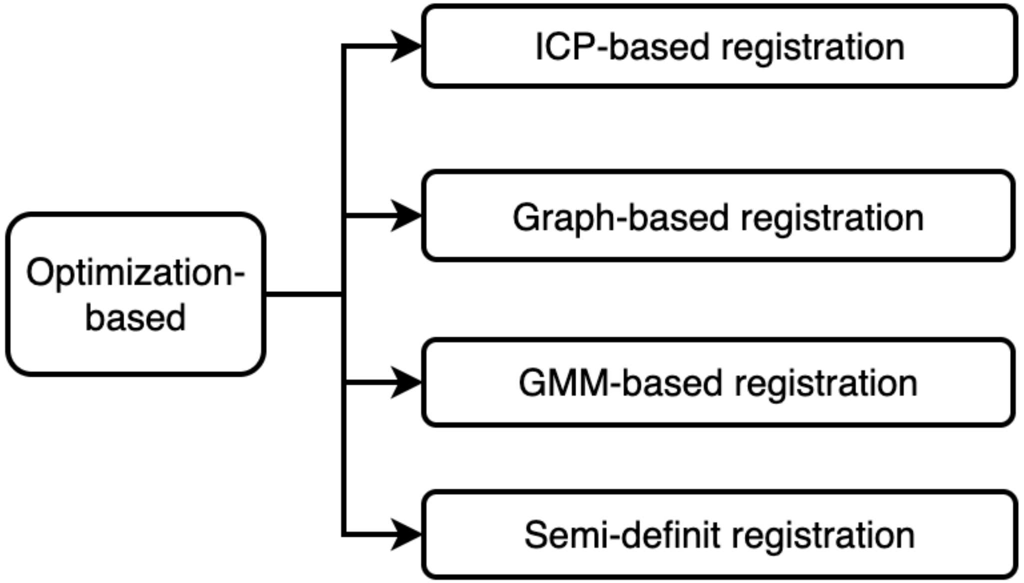
2.1: Optimization-based Registration

Given two input point clouds, the correspondences and transformation between these point clouds are iteratively estimated.

Strength : Strict mathematical theory can ensure convergence without training data

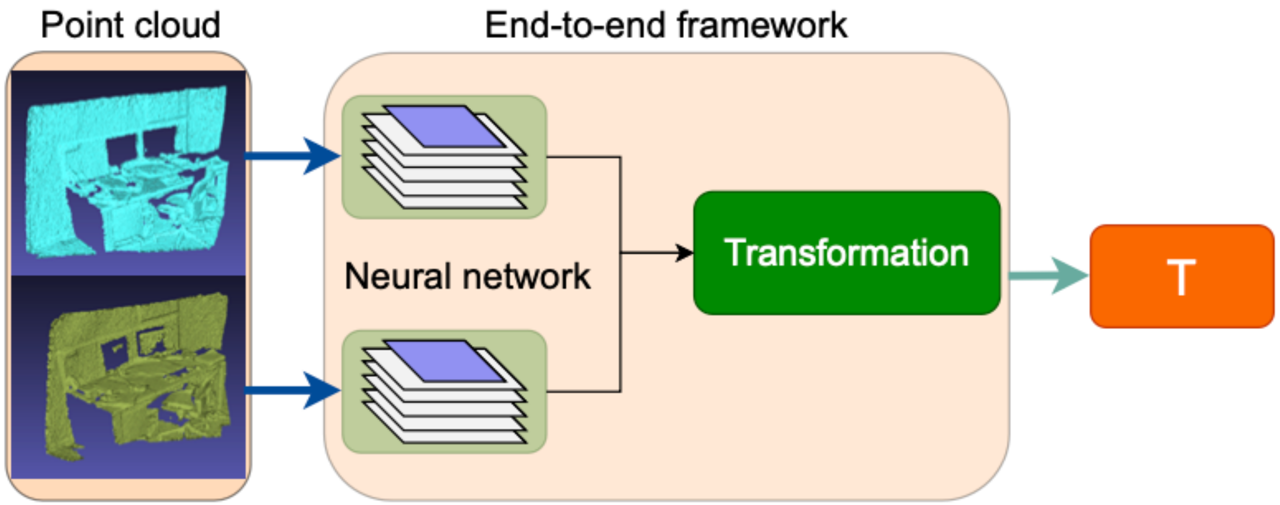
Weakness : Complex strategies are needed to overcome noise and outliers, and it is easy to fall into the local minimum





2.2 Feature-learning Registration

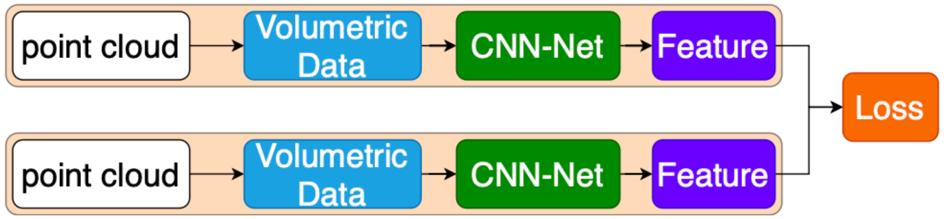
Given two input point clouds, the features are estimated using a deep neural network. Then, correspondence and transformation estimation run iteratively to estimate T



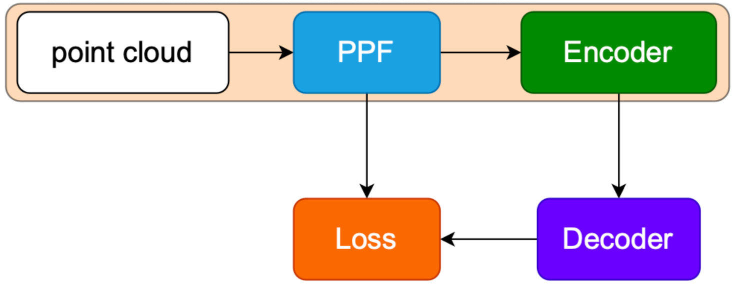
Strength：Feature extraction is the key to accurate estimation and registration using depth features

Weakness：It requires large training data, large memory and computing consumption

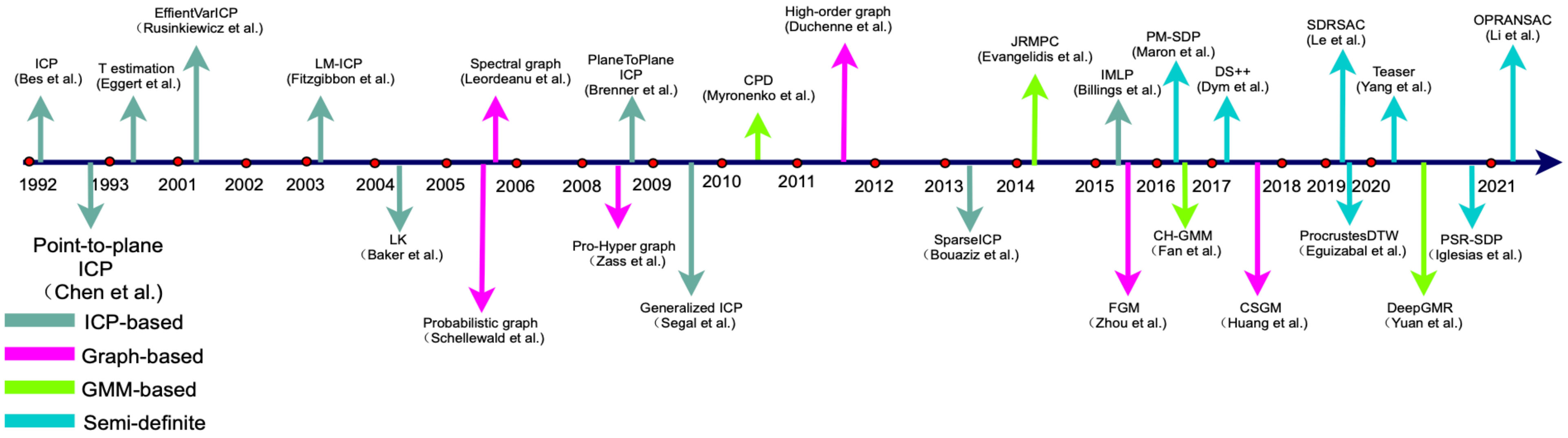
A. Learning on volumetric data：3DMatch



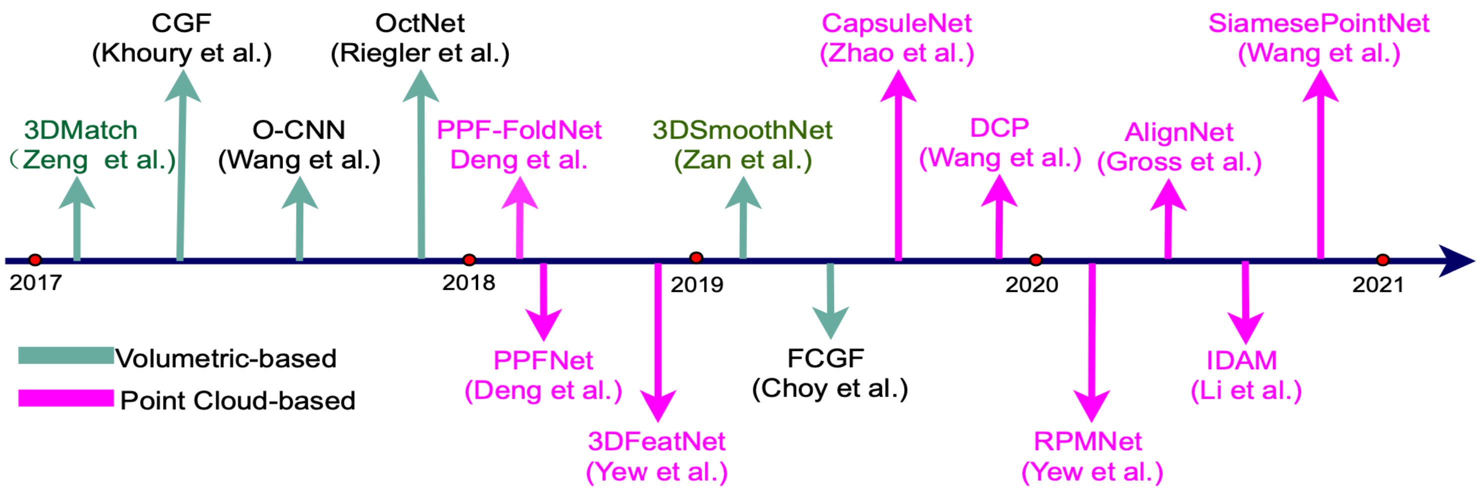
B. Learning on Cloud Point：PPF-Net



2.1 Optimization-based Registration Timeline

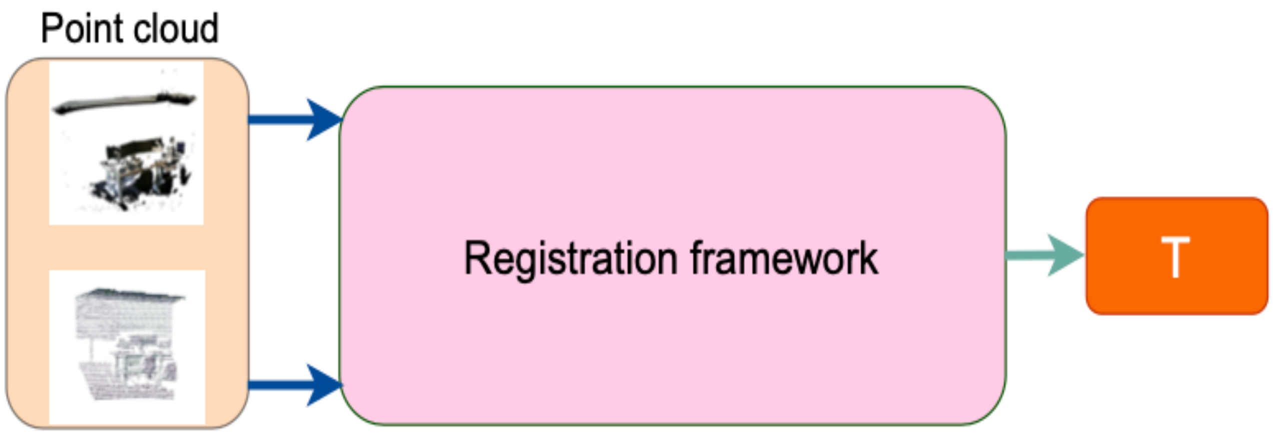


2.2 Feature-learning Registration Timeline



2.3 End-to-end learning Registration

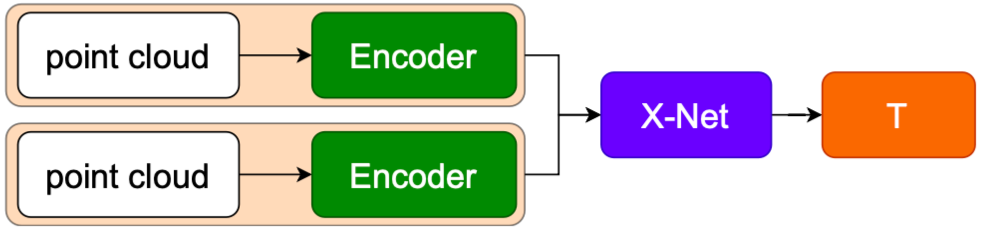
An end-to-end learning-based framework for point cloud registra- tion. Given two input point clouds, an end-to-end framework is used to estimate T



Strength： dimensionality reduced information directly input into the neural network directly obtain the T matrix

Weakness： need more training data to support the neural network. We can use network-based combine with optimization to obtain better results

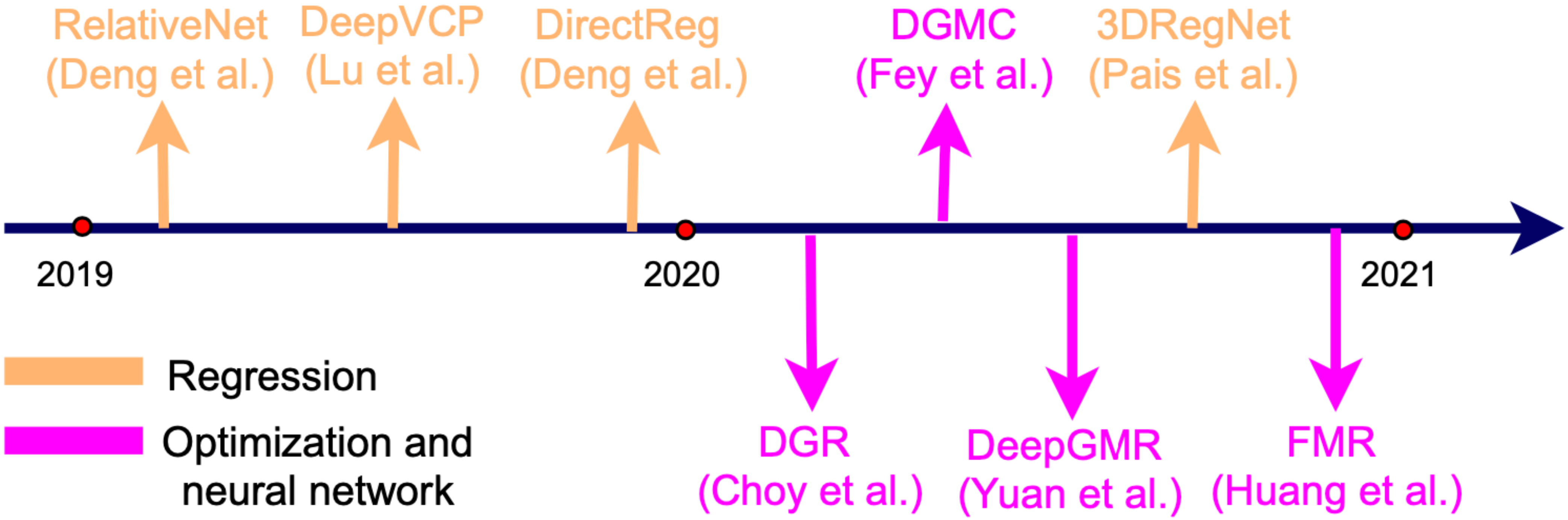
A. Registration by regression： *framework*



B. Relative Research：

DeepVCR、*PointnetLK*、*DeepGMR*、*etc*

2.3 End-to-End Learning Registration Timeline



2.4 Coarse Registration and Fine Registration

Coarse Registration:

LORAX:

Using neural network to compress and extract features or descriptor, then use descriptor position to do Coarse Registration

4-Points Congruent Sets for Robust Pairwise Surface Registration :

Find the best rigid transformation between two point sets, to get most points have the distance between two points is less than δ

Other Methods :

SK-4PCS (Semantic Keypoint 4-Points Congruent Sets

Super 4PCS(Super 4-Points Congruent Sets)

G-4PCS (Generalized 4-points congruent sets)

2.4 Coarse Registration and Fine Registration

Fine Registration: → **ICP(** Iterative Closest Point **):**

Correspondence Estimate:

——> Point-Point : $Nearest\ neighbor$

——> Point-Plane : $argmin\{\sum \omega_i || n_i \cdot (x_i - (Ry_i + t)) ||^2\}$

——> Plane-Plane : $argmin\{\sum || nx_i - (Rny_i + t) ||^2\}$

Transformation Estimate:

——> SVD-based : $get\ closed\ form\ solution\ but\ large\ consumption$

——> Lucas-Kanade Algorithm : $Gauss - Newton, Jacobian\ estimation$

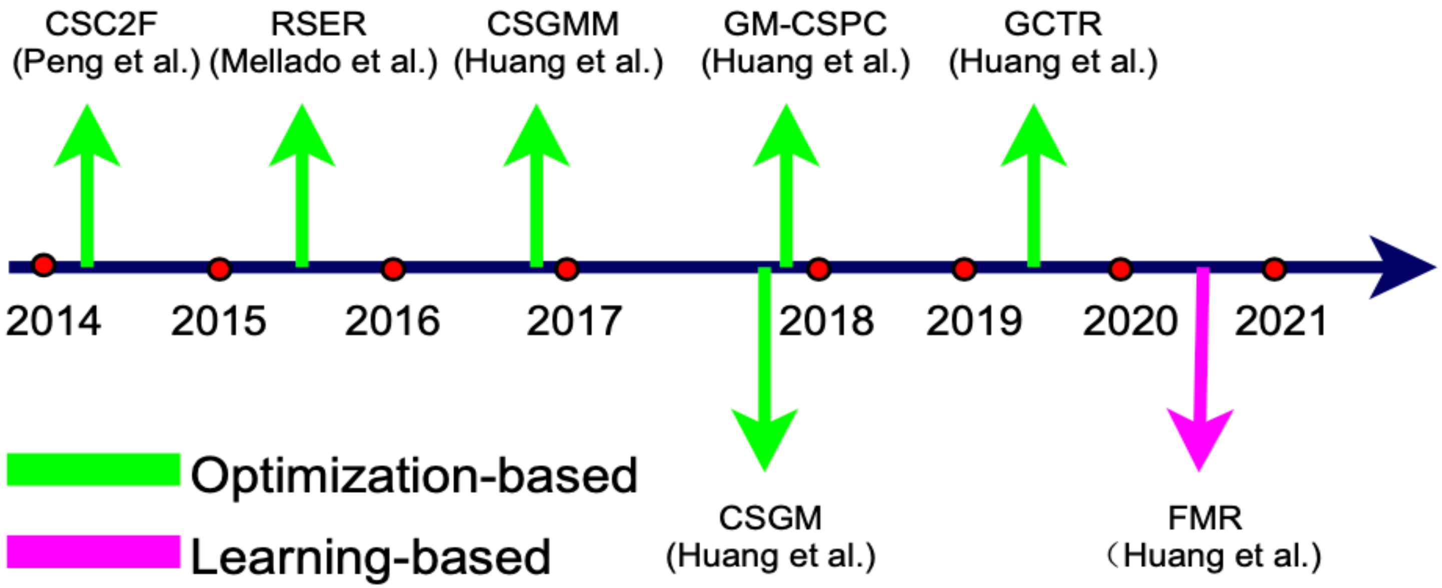
——> Procrustes Analysis : $highly\ dependent\ on\ corresponding\ searching$

2.5 Cross-Source Point Cloud Registration

Motivation:

Kinect generate dense point clouds but limited view range. Lidar has a long view range but sparse point clouds. Data fusion of these different kinds of 3D sensors combines their advantages and is a cross-source point cloud registration problem

Relevant cross-source methods:



3 The Survey on 3D Loop Closure

Introduction： 3D Loop Closure has always played a very important role in slam. Loop Closure is needed to increase constraints and reduce errors in real-time 3D reconstruction. Real-time 3D loop closure needs to be able to extract features and match them better and faster

This part will be：

——> 3.1 Loop Closure：2D vs 3D

——> 3.2 DBoW(Distributed Bag of World)

——> 3.3 VLAD(Vector of Local Aggregated Descriptors)、NetVLAD

——> 3.4 PointNetVLAD（2018）

3.1 Loop Closure : 2D vs 3D

Feature Correspondence :

(1) 2D-2D : It is usually used for slam system initialization;

(2) 3D-2D : operation stage of slam system. PNP used to solve image and 3D structure estimation;

(3) 3D-3D : used to correct of cumulative error，loop closure;

Loop Closure:

——> 2D Loop Closure：

——> Feature Extraction：SIFT、ORB、VALD + DWoB etc.

——> Learning Method：NetVALD etc.

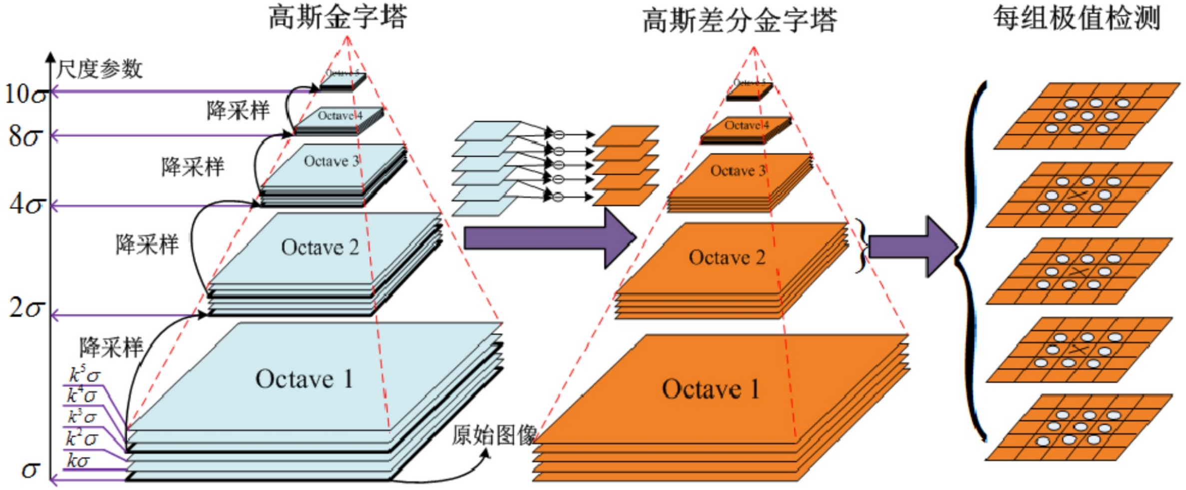
——> 3D Loop Closure：

——> Feature Extraction：3D SIFT + DWoB

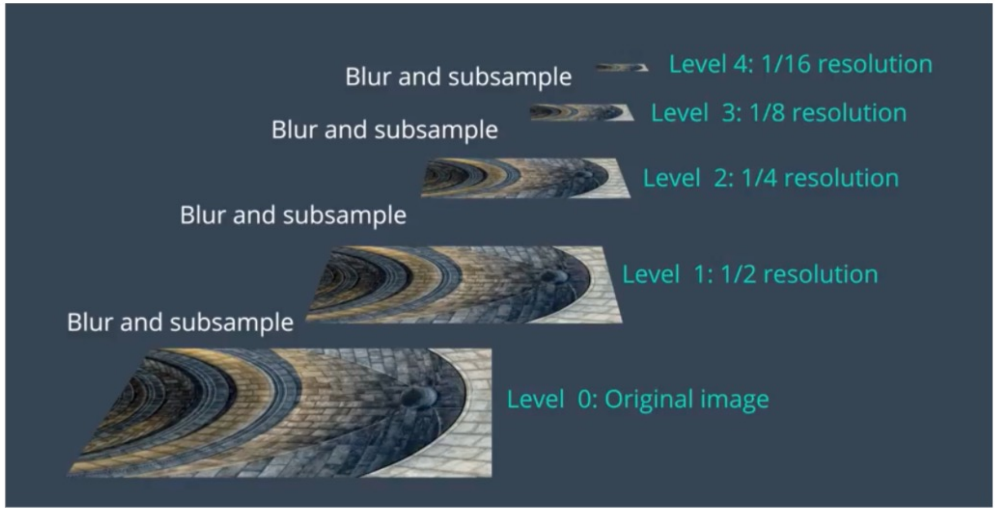
——> Learning Method：PointNetVLAD

3.2 Distrbuted Bag of Word

SIFT(Scale Invariant Feature Transform)



ORB(Oriented Fast and Rotated Brief)



Extract Scale/rotation

invariant features

DBoW (Distributed Bag of Word)

Algorithm :

K-means Clustering

Data Structure :

K-ary tree、K-ary tree、Chou-Liu tree

3.3 VLAD——>NetVLAD

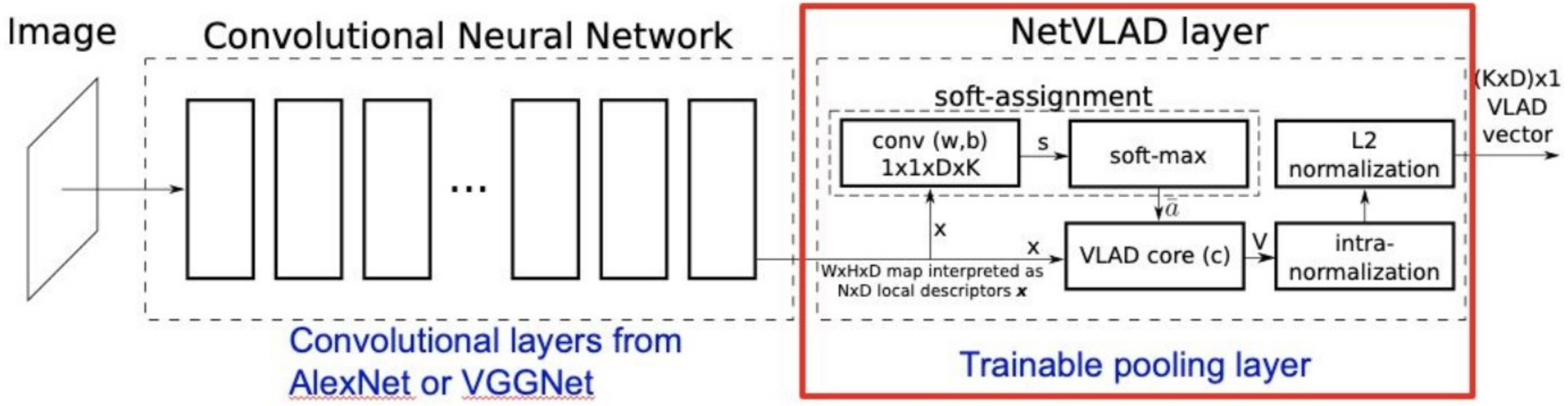
VLAD(Vector of Local Aggregated Descriptors)

Use neural network

Use VLAD function to get the global feature of Image

Also need to apply k-means clustering algorithm

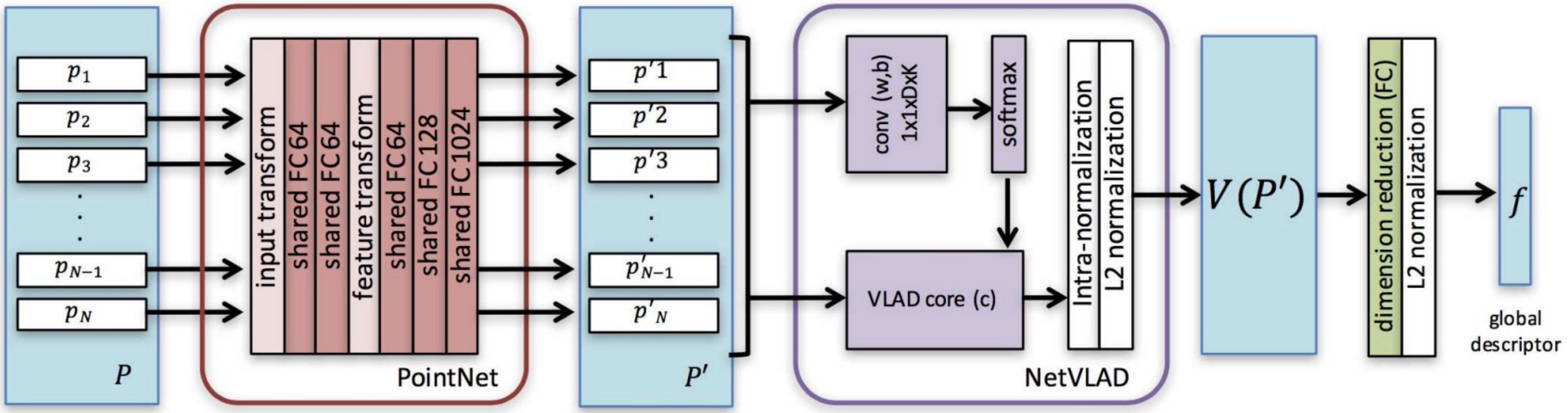
NetVLAD Structure :



3.4 PointNetVLAD

PointNetVLAD is a combination/modification of the existing PointNet and NetVLAD, which allows end-to-end training and inference to extract the global descriptor from a given 3D point cloud.

We create benchmark datasets for point cloud based retrieval for place recognition, and the experimental results on these datasets show the feasibility



4 Proposed Solutions

$How\ to\ fully\ use\ the\ semantic\ information?\ \ High - level\ and\ Abstract\ feature\,?$

(1) STFT DBoW + ICP-Semantic-based

——> **First Step :** $Based\ on\ semantic\ feature\ and\ position\ judge\ the\ same$

——> **Second Step :** $"\,Semantic\ Nearest\ Neighbor\,"\ —> ICP(try\ other\ methods)$

(2) Learning + End-to-End Network

——> **First Step :** $Based\ on\ semantic\ feature\ and\ position\ judge\ the\ same$

——> **Second Step :** $PointNetLK(try\ other\ methods)$

Reference

[1] A comprehensive survey on point cloud registration

[2] Convex Hull indexed Gaussian Mixture Model (CH-GMM) for 3D Point Set Registration

[3] SDRSAC Semidefinite Based Randomized Approach for Robust Point Cloud Registration Without

[4] 3DMatch_Learning_Local_CVPR_2017_paper

[5] 3DSmoothNet：Gojcic_The_Perfect_Match_3D_Point_Cloud_Matching_With_Smoothed_Densities

[6] PointNet_Deep_Learning_CVPR_2017_paper

[7] RPM-Net_Robust_Point_Matching_Using_Learned_Features_CVPR_2020_paper

[8] DeepVCP_An_End-to-End_Deep_Neural_Network_for_Point_Cloud_Registration_ICCV_2019_paper

[9]PointNetLK_Robust__Efficient_Point_Cloud_Registration_Using_PointNet_CVPR_2019_paper

[10]DeepGMR：Learning Laterned