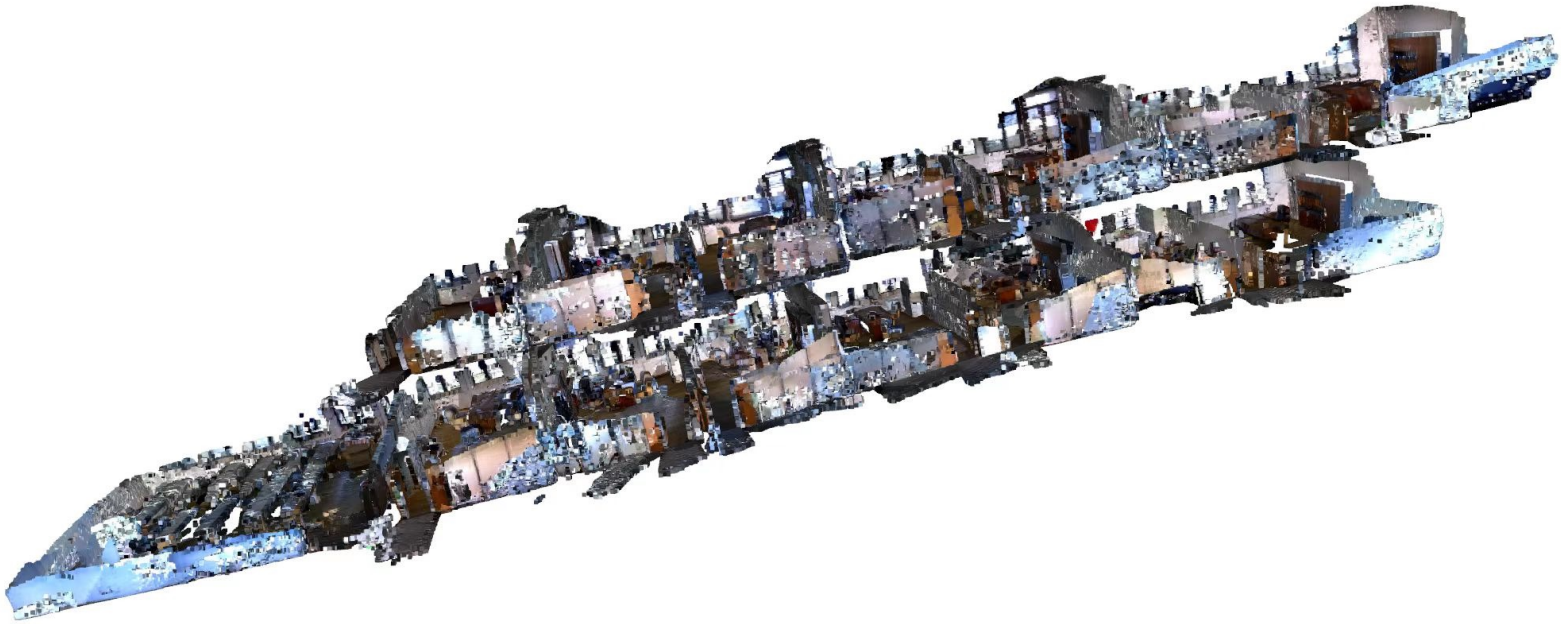
3D Reconstruction：Point Cloud & Loop Closure



姜一洲

jiang-yz19@mails.tsinghua.edu.cn

Contents

1 Project Question & Timeline

→ *Experiment Process? The total conclusion of past two weeks?*

2 Analysis of the Project Process

→ *Which Method or data to use? The details of the Results?*

3 Thoughts of the Project’s improvement

→ *How to better use semantic label? Less Consumption?*

1 Project Question

$\boldsymbol{Question\_1(Scannet)}$: $Specifically, given\ a\ pair\ of\ 3D\ models, each\ 3D\ model\ has\ per-point\ coordinates, RGB\ color, semantic\ labels\ and\ instance\ labels. The\ algorithm\ will\ output$:

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

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

$\boldsymbol{Question\_2(info)}$: $Given\ \ a\ group\ of\ 3D\ point\ clouds(mostly\ rooms)\ in\ different\ coordinate\ systems\ to\ reconstruct\ a\ complete\ scene: more\ accurate\ \&\ less\ time\ consumption$:

(1) $Firstly\ select\ one\ as\ the\ benchmark, then\ loop\ closure$.

(2) $If\ (1)\ find\ the\ corresponding\ two, transform\ \ to\ reach\ the\ same\ coordinates$.

1 Project Timeline

- ICP(only) to handle the point cloud registration(on data from Scannet)

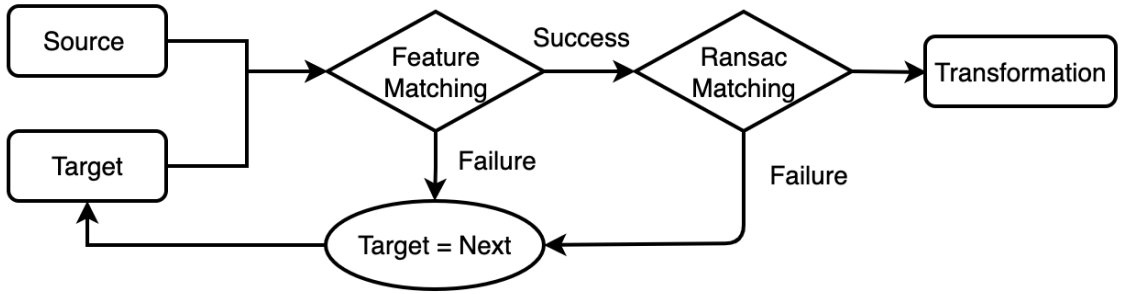
  Problem : (1) large time consumption each registration; (2) sensitive to the origin pose

- Firstly voxel_down point cloud; Secondly use RANSAC to handle the point Cloud registration;

  Randomly choose room, use fitness to judge the same and get the transformation

  Problem : (1) we get the O(N) complexity, large time cost;

- Use semantic label to generate feature vector, the total algorithm shows left



- Another dataset : to reconstruct the 3D scene

  Problem : (1) O(N^2) complexity; (2) Feature Matching cannot get the good result

2 Analysis of the Project Process

Introduction： This part will analyze the specific details of the two questions above, especially the algorithm, complexity analysis, failure rate etc.

This part will be：

——> 2.1 Compare ICP and RANSAC (Algorithm Principle)

——> 2.2 Analysis with Question_1 (Data Set： scannet)

——> 2.3 Analysis with Question_2 (Data Set： info_room)

——> 2.4 The Final Result

2.1 Compare ICP and RANSAC

1. $ICP:\ problem\ definition\ \&\ solution\ methods$

$Find\ two\ group\ of\ Point(nearest)$: $P = \{p_1, p_2 \dots, p_n\}$ $P' = \{p_1', p_2' \dots, p_n'\}$

$Solving\ Optimization\ Problems$: $min_{R,t}\, J = \frac{1}{2}\sum_{i=1}^{n} ||(p_i - (Rp_i' + t)||^2$

$Solution\ Methods$: → $(1)\ SVD; (2)\ Lie\ Algebraic\ Perturbation\ Model$

2. $RANSAC:\ problem\ definition\ \&\ solution\ methods$

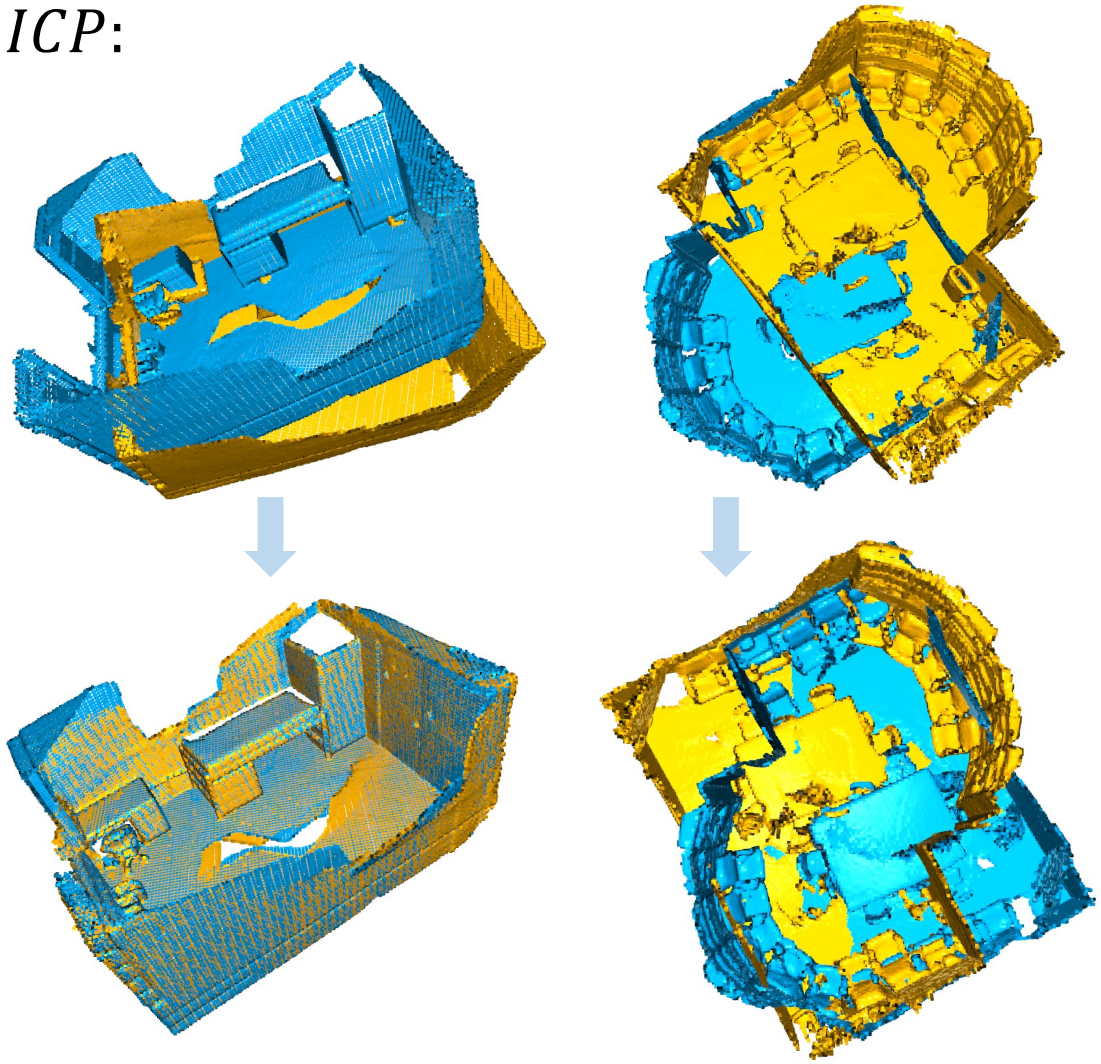
$Extract\ Feature\ from\ points\ cloud$: $FPFH\ feature$: $Source\ and\ Target$

$RANSAC\ iteration$: → $Finding\ nearest\ neighbor\ in\ feature\ space$

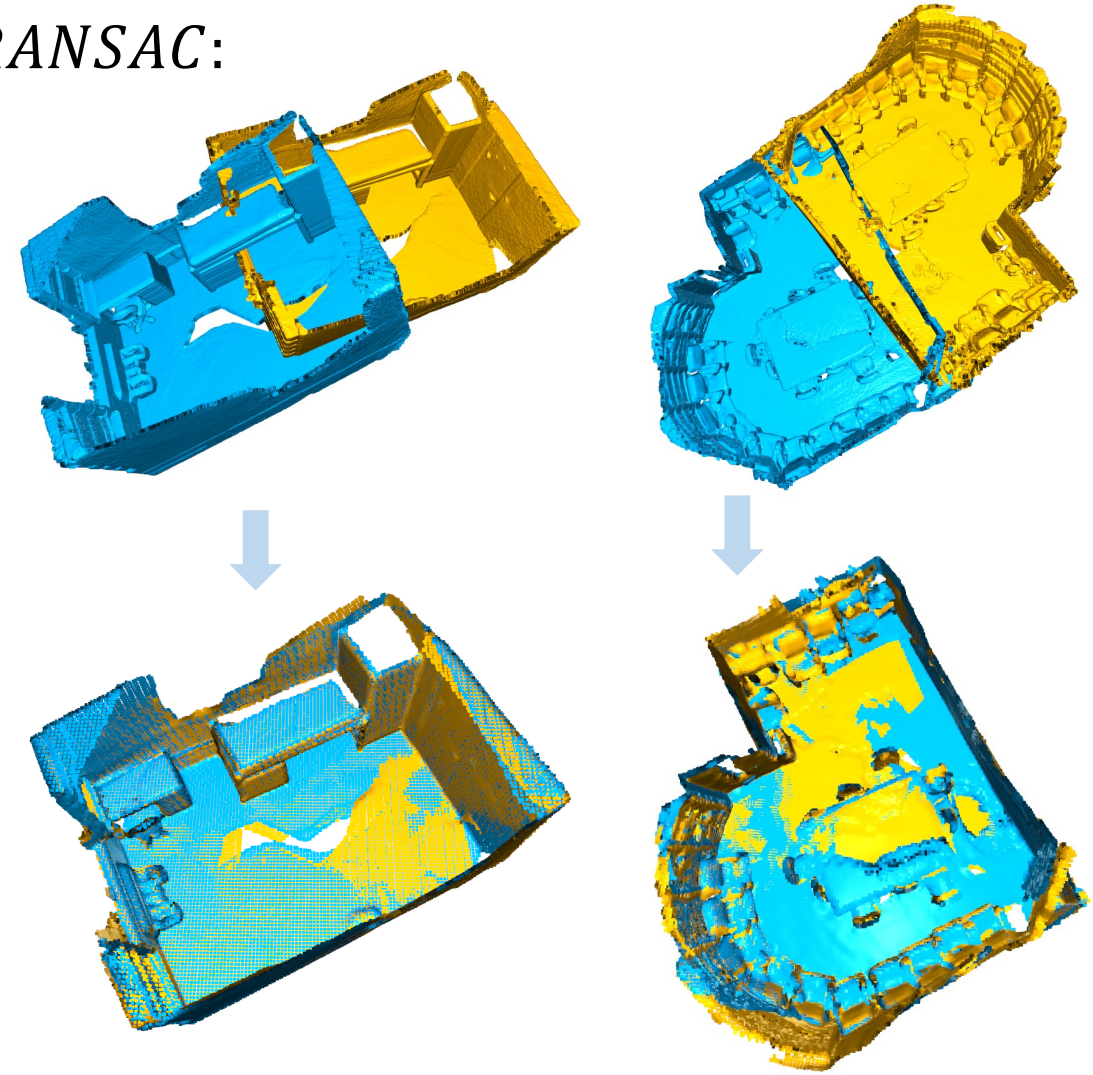
→ $Using\ pruning\ algorithm\ to\ speed\ up$

2.1 Compare ICP and RANSAC: initial pose

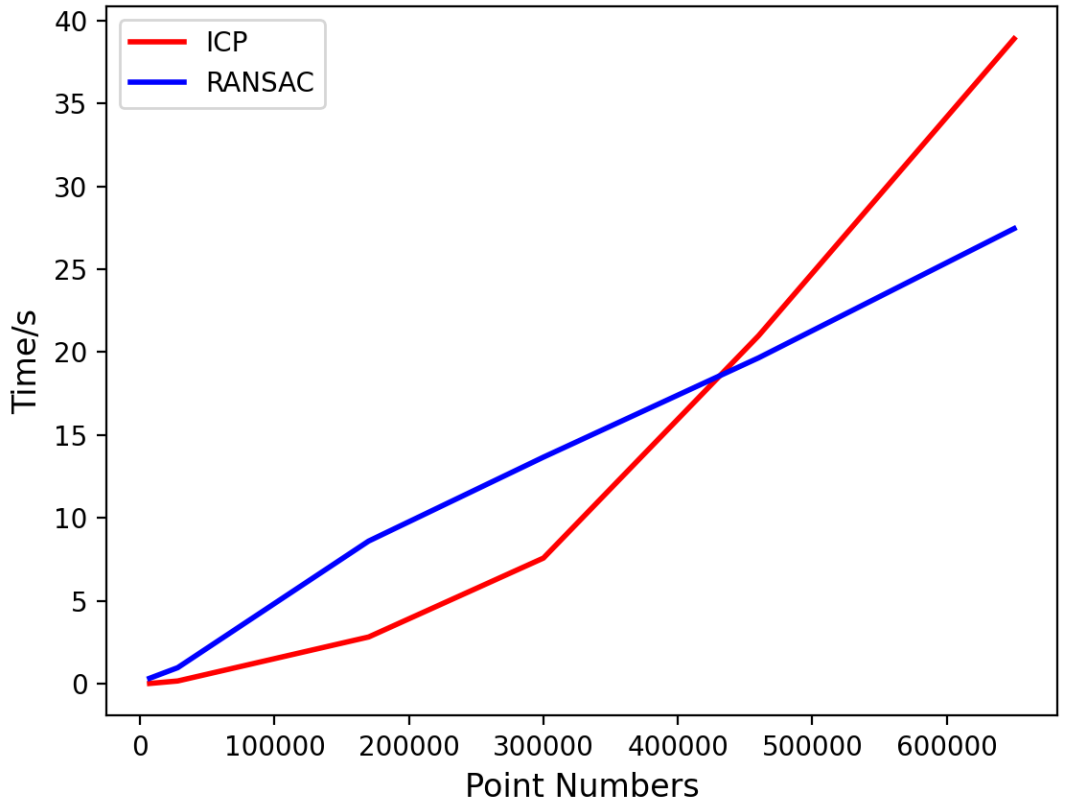
ICP:



$RANSAC$:



2.1 Compare ICP and RANSAC: time cost



1. Time Cost Comparing: As the number of points grow

   ICP time cost began to speed ≈ O(N^(3/2));

   RANSAC increase by the speed that ≈ O(N);

2. *Problem*:

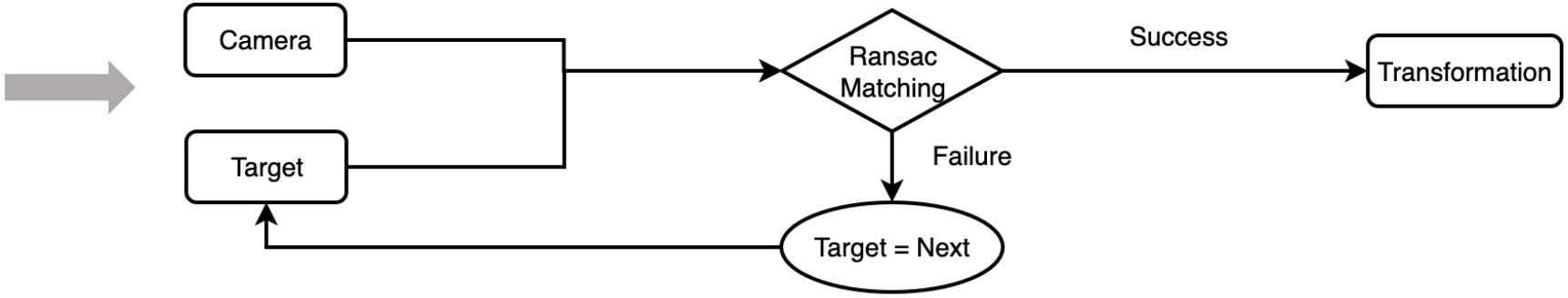
   number large, time cost large, need voxel down

   when voxel low, the result of ransac not good,

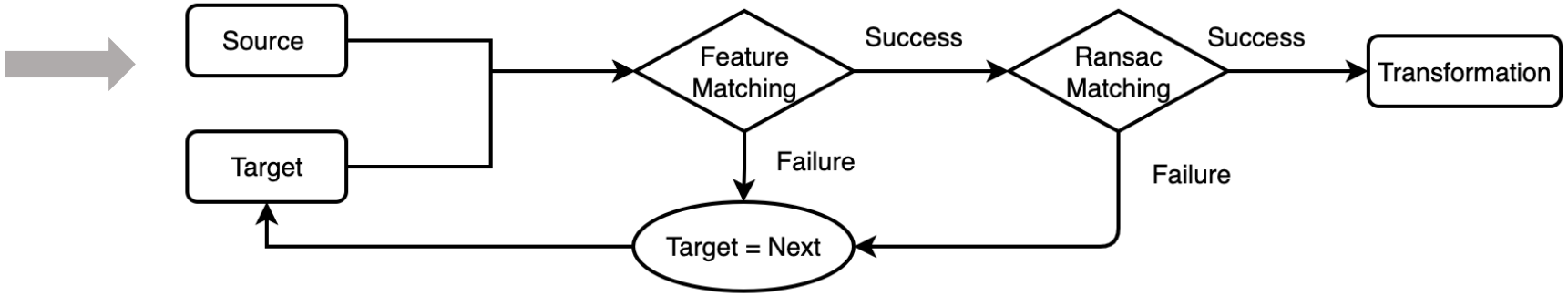
   because hard to extract the feature

2.2 Analysis with Question_1(based on scannet)

(1)Randomly choose room, use fitness to judge the same and get the transformation



(2)Randomly choose room, use fitness to judge the same and get the transformation



2.2 Analysis with Question_1(based on scannet)

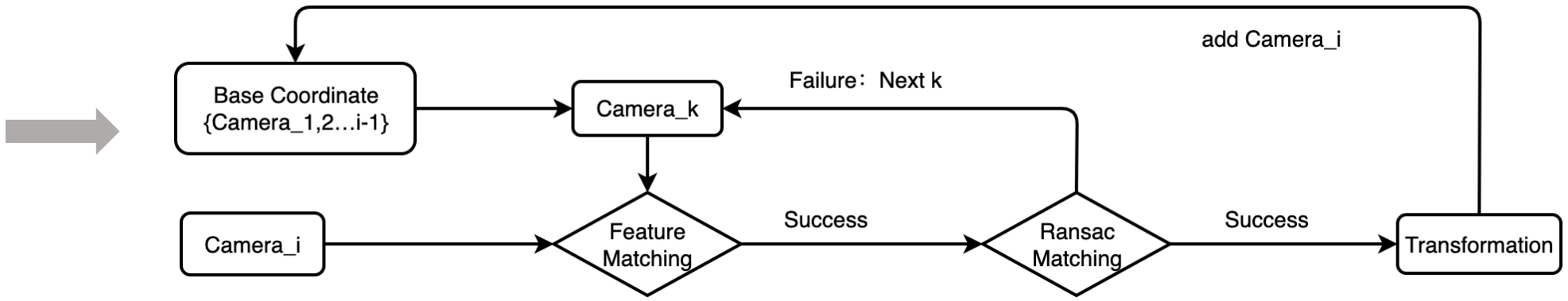
Compare the Ransac with Feature + Ransac on Question1

Method	Complexity	Time cost	Accuracy	Failure case
Only Ransac (fitness judge)	O(N) (have N scenes)	$average: \frac{N}{2}$	$\approx 100\%$	≈ 0
Ransac & Feature Match	O(1) (have N scenes)	< 1	$\approx 100\%$	≈ 0

2.3 Analysis with Question_2(based on info_room)

Another dataset : to reconstruct the 3D scene use the scale room in different coordinate

⟶ We have 10 group, each group has different coordinate, we should detect the same room, and transform to the global coordinates



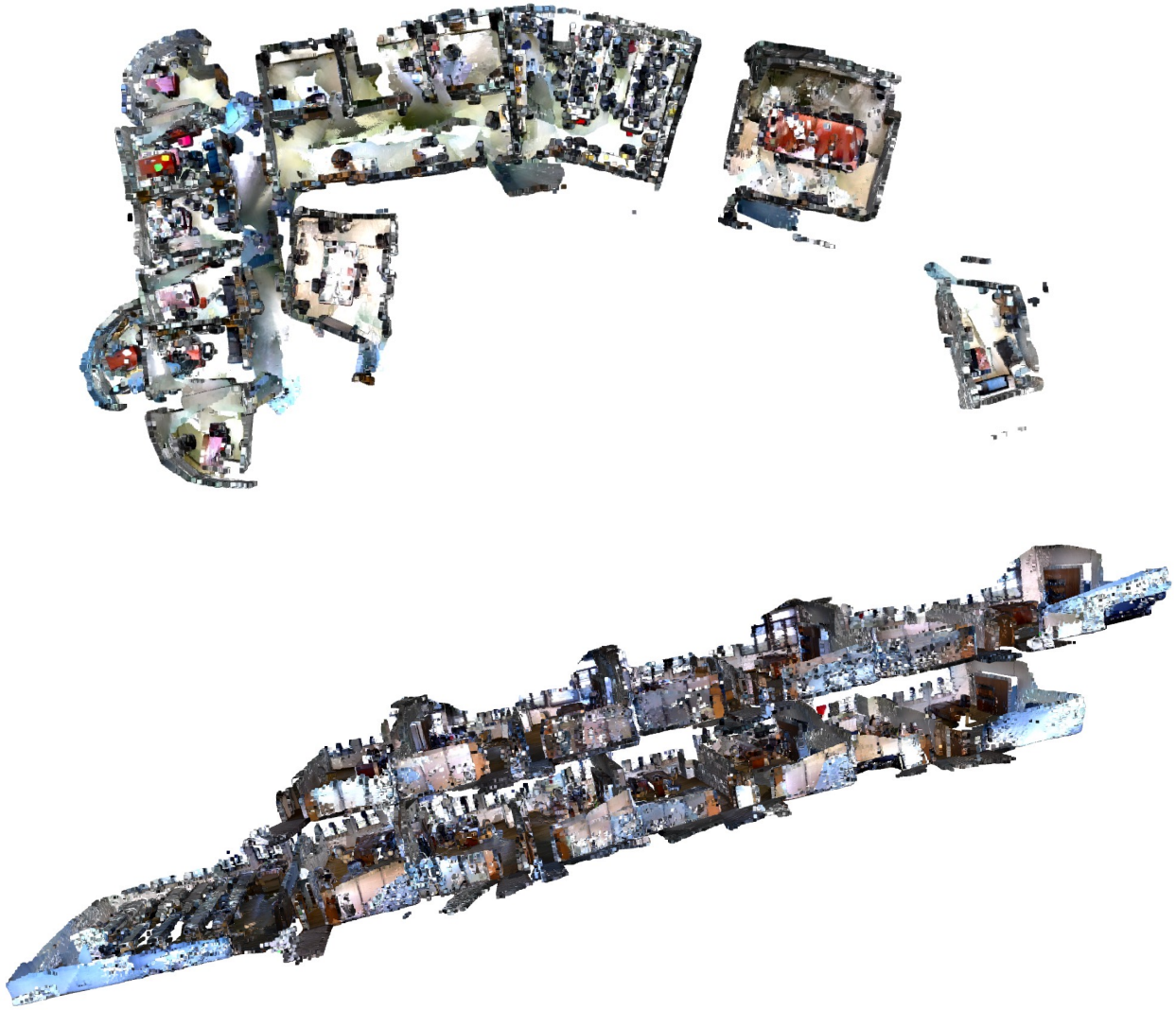
2.3 Analysis with Question_2(based on info_room)

Compare the Ransac with Feature + Ransac on Question2

Method	Complexity	Time cost	Failure case
Only Ransac (fitness judge)	$O(N^2)$ (have N rooms)	$average: \frac{N^2}{4}$	(1) Loop detected not matching; (2) Loop not detected matching;
Ransac & Feature Match	$\approx O(N^2)$ (have N rooms)	$average: \frac{N^2}{4} \times p$	Feature Match can reduce some failure case

* A useful method to reduce failure case: use ICP after the ransac

2.4 The Final Result



3 Thoughts of the Project’s improvement

Introduction：This part will show my thoughts of the project’s improvement, considering the time consumption and registration algorithm

This part will be：

——> 3.1 How to better organize semantic labels

——> 3.2 Can we find better point cloud registration

——> 3.3 Should we considering label generate time consumption

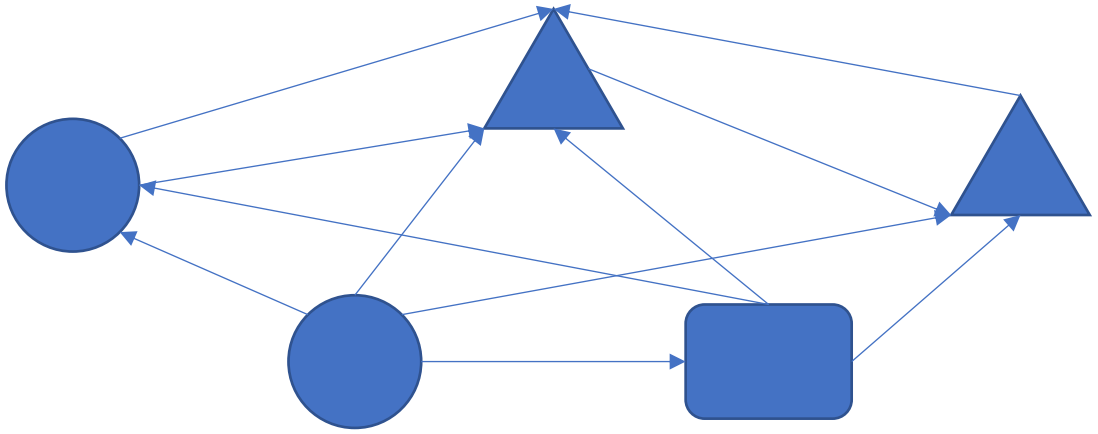
3.1 How to better organize semantic labels

In this algorithm, we just use instance semantic labels to generate the feature by quantity of objects and use 2-norm product to measure the similarity

How to better organize the semantic labels?

Maybe we can consider instance semantic & position, to construct a vector graph

Then use vector graph to define a new similarity(rotation invariant, Translation invariant…)



3.2 Can we find better point cloud registration

Idea1: combine RANSAC and ICP

RANSAC extract the feature and do the coarse registration(not sensitive to the initial pose)

ICP can get the point cloud registration more accurate(RANSAC give a good initial pose)

Bad effect: this will cause more time cost

Idea2: Can we mix semantic in point cloud registration

Semantic label on the scene is a high-level feature in a sense

RANSAC extract the FPFH feature, so can we combine the semantic with RANSAC

A consideration: can we get semantic label at the same time reach registration

3.3 Should we considering label generate time consumption

Now apply the algorithm based on the instance semantic label(already given), but if we cannot get the instance semantic label at first, or we should considering label generate time consumption

——> $estimate\ the\ complexity$: $each\ room\ just\ need\ one\ instance\ semantic\ segmentation, so\ the\ complexity\ will\ be\ \approx O(N)$

Also we should consider the idea2: Can we mix semantic in point cloud registration

Semantic label on the scene is a high-level feature in a sense

RANSAC extract the FPFH feature, so can we combine the semantic with RANSAC

A consideration: can we get semantic label at the same time reach registration

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