

Object-based Loop Closure with Directional Histogram Descriptor

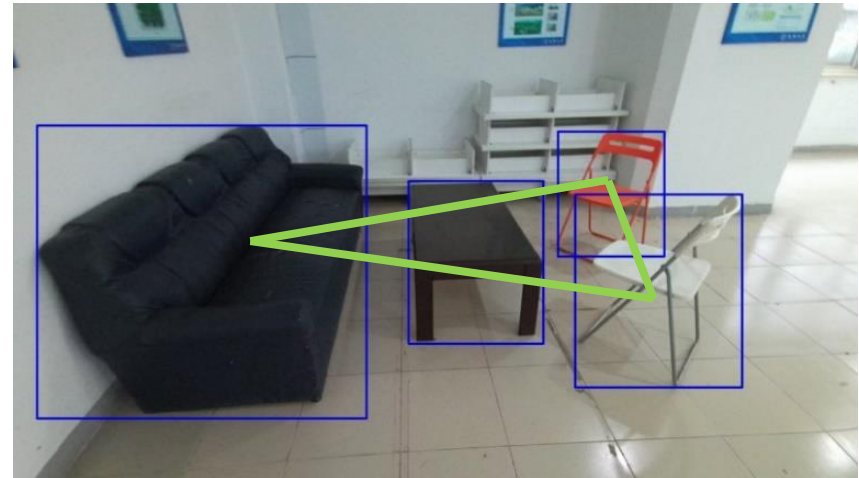
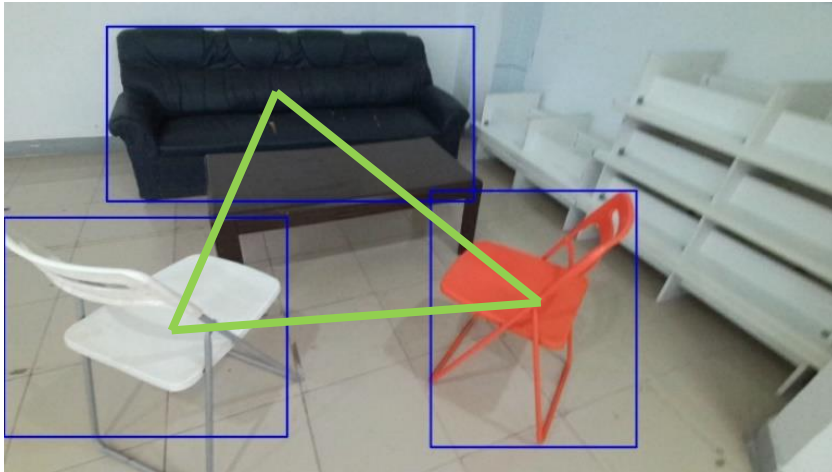
Benchun Zhou, Yongqi Meng, and Furmans Kai

Institute of Material Handling and Logistics, Karlsruhe Institute of Technology, Germany

2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)

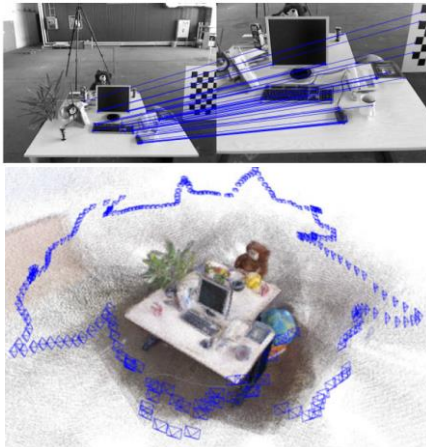
Background

- Mobile robots will produce inevitable accumulated errors because of sensor noise
- Traditional feature-based loop closure methods are sensitive to environment changes and failed to detect loop under large viewpoint difference.
- Research problem: How to effectively detect the loop under large viewpoint difference?

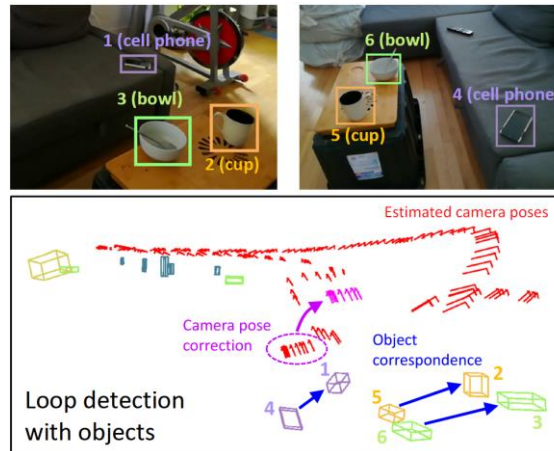


Related work

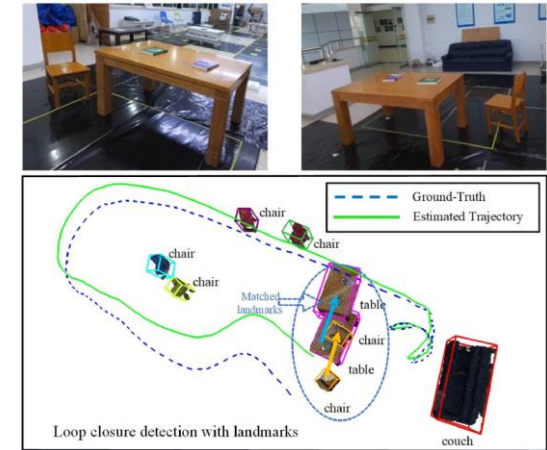
- Research problem: How to effectively detect the loop under large viewpoint difference?
 - feature-based loop closure: accurate, but sensitive to environment changes
 - object-based loop closure: robust to viewpoint changes but not accurate
 - our idea: feature-based + object-based loop closure



ORB SLAM 2: Loop detection with ORB features (Mur-Artal et al. 2017)



Loop detection with objects (Li, et al, 2020)

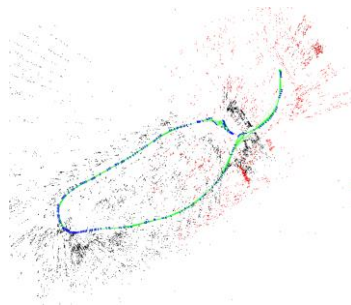
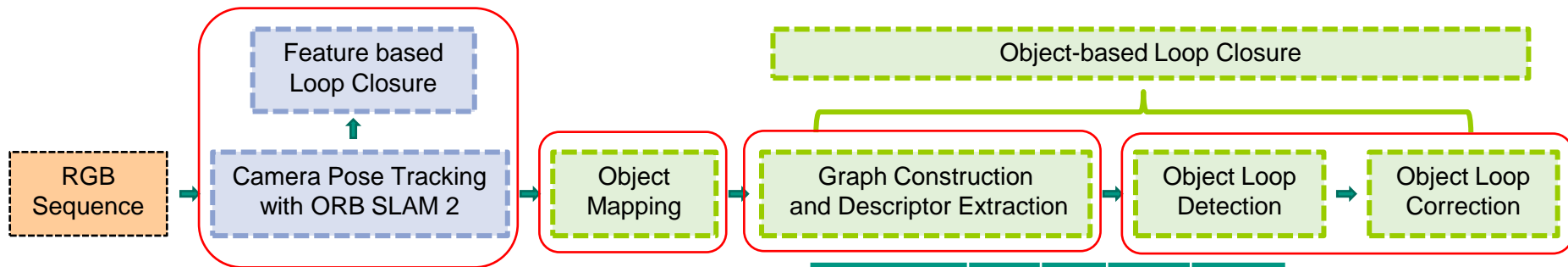


Loop detection with objects (Lin, et al, 2021)

Method: overall framework

Assumptions:

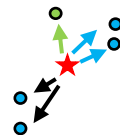
- 1) Use only RGB images.
- 2) Objects are static and, on the ground.
- 3) Objects can be viewed in different viewpoint.



Sparse Point Map

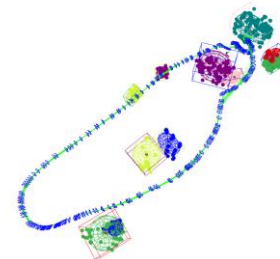


Sparse Object Map



Graph Example with Directional Histogram Descriptor

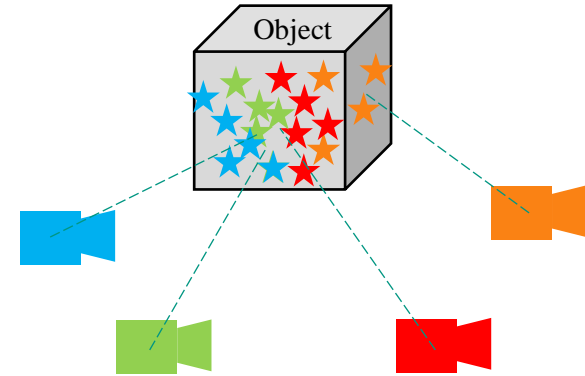
Direction Range	Chair	Sofa	Table	Book ...
$[-180^\circ, -135^\circ]$	1	0	0	0
$[-135^\circ, -90^\circ]$	1	0	0	0
...
$[0^\circ, 45^\circ]$	2	0	0	0
$[45^\circ, 90^\circ]$	0	0	0	0
$[90^\circ, 135^\circ]$	0	1	0	0
$[135^\circ, 180^\circ]$	0	0	0	0



Corrected the drift

Method: object mapping

- How to detect objects in the environment?
- Object Representation: 3D bounding box with class, translation, rotation, dimension, feature points
- Object Detection: YOLOv5, collect the feature points that belongs to the object.
- Object Association: In every frame, check detected objects has a large 2D IoU (>0.5) with mapped objects, if so, merge them, if not, add a new object to the map.



Object detection example

Method: graph construction

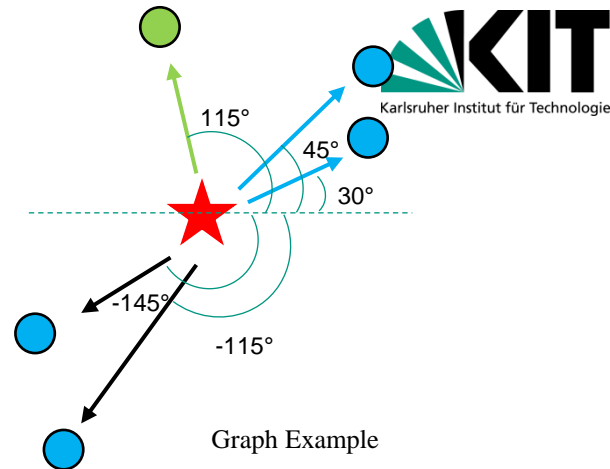
- How to represent the environment (not only the objects but also their relationship)?
- Graph $(G) = \{N, E\}$,
 - Node $(N) := \{ID, Class, Centroid, Descriptor, Others\}$,
 - Edge $(E) := \{distance_{ij}, yaw_{ij}\}$

Algorithm 2: directional histogram descriptor

Input: O: Object landmarks;

Output: V: directional histogram descriptor for O;

1. **for** o in objects **do**
2. neighbor \leftarrow find the closest 8 objects;
3. **for** n in neighbor of o **do**
4. direction: a \leftarrow calculate the direction(n,o);
5. label: ln \leftarrow record the object label;
6. histogram \leftarrow the histogram cell $V_o(a, ln) + 1$;
7. **end for**
8. add V_o to into V;
9. **end for**



Yaw Angle Range	Chair	Sofa	Table	Book
$[-180^\circ, -135^\circ]$	1	0	0	0
$[-135^\circ, -90^\circ]$	1	0	0	0
...
$[0^\circ, 45^\circ]$	2	0	0	0
$[45^\circ, 90^\circ]$	0	0	0	0
$[90^\circ, 135^\circ]$	0	1	0	0
$[135^\circ, 180^\circ]$	0	0	0	0

Directional Histogram Descriptor

Method: loop detection

- How to detect loop with object information?
- General check: same class and ID difference > 3

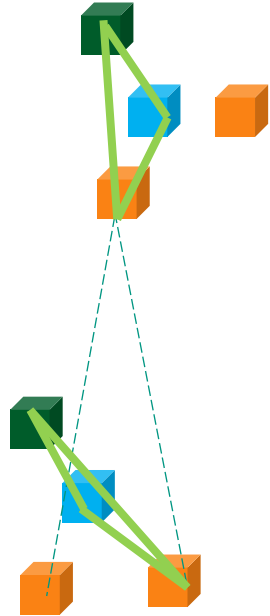
$$\eta(a, b) = |ID_a - ID_b| \quad (a \in L_a, b \in L_b)$$
- Graph check: the similarity score of their directional histogram descriptor > 0.8

$$\text{Score}(A, B) = \frac{\sum_{d=1}^{n_d} A_d \times B_d}{\sqrt{\sum_{d=1}^{n_d} (A_d)^2} \times \sqrt{\sum_{d=1}^{n_d} (B_d)^2}}$$

- Geometry check: at least 3 objects are matched, and their scale factor < 0.01

$$s_{12} = |dis_{a_1, a_2}| / |dis_{b_1, b_2}| \quad \delta = \sqrt{(s_{12} - \bar{s})^2 + (s_{23} - \bar{s})^2 + (s_{13} - \bar{s})^2}$$

- Duplicated check:



Geometry verification example

Method: loop correction

- How to correct the drift after the object loop is detected?
- General idea: compute a similarity transformation that satisfy:

$$O_a = sR * O_b + t$$

- However, use only object information may cause errors, because the object detection is not so accurate, instead, we take the associated map points and compute a scaled ICP.
- The scale ratio is computed by:

$$s = \frac{1}{n} \sum \frac{|dis_{a_1, a_2}|}{|dis_{b_1, b_2}|} + \dots + \frac{|dis_{a_n, a_1}|}{|dis_{b_n, b_1}|}$$

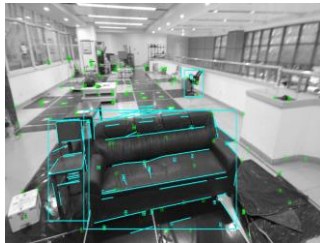
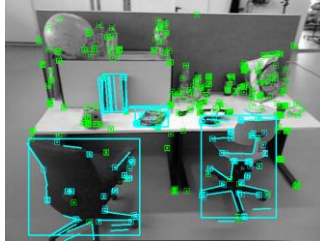
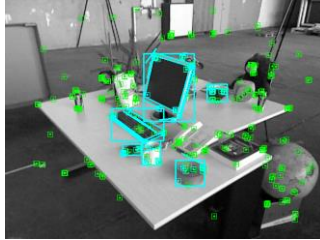
- Then, we scale the associated map points and compute the rigid transform (R, t) by minimizing the sum of squared error

$$(R^*, t^*) = \operatorname{argmin}_{R, t} \sum_{k=1}^{N_p} \|p_{a,k} - sR * p_{b,k} - t\|^2$$

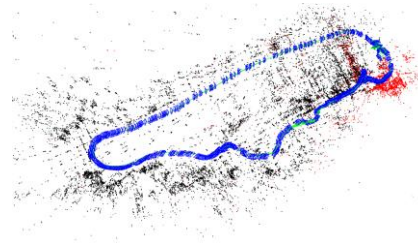
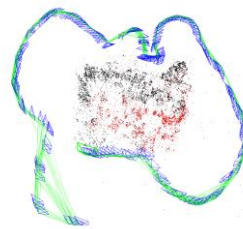
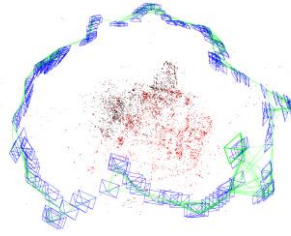
Experiments

- Datasets:
 - TUM RGB-D Dataset
 - USTC RGB-D Dataset
- Metrics:
 - Object mapping performance
 - Loop detection performance
 - Localization performance

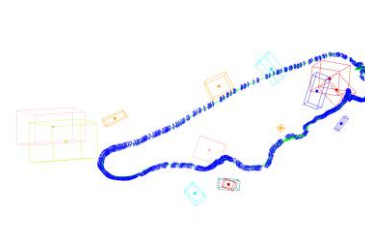
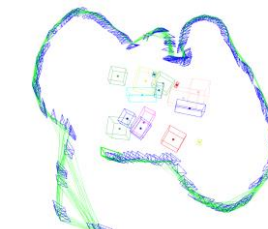
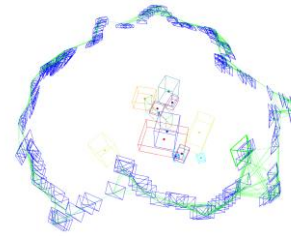
Experiments – object mapping performance



images

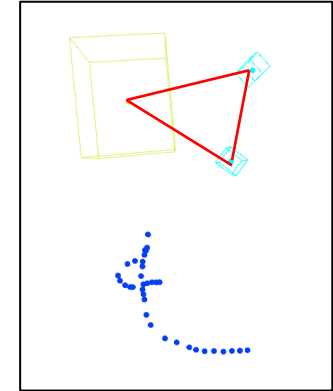
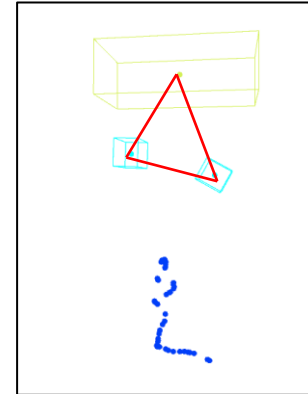
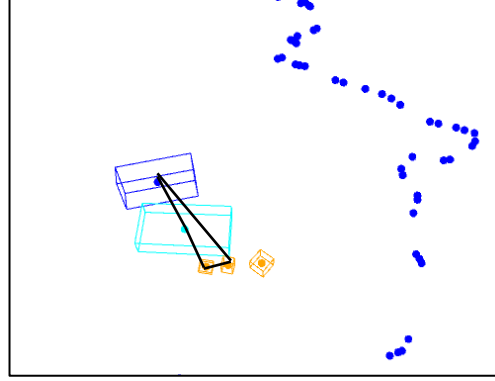
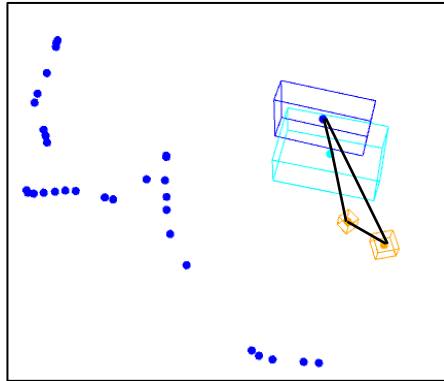
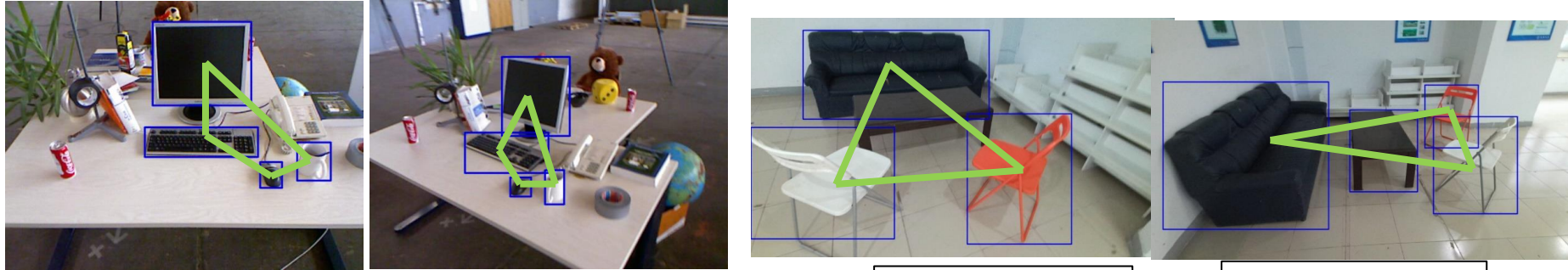


feature point map



object oriented map

Experiments – loop detection performance



fr2_desk sequences

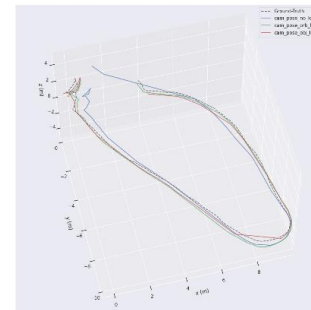
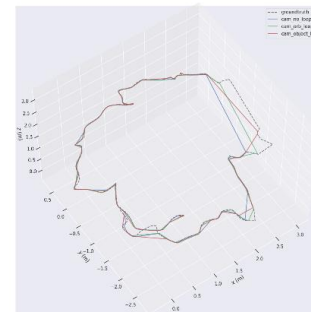
ustc_07 sequence

Experiments – offline video



Experiments – localization result

Sequence	ORB SLAM without loop closure (m)	ORB SLAM with feature-based loop closure (m)	Ours (object and feature-based loop)
fr2_desk	0.012499	0.007723	0.008851
fr3_long_office	0.034488	0.012362	0.013587
ustc_01	1.515137	1.515137	0.268989
ustc_02	0.733596	0.733596	0.191738
ustc_03	4.399249	4.399249	0.742086
ustc_07	0.733087	0.733087	0.498279



Conclusion

- We present a monocular SLAM system that can build a semantic map with feature points and object landmarks.
- An object-based loop closure method based on semantic graph matching is proposed, which is robust to the change of viewpoint.
- We propose a loop correction method to align the matched objects, which can effectively correct the drift error.

- Future work
- Object detection method with other sensors
- Field experiments in logistic environment



Karlsruher Institut für Technologie (KIT) Institut für Fördertechnik und Logistiksysteme (IFL)

Besucheranschrift

Gotthard-Franz-Str. 8 Geb. 50.38
76131 Karlsruhe

Telefon

+49 (721) 608-48600

Telefax

+49 (721) 608-48609

E-Mail

info@ifl.kit.edu

Homepage

<http://www.ifl.kit.edu>

