Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen
International Center for
Information Technology

Semantic Mapping for Enhanced Localization in Indoor Environments

R&D Defense
December 01, 2023

Allen Isaac Jose

Advisors
Prof. Dr.-Ing. Sebastian Houben
Deebul Nair, M.Sc.

Introduction

Fig 1. Indoor environment [1]

- Simultaneous localization and mapping (SLAM).
- Mapping – given robot pose, unknown environment.
- Localization – given map, unknown robot pose.
- Focus on the indoor environment.
- Semantics of objects.

[1] https://images.app.goo.gl/48M9ACyDLVaTnRmN6

2D SLAM

Fig 2. Occupany grid map[1]

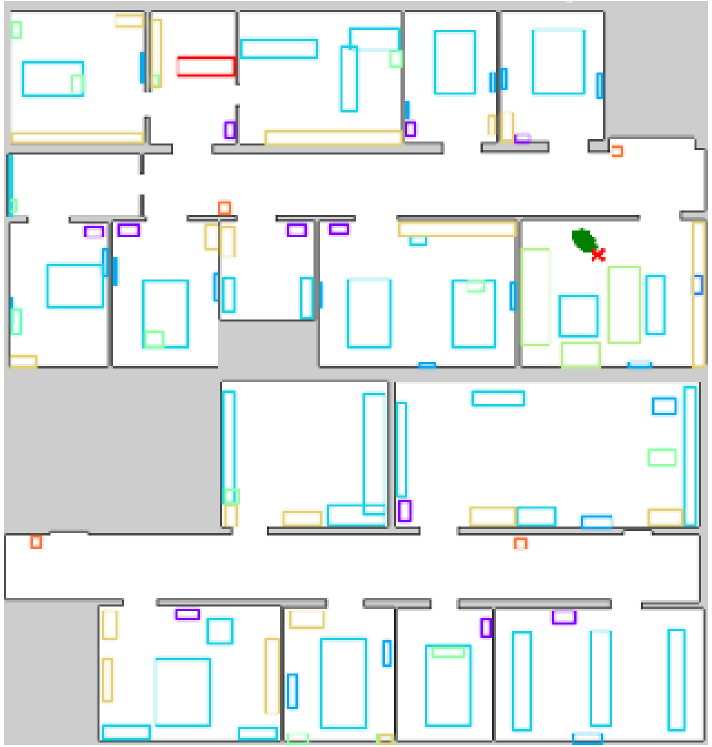


Fig 3. 2D Semantic map[2]

[1] https://images.app.goo.gl/Zeatfe5GhCLQtaCc8
[2] “Long-Term Localization using Semantic Cues in Floor Plan Maps” - https://arxiv.org/pdf/2210.01456.pdf

3D SLAM

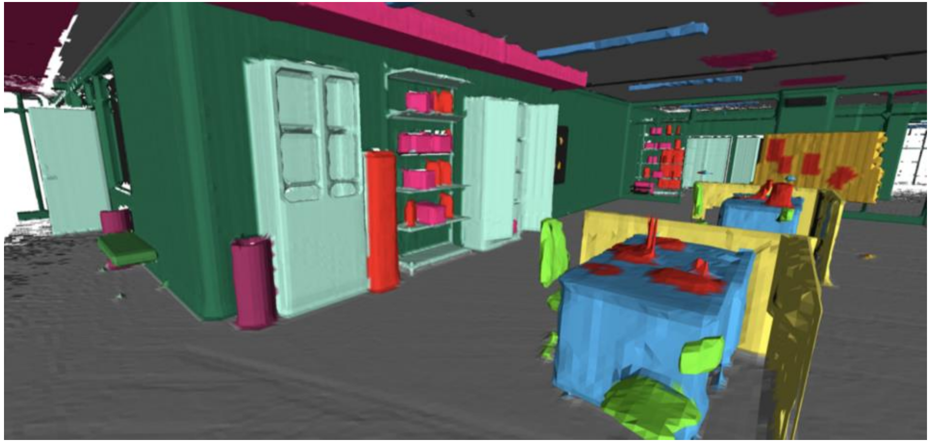


Fig 4. Kimera dense 3D map[1]

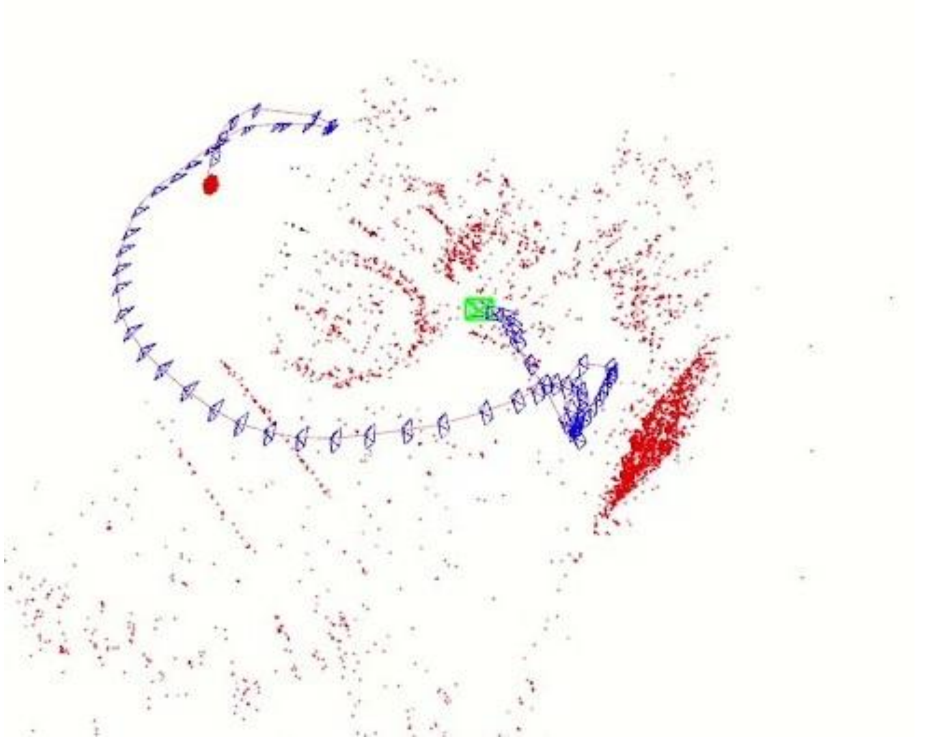


Fig 5. ORB-SLAM sparse 3D map[2]

[1] Kimera - https://github.com/MIT-SPARK/Kimera
[2] https://images.app.goo.gl/M1drAPXjd341vR5P9

Research question

- How to add semantic information to a 3D sparse map?
- How does the localization capability improve by incorporating abstract object models?
- How to perform comparative evaluation on the mapped objects quantitatively?
- How is the real-time performance of the 3D sparse semantic mapping?
- How robust is the SLAM algorithm to the noise in object detection?

Ellipsoidal mapping

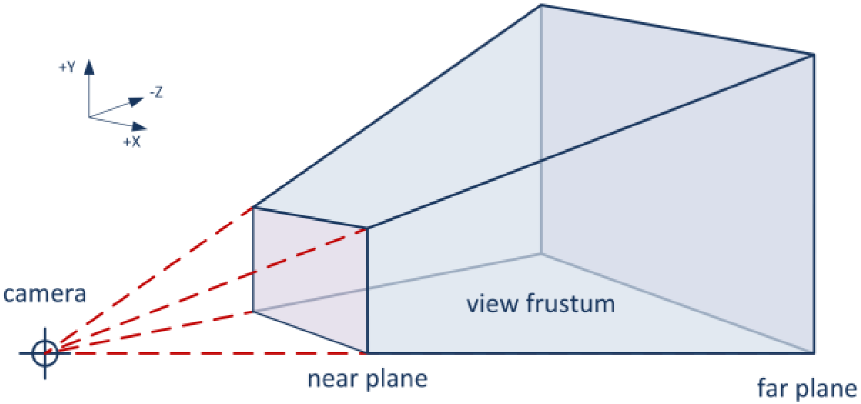


Fig 6. Camera frustum[1]

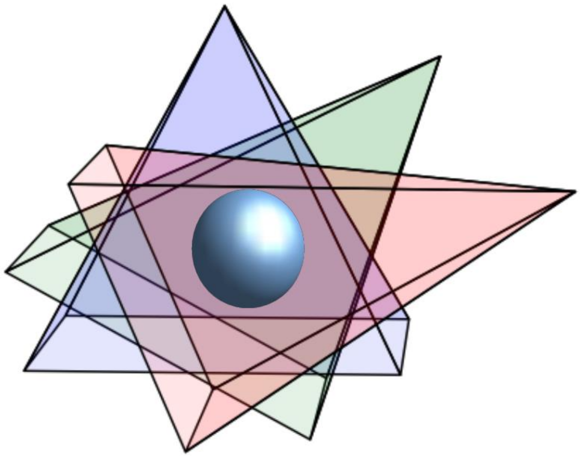


Fig 7. Constrained space

9 unknown parameters – 3 each for the centroid, orientation and size

[1] https://learnopengl.com/img/guest/2021/Frustum_culling/VisualCameraFrustum.png

GraphSLAM

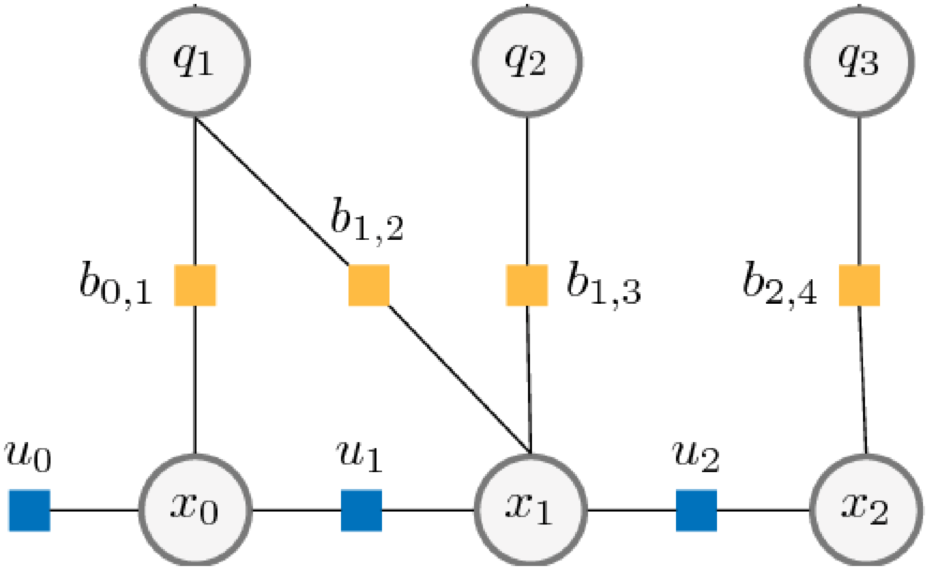
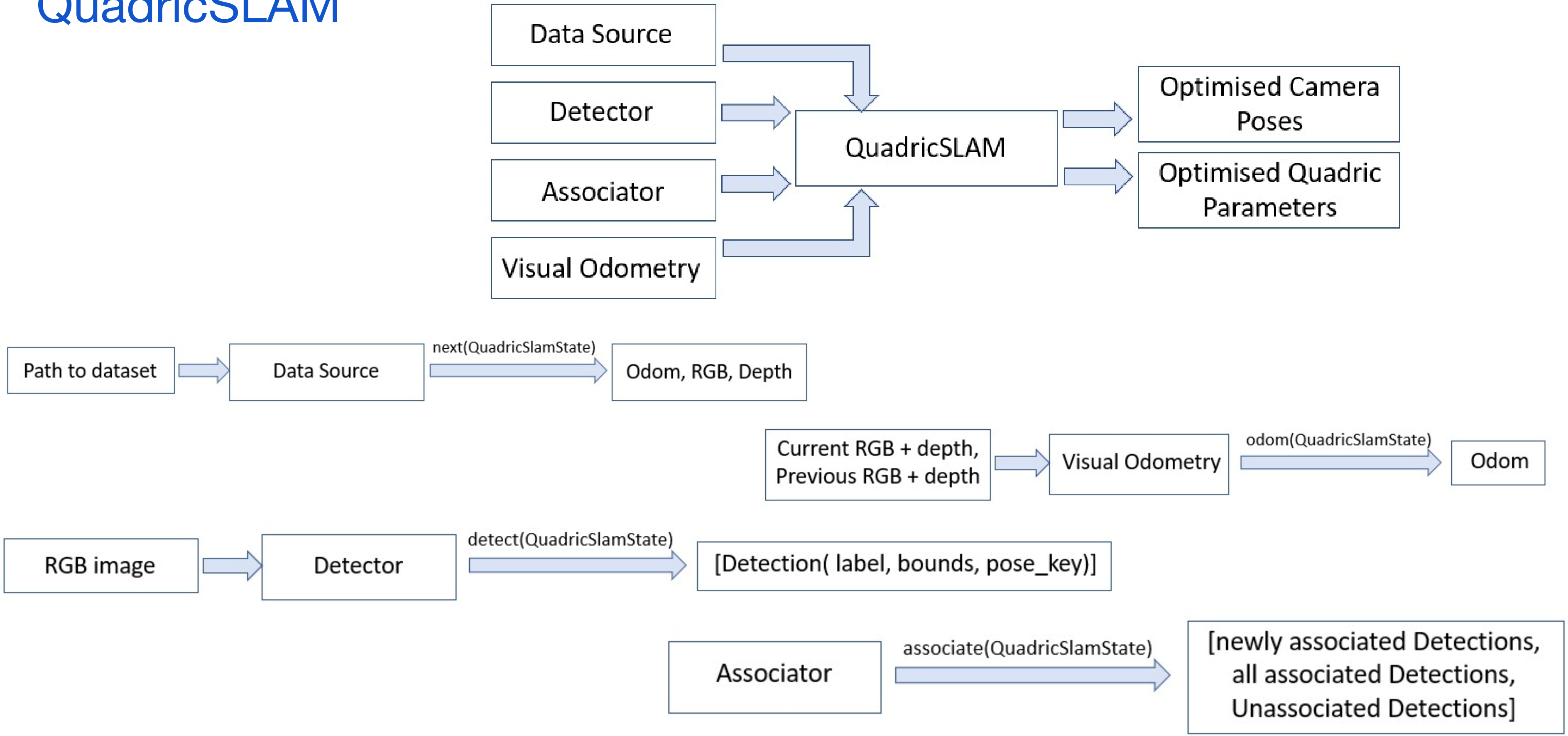


Fig 8. Factor Graph[1]

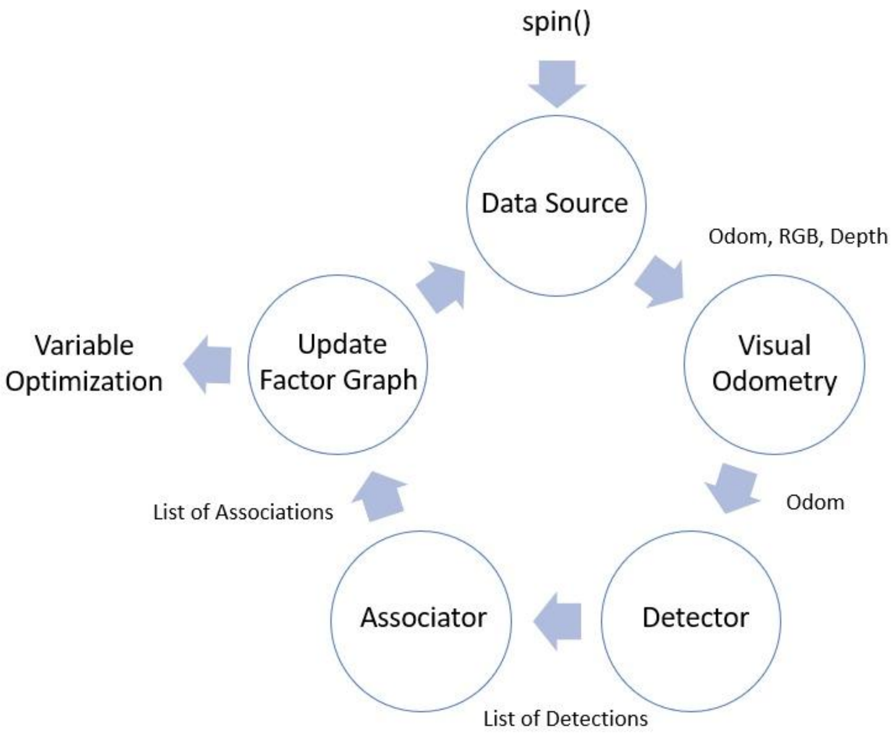
- u – odometry factor
- b – bounding box factor
- x – robot pose
- q – quadric/ellipsoid parameters

[1] "An Orientation Factor for Object-Oriented SLAM" - https://arxiv.org/pdf/1809.06977.pdf

QuadricSLAM



QuadricSLAM



Object Aided-SLAM (OA-SLAM)

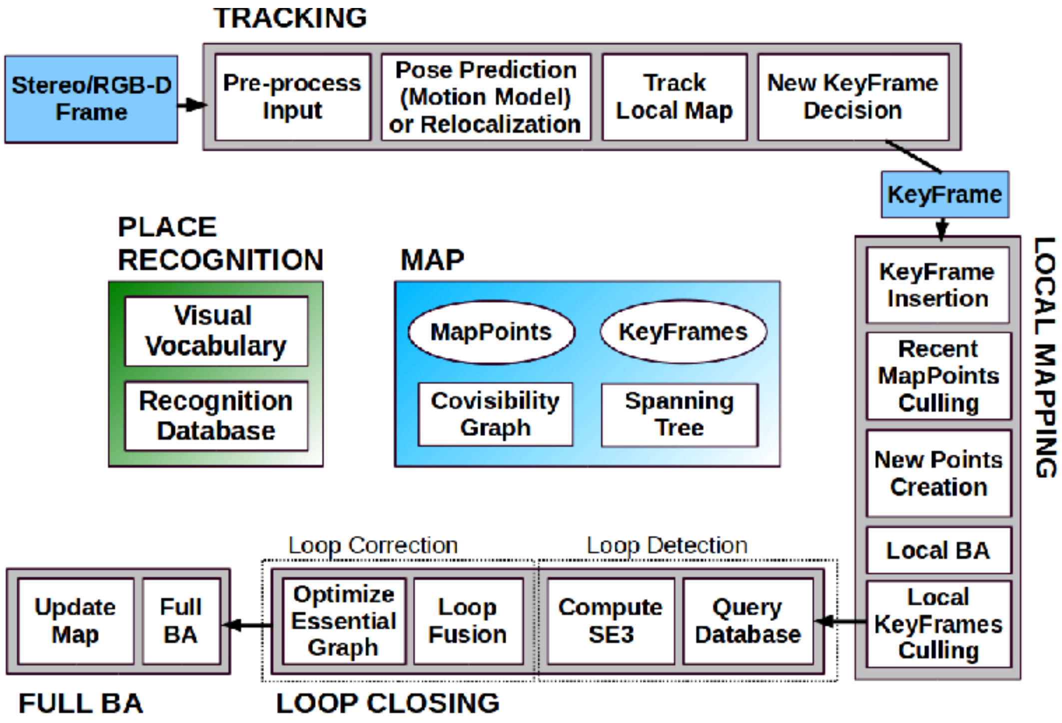


Fig 9. ORB SLAM2[1]

[1] ORB-SLAM2 - https://arxiv.org/pdf/1610.06475.pdf

Object Aided-SLAM (OA-SLAM)

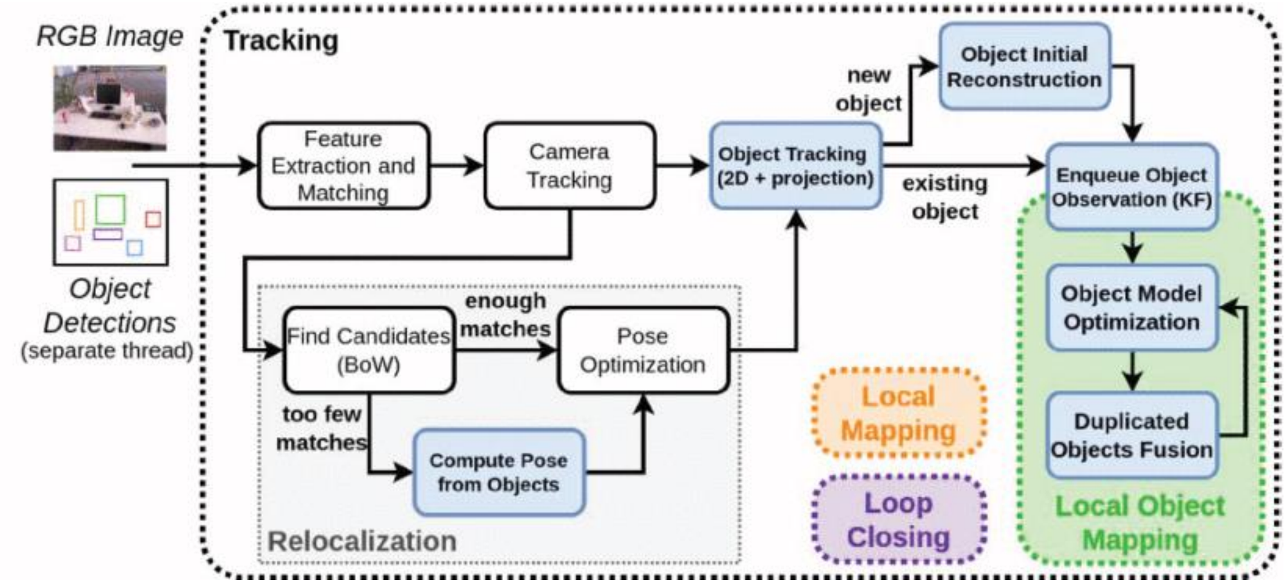


Fig 10. OA SLAM[1]

[1] OA-SLAM - https://ieeexplore.ieee.org/document/9995573

Synthetic dataset generation

Fig 11. Sample image from scene

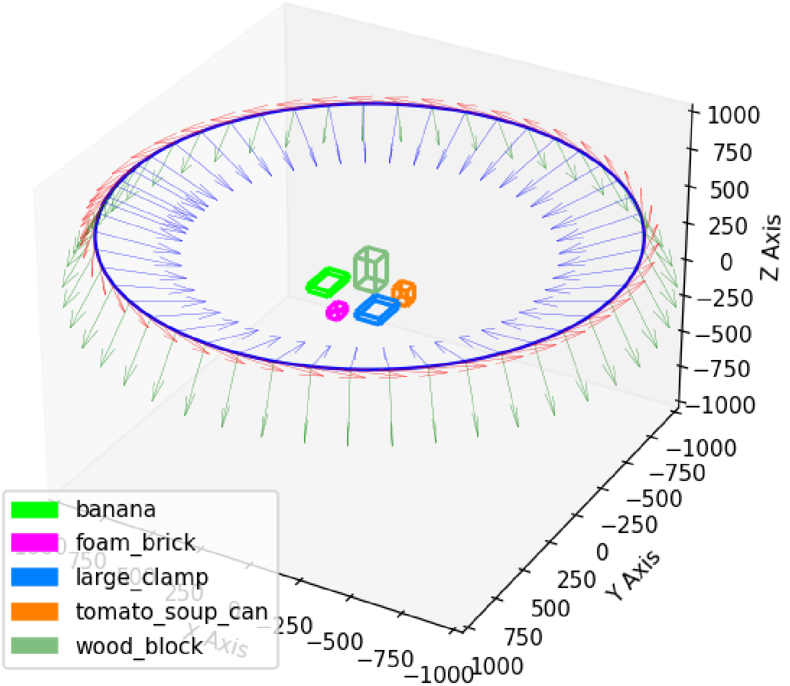


Fig 12. Visualization of dataset

Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Evaluation metrics

Camera trajectory metrics

- Camera trajectory error
- Camera rotation error
- Procrustes analysis
- Fréchet distance
- Chamfer distance

Object mapping metrics

- Object centroid error
- Object rotation error
- Object IoU comparison

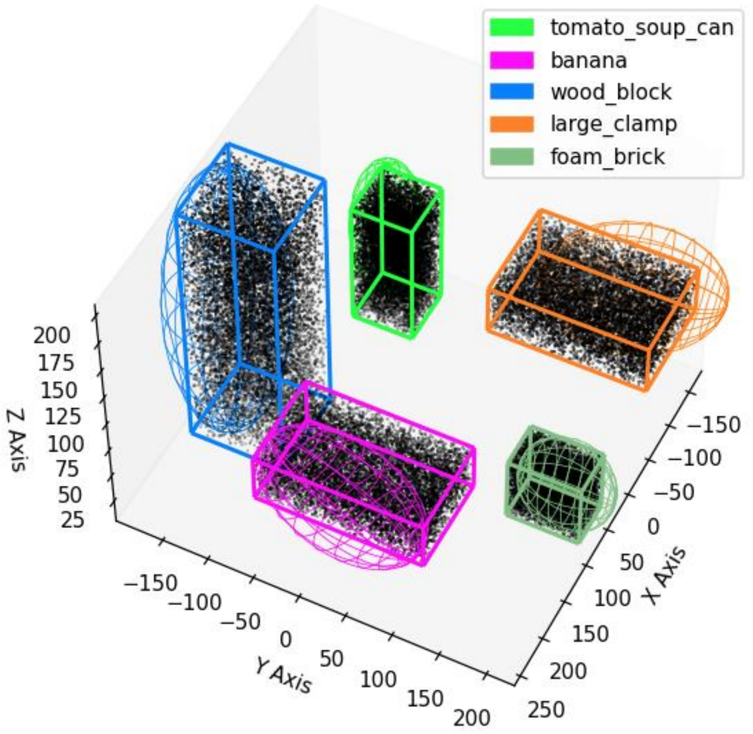


Fig 13. Object IoU comparison method

Evaluation results on scene 9

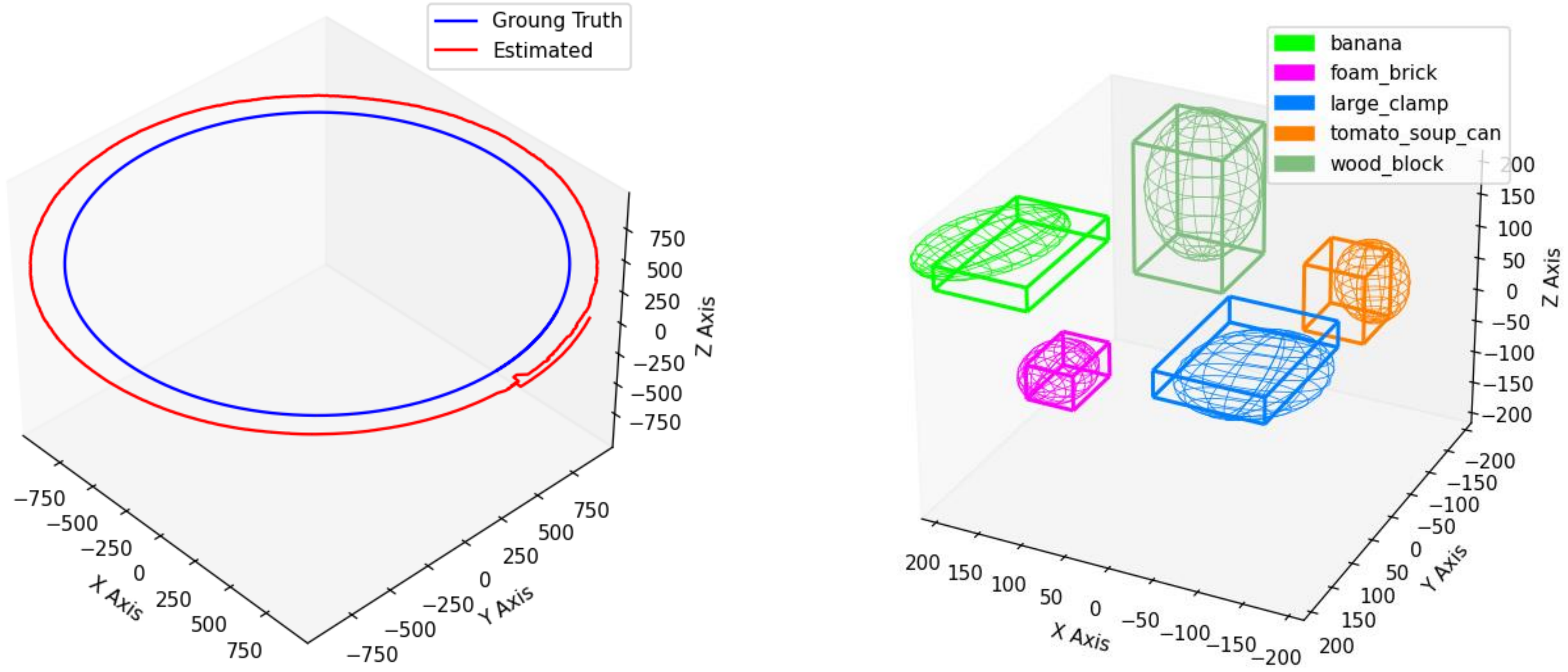


Fig 14. OA-SLAM output

Evaluation results on scene 9

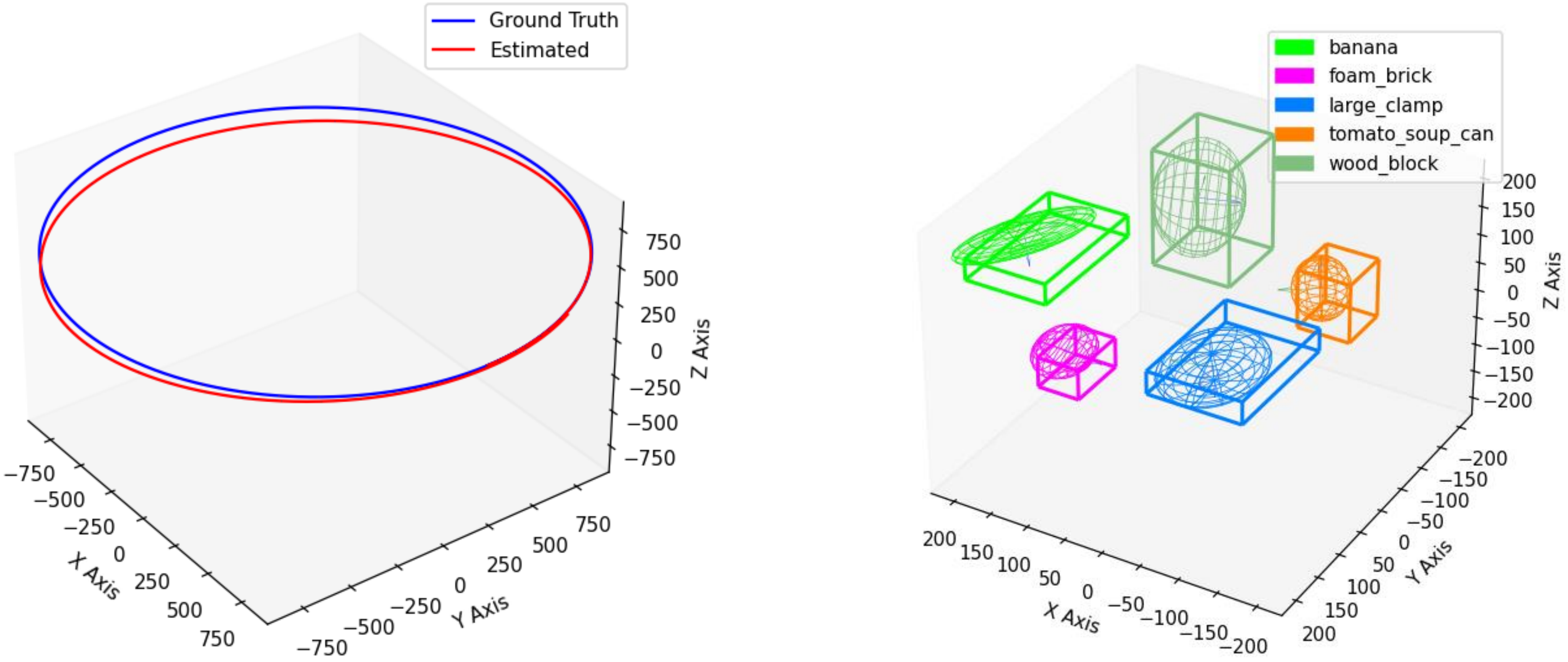
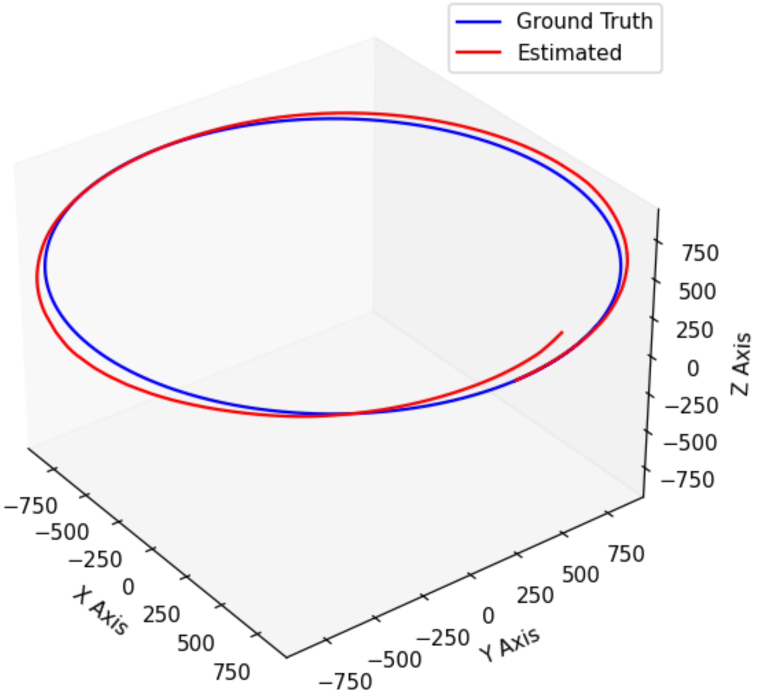


Fig 15. QuadricSLAM batch mode output

Evaluation results on scene 9



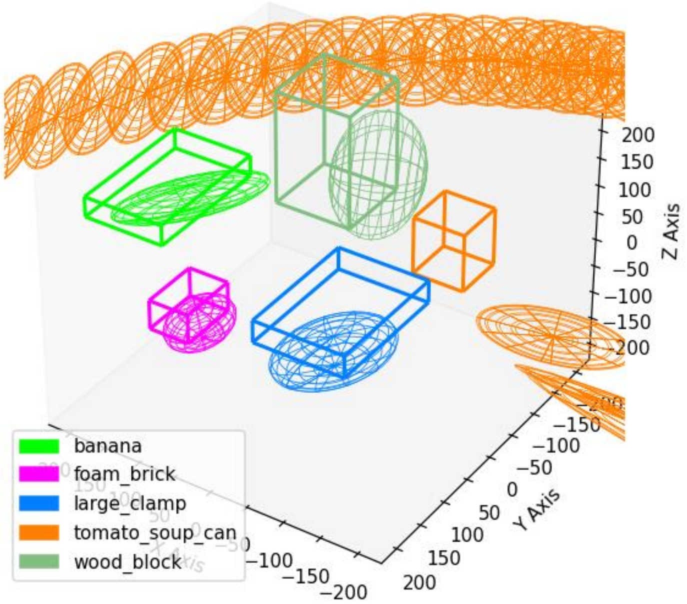


Fig 16. QuadricSLAM increment mode output

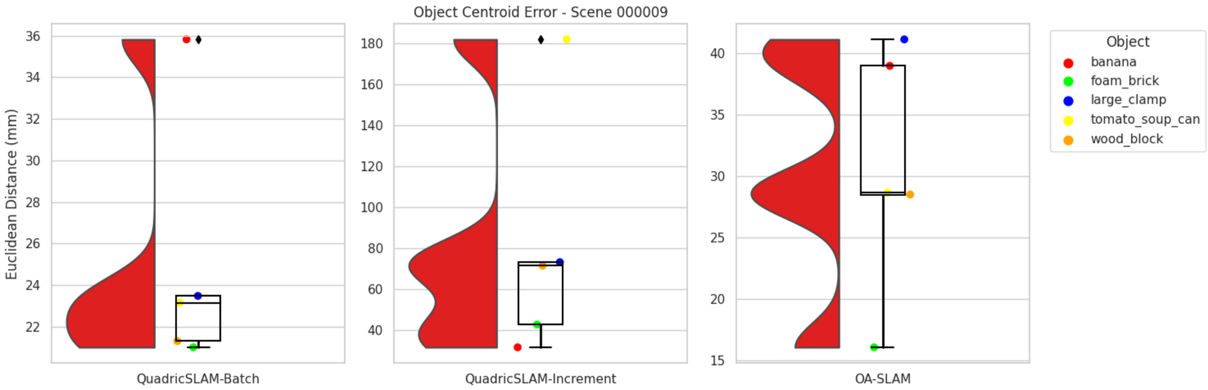
System profile metrics

	OA-SLAM	QuadricSLAM - batch	QuadricSLAM - incremental
Average CPU Utilization	30.59%	16.01%	58.97%
Min CPU Utilization	10.86%	10.61%	10.24%
Max CPU Utilization	35.96%	75.85%	71.43%
Average Memory Utilization	645.09 MB	209.43 MB	261.52 MB
Min Memory Utilization	128.53 MB	205.05 MB	211.66 MB
Max Memory Utilization	744.55 MB	234.16 MB	328.04 MB
Overall Time Taken	68.13 s	49.06 s	532.58 s
FPS	22	30	2
Map Size	5.28 MB	444 KB	612 KB

Table 1. System profile – scene 9

Evaluation metrics – object mapping

Fig 17. Object centroid error – scene 9



Evaluation metrics – object mapping

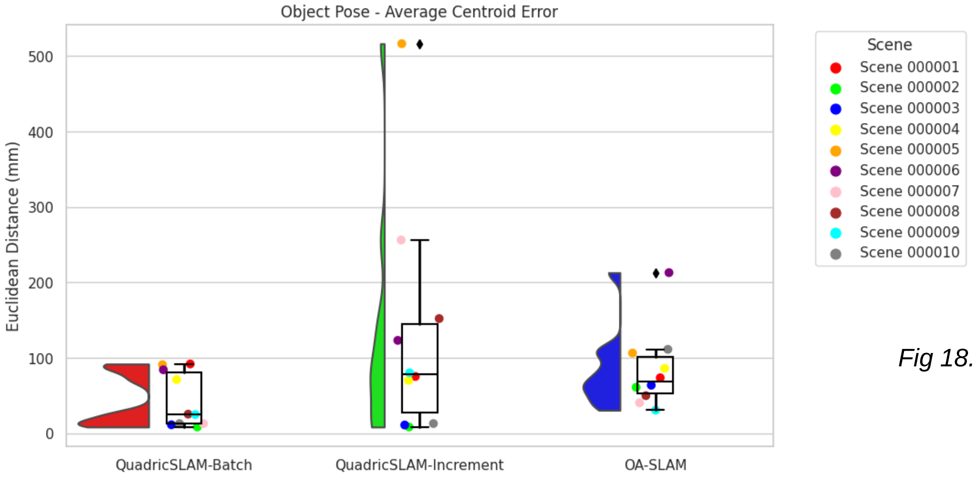
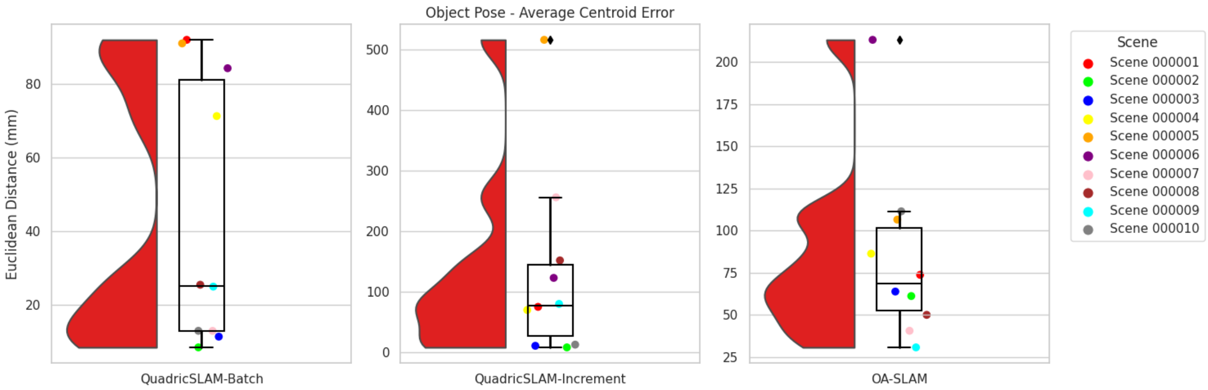


Fig 18. Object centroid error – all scenes



Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Evaluation metrics – object mapping

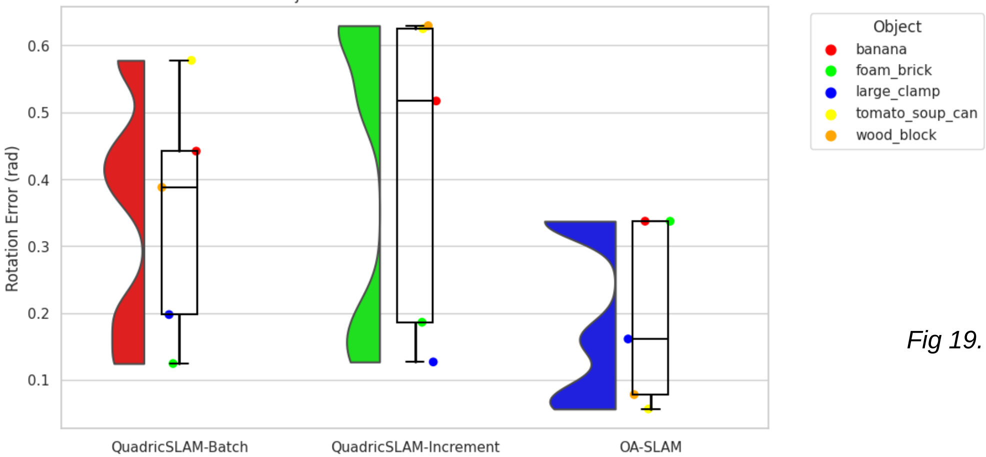
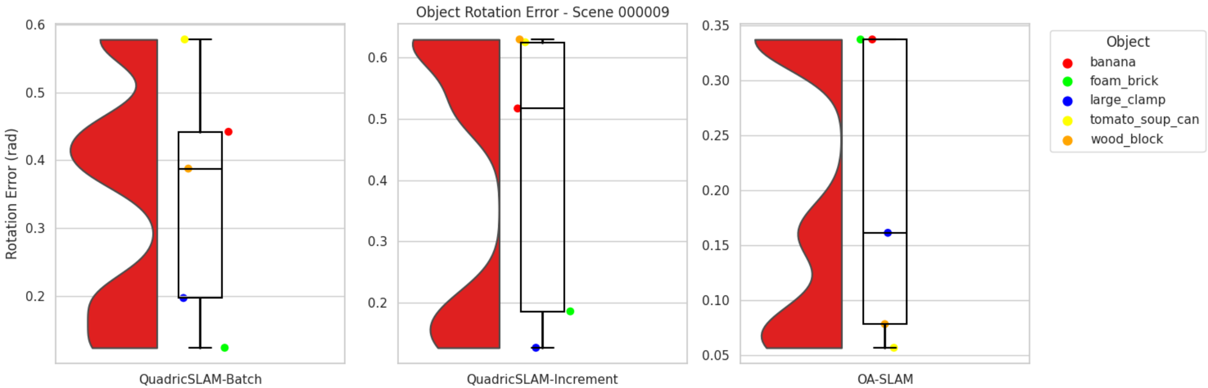


Fig 19. Object rotation error – scene 9



Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Evaluation metrics – object mapping

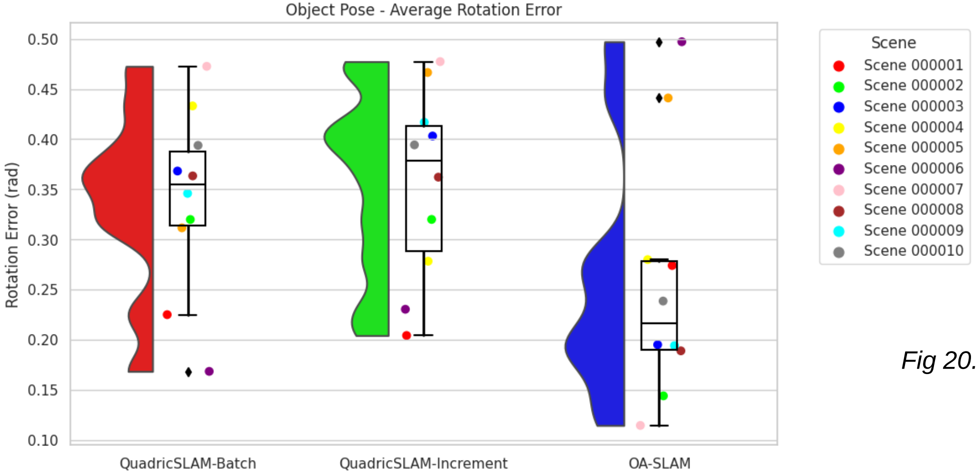
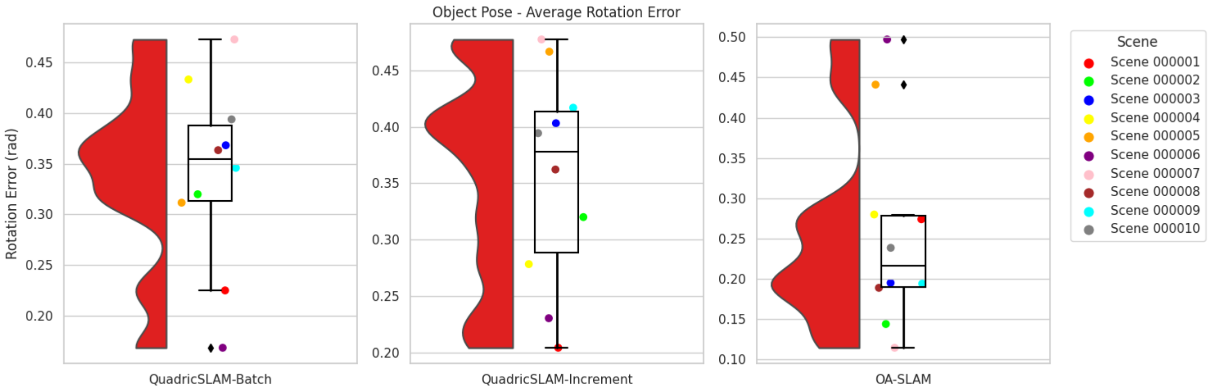


Fig 20. Object rotation error – all scenes



Evaluation metrics – object mapping

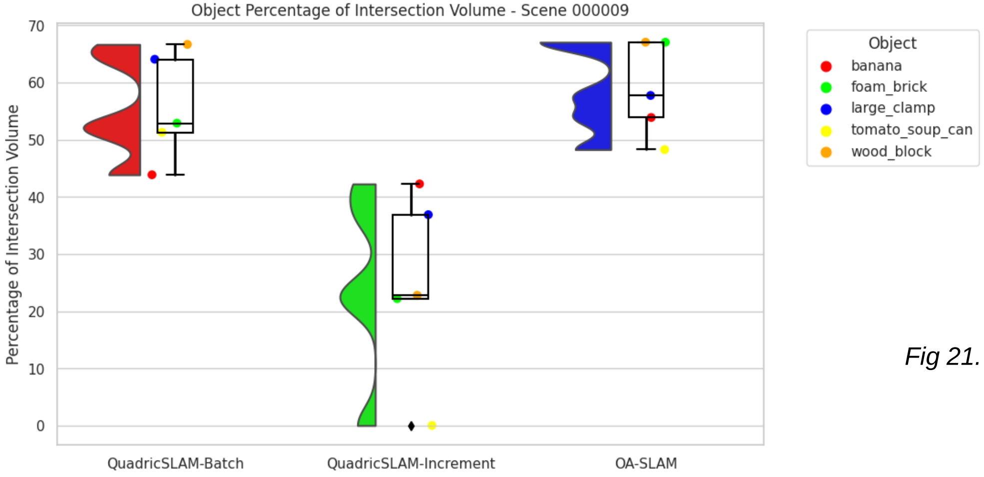
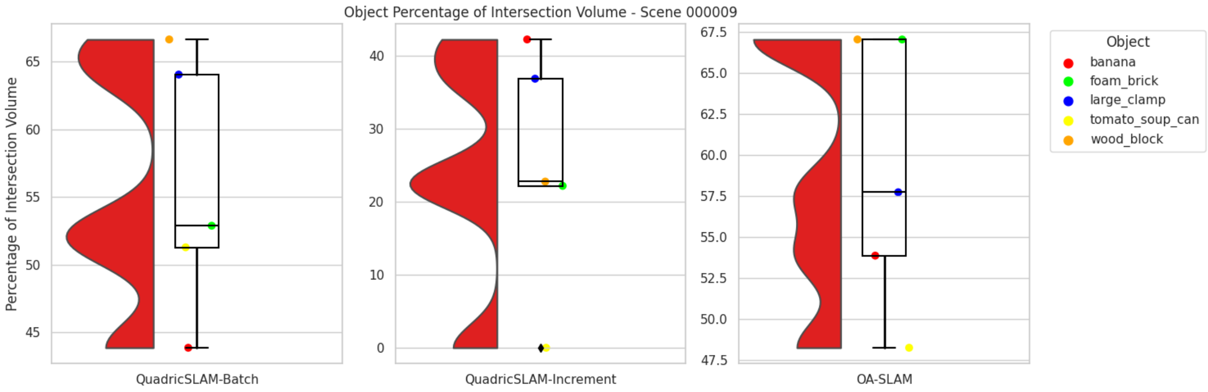


Fig 21. Object overlap percentage – scene 9



Evaluation metrics – object mapping

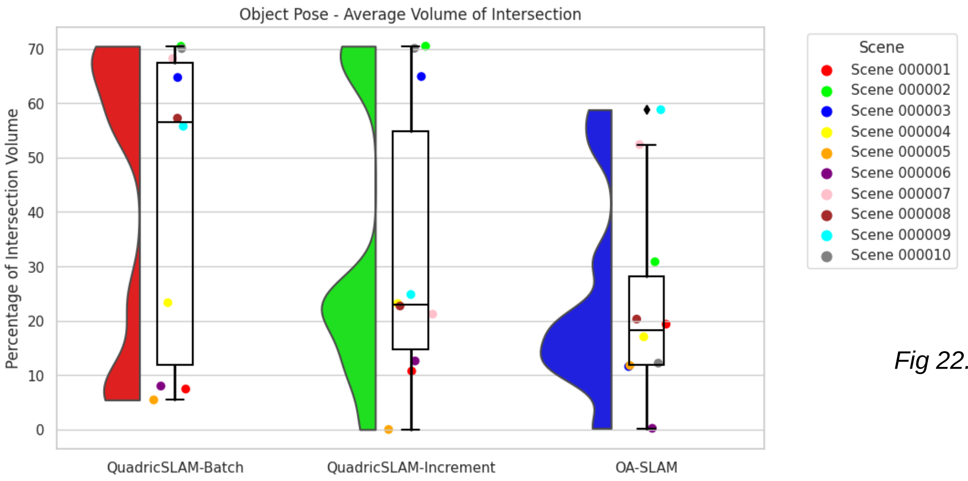
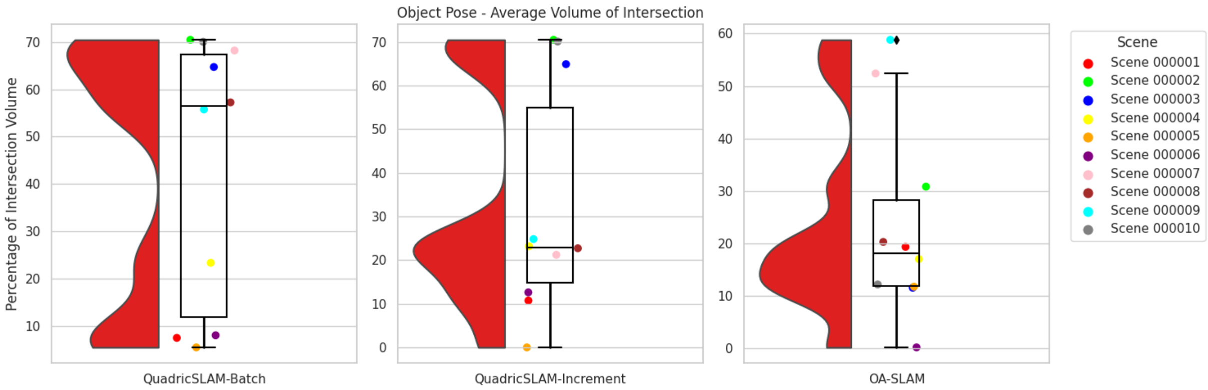


Fig 22. Object overlap percentage – all scenes



Evaluation metrics – object mapping

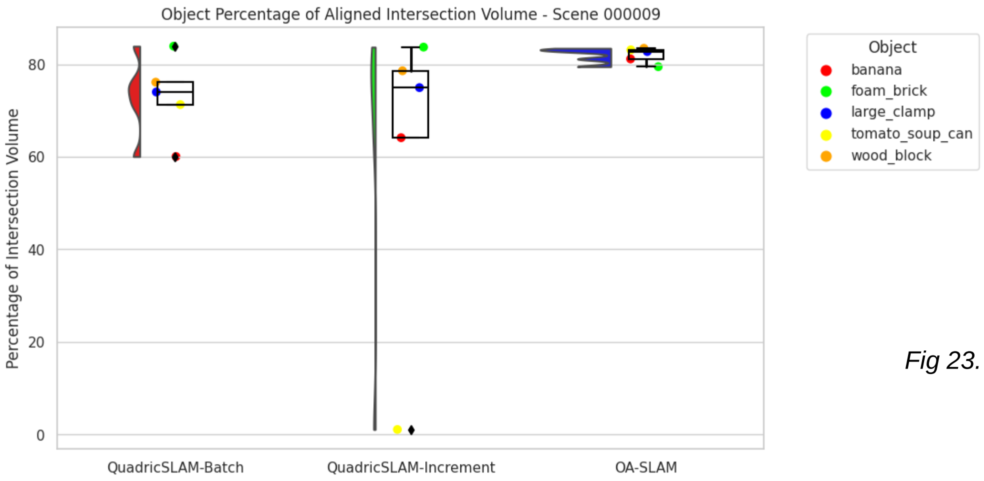
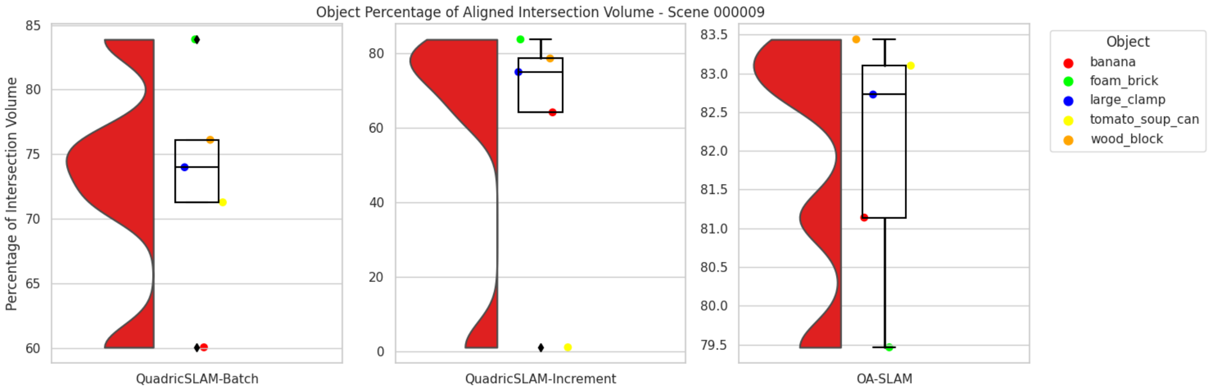


Fig 23. Object aligned overlap percentage – scene 9



Evaluation metrics – object mapping

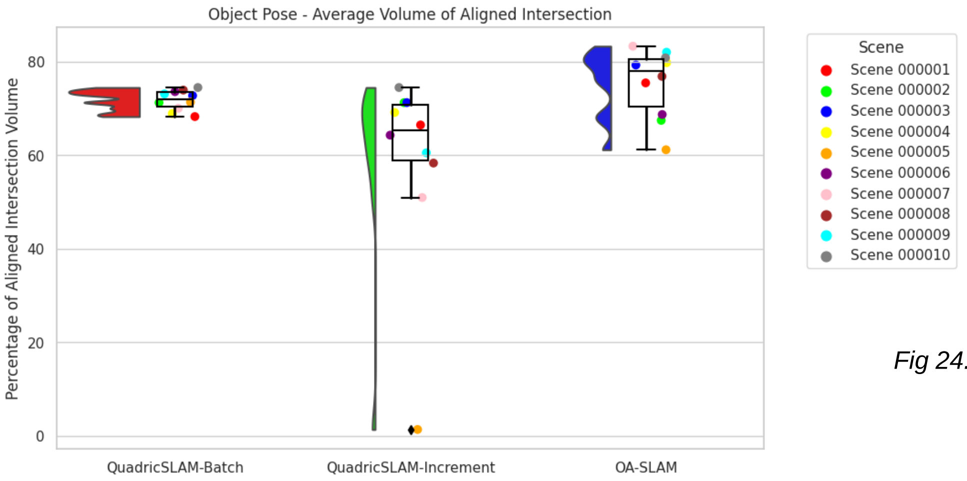
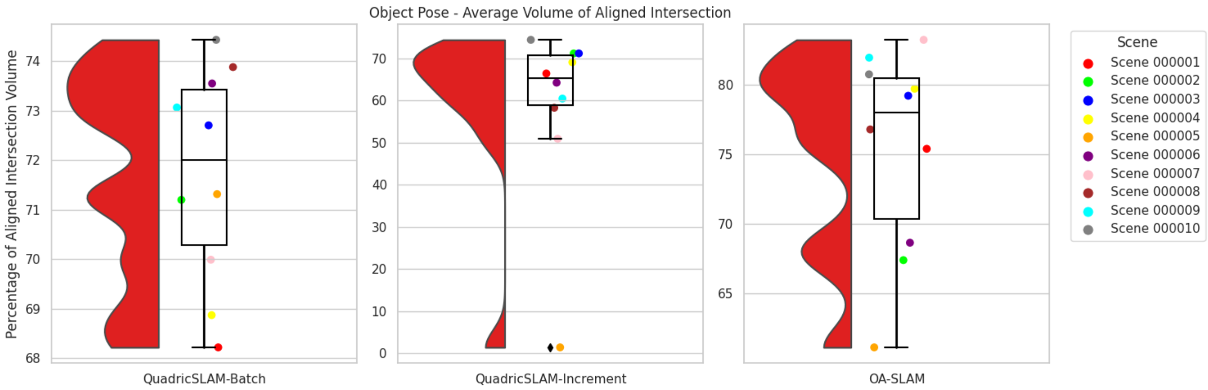


Fig 24. Object aligned overlap percentage – all scenes



Evaluation metrics – camera trajectory

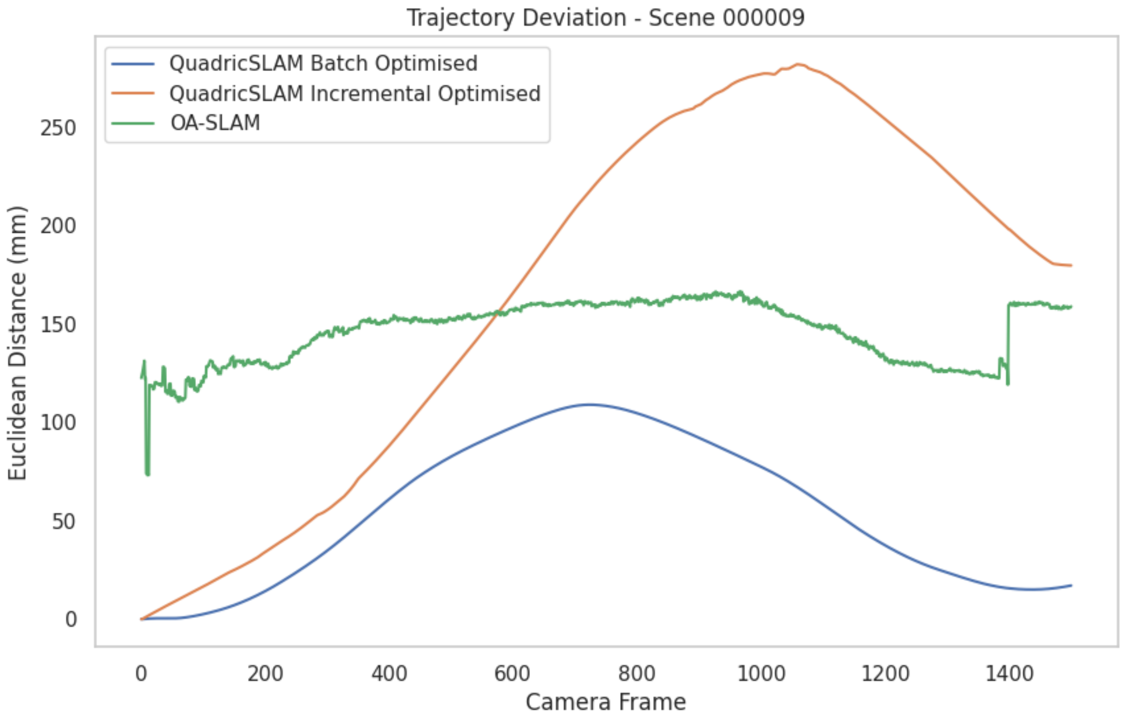


Fig 25. Camera trajectory error comparison – scene 9

Evaluation metrics – camera trajectory

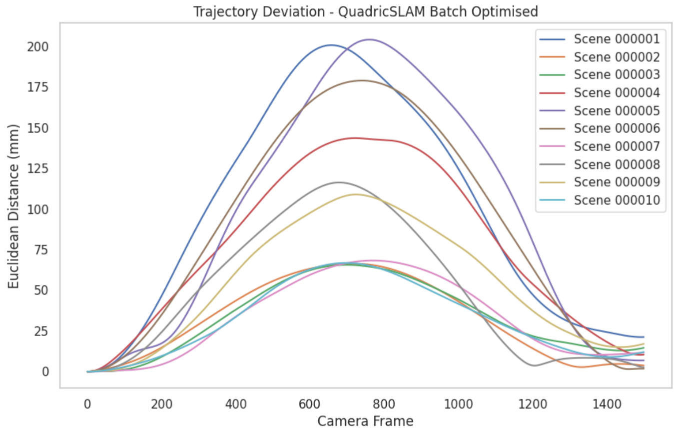
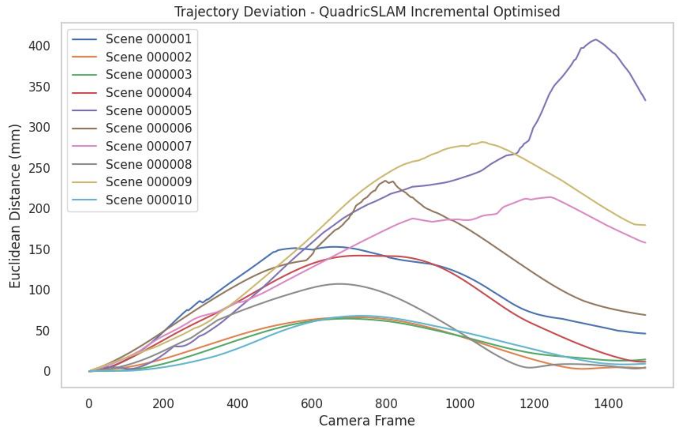
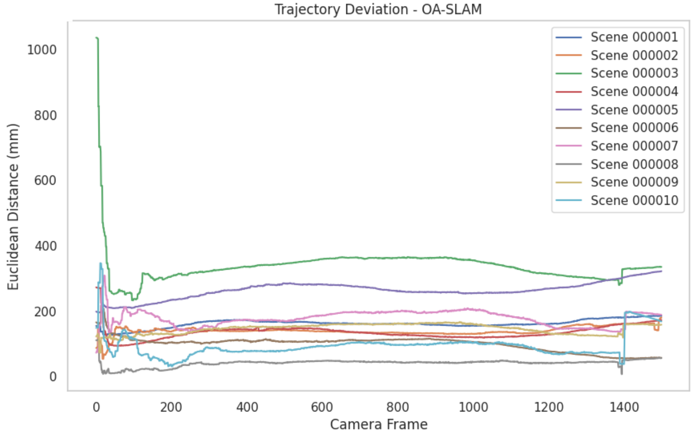


Fig 26. Camera trajectory error comparison – all scenes





Evaluation metrics – camera trajectory

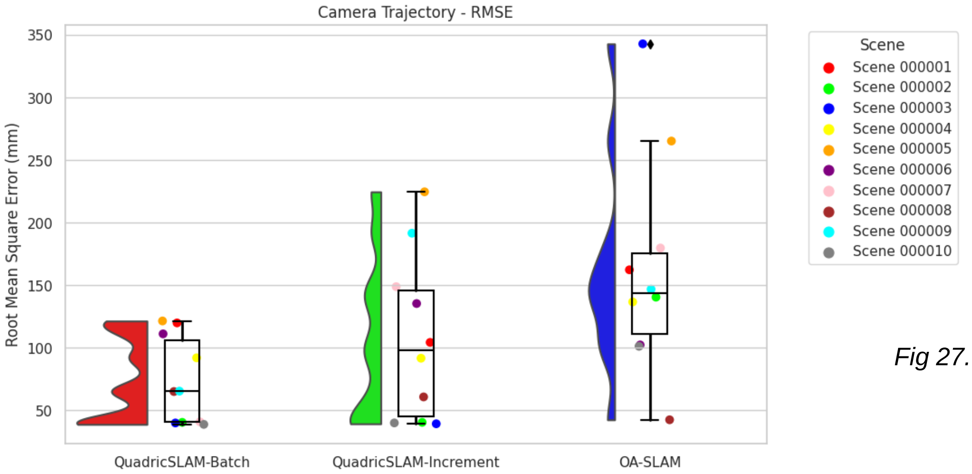
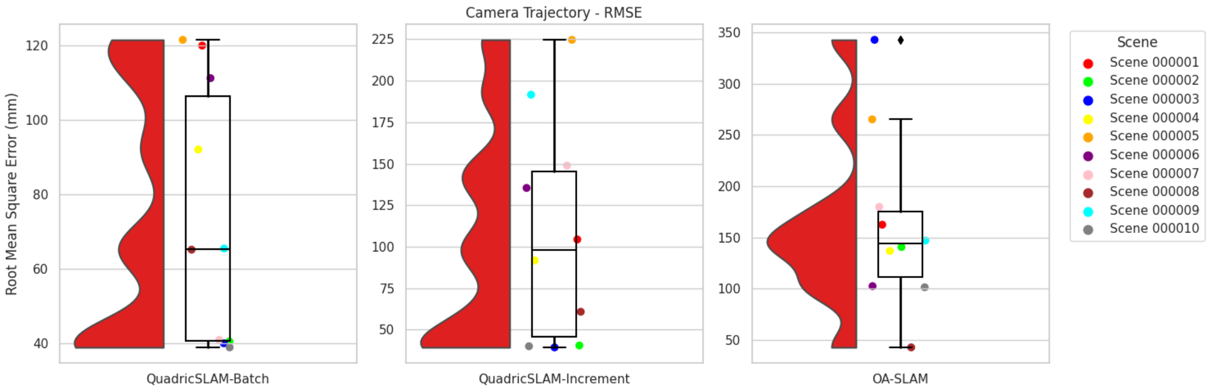


Fig 27. Camera trajectory RMSE comparison – all scenes



Evaluation metrics – camera trajectory

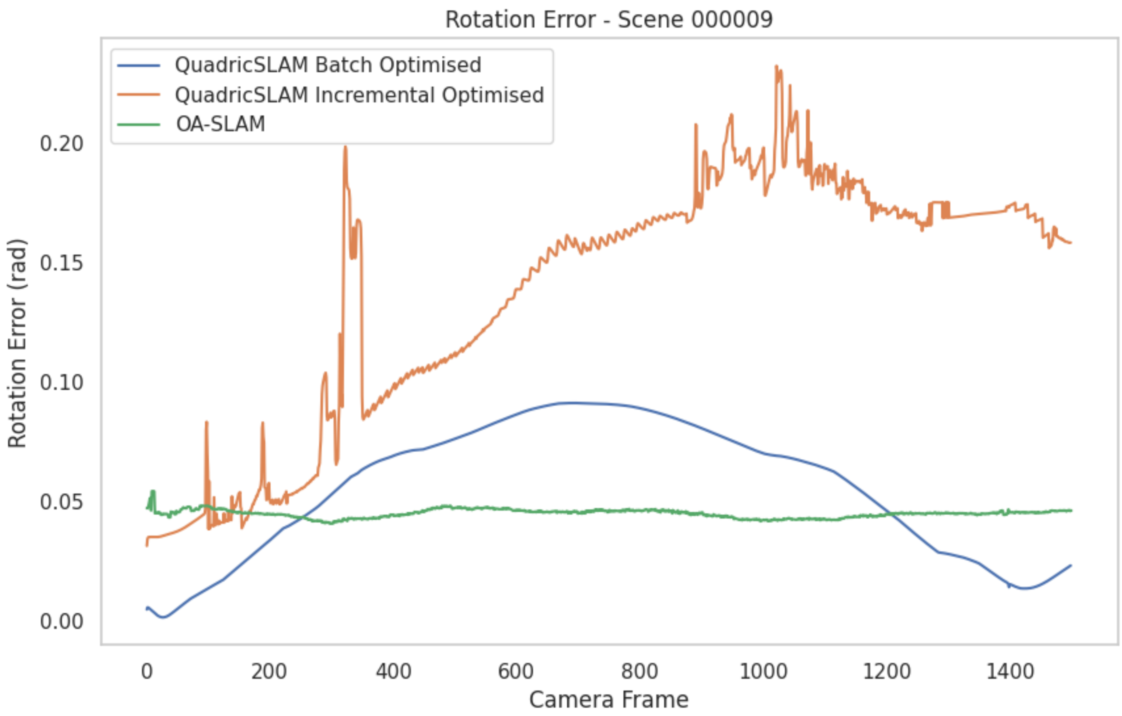


Fig 28. Camera rotation error comparison – scene 9

Evaluation metrics – camera trajectory

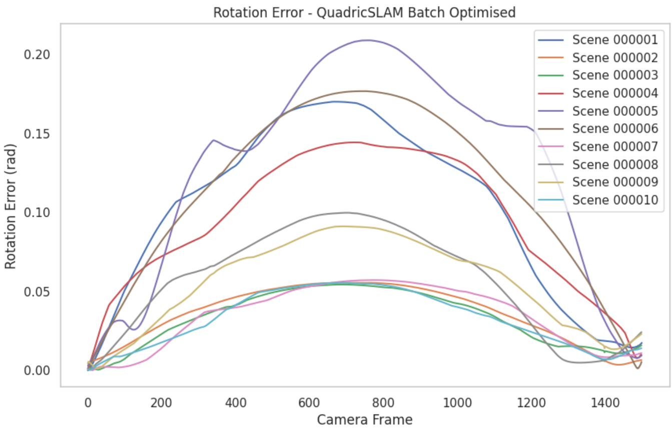
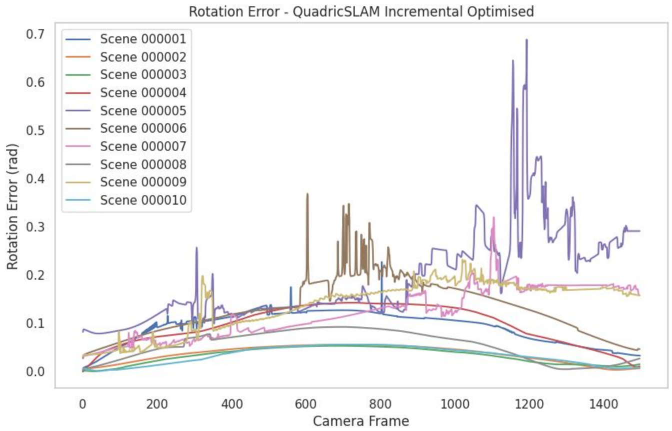
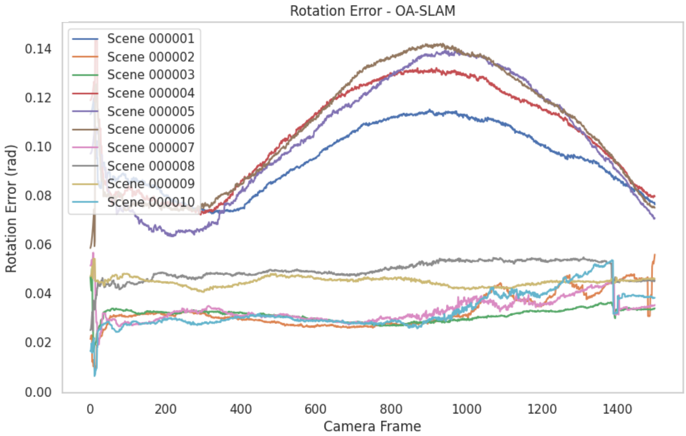


Fig 29. Camera rotation error comparison – all scenes





Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Evaluation metrics – camera trajectory

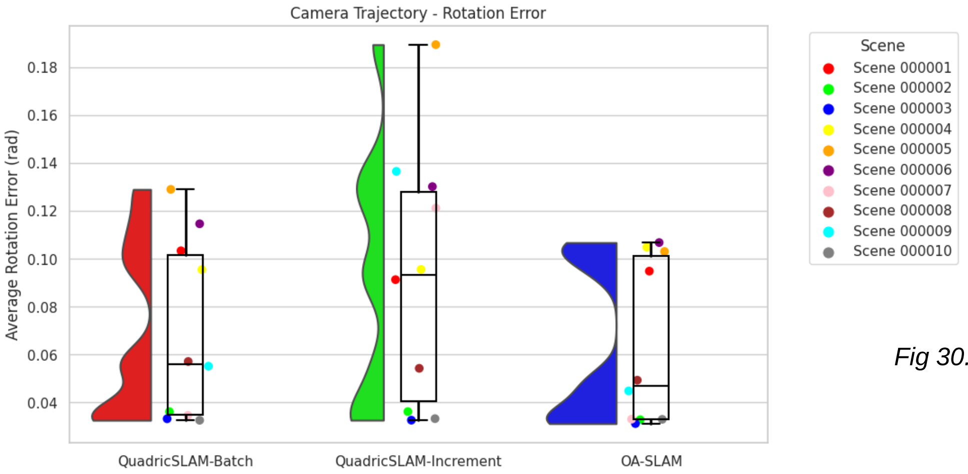
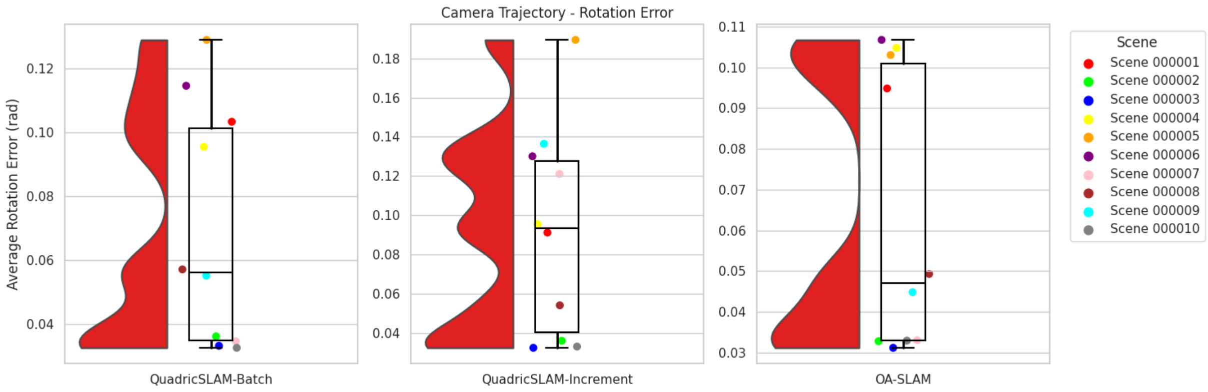


Fig 30. Camera average rotation error comparison – all scenes



Evaluation metrics – camera trajectory

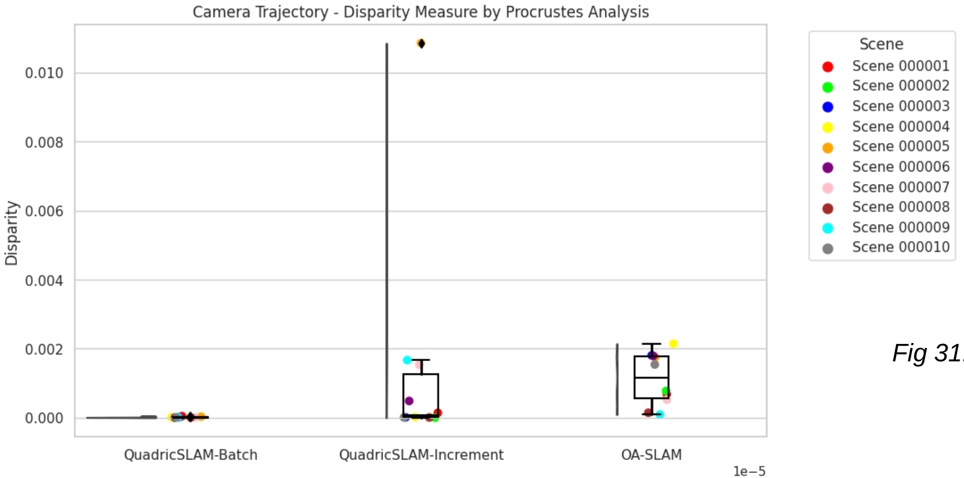
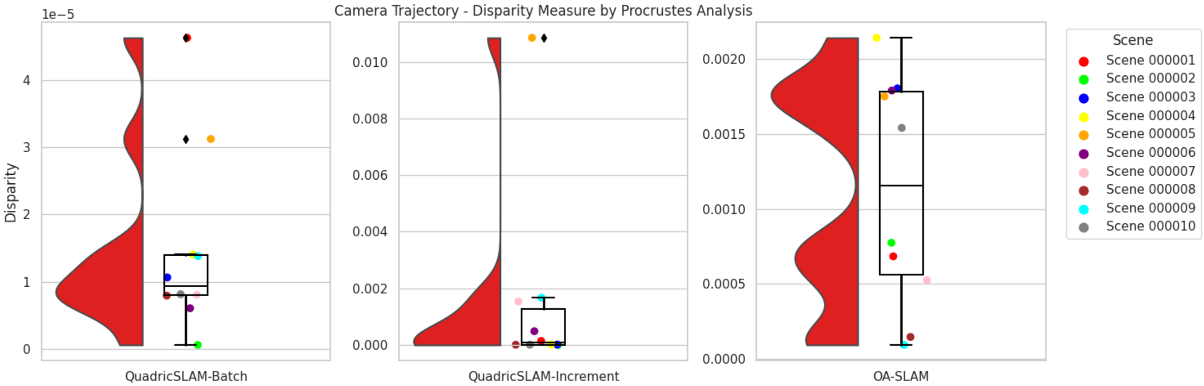


Fig 31. Procrustes analysis comparison – all scenes



Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Evaluation metrics – camera trajectory

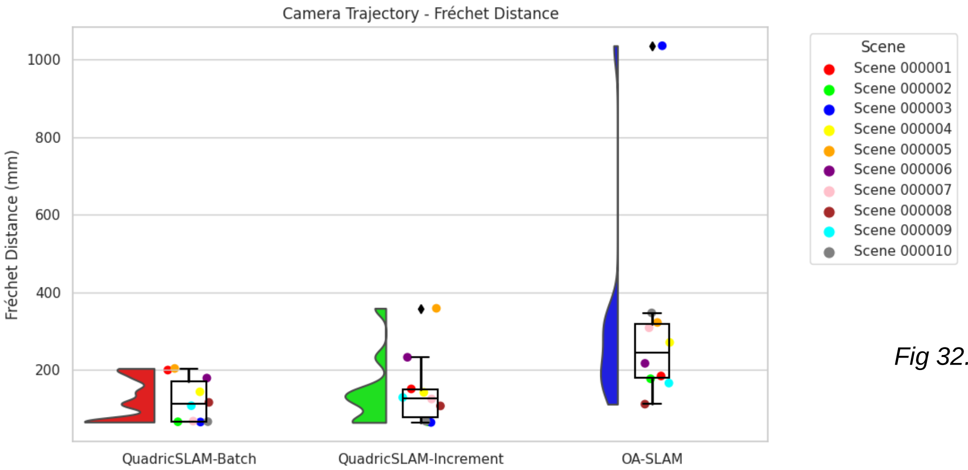
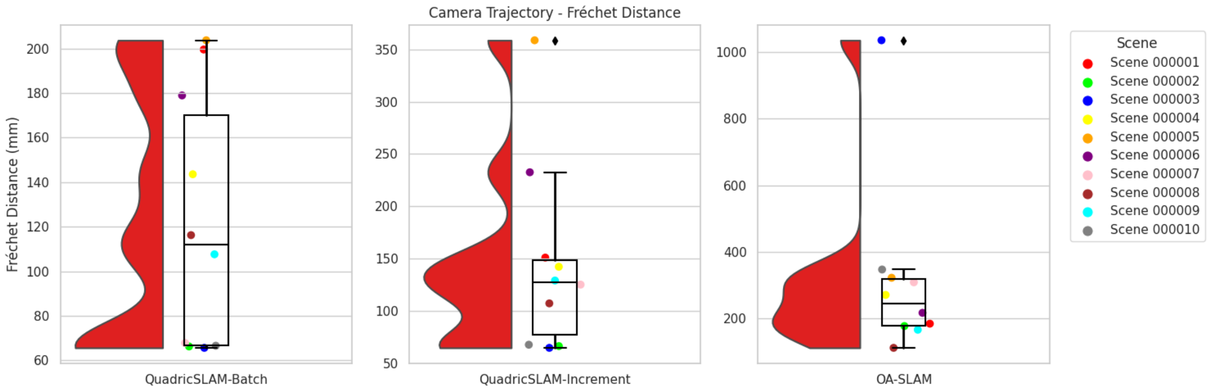


Fig 32. Frechet distance comparison – all scenes



Evaluation metrics – camera trajectory

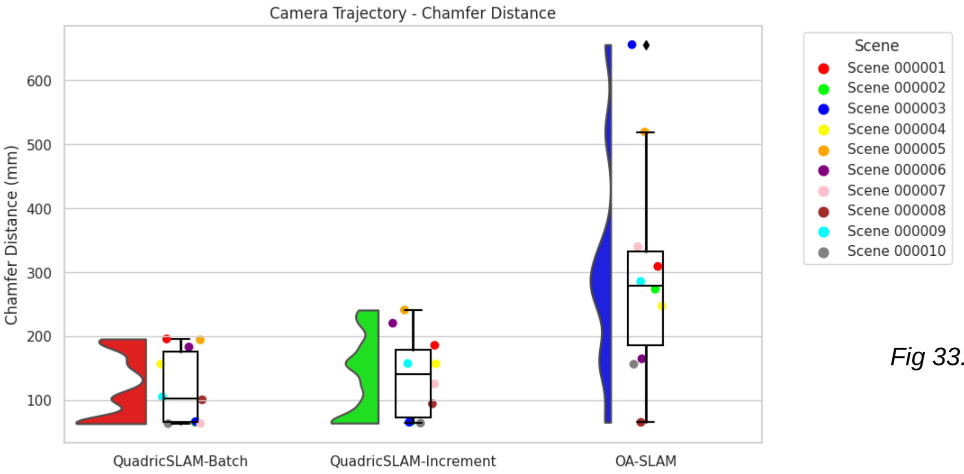
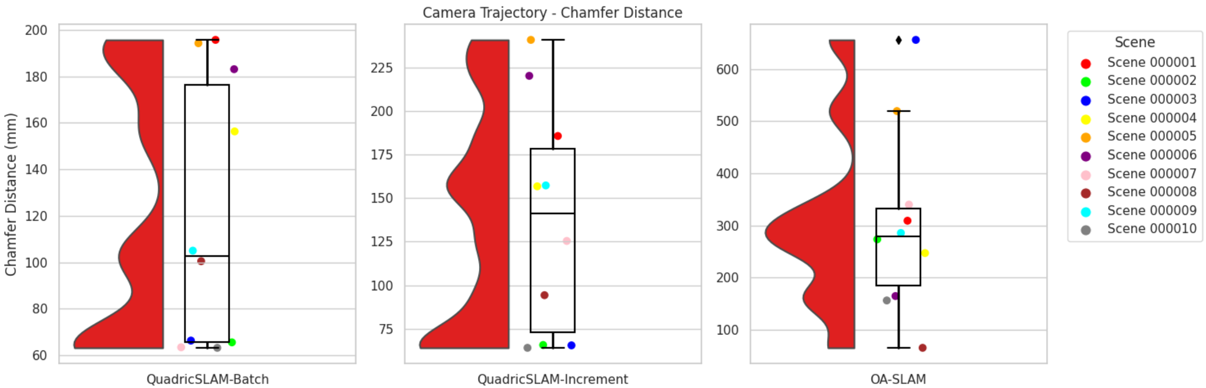


Fig 33. Chamfer distance comparison – all scenes



Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Results of evaluation on scene 9 corrupted with noise

Bbox corruption

- Changing box center x.
- Changing box center y.
- Changing box height.
- Changing box width.
- Deleteing 1 object’s bbox.

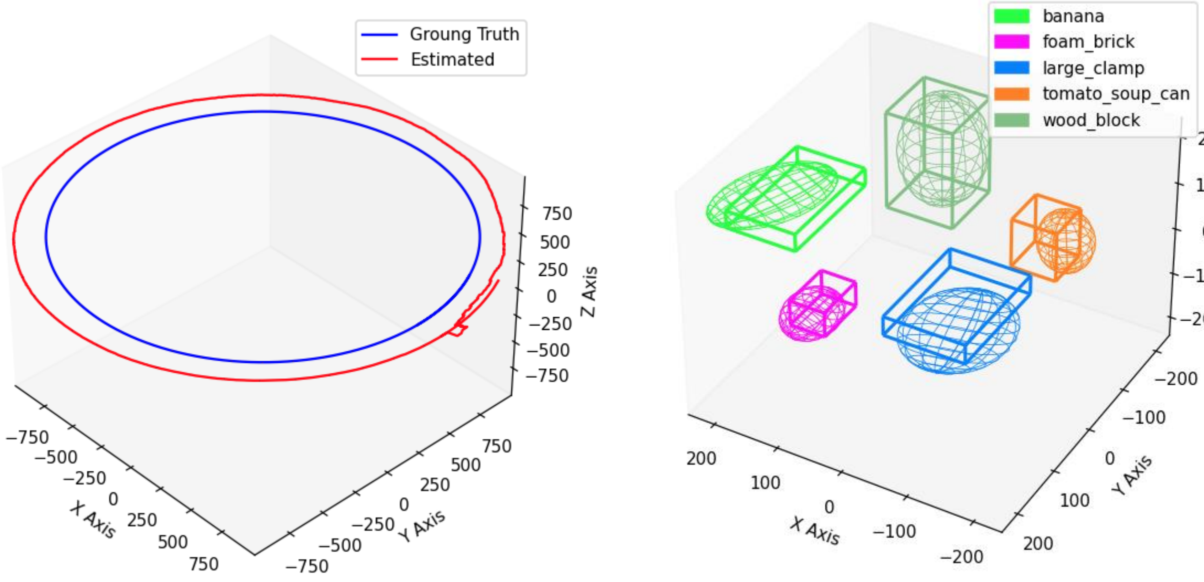
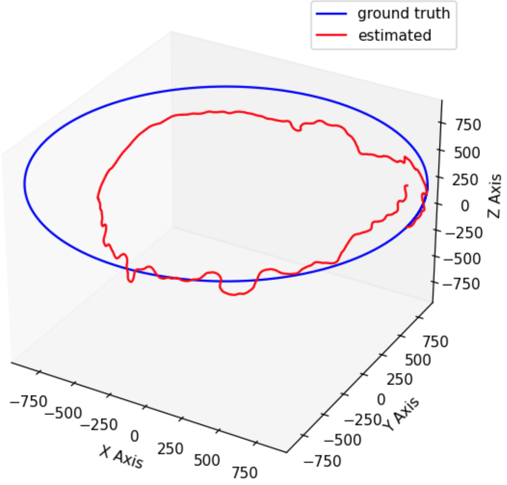


Fig 34. OA-SLAM output

Results of evaluation on scene 9 corrupted with noise



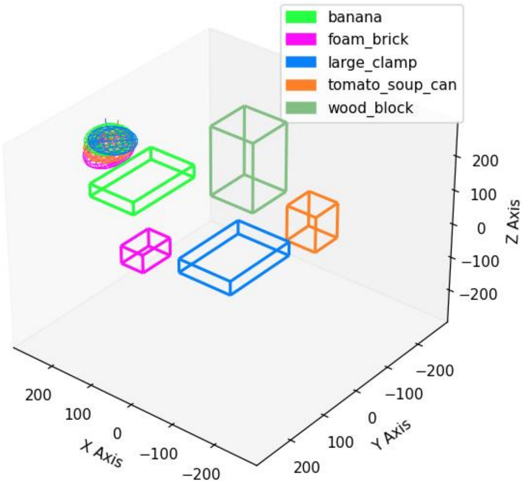
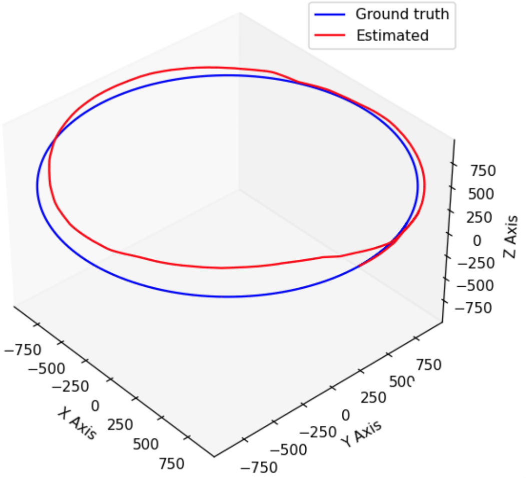


Fig 35. QuadricSLAM batch mode output



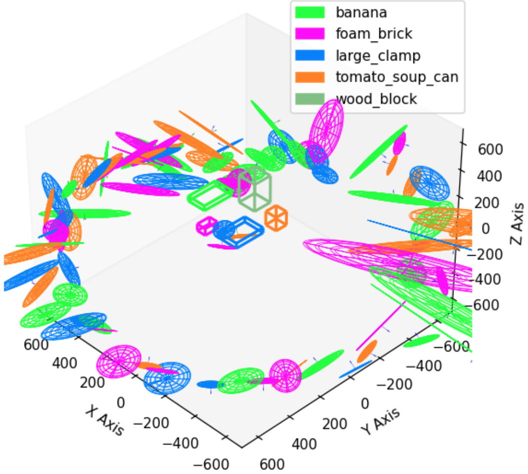


Fig 36. QuadricSLAM increment mode output

Results of localization task on scene 9 – full map

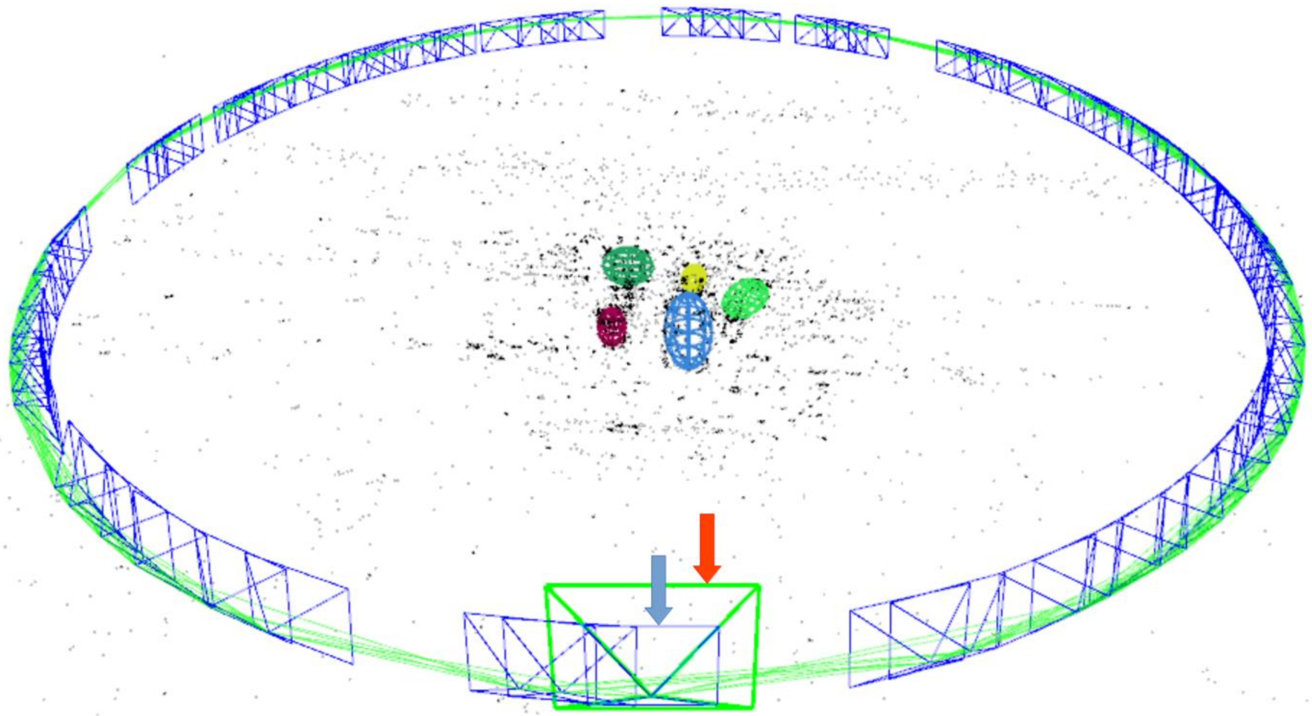


Fig 37. Localization using objects-only mode at a previously seen viewpoint

Results of localization task on scene 9 – half map

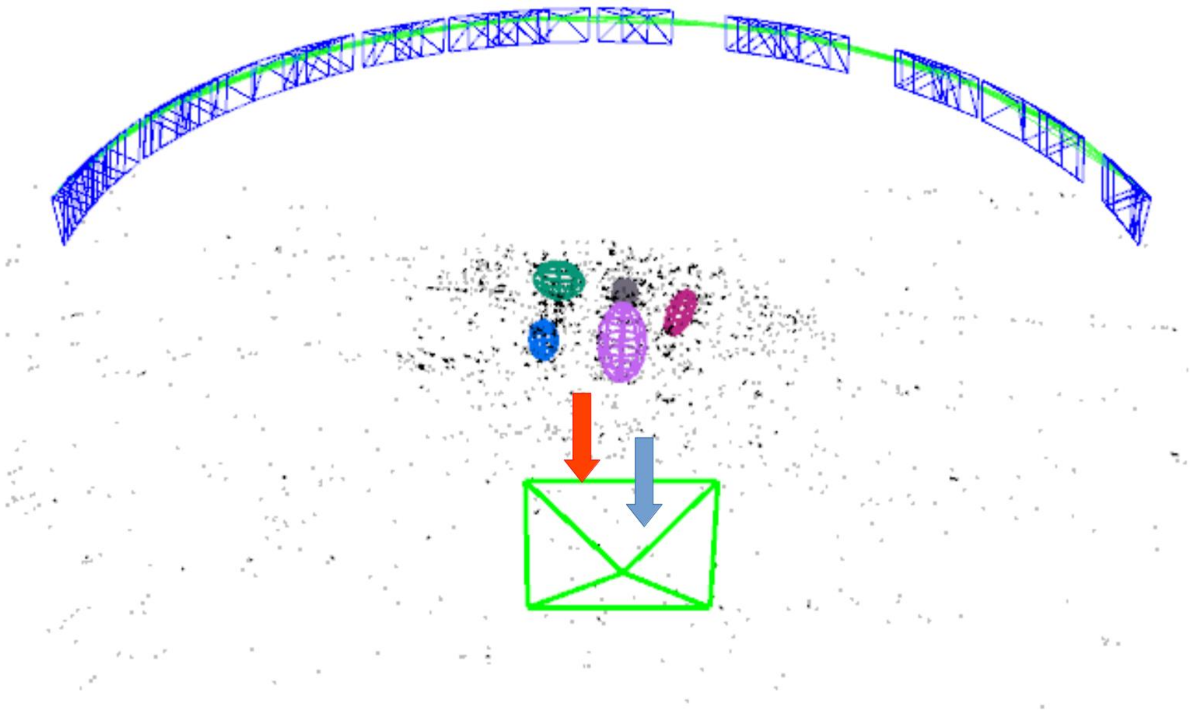


Fig 38. Localization using objects-only mode only at an unseen viewpoint

Summary

- OA-SLAM is more robust to noises in a bounding box.
- OA-SLAM is able to capture the size and orientation of the object more accurately.
- QuadricSLAM has a lower object centroid error and camera trajectory error as it has the true depth info.
- Localization could be achieved within a single RGB image using the pnp algorithm from an unseen viewpoint.
- Point constraints are more accurate for mapping and objects are useful in global relocalization.
- The parallel threads in the backend and the redundant mappoints and keyframes culling help in achieving real-time operation and efficient storage of the scene.

Future work

- Along with point and object constraints, add plane constraints to the factor graph to improve the mapping accuracy and also to provide a meaningful representation of the map.
- The addition of planes such as wall, floor and ceiling can help to extend the application of OA-SLAM to semantic navigation.

Thank you! Questions?