3D Indoor Modeling Using RGB-D Camera

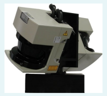
Weilun Wu, Daniel Huber
Carnegie Mellon University, Pittsburgh, Pennsylvania

Motivation

➢Detailed 3D models of indoor space, from walls to floors to objects and their configurations, can provide extensive knowledge about the environments.

➢People used to use accurate, expensive laser scanners for 3D indoor modeling. Can we use a lower cost RGB-D sensor instead? RGB-D sensors like Kinect are much cheaper than laser scanners and can provide not only depth measurements, but also color information.

➢Current existing RGB-D methods are limited in their size since they store raw point cloud data. The memory requirement becomes prohibitive after a while.

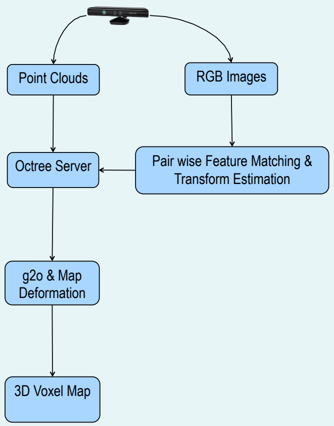


Objective

➢The objective of this project is to reconstruct indoor scenes using a Kinect by combining pose graph optimization with "as-rigid-as-possible" space deformation. We efficiently store the map in an octree structure and use the non-rigid deformation for octree map correction in a pose graph optimization framework.

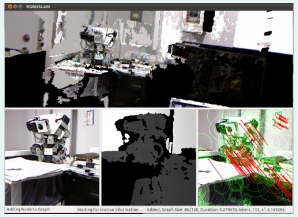
➢The algorithm should be robust to sensor noise in Kinect, memory efficient compared with other existing algorithms.

Overall Framework

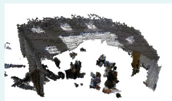


Overview of the Algorithm

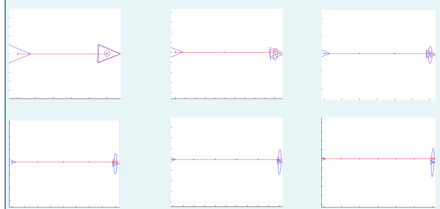
➢Incrementally register point cloud using pairwise SURF feature matching [1].



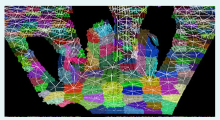
➢Store the map in an octree data structure.



➢Create a new submap when accumulated uncertainty is high or closing a loop.



➢Using supervoxel clustering to get supervoxels and use center of supervoxels as deformation graph nodes. Edges generated from supervoxel adjacency information.



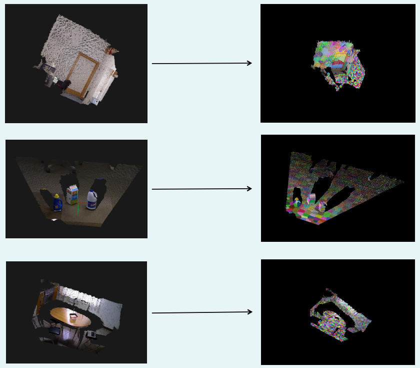
➢Get transformation associated with each node from Global Pose Graph Optimization (g2o) framework and use it as user edit constraints for map deformation.

➢Compute deformed position of supervoxel center by as-rigid-as-possible shape deformation (enforcing rigidity, smooth & user constraint).

Theory & Preliminary Result

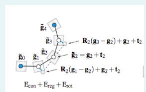
➢VCCS Supervoxel Clustering [2].

- ➢Construct an adjacency graph of occupied voxels.
- ➢Initialize supervoxels by uniform seeding and filter out seeds caused by noise.
- ➢Compute edge weight in adjacency graph.
- ➢Growing the seeds by running breadth-first-search of the adjacency graph.

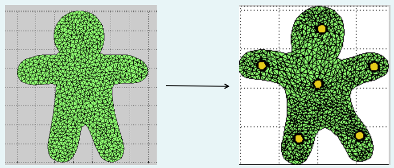


➢As-rigid-as-possible Map Deformation [3].

- ➢Construct deformation graph from uniform sampling of the model surface.
- ➢The influence of individual graph nodes on any 3D points is smoothly blended by a weighted sum of its position.

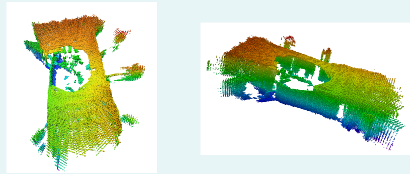


- ➢Associate affine transformation with each graph node.
- ➢Optimization (Rotation, Regularization, Constraints).

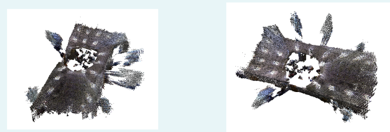


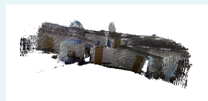
➢Voxel map obtained using only feature matching & visual odometry.

- ➢Smith Hall 2nd Floor Dataset



**Note that drift error accumulates after a while and the map becomes distorted.





**RGB-D SLAM result looks good overall. But surface lacks smoothness after pose graph optimization.

Further Work

➢Integrate code in each module into pipeline and make a working 3D indoor modeling system.

➢Collect more data sets for testing.

References

1. F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers and W. Burgard, "An evaluation of the RGB-D SLAM system", ICRA 2012
2. J. Papon, A. Abramov, M. Schoeler and F. Worgotter, "Voxel Cloud Connectivity Segmentation - Supervoxels for Point Clouds", CVPR 2013
3. R. Summer, J. Schmid and M. Pauly, "Embedded Deformation for Shape Manipulation", ACM SIGGRAPH 2007

Acknowledgements

Thanks to Daniel Huber for supporting this work and providing guidance. Thanks to my lab colleagues, Shaurya, Gaurav, Danfei and the RISS Program.