Master Thesis Specification

Michal Staniaszek (michalst@kth.se)
Supervisor: John Folkesson

February 20, 2015

1 Problem Definition

In this project, we will study techniques for searching point clouds generated using RGB-D sensors such as the Microsoft Kinect. We are interested in being able to extract query objects from large clouds, such as those generated by a robot wandering around a corridor with offices off to the sides. An important part of the task is that "interesting" objects are unknown. Thus, when analysing the cloud it will be necessary to find objects without any prior knowledge, which may require some form of segmentation. Since we don't want to waste time defining walls and floors as objects, they would be removed in a pre-processing step. To retrieve objects from the cloud it will be necessary to describe objects in some way and add them to a database in so that it is possible to distinguish between different objects. Whether this is done by representing full objects or partial objects is still to be decided. The specifics of query objects are variable, but they could be either a single frame, from which the query object must be extracted, or a full cloud generated by scanning the object in a controlled environment. The former would be the ideal case but may complicate things. The query objects will probably be extracted manually from clouds, but this process could also be automated using segmentation. The final goal of the project is to have a system which can get large clouds into a form where only objects are present, extract features of some kind from the objects in the cloud, add them to a data structure, and then query the structure for features similar to the ones generated from a query object. The simplest way of demonstrating the system is to select some objects present in the main cloud, and see how well the objects are retrieved when the system is given multiple frames of the objects in question seen from different viewpoints.

2 Literature Study

The literature study will use an initial pool of the papers cited in this section, and will expand its scope to other relevant papers found in the process. We will focus on 2D and 3D feature descriptors which can be used to compress the point clouds into more manageable sizes, and to use later in object queries [1, 2, 3, 4, 5]. Segmentation methods will also be studied in order to better understand the available techniques, but whether or not the project will make use of segmentation is unknown. It may be possible to combine 2D and 3D segmentation methods for better results [6, 7, 8]. In addition, we will look at methods for efficient storage and retrieval of feature descriptor data [9, 10].

3 Problem Solving Methods

For segmentation, some sort of oversegmentation and recombination or clustering methods such as k-means will be used. As mentioned, it may not be necessary to do the segmentation directly. Plane extraction can be done using built in ROS libraries which use RANSAC. It may be possible to use existing code for parts of the project which are not the main focus in order to save time.

4 Required Resources

- Ubuntu 14.04 machine with ROS install
- Dataset with large registered point cloud map, along with individual point clouds used to construct the map
- Access to objects present in the map, so that multiple viewpoint clouds can be constructed to use in experiments
- Access to Kinect/PrimeSense (maybe)

Week #	Start date	Activity
1	Jan 19	Gathering papers, developing understanding of project aims
2	Jan 26	
3	Feb 2	Literature study (writing introduction and background sections of final report)
4	Feb 9	
5	Feb 16	
6	Feb 23	
7	March 2	Dataset acquisition, system setup, understanding existing code (if any)
8	March 9	Extraction of objects from the full cloud, initial naive method implementation
9	March 16	Finish naive method, do some experiments
10	March 23	Write up naive method in report, start more complex method
11	March 30	Continue with more complex method
12	April 6	complex method
13	April 13	complex method
14	April 20	complex method
15	April 27	finish complex method, experiments
16	May 4	Write up complex method
17	May 11	report and presentation, method tweaks/improvements (complete draft)
18	May 18	report and presentation, method tweaks/improvements
19	May 25	report and presentation, method tweaks/improvements
20	Jun 1	report and presentation, method tweaks/improvements

Table 1: Timetable for thesis work

5 Boundaries

The project will be limited to finding query objects in the map. While it could be possible to put the system onto a working robot, this will probably not be necessary as the retrieval is the main focus of the project.

References

[1] Kate Saenko et al. "Practical 3-D object detection using category and instance-level appearance models". In: *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on.* IEEE. 2011, pp. 793–800.

[2] Kevin Lai et al. "A Scalable Tree-Based Approach for Joint Object and Pose Recognition." In: *AAAI.* 2011.

[3] Walter Wohlkinger and Markus Vincze. "Shape-based depth image to 3D model matching and classification with inter-view similarity". In: *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on.* IEEE. 2011, pp. 4865–4870.

[4] C.A. Mueller, K. Pathak, and A. Birk. "Object recognition in RGBD images of cluttered environments using graph-based categorization with unsupervised learning of shape parts". In: *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on.* 2013, pp. 2248–2255.

[5] Liefeng Bo, Xiaofeng Ren, and Dieter Fox. "Hierarchical matching pursuit for image classification: Architecture and fast algorithms". In: *Advances in neural information processing systems.* 2011, pp. 2115–2123.

[6] K. Haris et al. "Hybrid image segmentation using watersheds and fast region merging". In: *Image Processing, IEEE Transactions on* 7.12 (1998), pp. 1684–1699.

[7] H Woo et al. "A new segmentation method for point cloud data". In: *International Journal of Machine Tools and Manufacture* 42.2 (2002), pp. 167–178.

[8] Tahir Rabbani, Frank van den Heuvel, and G Vosselmann. "Segmentation of point clouds using smoothness constraint". In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.5 (2006), pp. 248–253.

[9] David Nister and Henrik Stewenius. "Scalable recognition with a vocabulary tree". In: *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on.* Vol. 2. IEEE. 2006, pp. 2161–2168.

[10] James Philbin et al. "Object retrieval with large vocabularies and fast spatial matching". In: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on.* IEEE. 2007, pp. 1–8.