Initial Plan: View Planning for 3D Object Reconstruction using a Robotic Manipulator

Author: Ivo Kosa
Student Number: C1928033
KosaI@cardiff.ac.uk

Project Supervisor:
Dr Juan Hernandez Vega
HernandezVegaJ@cardiff.ac.uk

1 Project Description:

Methods of object reconstruction from images have been around almost since the invention of photography[1] and huge advances have been made in this field with the introduction of digital image gathering and processing methods. However, these methods usually still typically rely on a manual aspect, such as taking photos of the object by hand from different angles. There are some more automated methods for example the use of a turn-table used for photography[2], although this may have limitations for larger objects. There are also various machine learning methods for object reconstruction for both single and multiple views, using large-scale repositories of 3D objects (such as ShapeNet) as training data[3]. Another method for gathering image data for object reconstruction can be found within the field of robotics. This is where so-called next-best-view planning (NBV) can be used in conjunction with a robot manipulator to effectively gather the information needed to reconstruct an object with optimal movement and minimal images taken.

The aim of this project is to develop an optimised method/ pipeline to fully and accurately reconstruct an object in 3D using images/ video taken by a robotic manipulator. This project is based around the 'Franka' robotic manipulator arm, which will be used in conjunction with a next-best-view planner in order to gather the necessary visual data. A key consideration for the success of this project will be the accuracy of the reconstructed object. As such a method of determining this accuracy will also need to be implemented.

This project will primarily be carried out in simulation using the Gazebo simulation software. If time allows during the project, the physical 'Franka' robot can also be used in order to test the progress of the work outside of simulation. The software will be written using the Robot Operating System (ROS) framework, which provides the infrastructure to run all the different aspects of the system. Individual programs will be written in python and C++ and will make use of various libraries, including the Open Motion Planning Library (OMPL) and MoveIt, for motion planning and control of the robot respectively.

This project is in conjunction with a PhD student working on object reconstruction and view planning.

2 Aims and Objectives:

The overall aim of this project is to utilise a robotic manipulator to gather images and/ or video from a scene by moving a camera around an object. This image data of the scene, taken from different angles, will then be used to reconstruct the object(s) as a 3D model. The key objectives are as follows:

Objective (i): To perform a literature review of methods of object reconstruction, NBV planning and 3D object comparison as well as to familiarise myself with the different tools necessary for this project such as ROS, C++, OMPL and MoveIt.

Objective (ii): To create a system which allows for images/ video taken by the robot to be fed into 3D reconstruction software to generate a 3D model.

Objective (iii): To create a system for optimal NBV planning, so that overall the minimum number of images will need to be taken of the scene.

Objective (iv): Find a method in order to benchmark the accuracy of the reconstructed object in direct comparison to the real (or simulated) object.

A primary risk of **objective (ii)** is an issue of compatibility, where it may not be possible to directly integrate the object reconstruction program with the robot control system. Were this to be the case, it would not be a significant issue as the images gathered could simply be standardised and manually exported to the program in question. However, the ideal outcome would involve a fully integrated/ automatic system.

The risks involved with **objective (iii)** and some aspects of **objective (ii)** are primarily focused on the requisite knowledge needed for its execution. OMPL is a C++ based library, which is a language that I don't have much experience in. This knowledge will need to be gained throughout the course of the project. In a worst case scenario the OMPL python API can be used, although this has reduced functionality it should still suffice for the scope of the project.

The primary risk for **objective (iv)** involves how data is gathered from the 3D model. In order to programmatically assess the attributes of a model, a program will need to directly interact with the files themselves to pull data. I have never used 3D models in an automated way such as this which could pose significant challenges, however there are a variety of resources available for tasks such as this. There are several open-source programs available which can perform these tasks[4], as well as many papers detailing different methods of 3D object similarity measurement[5].

Finally, another key risk of the project as a whole involves integration of the NBV planner, with the motion planning and the object reconstruction program. In an ideal scenario all of the mentioned systems will be integrated together, which may not be possible given the time-frame of the project.

3 Feasibility:

This project will not require any ethical approval as there is no data being gathered from any participants. If necessary, the use of the Franka robot manipulator arm will be provided by the Cardiff University Human Computer Interaction (HCI) Lab. The use of this robot may also involve filling out a risk assessment form. For any technical issues faced when using the Franka robot, I will be in contact with the Cardiff University Computational Robotics (CUCR) Team who will be able to provide assistance if necessary.

4.1 Work Plan - Tasks:

Literature review: reading and analysing different resources with the aim of finding ways to implement objectives **(ii)**, **(iii)** and **(iv)**. These resources will consist primarily of academic papers, but will also include videos on the topics as well as analysis of publicly available programs such as those found on GitHub.

Familiarisation with ROS and related tools: Working through various online tutorials to gain a better understanding of the use of the ROS framework, C++ in conjunction with ROS, OMPL, MoveIt as well as other simulation infrastructure available from the Cardiff University Computational Robotics (CUCR) research team.

Deliverable (1): Implementing object reconstruction

a) Initial Implementation, **Milestone (1)**: Finding and testing various methods of 3D object reconstruction using a standard set of images. Data gathered here can later be used for comparison of reconstruction methods.
b) Testing with Simulation Data, **Milestone (2)**: Once object reconstruction methods have initially been set up, they can be tested with image data gathered using the simulation.
c) Integration, **Milestone (3)**: Integrating the object reconstruction methods directly with the existing simulation. Ideally, create infrastructure which allows for easy interchangeability of the different reconstruction methods.

Deliverable (2): Implementing next-best-view planning

a) Initial Implementation, **Milestone (4)**: Selecting and implementing various methods for NBV planning within ROS.
b) Integration, **Milestone (5)**: Integrating the NBV planners within the existing simulation.

Deliverable (3): Implementing 3D object benchmarking

a) Initial Implementation, **Milestone (6)**: Finding and testing a suitable 3D object comparison tool using a clean 3D model of the object and the 3D reconstruction of this object gathered from **Milestone (2)**.
b) Integration, **Milestone (7)**: integrating this tool directly with the rest of the program, allowing for seamless benchmarking of reconstructed objects.

Evaluation: Producing a written evaluation of the project. This will primarily focus on the success of the reproduced 3D objects, but will also consider the implementation and efficiency of the NBV planner as well as the integration of the system as a whole.

Final Report: Producing a written report detailing all the findings of the project and presenting and drawing conclusions from the evaluation. The writing of the report will largely be ongoing throughout the project, focusing primarily on the literature review in the earlier weeks.

4.2 Work Plan - GANNT chart:

Tasks	**Weeks**												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**ER**	**9**	**10**	**11**	**12**
Literature review													
Familiarisation with ROS and related tools													
Implementing object reconstruction:													
a) Initial Implementation													
b) Testing with Simulation Data													
c) Integration													
Implementing next-best-view planning:													
a) Initial Implementation													
b) Integration													
Implementing 3D object benchmarking:													
a) Initial Implementation													
b) Integration													
Evaluation													
Final Report													

5 References:

[1] J. S. Aber, I. Marzolff, J. B. Ries. Small-Format Aerial Photography, Chapter 3 - Photogrammetry. 2010. ISBN 9780444532602 [Accessed 26/01/2024]

[2] V. Fremont, R. Chellali. Turntable-Based 3D Object Reconstruction. 2004 [Accessed: 26/01/2024]

[3] C. B. Choy, D. Xu, J. Gwak, K. Chen, S. Savarese. 3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction. 2016 [Accessed: 27/01/2024]

[4] K. Žarn, Reconstruction-Evaluation, 2019. https://github.com/KristianZarn/Reconstruction-Evaluation [Accessed: 28/01/2024]

[5] H. Shum, M. Herbert, K. Ikeuchi. On 3D Shape Similarity, 1995 [Accessed: 28/01/2024]