# Autonomous Active Perception Framework for Large Object Recognition Tasks

*A thesis submitted in partial fulfilment of the requirements*

*for the degree of*

Master of Engineering (Research)

*by*

**Thanh Long Vu**

*to*

School of Mechanical and Mechatronic Engineering
Faculty of Engineering and Information Technology
University of Technology Sydney
NSW - 2007, Australia

July 2022

# Certificate of Original Authorship

I, *Thanh Long Vu*, declare that this thesis is submitted in fulfilment of the requirements for the award of *Master of Engineering (Research)*, in the *Faculty of Engineering and IT* at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signed: Production Note:
Signature removed prior to publication.

Date: 31/7/2022

# Autonomous Active Perception Framework for Large Object Recognition Tasks

by

Thanh Long Vu

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Engineering (Research)

# *Abstract*

In recent decades, the technology development in hardware and software has stimulated robotics systems' implementation in various fields. Robots are most commonly deployed in dangerous or mundane working environments, significantly reducing accidents, injuries and casualties in the workforce. Ideally, robots would be designed to perform the tasks autonomously, conducting calculations and making decisions based on sensory data. However, although robotics research has continuously advanced throughout the last half a century, there are still many complicated tasks where robots cannot achieve full autonomy yet. In these scenarios, interaction from a human supervisor may be required to make control decisions either locally or remotely. The quality of the decisions made by the supervisor or operator depends heavily on the available sensory feedback from the robotics system, which helps the human perceive the environment where the robot is operating. Therefore, perception capabilities are required for all autonomous and semi-autonomous robotics systems to process and make sense of the received data, so the system or the operator can perform necessary actions.

This thesis focuses on developing an active perception framework for robots working remotely, where the human operator cannot directly perceive the surrounding environment. The specific modality of sensor data received from the remote system is coloured Three-Dimensional (3D) point cloud data obtained from sensing devices such as Light Detection

and Ranging (LiDAR) or depth cameras. Additionally, this thesis investigates the practicality and benefits of utilising Virtual Reality (VR) as a tool to visualise the data obtained from a remote system.

This thesis firstly reviews multiple frontier detection algorithms for Two-Dimensional (2D) exploration, comparing these algorithms with other notable frontier detection algorithms. The algorithms are implemented to enable insight to be gained about the performance of the inspected algorithms compared with the other notable algorithms. A 3D interactive and active mapping framework for a mobile manipulator platform based on dynamic Gaussian Probabilistic Implicit Surfaces (GPIS) was investigated and implemented to validate its efficiency while simultaneously exploring and interacting with a large pile of objects. The framework is shown to perform near real-time map updates for a dynamically changing environment due to its probabilistic nature. Two perception systems are presented that employ the point cloud data in the framework to perform object detection, pose estimation and scene overlay annotation. Finally, a framework for sensor data visualisation in VR environments is presented, which acquires, transmits and renders real-time RGB-D sensor data with integrated automatic perception annotations in a VR environment. This enables various data modalities and limiting factors to be investigated and optimised to improve the subjective cognitive workload.

Experiments are conducted in both simulated and real-life scenarios. The two real robotic platforms used in the experiments are mobile manipulators. The first platform is composed of a 6 Degree of Freedom (DOF) manipulator and a commercial mobile base platform and was mainly used to validate exploration algorithms. The second mobile manipulator platform comprises a custom-designed mobile base and a 5 DOF manipulator, equipped with a camera system consisting of two RGB-D calibrated cameras and a real-time VR interface. This platform was used to conduct and investigate the optimal configuration for operator performance during collaborative autonomy tasks.

# Contents

# List of Figures

# List of Tables

# Acronyms & Abbreviations

| | |
|---|---|
| **2D** | Two-Dimensional |
| **3D** | Three-Dimensional |
| **AR** | Augmented Reality |
| **BFS** | Breadth-First Search |
| **DOF** | Degree of Freedom |
| **EWFD** | Expanding-Wavefront Frontier Detection |
| **FFD** | Fast Frontier Detection |
| **FOV** | Field of View |
| **FTFD** | Frontier-Tracing Frontier Detection |
| **GDP** | Gross Domestic Product |
| **GP** | Gaussian Process |
| **GPIS** | Gaussian Probabilistic Implicit Surfaces |
| **GPOM** | Gaussian Processes Occupancy Maps |
| **HALO** | High Access Localised Operations |
| **ICP** | Iterative Closest Point |
| **IG** | Information Gain |
| **IoU** | Intersection Over Union |

**LiDAR** Light Detection and Ranging

**MBZIRC** Mohamed Bin Zayed International Robotics Challenge

**NBV** Next Best View

**PCA** Principal Component Analysis

**PCL** Point Cloud Library

**RANSAC** RANdom Sample And Consensus

**ROS** Robot Operating System

**SLAM** Simultaneous Localisation and Mapping

**UTS** University of Technology, Sydney

**VR** Virtual Reality

**WFD** Wavefront Frontier Detector

**WFD-INC** Incremental Wavefront Frontier Detector

# Glossary of Terms

| | |
|---|---|
| Autonomous | Without human intervention. |
| Active and passive perception | Active perception involves actively seeking and processing sensory information through exploration and interaction with the environment, while passive perception involves receiving sensory information without any active involvement on the part of the system. |
| Collaborative Autonomy | Refers to a mode of operation in which humans and robots work together as a team to achieve a common goal, with each entity contributing their unique strengths and abilities. This approach leverages the strengths of both humans and robots to enhance performance and efficiency beyond what either entity could achieve individually. |
| Human operator | A person who manipulates the robot's action at a remote location. |
| Human supervisor | A person who oversees the robot's operations and provides the robot with instructions when needed. |
| Manipulator | Robotic arm |
| Mobile base | A robot that exclusively navigates on horizontal planes. |
| Occupancy map | A discretised grid representing the environment is made up of cells with values that indicate the likelihood that a given point in space is obstructed. |
| Voxel | Volumetric Pixel represents a 3D cube-like volume in Euclidean space. |