論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Utilization of Environmental and Robotic Characteristics for Environment Mapping towards Fukushima Daiichi Decommissioning
著者(和文)	汪 振宇
Author(English)	Zhenyu Wang
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12356, 授与年月日:2023年3月26日, 学位の種別:課程博士, 審査員:遠藤 玄,鈴森 康一,松永 三郎,木倉 宏成,髙山 俊男
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12356, Conferred date:2023/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

令和 4 年度　博士論文

Utilization of Environmental and Robotic Characteristics for Environment Mapping towards Fukushima Daiichi Decommissioning

東京工業大学　工学院

機械系　機械コース

WANG ZHENYU

指導教員　遠藤 玄 教授

副指導教員　鈴森 康一 教授

Abstract

The decommissioning works of Fukushima Daiichi disaster are of great significance, and the preliminary investigations require specialized robotic investigation systems for the harsh environment. The preliminary investigation stages of fuel debris retrieval phases gave rise to need of remotely-operated robotic investigation systems with robust environment mapping methods.

However, the feasibility of conventional robotic investigation systems was limited by lacking robust underwater investigation ability due to visibility and limitation, and limitations in manipulator weight, etc. Besides, conventional environment mapping methods were not adaptable to specific environment and robot, thus suffering from compatibility degradation. In order to solve these problems, this thesis proposed a new mobile robotic investigation system "RhinoUS-II" with more robust underwater investigation ability from a 2D phased array ultrasonic sensor; and proposed a long-reach light-weighted flexible wire-drive manipulator "BWD Robot" with compact and light-weighted design. On the basis of the characteristics from the specific environment (grating lattices with known size and regular shape) and the specific robot (motion patterns' features), three improvement methods for conventional environment mapping methods Visual Odometry and Structure from Motion were also proposed.

The improvement methods include (1) a perspective transformation method to provide intuitive top view mapping by using homography matrix; (2) a motion-based compensation method to correct map areas with feature tracking failures for Visual Odometry; and (3) a dataset creation method to improve the compatibility between Structure from Motion and flexible manipulator by using motion patterns of the manipulator. The effectiveness of proposed three improvement methods were quantitively confirmed through experiments with the hardware prototypes. The refined Visual Odometry method with the proposed perspective transformation method could generate top view maps and count the number of past grating lattices, and achieved localization accuracy with error rate of 12.5mm/1500mm in the front-back direction, 40.5mm/1500mm in the left-right direction; The compensation method improved the smoothness and consistency of generated maps from 62% to 90%; The proposed dataset creation method were evaluated by using a widely-used Structure from Motion algorism (colmap) together with two dense reconstruction software, colmap_rob and instant-ngp. By using the proposed dataset creation method, the average reconstruction accuracy improved at 3.7% for colmap_rob, 14.7% for instant-ngp; completeness dropped at 1.3% for colmap_rob, 2.7% for instant-ngp; runtime reduced at 22.0% for colmap_rob, 9% for instant-ngp.

As the conclusion, proposals of two robotic investigation systems provide new insight of using ultrasonic sensor/ flexible manipulator for future Fukushima investigations. Implementation of three proposed improvement methods allow conventional environment mapping methods to

utilize environmental and robotic characteristics as additional prior knowledge, therefore may enhance the feasibility under more realistic and harsher environment.

Table of Contents

Figures

Tables

List of Symbol

a	Width of one element
d	Inter-element spacing
θ	Steering angle
Δt	Time delay
c	Speed of sound
f_{Di}	Doppler frequency observed at the i_{th} channel
f_0	Frequency of the emitted pulse echoed by the particle
$\boldsymbol{e_e}$	Unit vector in the direction of the measurement line
$\boldsymbol{e_l}$	Unit vector in the direction from the particle to the i_{th}-channel element
$\boldsymbol{V}$	Particle velocity
P_n	Matched point group in the n-th frame
$P_n[i]$	The i-th point in the matched points pair
$\boldsymbol{C_n}$	Estimated camera motion at n-th frame
$\boldsymbol{A}$	Decomposed rotation matrix
$\boldsymbol{b}$	Decomposed translation vector
$\theta_{c,n}$	The instant rotation angle in the n-th frame
$x_{c,n}$	The instant translation in x-direction in the n-th frame
$y_{c,n}$	The instant translation in y-direction in the n-th frame
O_g, X_g, Y_g	The coordinate system defined with sides of the grating lattice
O_r, X_r, Y_r	The coordinate system defined with sides of the robot
O_c, X_c, Y_c	The coordinate system defined with sides of the checkerboard
x_i, y_i	The coordinate of a grid point in the initial plane of perspective transformation
$x`_i, y`_i$	The coordinate of the corresponding grid point in the transformed plane of perspective transformation
$\boldsymbol{H}$	Homography matrix
h_{ij}	The element at (i, j) of homography matrix
F_n	All the extracted feature points group in n-th frame
R_n	The rate of successfully matched feature points in n-th frame
$\boldsymbol{\xi_r}$	Posture vector of robot position on plane
x_r, y_r	Coordinate of mass center of the robot chassis
$\boldsymbol{R_c}$	Vector connecting the robot mass center to the camera center
$\boldsymbol{\xi_c}$	Posture vector of camera plane on plane

$x_c y_c$	Coordinate of camera center
${}^{R}\boldsymbol{V_r}$	Relative velocity of robot under the robot coordinate system
${}^{R}\boldsymbol{V_c}$	Relative velocity of camera under the robot coordinate system
ω_r	Angular velocity of the robot center
v_x, v_y	Velocity vector component in X and Y axis
F_t	Flag quantity to judge the magnitude difference between velocity components
θ_j	Rotation angle of the joint
θ_{mw}	Rotation angle of the winding up motor
θ_{mf}	Rotation angle of the feeding motor
D_{motor}	Diameter of the pulley
D_{joint}	Diameter of the joint
N_{acc}	Number of accurate points in the reconstructed point cloud
N_{rec}	Number of points in the reconstructed point cloud
N_{com}	Number of complete points in the ground truth point cloud
N_{truth}	Number of points in the ground truth point cloud
Com	Completeness benchmark
Acc	Accuracy benchmark
β	The singular value feature
Ω	Area of a digital image
Ω_b	Area of the blurry regions in a digital image
k	Weight constant
D	Benchmark to estimate the degree of blur

List of Abbreviation

TEPCO	Tokyo Electric Power Company Holdings
RPV	Reactor Pressure Vessel
PCV	Primary Containment Vessel
CRD	Control Rod Drive
IRID	International Research Institute for Nuclear Decommissioning
BWD	Bundled Wire Drive
DoF	Degree of Freedom
SfM	Structure from Motion
MVS	Multi View Stereo
NVS	Novel View Synthesis
NeRF	Neural Radiance Fields for View Synthesis
NGP	Neural Graphics Primitives
IMU	Inertial Motion Unit
UVP	Ultrasonic Velocity Profiler
CAN	Controller Area Network
ROS	Robot Operating System
SLAM	Simultaneous Localization and Mapping
VO	Visual Odometry
URDF	Unified Robot Description Format
SVD	Singular Value Decomposition

Chapter.1 Introduction

1.1 Background

On March 11th 2011, the Great East Japan Earthquake occurred in the north-western Pacific Ocean, leading to a series of severe problems in east Japan, including a massive tsunami and the most significant nuclear incident since the Chernobyl disaster, known as the Fukushima Daiichi Nuclear Disaster. When the earthquake occurred, in the Fukushima Nuclear Power Plants, Unit 1 to Unit 3 reactors were operating with fuel rods loaded, Unit 4 to Unit 6 reactors were under periodic inspection. When detecting the earthquake, the active reactors were automatically shut down. Soon, these shutdowns and regional power supply shortage resulted in failure of the electricity supply to the underground diesel generators as the emergency power supplies. While, these generators were crucial in providing power to the pumps that circulate coolant through the reactors' cores, as an essential process for removing residual decay heat. However, the tsunami that followed the earthquake, measuring 14 meters in height, overwhelmed the plant's seawall and flooded the lower levels of the reactor buildings. This flooding caused the failure of the emergency generators and loss of power to the circulating pumps, leading to nuclear meltdowns, hydrogen explosions, and the release of radioactive contamination in Units 1, 2, and 3. Additionally, the spent fuel pool of the previously shut down reactor 4 also experienced an increase in temperature due to decay heat from newly added fuel rods. After the disaster, in order to reduce and avoid further expansion of the radiation hazard, a long-termed decommissioning process had started.

Overall, the decommissioning of the Fukushima nuclear power plant presents an extremely challenging task that requires careful planning and execution to ensure the safety of workers and the public. The roadmap of decommissioning tasks is shown in Figure 1.1. Currently, relevant workers and researchers from Tokyo Electric Power Company Holdings (TEPCO) and International Research Institute for Nuclear Decommissioning (IRID) have been conducting the Phase 2 of the roadmap, as the period until start of fuel debris retrieval.

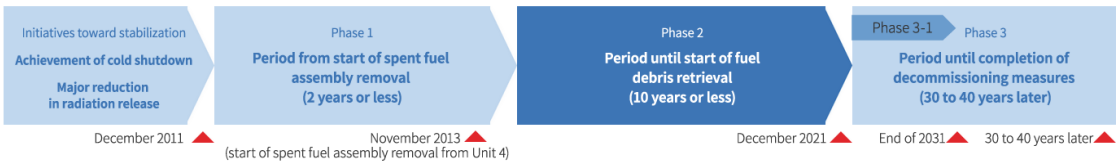


Figure 1.1 The roadmap of decommissioning tasks. [1]

The retrieval processes of fuel debris are of great significance. Fuel debris refers to the melted fuel and other substances after natural cooling and re-solidification. A general procedure of fuel debris retrieval should consist of several stages including:

- Investigation stage. As the prerequisite of the following stages, the investigation stage aims to acquire the necessary prior knowledge for the subsequent stages, including the internal

state of reactors, intuitive and accurate estimation of fuel debris distributions, and available routes for further investigations. Therefore, the investigation stage requires suitable investigation strategy to fit with the specific states of the target reactors, which should possess intuitive environment mapping ability and abundant robustness against the harsh environment conditions.

- Removal stage. After the preliminary investigation on the fuel debris distributions, the removal stage aims to approach the fuel debris, and remove them out of the reactor for the subsequent stages. The most commonly-utilized retrieval method is the submersion method, which will submerge the whole containment vessel to avoid radiation diffusion and then conduct the retrieval. Therefore, considering the possibly existing water leakages inside the reactor, it's impossible to conduct the submersion method before localizing and fixing all the water leakages. In terms of platform requirements, the removal stage demands high remote operation accuracy and prior knowledge on the physical properties of the fuel debris.
- Harmless and cleanup stage. At last, the retrieved fuel debris should be transported to special storage facilities with radiation-resistant containers. Then, further treatment processes to make it harmless should be conducted.

Currently, the process of fuel debris retrieval is at the investigation stage. In the investigation stage, concerned topics can be concluded as: (1) Access strategy to the target reactor; (2) Development of remotely-operated robotic platform; (3) Radiation resistance processing for robots and sensors; (4) Fuel debris and water leakage localization technique; (5) Autonomous ego-localization and environment mapping technique for robot. Within these topics, this thesis aimed to solve three key problems, listed below:

- The conventional robotic investigation system "PMORPH2"[7] with only underwater camera could not provide intuitive and robust images for spatial awareness of underwater fuel debris and water leakage localization process.
- Designed for future investigations of Unit 2, performance of the tough manipulator "Veolia-MHI Arm"[12] was limited by heavy arm weight, weak collision resistance, and difficulty in joint failure recovery problems.
- Conventional environment mapping methods, Visual Odometry (VO) and Structure from Motion (SfM) could not adapt to the specific environment and robot by taking advantage of their characteristics, therefore suffering from compatibility degradation.

The following sections are organized as follows. Section 1.2 presents the previous investigation results by PMORPH2 in Unit 1 and discusses the low visibility problem when using underwater camera to conduct fuel debris and water leakage; Section 1.3 describes the existing problems of Veolia-MHI arm in details. Then, Section 1.4 discusses the characteristics of the

specific environment and robot in reactor investigation and conventional methods' inability to fully utilize them. Section 1.5 concludes these existing problems and introduces proposals with research objectives. At last, Section 1.6 serve as a chapter summary and a thesis reminder.

1.2 Conventional Robotic Investigation System "PMORPH2" Utilized in Previous Investigation of Unit 1

For the reactors which suffered from serious core damage like Unit 1 and Unit 2, a common problem is the possibility of fuel debris exceptional spread [2] from the Reactor Pressure Vessel (RPV). Especially, the assumed distributions of fuel debris in Unit 1 and Unit 2 turned out to be quite different according to the degrees of damage. This section will present the situation of PMORPH2 in the flow of the access strategy (Section 1.2.1), the previous investigation method (Section 1.2.2), and the existing problem (Section 1.2.3).

1.2.1 Internal State and Access Strategy to Bottom Pedestal

Access strategy refers to the plan and methodology used to gain entry to the site of the incident, as well as to collect and preserve relevant information. In order to minimize the exposure to radiation, a suitable access strategy should be well planned on the basis of the internal state of the target reactor, including the possible fuel debris distribution and the available entry route. Figure 1.2 shows the structure of nuclear reactors in Fukushima Daiichi Power Plants [3]. The light green part indicates the shape of the Primary Containment Vessel (PCV). Fuel rods were supposed to be stored in the RPV, which is in the upper areas of PCV. Below the RPV, the deep pink part indicates the supporting structures and the Control Rod Drive (CRD) housing. Below the CRD areas, there is a grating floor around the pedestal which serves as the supporting platform. The grating floor is connected to outside through X-6 penetration and X-100B penetration, the yellow and orange part in the figure, respectively. The light blue part in bottom shows the suppression pool, connecting with the PCV bottom through the vent pipe. While, due to the overheating and core damage in the RPV, the fuel debris may have dropped to the bottom of the PCV. In order to grasp prior knowledge of the distribution of fuel debris, in April 2015, a pre-investigation on the grating around the pedestal inside Unit 1 PCV was conducted by IRID [4], also known as "Investigation B1".

In investigation B1, the X-100B penetration was utilized as the entry to the PCV. The preliminary investigation showed that there were no obvious obstructions around the X-100B penetration and the opening to access the PCV bottom, which indicated that the penetration could serve as entry route for further investigation to approach the PCV bottom. After passing through the X-100B penetration, the remotely-operated robot platform used the grating floor at the 1st floor inside PCV as the supporting ground, and moved with designed routes around the central

pedestal. Temperature and radiation dose at each investigation point was measured with the equipped sensors on the robot platform. The radiation level on the grating floor was from 4.7 Sv/h to 9.7 Sv/h, slightly lower than expectation. While the temperature was from 17.8 degree to 21.1 degree [4]. The findings of investigation B1 revealed the possibility of accessing to the basement pedestal of the PCV, where the fuel debris may have dropped.

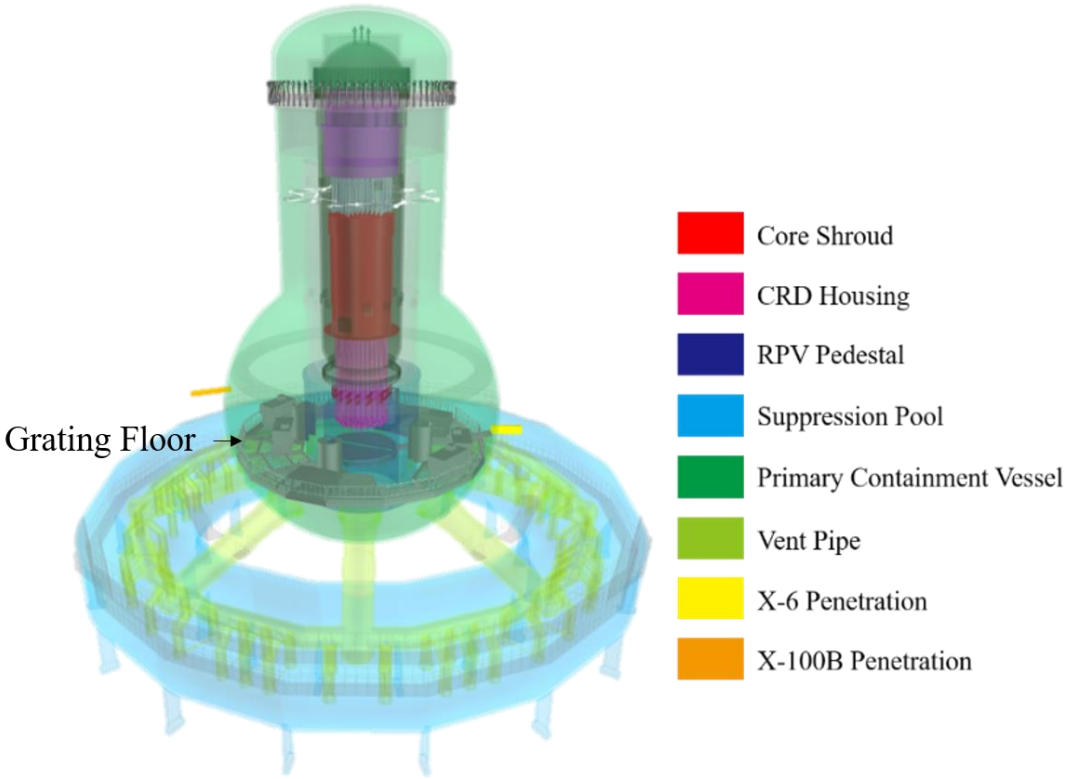


Figure 1.2 The 3D structure of nuclear reactors. [3]

In March 2017, IRID and Hitachi conducted the further investigation aiming to access the basement floor of pedestal, named as “Investigation B2” [5]. As the results, the existence of sediments like pipes was confirmed on the bottom pedestal of PCV. In terms of radiation level, the dose rates dropped when entering the water, but increased when approaching the bottom, which indicated the possibility that the sediments on the bottom pedestal were source of radiation [5]. Through the previous investigations of the internal state of Unit 1 PCV, Figure 1.3 shows the assumed distribution of fuel debris.

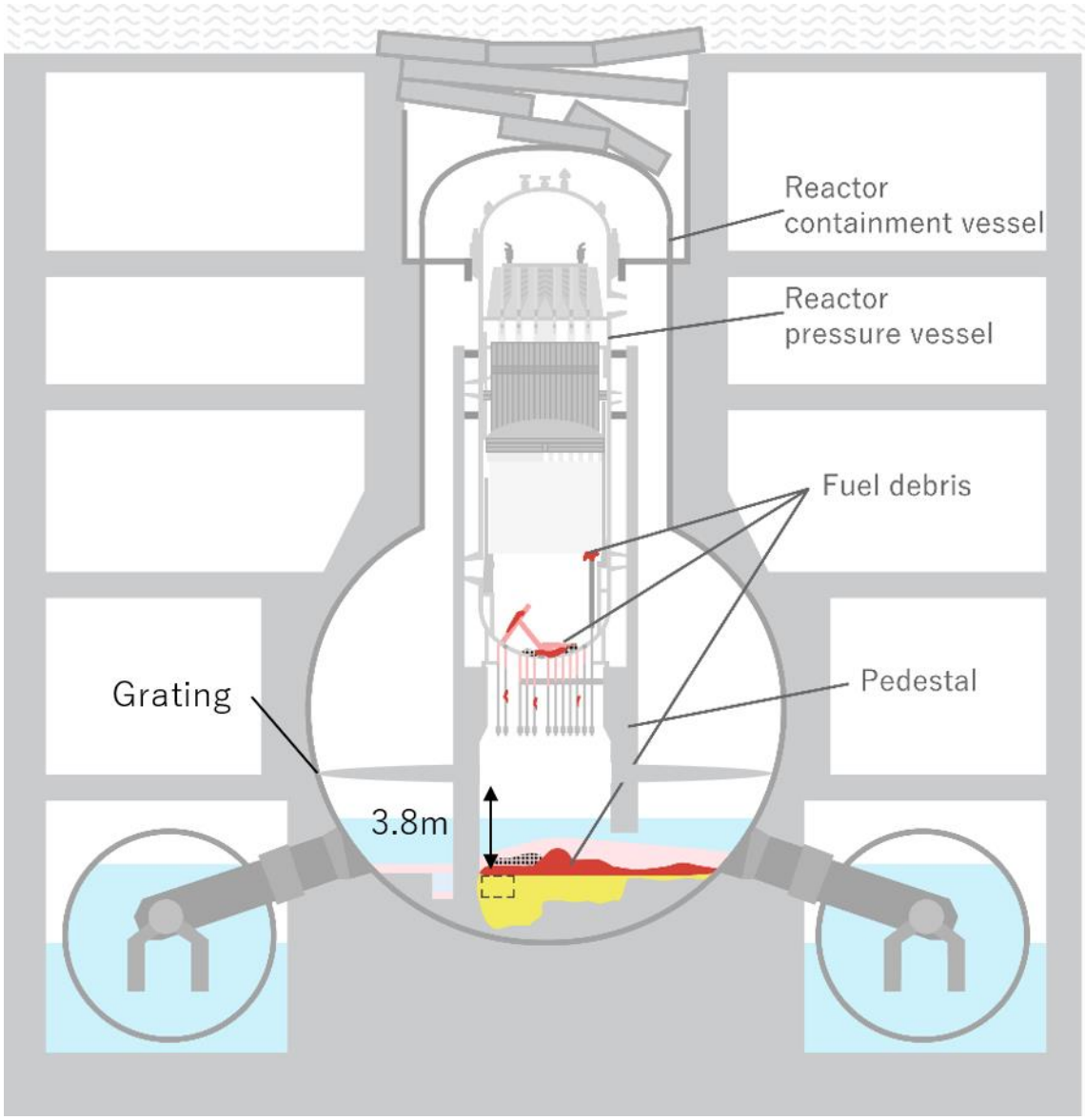


Figure 1.3 The assumed distribution of fuel debris in Unit 1 reactor. [6]

During the overheating accident, the fuel and the metal cladding (the outer jacket of fuel rods) in the RPV were assumed to melt through the RPV's bottom, fusing the below supporting structures and CRD together, and re-solidified as fuel debris at the bottom of PCV pedestal. Therefore, the preliminary assumption indicated that nearly all fuel had melted into the containment vessel, and might have widely spread on the basement pedestal. Apart from the fuel debris, there was also the danger of polluted water outflow from the unknown water leakages in the pedestal, which hindered implementation of the submersion method. In the next section, details of the robotic investigation system PMORPH2 utilized in investigation B2 will be discussed.

1.2.2 PMORPH2 in Investigation B2

For a highly contaminated site, it is way too dangerous for humans to enter the damaged reactor due to the high levels of radiation. The use of remotely-operated robotic platforms allows investigators to access the site and collect data and samples without exposing themselves to radiation. Besides, remotely-operated robotic investigation system can be equipped with specific sensors that can collect data and images from the site, providing investigators with a more detailed and accurate picture of the internal reactor. According to the concrete situations, the robotic platform needs to fit with the requirements of access strategy like limiting its size or weight. In the last section, the previous investigations of the internal state of Unit 1 reactor showed that the fuel debris may concentrate on the bottom pedestal of the PCV. Besides, the X-100B penetration served as the available route to internal PCV, and the grating floor served as the supporting platform with an opening to the bottom pedestal of the PCV. On the basis of these findings, the investigation B2 was successfully conducted to reach the bottom pedestal. The target of the investigation B2 was to gather information of the possible distribution of fuel debris and measure the radiation levels near the bottom pedestal. Taking account of the size of X-100 penetration (100mm), and the height difference between the grating floor and the bottom pedestal (3.8m), Hitachi developed a configure-changeable robotic platform equipped with a winch mechanism named as PMORPH2 [7], which was utilized in “Investigation B2”. Figure 1.4 shows the appearance of PMORPH2 on a mock-up grating floor.

Figure 1.4 Configuration changeable robot PMORPH2 by Hitachi. [7]

PMORPH2 is a coined word of “PCV” and “metamorphose”, indicating that configuration changeability is one of its core features. The structure of PMORPH2 can be generally divided into two side crawlers, and a winch part. In Figure 1.4, PMORPH2 was set to the “U-shape”, with the two side crawlers folded. The silver color part refers to the winch part, which utilizes a winch mechanism to carry the sensor unit. When passing through the narrow penetration, PMORPH2 expanded the side crawlers along the direction of the winch part, forming “I-shape” to reduce its width. The size of PMORPH2 with the two configurations is shown in Table 1.1.

Table 1.1 The size of PMORPH2 with two configurations. [7]

Configurations	Size	Value
Guide pipe walking (I-shape)	Length (mm)	699
	Width (mm)	72
	Height (mm)	93
Grating walking (U-shape)	Length (mm)	316
	Width (mm)	286
	Height (mm)	93

After entering the PCV through the narrow penetration, PMORPH2 utilized the grating floor as the supporting plane, and moved to the designed investigation points in order. When it reached the opening to the bottom pedestal, the robot stayed on the 1st grating floor, and sunk the investigation sensor into water between the grating floor and the bottom pedestal.

Before discussing the selection of sensing device and method, it’s necessary to define the concerned tasks of investigation. Exploration of an unknown environment with remotely-operated robotic platforms has been a popular scenario in multiple fields of study, including path planning, robotics, computer vision, etc. The core problems in this scenario can be naturally split into two problems of state estimation: tracking the movement of robotic platform; and providing a comprehensive measurement or reconstruction of the environment (external sensing). However, these two problems are like a chicken-and-egg relationship. In the actual investigations, preference will also vary from the concrete investigation contexts. In this thesis, the following discussion will also follow the bisectional structure of “Ego Localization” and “Environment Mapping”.

Furthermore, in the case of investigating the bottom pedestal of Unit 1, the investigation areas could be divided into two aspects, according to the spatial relationship. Concretely speaking, when moving on the grating floor, the robot platform needed to localize itself, and measure the rough map of the grating, thus to reach the planned investigation points correctly. This investigation area can be defined as “Pedestal Grating”; On the other hand, after reaching the opening to the bottom pedestal, the robot needed to grasp the depth of the sunk sensor, and provide visual information of the bottom pedestal, thus to localize possible fuel debris and water leakages.

This investigation area will be defined as “Water under Pedestal Grating”. The corresponding relationships of the investigation tasks of each area can be referred in Table 1.2. Especially, selection of suitable investigation sensor and the corresponding sensing method is of great importance to provide intuitive results with robustness in the harsh environment.

Table 1.2 The corresponding relationships of investigation tasks of the concerned areas.

Investigation Areas	**Ego localization**	**Environment Mapping (External)**
Pedestal Grating	Localize the position of robot on the grating floor	Reconstruct the map of the grating floor
Water under Pedestal Grating	Grasp the position (depth) of the underwater sensor	Gather information of possible fuel debris and water leakages

As for the equipped sensors, PMORPH2 carried 5 cameras and a dosimeter. In the 5 cameras, two cameras were installed on the crawlers for the area of “pedestal grating”. Also, two cameras were utilized for the area of “water under pedestal grating”. One camera was installed on the tip of the winch mechanism to confirm the winch depth; an underwater camera combined with the dosimeter was hanged by the winch mechanism, serving as the investigation sensor to capture images when approaching the bottom pedestal [7]. At last, a camera on the tip of the crawler was used to provide assisting view when passing through the penetration.

1.2.3 Underwater Camera Problems of PMORPH2

As has been indicated in Section 1.1, the underwater camera utilized by PMORPH2 suffered from the low visibility problem due to turbid water, as Figure 1.5 shows.

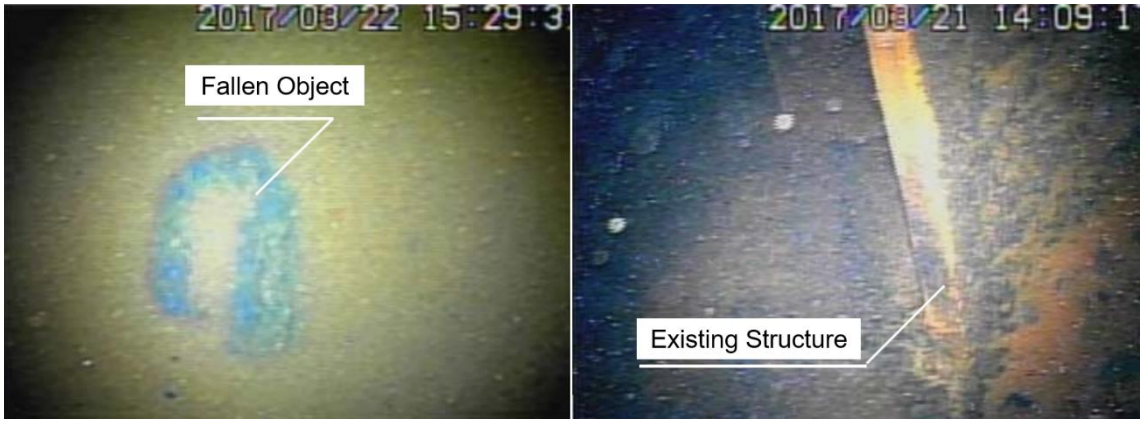


Figure 1.5 The underwater view influenced by turbid water. [5]

In this figure, the underwater views appear quite blurry, making it hard to recognize the fallen object or the existing structure marked with arrows. The visibility of the underwater camera might

have been greatly reduced by the high concentration of suspended particles in the turbid water. The turbidity also might scatter light, causing distortion and reduce the overall images' quality.

Besides, the previous investigation showed the increasing tendency of dose rate when approaching the bottom, which might limit the working time of underwater camera considering its radiation resistance. Therefore, these problems can be concluded as: previous Unit 1 investigation result showed that the conventional robotic investigation system "PMORPH2" with only underwater camera could not provide intuitive and robust images for spatial awareness of underwater fuel debris and water leakage localization process.

1.3 Conventional Tough Manipulator "Veolia-MHI Arm" Designed for Future Unit 2 Investigation

Similarly, the introduction of the conventional tough manipulator "Veolia-MHI Arm" will also follow the flow of internal state (Section 1.3.1), arm details (Section 1.3.2), and existing problems (Section 1.3.3). Especially, as for the investigation context in Unit 2 reactor, the internal state is different from that of Unit 1 reactor, which gave rise to new requirements on the robot platform and the corresponding sensing setup.

1.3.1 Internal State and Previous Investigations

Similar with the situation of Unit 1, Unit 2 reactor also suffered from serious core damage inside the RPV, thus the possible spread of fuel debris. Investigations inside the pedestal chose the access strategy of entering from X-6 penetration, and move on the CRD rail. From Jan 2017 to Feb 2019, aiming to confirm the state of fallen debris from CRD housing and internal structures, three investigations towards the inside areas of pedestal were conducted, named as "Investigation A2". The first investigation conducted in Jan 2017 had the similar concept with that of Unit 1 reactor, including usage of a self-propelled investigation device on the grating platform inside the pedestal [9]. Figure 1.6 shows the scorpion-shaped investigation device by Toshiba. Similar with PMORPH2, it could also transform its shape from line to scorpion. A camera was installed on the end of the movable scorpion "tail"; thus, it could capture when moving with the "tail". It entered the PCV by passing through the X-6 penetration, and moved on the CRD rail to approach the internal pedestal. Unfortunately, the left crawler belt of the scorpion robot got stuck and failed to move forward any more. Hence, with its cable cut, the scorpion robot was left to side of the CRD rail to avoid blocking the following investigation.

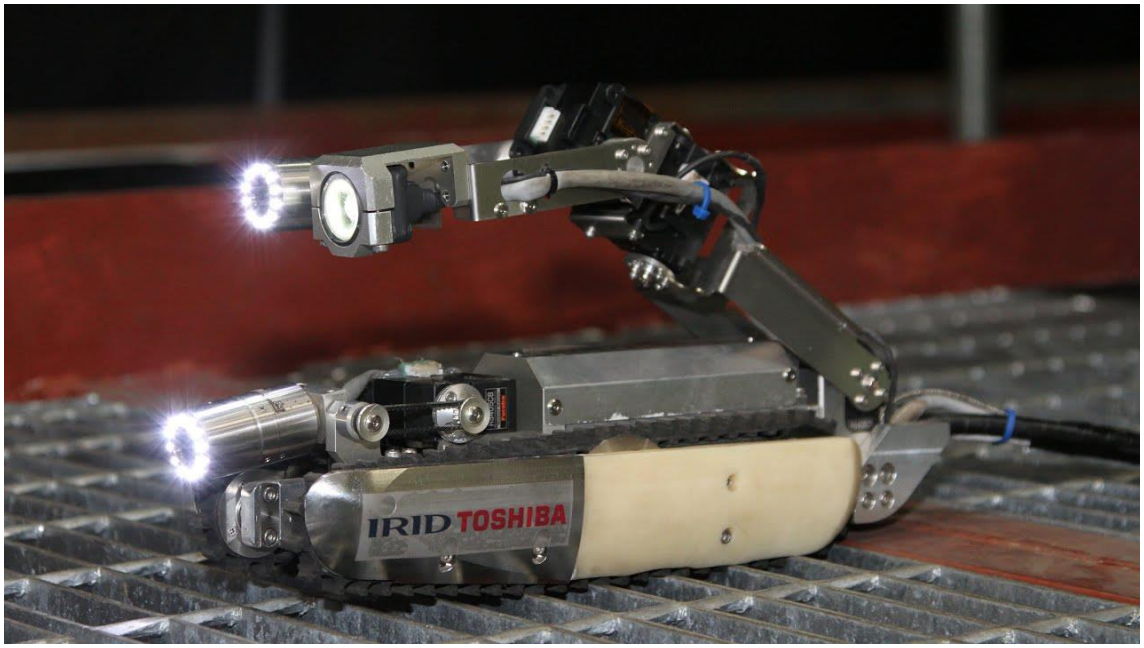


Figure 1.6 The scorpion-shaped self-propelled investigation device by Toshiba [9].

Still, it gathered precious information around the pedestal aperture with the equipped camera, dosimeter, and thermometer. Then, in Feb 2018 and Feb 2019, the subsequent investigations successfully utilized a fishing rod-shaped access device to reach the internal pedestal by hanging down the investigation sensor through the platform, named as "Investigation A2`". The investigation setup of investigation A2` is shown in Figure 1.7. The slim part in yellow indicates the concerned fishing rod-shaped access device. It passed through the X-6 penetration to enter the PCV, and used the retractable pipe on the front of the guide pipe to approach the pedestal. On its front end, the sensor combination included a camera, a dosimeter and a thermometer [10].

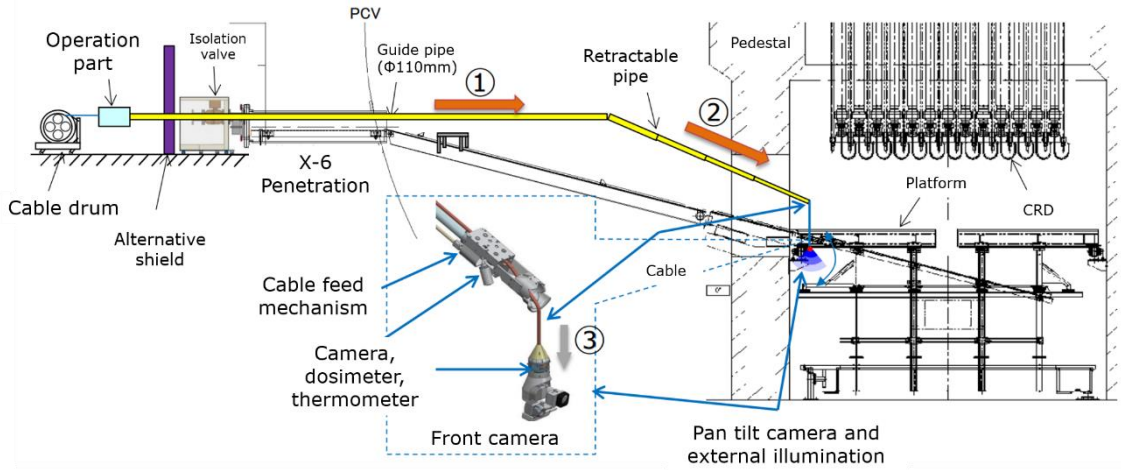


Figure 1.7 The investigation setup of investigation A2`.

On the bottom of the pedestal, sediments in shape of pebbles and clays were captured, which might be the dropped fuel debris from the RPV. Besides, by approaching the sediments with the

equipped camera, the preliminary information of the approximate size of sediments were acquired. As for the dose rate, the tendency of slightly increasing dose rate when approaching the bottom was confirmed [11]. On the other hand, the average dose rate inside the pedestal was lower than that outside the pedestal, which indicated that the fuel debris distribution might be different from Unit 1 reactor. As the investigation results, on the pedestal grating, fallen melted structures were not confirmed. Besides, the water level in Unit 2 is also slightly lower than that of Unit 1.

1.3.2 Veolia-MHI Arm Developed by IRID

In the previous section, the internal state of Unit 2 has been described, especially the difference of the further robot selection with that in Unit 1. In the Unit 2 situation, the subsequent investigations tend to utilize long-reach arm as the main access device instead of self-propelled robot platform. Therefore, the base of the long-reach arm could be deployed outside, while using the long arm to enter the PCV. The changes of robot type also brought new challenges to the design of robot platform. Naturally, the access strategy of entering the PCV by passing through X-6 gave rise to the basic condition that the robot size should be compact enough to pass through the X-6 penetration, which is merely 0.5m.

When it comes to passing through a narrow guide pipe, it's natural to attach importance to precise operational controllability, thus to avoid the possible collision between the guide pipe and the robot. Valuing motion accuracy, IRID, Oxford Technologies Ltd, and Mitsubishi Heavy Industries have developed a prototype access device (Veolia-MHI Arm), as Figure 1.8 shows. In terms of arm structure, it includes a carriage part, several expandable links, and a telescopic front end. The specification of the prototype device is shown in Table 1.3.

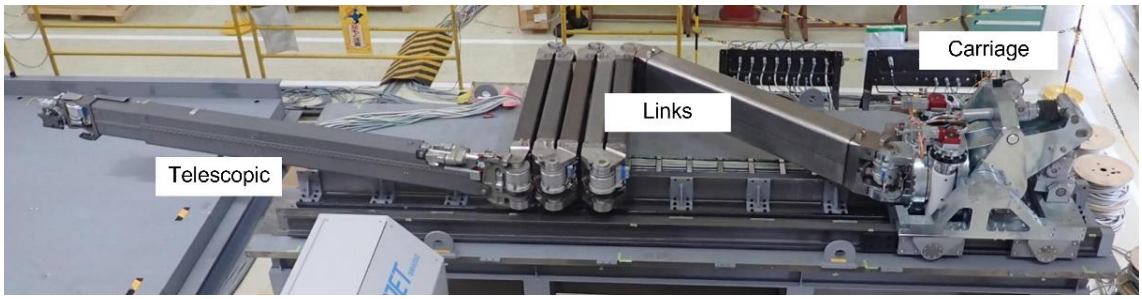


Figure 1.8 The prototype access device "Veolia-MHI Arm" developed by IRID. [12]

Table 1.3 Specification of the prototype Veolia-MHI Arm. [12]

Metrics	Value
Length	22m
Weight	4.6ton
Number of Axis	18

Based on the conventional robot arm design idea, most of the actuators were installed on the joints of the arm. In order to reduce the deflections due to weight, stainless steel and aluminum

with high strength was selected as the links' materials. Therefore, the arm part turned out to be quite heavy.

1.3.3 Existing Problems of Veolia-MHI Arm

Various experiments have been conducted to evaluate the performances of the access device, including deflection, repetitive positional accuracy, and adequacy. For the deflection measurement, a 3D laser scanner was utilized to measure the actual positions of links, and the preliminary experimental results on the deflection are shown in Figure 1.9.

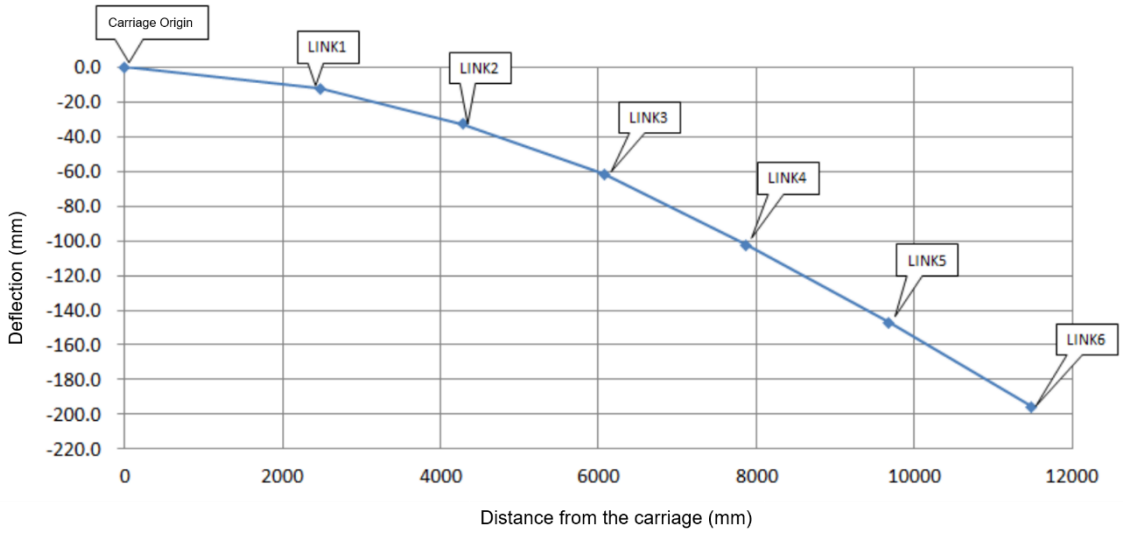


Figure 1.9 The preliminary experimental results on the deflection of the prototype access device by IRID. [13]

The maximum deflection of the front end reached about 200 mm when the arm was fully stretched. Furthermore, based on the deflection results, IRID adjusted the necessary angle tilt of the arm in the experiment of passing through a mock-up X-6 penetration. Therefore, the access device successfully passed through the mock-up X-6 penetration with about 1.12-degree tilt, but still the clearance got limited to about 15 mm. The experiment setup of passing through mock-up X-6 penetration is shown in Figure 1.10.

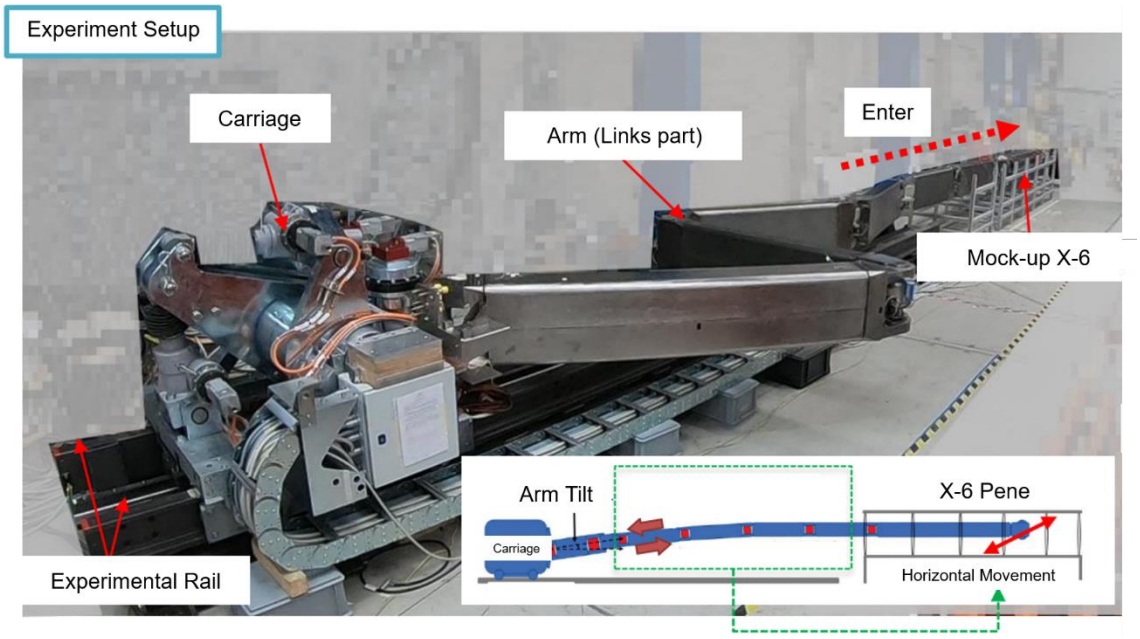


Figure 1.10 The experiment setup to pass through X-6 penetration. [14]

In this section, the development idea and the preliminary experiment results of the Veolia-MHI Arm has been presented. Here, it can be concluded that the feasibility of this arm is limited by these existing problems listed below:

- Heavy weight. Experiments have shown that the heavy weight may lead to deflection of the front end, component deformations, and positional misalignments. Although it's possible to reduce the influences of deflection with angle tilt, the available clearance will be limited.
- Installing actuators on joints. When passing through the narrow X-6 penetration, the size of the actuators can be limiting a factor, thus limiting the available torques of the actuators. As the result, the available speed of the arm would be influenced. Besides, entering the PCV with the arm indicates more harsh requirements on the radiation resistance of the actuators.
- Hard to retrieve from joint failure. When one of the joints fails, the complexity and weight of the arm will make it hard to recover, or retrieve the failed joint out of the inaccessible areas for repairs.
- Vulnerable against collisions. The high joint stiffness guaranteed precise position control while making the arm relatively vulnerable to collisions.

1.4 Problems of Conventional Environment Mapping Methods: Visual Odometry and Structure from Motion

Environment mapping methods have a vital position in the internal investigation of damaged reactor since they are utilized to increase the spatial awareness of operators. Especially, the vision-based environmental mapping methods refer to methods which use captured images or video as input, and provide detailed reconstruction of the surrounding environment for reference. Compared with the other methods, vision-based methods are able to provide intuitive and comprehensive models of the environment, like 2D grid map or 3D point cloud with textures.

In the approach groups of vision-based environment mapping methods, the group of Structure from Motion (SfM) methods and Visual Odometry (VO) has a common advantage that they can achieve environment mapping by only using image sequence as the input source. By extracting feature points from the image sequence, they could track these features and estimate the camera poses, realizing a simplified ego-localization. Such an advantage allows them to accept unordered images captured by different cameras, and even images captured under different illumination conditions. These approaches include the great work of Agarwal et al. [22] on reconstructing Rome by unordered photographs captured by tourists, and a general-purpose SfM framework COLMAP proposed by Schönberger et al [23].

However, the pose estimation processes are highly dependent on the available feature points extracted from the input sequence. Especially in the scene with texture-less component, like a white wall or empty desk, SfM/VO methods may not work due to feature points shortage. Besides, because the classic feature extraction processes could only provide the poorly-localized features, when there are repetitive features in the scene, the feature tracking processes may fail to track the correct features due to visual similarity. Recently, Lindenberger et al. [6] refined key steps of feature extraction in COLMAP framework with a convolutional neural network (CNNs), which improved the robustness of SfM processes in challenging conditions. Back to the situation of Fukushima decommissioning investigation, internal situations of these damaged reactors are harsh, complex, and lacking prior knowledge, forming significant challenges to conventional vision-based environment mapping methods that were designed for ideal environments. Therefore, to enhance its compatibility with the special environment, it's necessary for the conventional environment mapping methods to adapt by utilizing available prior knowledge from the environment and robot.

In terms of the available prior knowledge from the environment, it's feasible to conduct a preliminary scene classification to the environment inside damaged reactors as a scene containing both regular artificial objects and irregular damaged objects. Regular artificial objects, such as

buildings, infrastructures and equipment, typically have regular shape and known size. Previous investigations have revealed that the grating floor inside reactors is relatively intact and could serve as a supporting platform for robots (like the investigation B1 and B2). Therefore, the grating lattices with known size and regular shape can be considered as a unique characteristic from the specific environment as prior knowledge. However, the conventional environment method Visual Odometry could not fully utilize it due to the inclined view.

Visual Odometry is usually known as the front end of visual Simultaneous Localization and Mapping (SLAM), which deals with the problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within it. Currently, SLAM has been applied in various aspects like cleaning robot and automatic driving. Visual Odometry (VO) is the process of incrementally estimating the ego-motion of an agent (e.g., vehicle, human, and robot) using only the input of a single or multiple cameras attached to it. In terms of localization on the pedestal grating, developers for PMORPH2, Konishi T, et al proposed a localization workflow including a stereo method for position initialization and a repetitive construction tracking method for displacement estimation [8]. Through experiments conducted in a mock-up environment, the localization workflow for grating could provide error rate within the required 100mm/1.5m displacement, which was also the width of the standard grating lattice. Figure 1.11 shows the inclined view by the cameras on the pedestal grating areas.

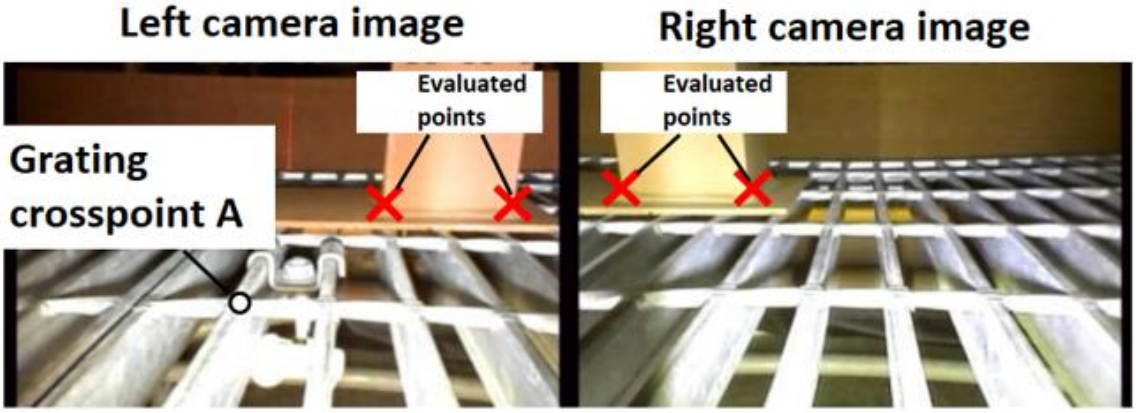


Figure 1.11 The inclined view by stereo cameras on crawlers. [8]

As Figure 1.11 shows, although the grating lattices on the floor have highly-repetitive structure with known size and regular shape, the inclined view was not intuitive enough to allow counting the number of past grating lattices. By counting the number of past grating lattices, it will be much easier and more robust to assist the localization on grating floor. Therefore, to fully utilize the characteristics of known grating size, it's necessary to provide more intuitive top view instead.

On the other hand, when considering the available prior knowledge from the robot, it's natural to start from the robot-mounted camera. Compared with the typical situation of mentioned

environment mapping methods, like building Rome with images from tourists, the image sequence from robot-mounted camera is highly ordered and sharing same intrinsic. Besides, the overall trajectory of robot-mounted camera should be spatial continuous and smooth. According to the specific motion patterns of the robot, the camera trajectory should also inherit the features of motion patterns. Especially, under a remote-operated situation, the camera trajectory can be pre-designed and adjusted, which form potential prior knowledge available. However, currently VO/SfM is not aware of the trajectory continuity and regularity from specific motion patterns, which makes it hard to fully utilize the characteristics to improve its robustness.

One of the typical problems is the feature tracking failure for such feature-based methods. Since most of the subsequent processes are based on extracted features and internal tracking, situations with few available features or low overlapping areas will suffer from feature tracking failure and lead to serial failure in the whole mapping. Besides, due to the high requirement on the image sequence quality on available features, these methods usually need stable and accurate camera poses, leading to low compatibility with flexible manipulators which may introduce positional deviation and motion blur. Therefore, to solve these problems, it's necessary to refine the conventional methods to allow them to utilize robot's motion pattern information to improve robustness and compatibility.

1.5 Research Objectives

In the previous sections, the existing problems have been briefly introduced from three aspects. Once again, this thesis aimed to help solve these three problems, listed below:

- The available awareness of fuel debris or water leakage provided by underwater camera of PMORPH2 was limited by low visibility and radiation level problem in Unit 1 situation.
- Feasibility of using tough manipulator "Veolia-MHI Arm" in Unit 2 situation was limited by heavy arm weight, weak collision resistance, and difficulty in joint failure recovery.
- Compatibility of Visual Odometry (VO) and Structure from Motion (SfM) with the specific environment and robot was limited, due to their inability to fully utilize characteristics as prior knowledge, such as grating lattices with known size, and specific robot motion patterns.

To solve these problems, the research objectives together with the contributions in this thesis can be divided into three aspects:

- A mobile robotic investigation system "RhinoUS-II" carrying a 2D phased array ultrasonic sensor is proposed, in order to provide more intuitive and robust images for underwater investigation. Especially, the newly-implemented ultrasonic sensor can conduct long-term shape & flow measurements in radioactive turbid water. Verification of proposed system's effectiveness through experiments with hardware prototype is one of the research objectives.

- A light-weighted, long-reach, flexible wire drive manipulator “BWD robot” is proposed, as an alternative option of tough manipulator “Veolia-MHI Arm” in investigation stage. Especially, the proposed manipulator can install actuators at base, retrieve from joint failure with ease, and have high tolerance to collisions thanks to the flexible feature. Verification of the proposed system’s effectiveness through experiments with hardware prototype is one of the research objectives.
- Three improvement methods for VO and SfM are proposed, to allow them to utilize environmental characteristics (grating lattice) and robots’ characteristics (specific motion pattern) as prior knowledge. Concretely speaking, proposed improvement methods include (1) a perspective transformation method to count the number of grating lattices under transformed top view maps; (2) a feature tracking failure compensation method to utilize RhinoUS-II’s motion patterns to conduct trajectory clustering, and then correct erroneous map areas by using average expectation in the clustered trajectory; (3) an image dataset creation method to utilize BWD robot’s motion patterns to avoid image quality degradations from multiple joints interference and vibration motion blur, by creating image sequences with overlapping rotation trajectories and filtering highly blurry images as pre-processing. Verification of proposed improvement methods’ effectiveness through experiments is one of the research objectives.

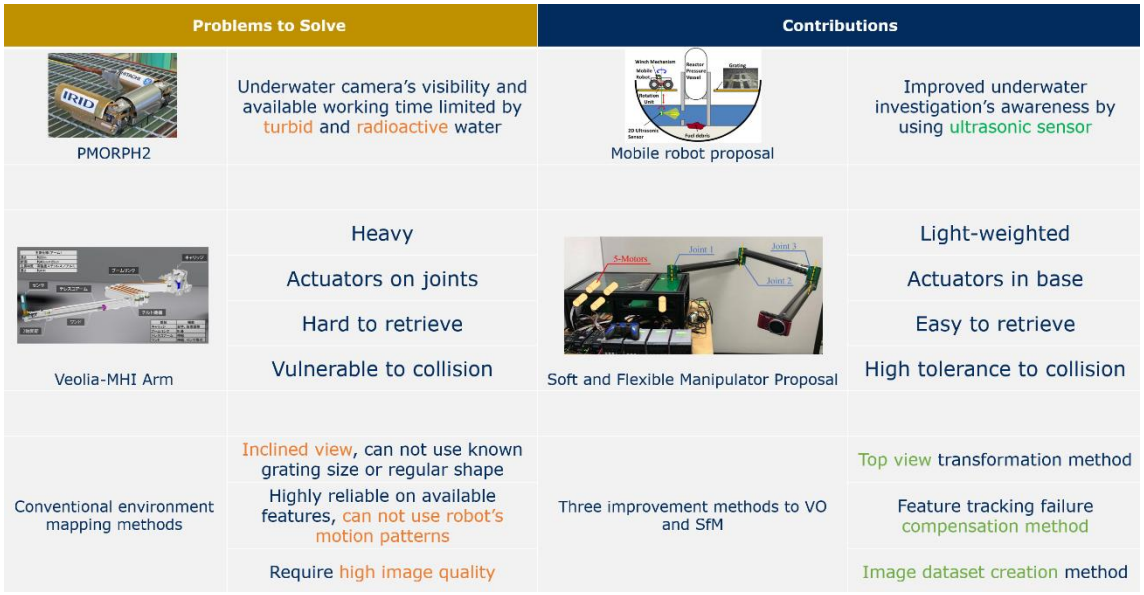


Figure 1.12 The structure of this thesis.

The corresponding relationships between the problems to solve and contributions can be checked with Figure 1.12. Especially, the concepts of the three improvement methods are presented with Figure 1.13.

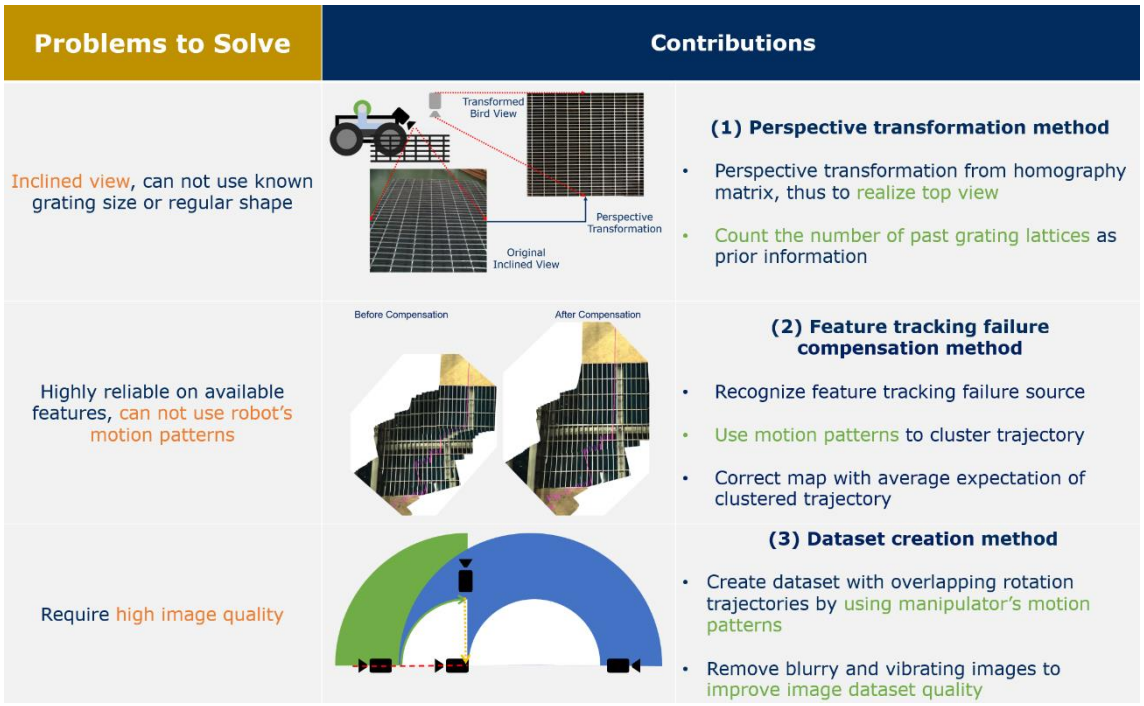


Figure 1.13 Details of the three improvement methods for VO and SfM.

1.6 Summary

In this chapter, the background of this research is introduced, as the decommissioning works of Fukushima nuclear power plants. Currently, the decommissioning works have come to the phase of preparing for fuel debris retrieval. The research scope of this thesis has been clarified as the preliminary investigation stage of the damaged reactors of Fukushima decommissioning. Within the concerned topics, this thesis aimed to solve three problems:

The first problem is that the available awareness of fuel debris or water leakage provided by underwater camera of PMORPH2 was limited by low visibility and radiation level problem in Unit 1 situation. To address this problem, a mobile robotic investigation system with 2D phased array ultrasonic sensor is proposed. The second problem is that feasibility of using tough manipulator "Veolia-MHI Arm" in Unit 2 situation was limited by heavy arm weight, weak collision resistance, and difficulty in joint failure recovery. As a light-weighted alternative option for the investigation stage of Unit 2, a long-reach wire-drive manipulator which can install actuators at base, retrieve from joint failure with ease, and have high tolerance to collisions is proposed. The third problem is that compatibility of Visual Odometry (VO) and Structure from Motion (SfM) with the specific environment and robot was limited, due to their inability to fully utilize characteristics as prior knowledge, such as grating lattices with known size, and specific robot motion patterns. Three improvement methods are proposed by allowing them to utilize characteristics as prior knowledge, including a top view transformation method to utilize grating

lattices information; a feature tracking failure compensation method to utilize RhinoUS-II's motion patterns; an image dataset creation method to utilize flexible manipulator's motion patterns. The rest of this thesis aimed to experimentally verify the effectiveness of the proposals.

The reminder of this thesis is as follows: Chapter.2 presents the proposal of a mobile robotic investigation platform "RhinoUS-II" as an ultrasonic sensor carrier. Chapter.3 presents two improvement methods for VO, including a top view transformation method, and a compensation method to improve robustness by utilizing RhinoUS-II's motion patterns. Chapter.4 presents the proposal of a light-weighted wire drive manipulator "BWD Robot". Similarly, Chapter.5 presents another improvement method for SfM, including a dataset creation method to improve compatibility between SfM and flexible manipulator by utilizing BWD robot's motion patterns. Chapter.6 concludes the concerned works in this thesis, and discusses their significances upon the Fukushima decommissioning works. The flow chart of this thesis is concluded with Figure 1.14.

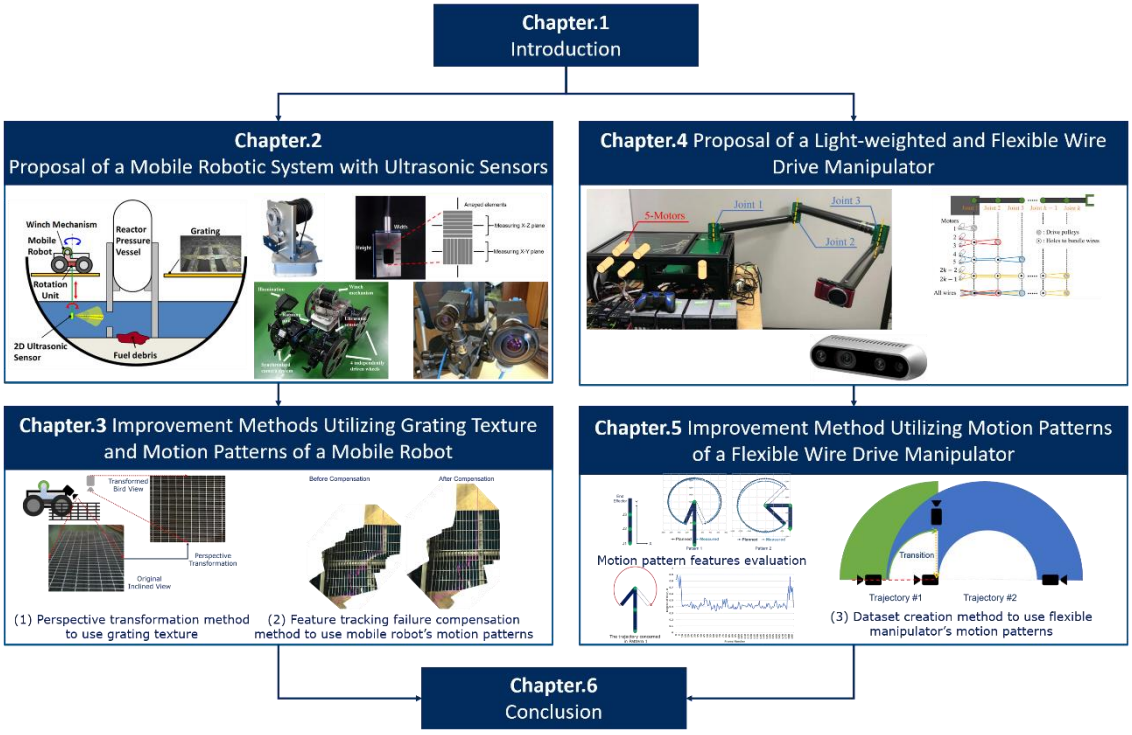


Figure 1.14 Flowchart of this thesis.

Chapter.2 Proposal of a Mobile Robotic System with Ultrasonic Sensors

In order to enhance the underwater investigation awareness, development of ultrasonic sensors has acquired great progress in terms of shape detection and flow mapping. On the basis of a 2D phased array ultrasonic sensor, a new mobile robotic system named “RhinoUS-II” has been proposed as the ultrasonic sensor carrier. The concept of the system is presented in Figure 2.1.

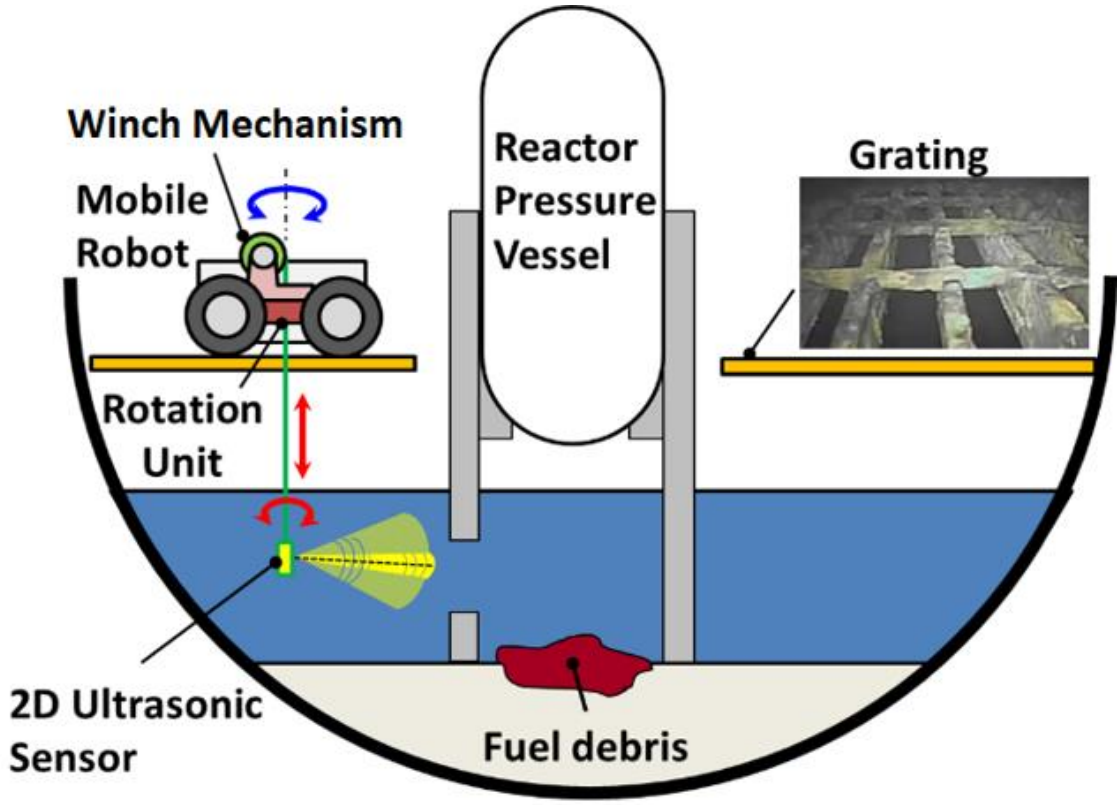


Figure 2.1 The concept of the proposed robotic system as the ultrasonic sensor carrier.

As Figure 2.1 shows, the mobile robotic system consists of a wheeled robot equipped with a winch mechanism, a 2D phased array ultrasonic sensor for underwater investigation, and a synchronized camera system to assist grating movement. The wheeled robot moves on the pedestal grating while using a monocular camera of the synchronized camera system to track the robot trajectory, and generating a map of the measurement locations on the grating floor. When the robot reaches the target point, it uses the equipped winch mechanism to adjust the orientation of the carried 2D phased array ultrasonic sensor into water under pedestal grating, then detecting the distribution of the fuel debris and possible water leakages.

In this chapter, the details of the proposed mobile robotic system will be introduced in the flow of hardware prototype overview (Section 2.1), 2DoF winch mechanism (Section 2.2), 2D phased array ultrasonic sensor (Section 2.3), synchronized camera system (Section 2.4), remote controlling (Section 2.5).

2.1 “RhinoUS-II” Hardware Prototype Overview

Figure 2.2 shows the hardware prototype robotic platform of the investigation system, RhinoUS-II. The robot is a wheeled mobile robot with 4 independently driven wheels, which serves as the carrier of the investigation sensor and other necessary devices like a winch mechanism at the chassis center.

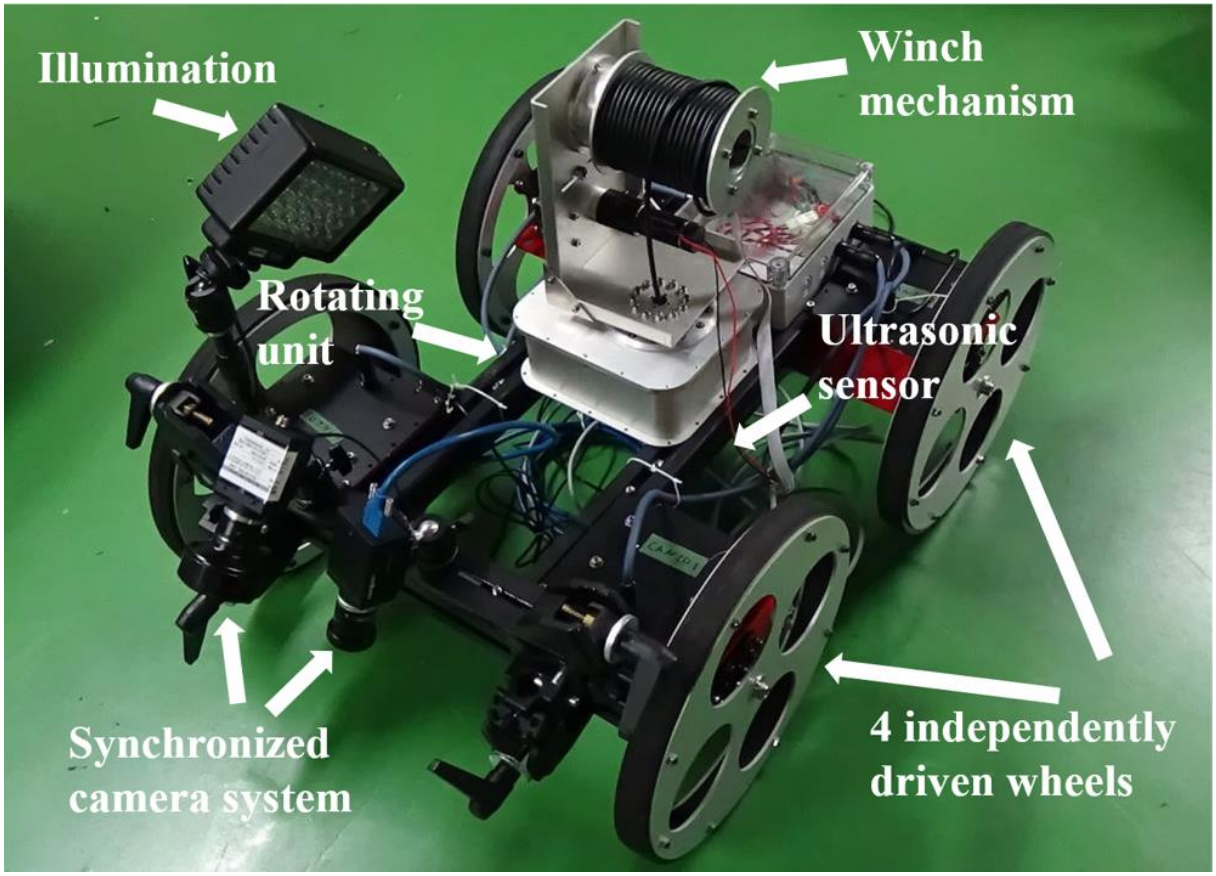


Figure 2.2 The hardware prototype of test platform RhinoUS-II.

Table 2.1 shows the specification [31]. A brushless motor EC60flat (Φ68mm, 24V, 100W,) made by Maxon serves as the actuator of each wheel. Correspondingly, the programmable controllers EPOS2 P24/5 made by Maxon are applied as motor controllers. The main body of the robot is a configurable chassis that composes of universal MISUMI frames, thus it’s easy to adjust the robot size. The overall design of RhinoUS-II is similar with that of PMORPH2 but without the configuration changing design, taking account of the winch mechanism for sinking the investigation sensor through the grating lattice. However, since RhinoUS-II was designed to be a test platform for the proposed investigation concept instead of the on-site investigation, the size requirements of passing through X-100B penetration and necessary radiation resistance for

operation were not included in the design. Generally, the test platform was supposed to move on the mock-up grating floor, and reach the target position, then sinking the investigation sensor with the winch mechanism.

Table 2.1 The parameters of RhinoUS-II. [31]

Parameter	Value
Length (mm)	500
Width (mm)	700
Height (mm)	300
The diameter of wheels (mm)	300
Weight (kg)	19
Degree of Freedom	4
Output (w)	400

Regarding the supposed application situation of RhinoUS-II, various functional components were installed on the main chassis, listed as below:

- A winch mechanism with a rotation unit for investigation sensor carrier, and sensor orientation adjustment.
- A phased array ultrasonic sensor carried by the winch mechanism, as the “water under pedestal grating” sensor.
- A synchronized camera system installed on the front end, as the “pedestal grating” sensor.
- A LED lightning device installed on the side for following illumination.

In the following sections, I will present the details of these functional components by following the listed flow.

2.2 Winch Mechanism with Rotation Unit

At the center of the robot, some space has been preserved for a rotation unit and a winch mechanism, which carries the ultrasonic below the chassis. Figure 2.3 shows the concept of the winch mechanism (left), and the developed prototype (right). The structure of the winch mechanism can be generalized as the cable winding spool, the DC motor to rotate the spool, a rotation unit to allow the carried sensor to rotate around the cable axis, a slip ring to allow infinite rotation.

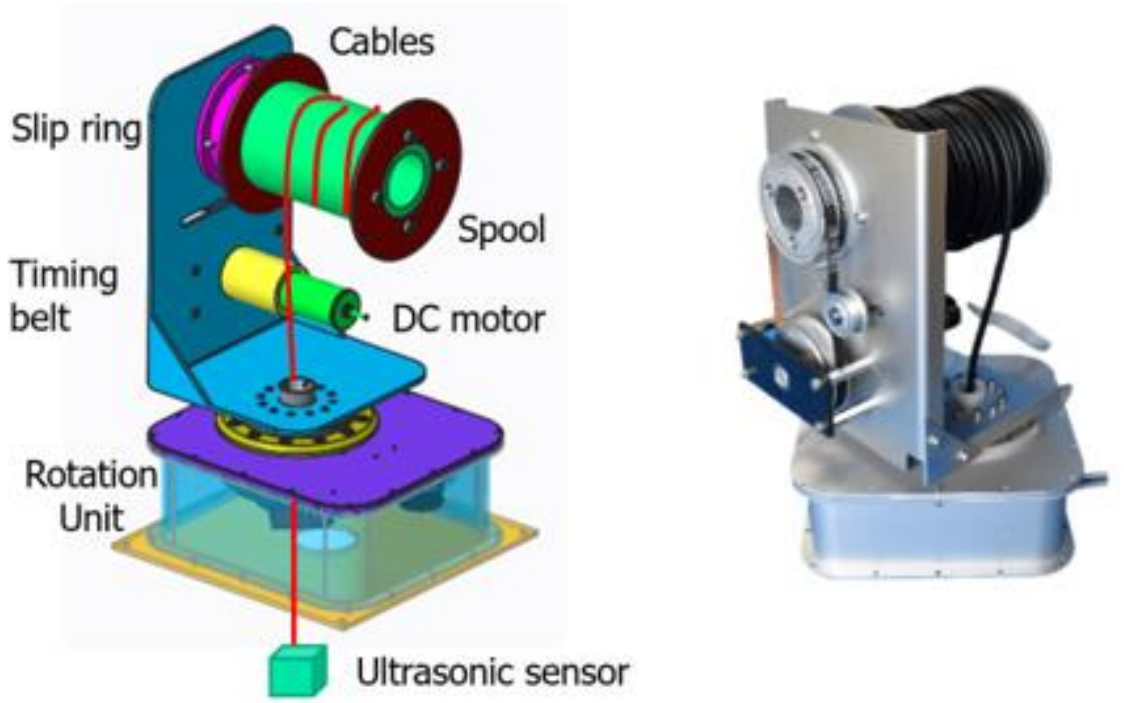


Figure 2.3 Left: The concept of the winch mechanism; Right: The developed winch mechanism prototype.

Compared with the winch mechanism utilized by PMORPH2, the developed winch mechanism added a swivel DoF around the sensor cable axis. When assuming the situation of sinking the investigation sensor through the narrow grating lattice (100mm x 30mm), I believe that the implementation of the extra DoF around the cable axis could allow the sensor to adjust its orientation, thus avoiding collision with the grating. Besides, since the height difference between the target pedestal bottom and the grating floor is 3.8m, I added the slip ring to allow infinite rotation of the reel while maintaining the electric connection. Furthermore, I have experimentally confirmed that insertion of slip ring between the ultrasonic sensor and the pulse receiver did not affect the measurement quality.

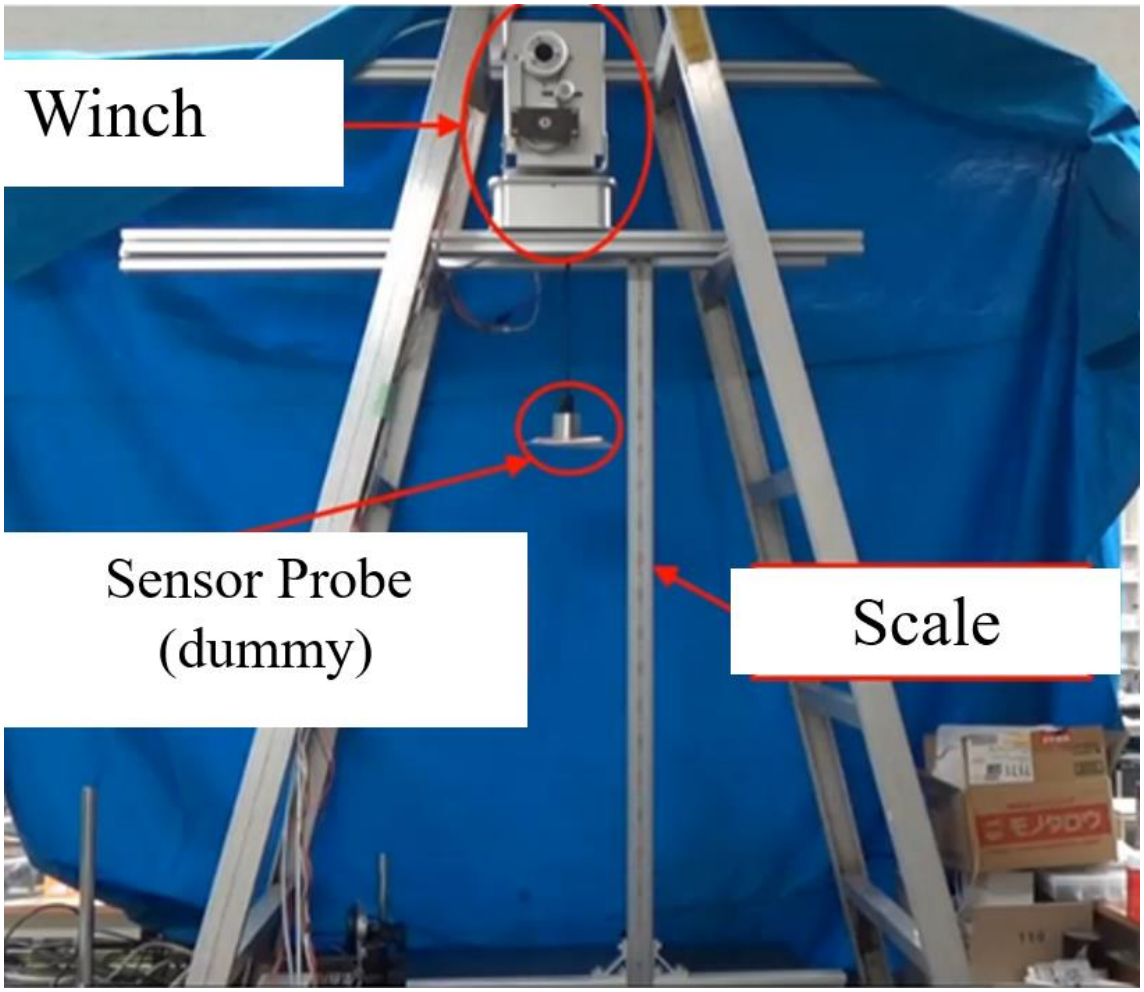


Figure 2.4 The experiment setup for winch mechanism positioning.

On the question of the basic performance of the winch mechanism, a preliminary positioning experiment of the winch mechanism was also conducted. Figure 2.4 showed the setup for the positioning experiment. The winch mechanism prototype was installed on a bracket, while the cable carries a sensor probe dummy. Parallel to the cable axis, a scale was set to measure the actual height of the sensor dummy. Therefore, I sent the command value of height to the DC motor, and measured the actual height with the scale. The comparison results of the commanded height and the measured height was shown in Figure 2.5.

The experiment results showed that the commanded height and the measured height maintained the basic linear relationship, but still about 25mm positioning error every 100mm. In order to solve the problem of changeable feed out length, a passive roller was installed on the winch mechanism. The cables from winch mechanism would pass through the installed passive roller and then hanging below. Therefore, instead of the rotation amount of the DC motor, the encoder readings of the passive roller would be utilized to keep the feed out length consistent. Figure 2.6 shows the passive roller image.

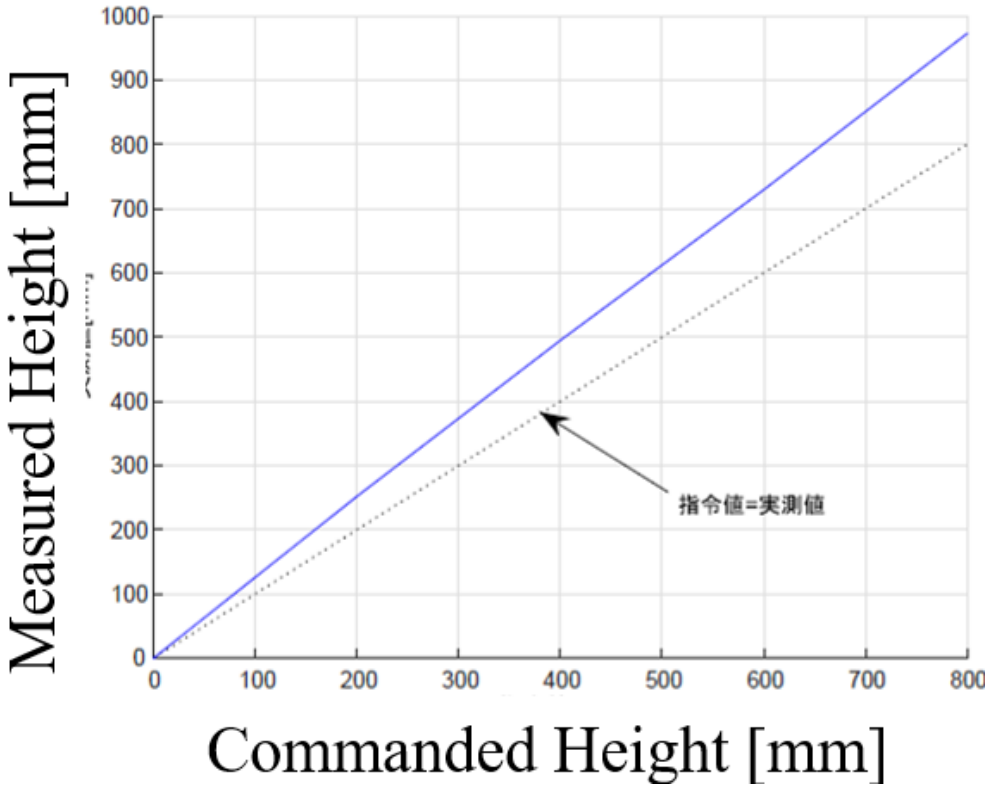


Figure 2.5 The comparison results of commanded height and measured height in winch positioning experiment (Blue line). Dotted line: The ideal line where the commanded value equals to measured value.

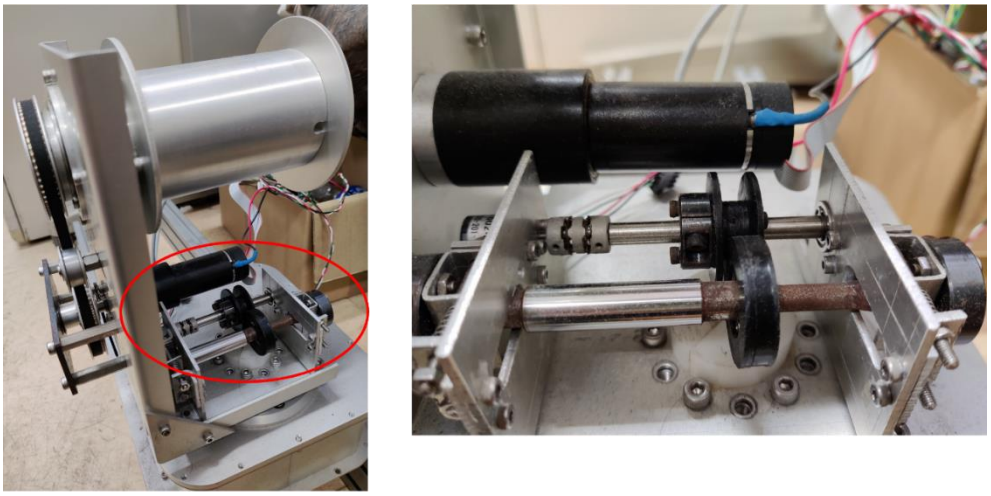


Figure 2.6 The passive roller to keep consistent feed out length.

Apart from the changeable feeding out length of cables, the axial winding of the cables may also influence the performance of the winch mechanism. To constrain the axial winding, a comparison study on the constrain effect between cable carrier, stand tube was conducted. The winch mechanism was commanded to feed out 1000mm length cables while equipping with these

winding constrain devices. Then, the winding angle of the cables would be measured and compared. As Figure 2.7 shows, the cable carrier showed decent effects of constraining the axial rotation of the cables.

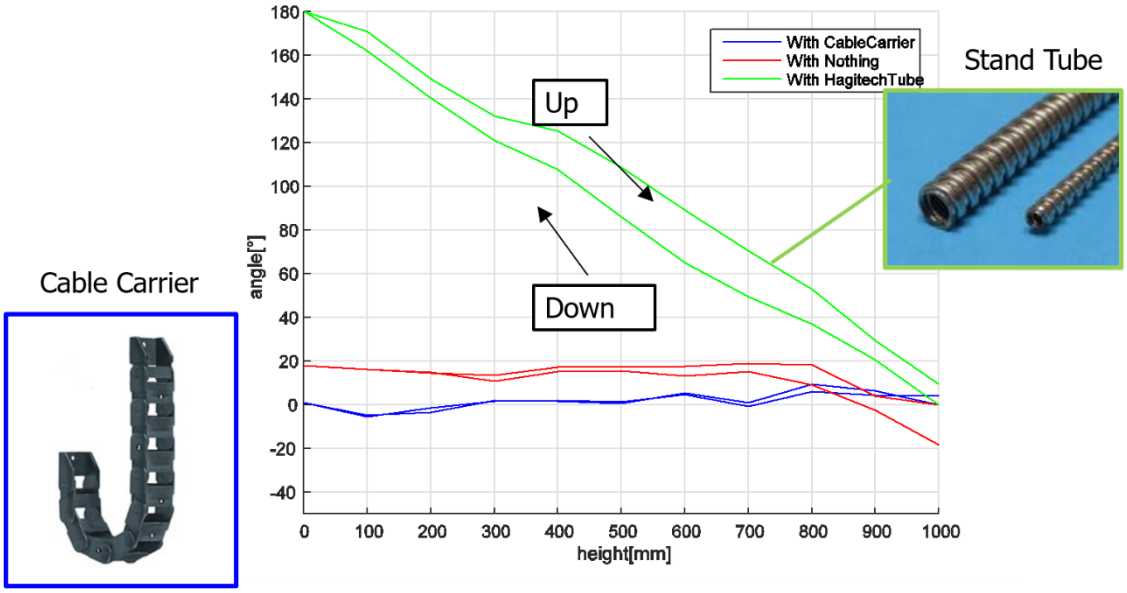


Figure 2.7 Comparison result between the axial rotation constrain effects.

2.3 2D Phased Array Ultrasonic Sensor

In this thesis, utilization of the newly-implemented 2D phased array ultrasonic sensor for underwater fuel debris/water leakage detection can be considered as the main difference between the proposed mobile robotic system “RhinoUS-II” and the conventional robotic investigation system “PMORPH2” which used underwater camera. The concept of the 2D phased array ultrasonic sensor is shown in Figure 2.8. Ultrasonic Velocity Profiler (UVP) method was utilized with a 2D phased array ultrasonic sensor to acquire a two-dimensional velocity profile of the turbid water, thus capable of localizing the water leakages. Furthermore, an aperture synthesis method was utilized to enhance the strength of reflected echo for detecting the shape of fuel debris.

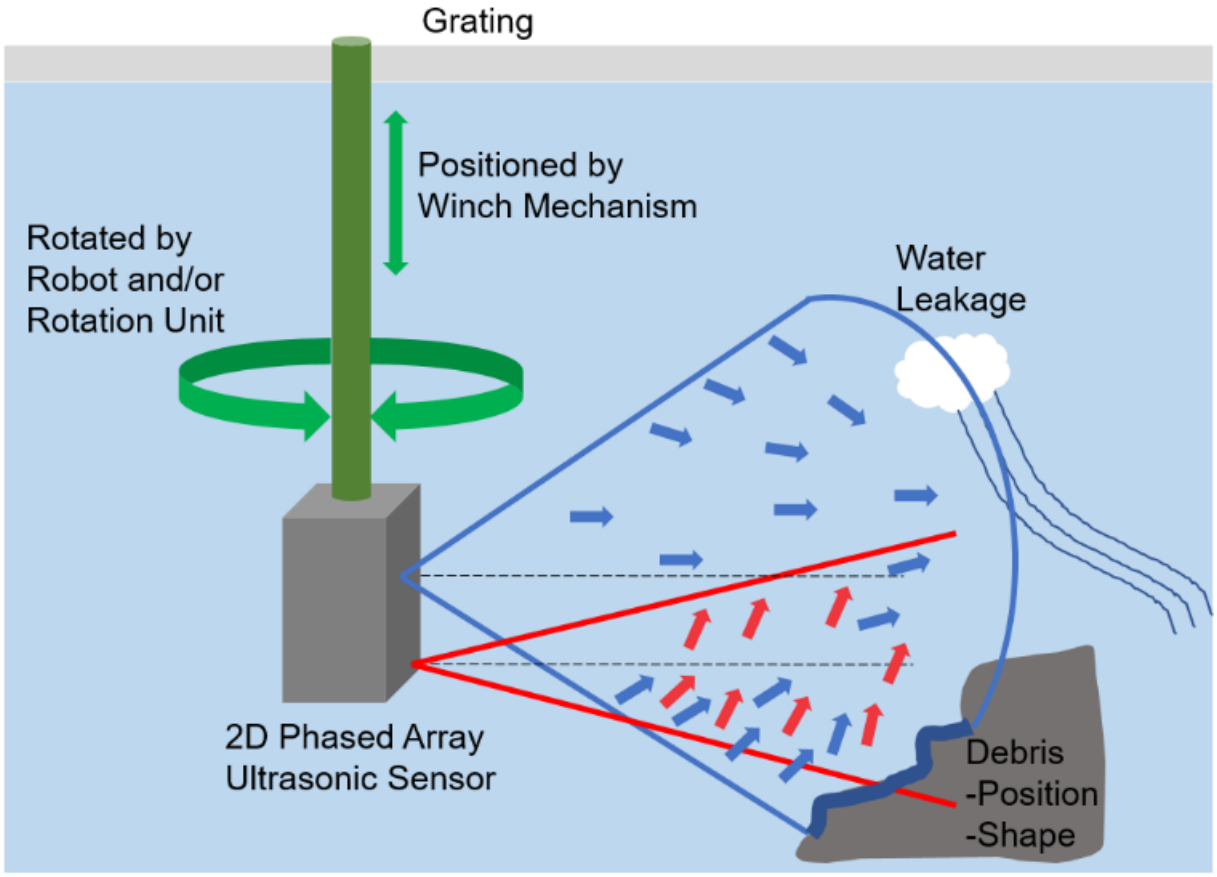


Figure 2.8 Conceptual image of the flow mapping-based measurement and debris localization with 2D phased array ultrasonic sensor.

This section is structured as follows: The reasons of selecting ultrasonic sensor (Section 2.3.1); Overview of the prototype phased array ultrasonic sensor (Section 2.3.2). The working principles and the verification experiments are attached in the Appendix AAppendix .

2.3.1 Selection of Ultrasonic Sensor

As has been stated in Section 1.2.2, the context of "water under pedestal grating" indicates the situation of passing through the narrow grating lattice (100mm x 30mm), and entering the turbid water to approach the pedestal bottom for fuel debris and water leakage localization. Besides, the previous investigations showed an increasing tendency of radiation level when approaching the pedestal bottom. Therefore, the underwater camera of PMORPH2 suffered from low visibility and radiation resistance problem due to the highly radioactive and turbid water, which is also one of the key problems that this thesis aimed to solve.

Taking account of specific conditions of this context, I would like to discuss the reasons of selecting ultrasonic sensor from three aspects: Size limitation to pass through grating lattice, measurement robustness against turbid influence, and resistance to radiation.

To reach the bottom pedestal, the current access strategy tended to reach the opening position to the basement, and pass through the grating lattice. While, the grating lattice is supposed to be in a regular rectangular shape with size of 100mm x 30mm, which limits the maximum sectional areas of the investigation sensor. In the situation of PMORPH2, the utilized investigation sensor consisted of a LED illumination, an underwater camera and a dosimeter, as Figure 2.9 shows. The investigation sensor was in cylindrical-like shape, with diameter of 20mm, length being 40mm. In order to limit the overall size of the sensor within 25mm, the underwater camera was deployed along the vertical axis of the investigation sensor, which limited the view range of the camera to the facing direction. While, ultrasonic sensor can be deployed along the side planes of the sensor, in other words, realizing measurements on more dimensions while keeping the compact size.

Compared with the ordinary environment where cameras are utilized, the underwater environment is much harsher in terms of vision occlusion and distortions. When checking the images captured during the previous investigations with underwater cameras, we could find that most images suffered from turbid water influences, including shadows when approaching to the bottom, and difficulty in recognizing the sighted objects [5]. Such influences in view may form obstruction in localizing the fuel debris and the water leakages. While, the measurement processes of ultrasonic sensor do not rely on visual information, which may ensure better robustness against turbid water.

Apart from the turbid water, the increasing radiation level underwater may influence not only the available working time, but also the visibility in image view in form of random white noise. In terms of the underwater camera of PMORPH2, the maximum allowable accumulated radiation dose of the camera was 1000 Gy, while within the activity areas of the underwater investigation, the maximum dose rate reached 10 Gy/h. In the case of decommissioning processes of Three-mile Island, the feasibility of utilizing ultrasonic sensor in high radiation environment had been confirmed [32]. The probe for the in-vessel investigations was required to tolerate radiation fields up to 10^4 Gy/h, and provide 40-mm lateral and 12-mm range resolution. Therefore, the designed probe consisted of arrayed ultrasonic transducers considering both adequate sonic coverage and mechanical reliability. On the basis of the ultrasonics-based underwater investigations, an intuitive image of the core cavity inside the reactor was successfully acquired, as Figure 2.10 shows. This image is a 360-degree panorama with a viewpoint perpendicular to the centerline of the core cavity. Therefore, the existence of piled debris near the core-former walls in places was confirmed. Furthermore, six objects hanging from the roof in the middle distance were confirmed to be axial power shaping rods and a series of broken fuel rods [32]. The successful experiences of utilizing arrayed ultrasonic sensors in Three-mile Island also contributed to the selection of ultrasonic sensor in this proposal.

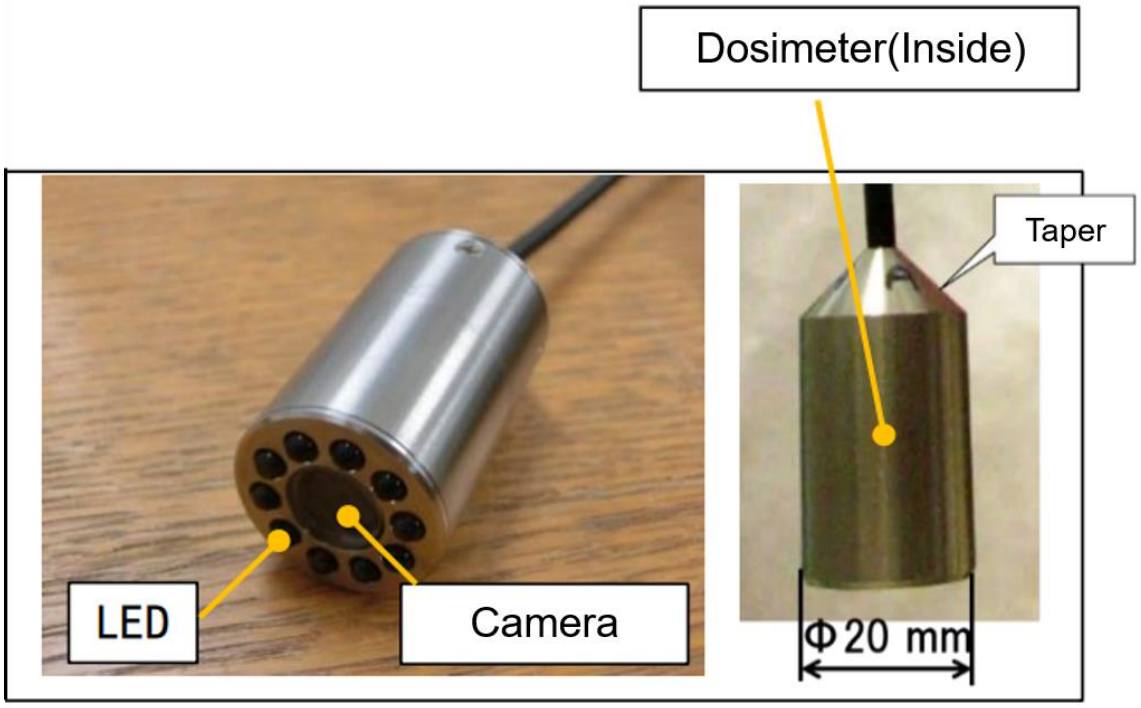


Figure 2.9 The investigation sensor of PMORPH2. [7]

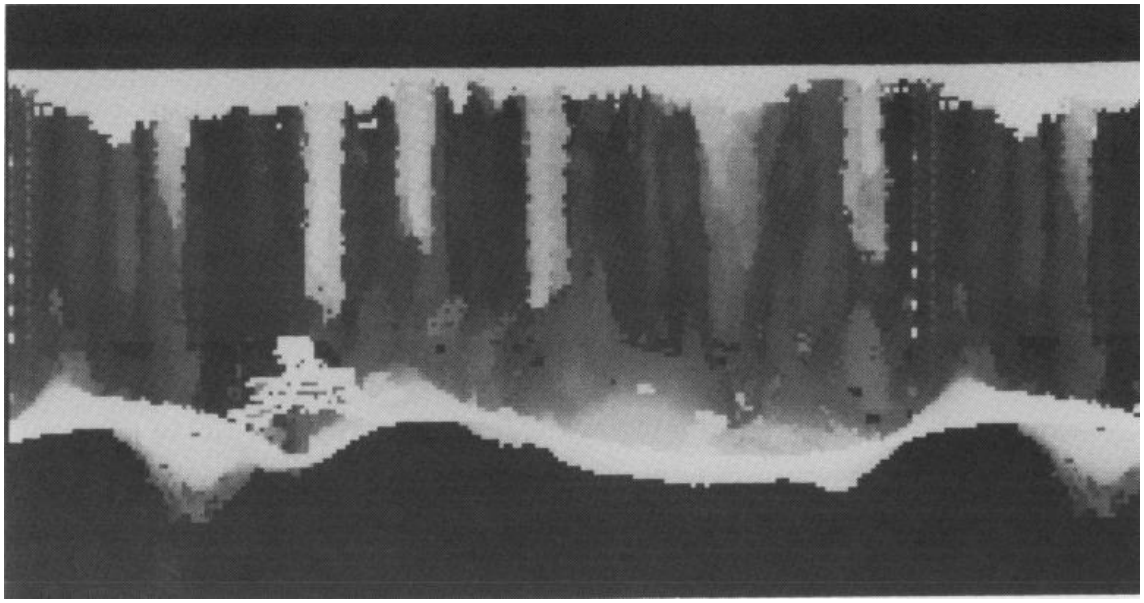
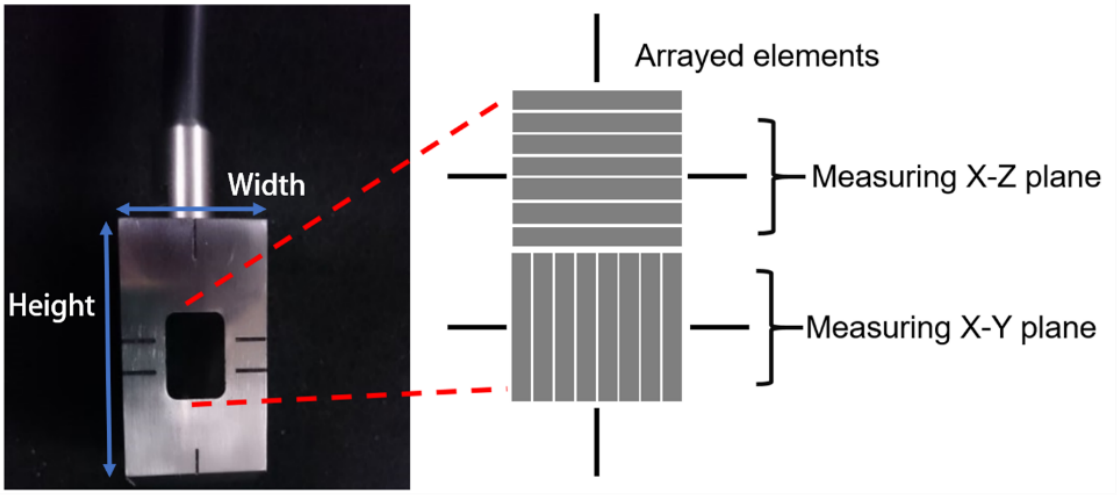


Figure 2.10 An ultrasonic computer image of the core cavity in Three-mile Island. [32]

2.3.2 Ultrasonic Sensor Overview



Parameters	Values
Sensor Size (height, width, length) (mm)	34 x 22 x 30
Sensor Weight (kg)	0.135

Specification of the arrayed ultrasonic sensor

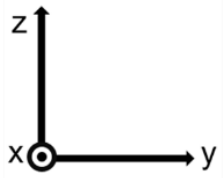


Figure 2.11 The prototype 2D phased array ultrasonic sensor.

Figure 2.11 shows the geometric appearance of the 2-dimensional phased array ultrasonic sensor, developed by our research cooperator laboratory of Kikura. The longer side of the sensor is defined as the Z direction and the shorter side as the Y direction. The array sensor can be divided into two parts by the central horizontal plane, so the upper part is utilized to measure the X-Y plane, while the lower part is for X-Z plane. Each array includes 7 elements. The size of the array sensor is 34mm x 22mm x 30mm, which meets the size requirements of passing through the grating lattice with size of 100mm x 30mm. Still, the orientation of the sensor should be adjusted to avoid collision to the lattice, which was realized by the implemented winch mechanism with rotation unit. A preliminary experiment on the radiation resistance ability had been conducted by exposing the sensor to gamma rays at a rate of 650 Gy/h by 60Co, and confirmed that the degradation of the sensor signal was less than 3% where the accumulated radiation dose had reached approximately 10000 Gy [33]. As for the working principles of the phased array technique, the flow mapping method, and the shape detection method are concluded in Appendix A.

2.4 Synchronized Camera System

Apart from the underwater localization with phased array ultrasonic sensor, the grating localization process is also of great importance. To provide spatial awareness of the grating

environment, the synchronized camera system was installed on the robotic platform to provide necessary images.

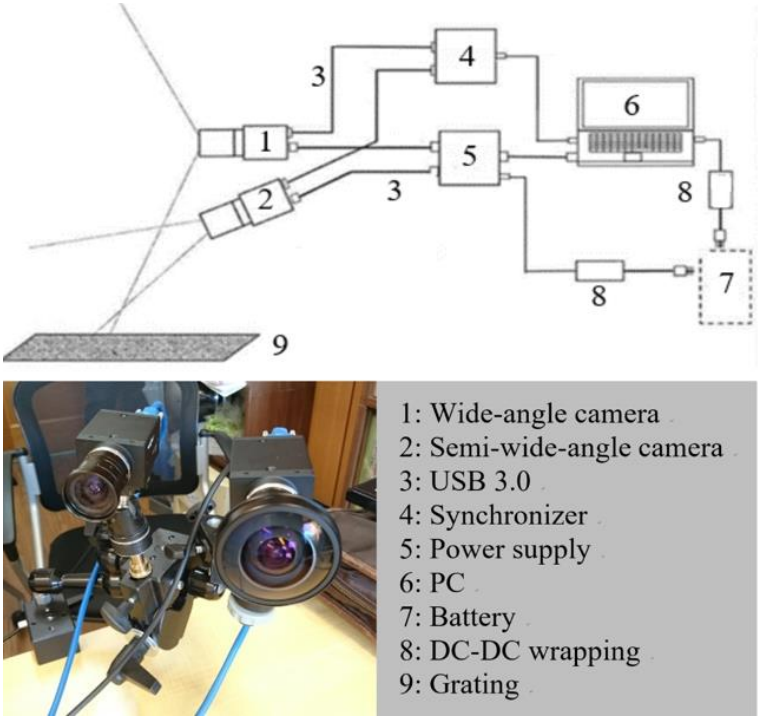


Figure 2.12 The synchronized camera systems.

To acquire the necessary information of the current situation of the robot when moving on the grating floor, we installed a synchronized camera system in the front of the robot, with a supporting frame, as Figure 2.12 shows. The supporting frame was installed at the front end of the robot, consisting of several links fixed with screws, while the height of the central link could change by moving the position of the fixing screw. A LED light is installed at the side of the supporting frame for illumination. The cameras were mounted on a camera holder, which allowed the camera to rotate. The camera system consisted of one wide-angle fish-eye camera, and one semi-wide-angle RGB camera.

The wide-angle fish-eye camera was utilized to serve as the main monitor, providing a necessary view of the front environment. Thus, the camera was set horizontal, facing the front of the robot. When performing autonomous localization, I mainly used the semi-wide-angle RGB camera, which was set inclined with the horizontal plane. In the front of the robot, a laser range finder (LRF) was installed as an auxiliary localization component. I used the conventional model Hokuyo UTM-30LX, also known as the Hokuyo Top-URG, which is a LiDAR that outputs data within 30 meters and 270 degrees range. The LRF was set horizontal, facing the front direction of the robot. The LRF could provide the distance information between the robot and the possible objects, and reconstruct a topological map including the planar features in vertical direction. While the view of camera was inclined, facing the grating floor. At the early stage of the

developing process, I considered to use both the camera and the laser range finder to localize the robot as mutual compensation. In the following experiments, I would like to compare the localization performance of LRF with the refined visual odometry method.

2.5 Remote Controlling

The system configuration of the test platform mobile robot system is shown in Figure 2.13. Each actuator was set with the motor and the motor drive, while the motor drives are connected with Controller Area Network (CAN). The power supply required 4 cables for the wheel drive, and 2 cables for the winch and base drive. 4 USB cables were connected with the PC, as the wheels controlling, the joystick, the base and winch controlling, and a laser range finder.

The controlling processes of the test platform was realized with Robot Operating System (ROS)[34], taking account of its adaptability to fast development and open-sourced community. ROS is a widely applied robotics middleware, providing services designed for a heterogeneous computer cluster such as hardware abstraction, low-level device control, implementation of commonly-used functionality, message-passing between processes, and package management. The main functions in the controlling system are listed below:

- Wheels driving and remote operation with a joystick.
- Winch controlling.
- Laser range finder usage.

For each motor driving the wheels, I created the corresponding joint controller based on Moveit! library in ROS, and actuated through EPOS support library. Through conversions, the position and the velocity of the motors including the wheels and the winch would be monitored in the GUI. The commanding data would be transformed into the target value of the wheels' velocity or winch's position, sending to the motor drivers as the outputted controlling signal. A node named /Joy was created to receive the signal of joystick and published the topic /joy to the /wheel_operate node, therefore controlling its movement, as the upper flow in Figure 2.14 shows. The Graphic User Interface (GUI) for remote controlling is shown in Figure 2.15. Then the velocity of the wheels would be transformed to robot odometry information and published to /tf for coordinate transformation.

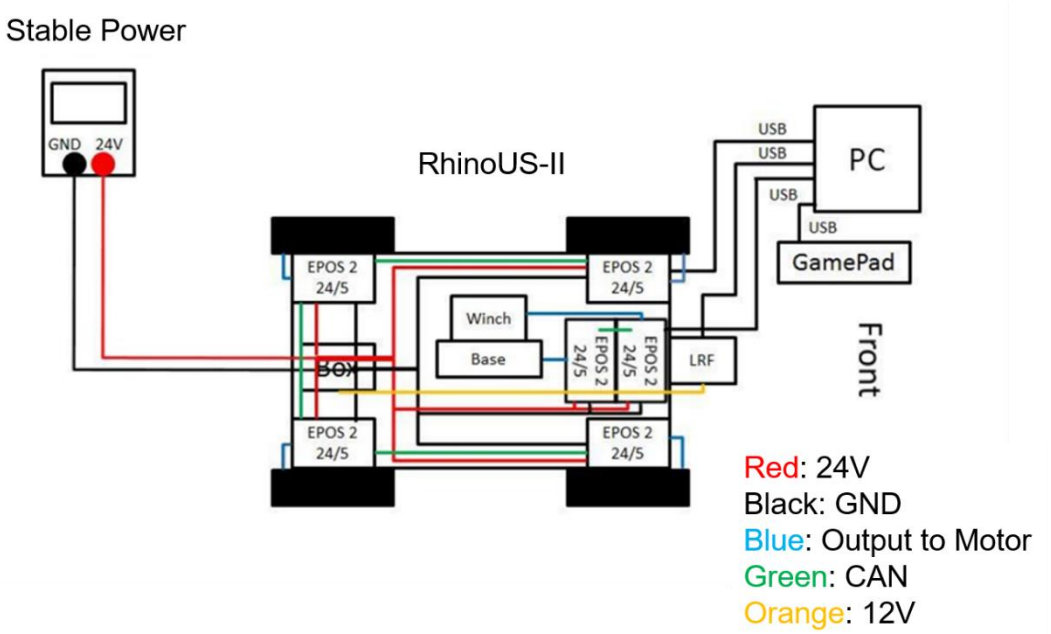


Figure 2.13 The configuration of RhinoUS-II.

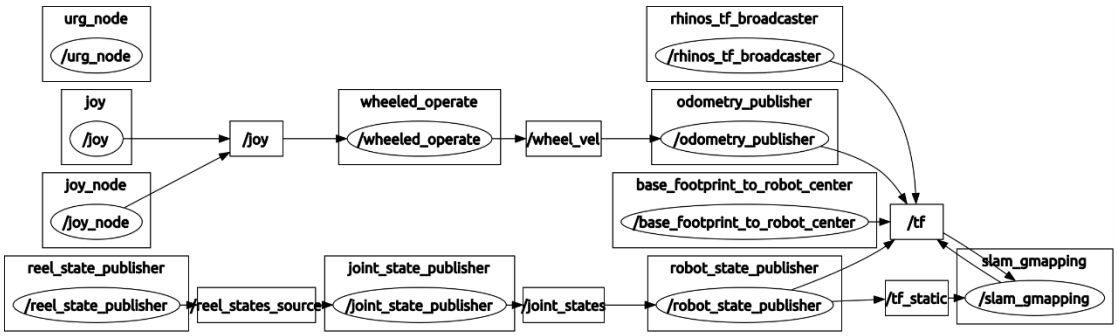


Figure 2.14 The node communication in ROS.

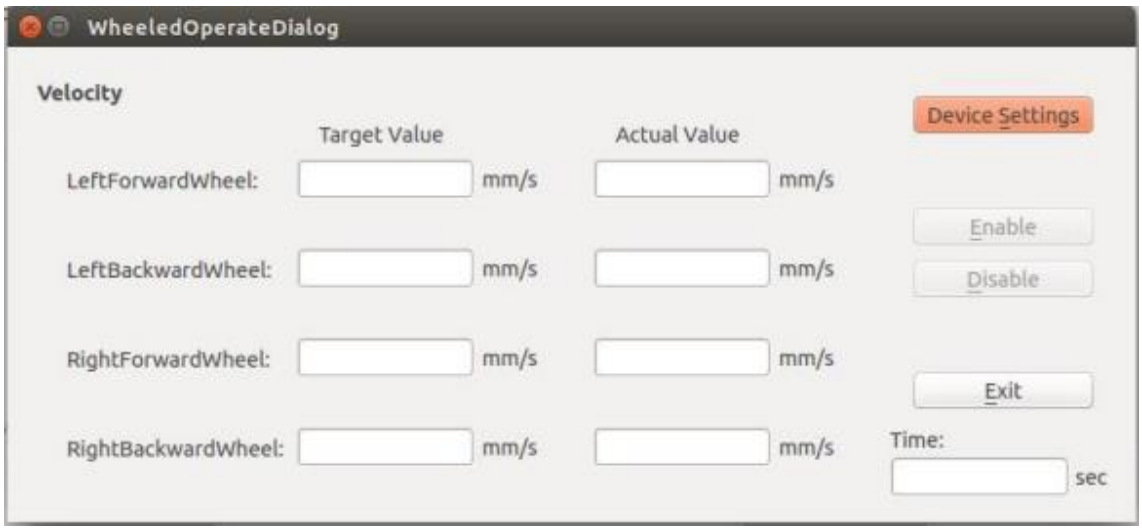


Figure 2.15 ROS node /wheel_operate GUI.

For the winch controlling processes, Figure 2.16 showed the GUI for ROS node /reel_operate. Here I could adjust the active operation mode for EPOS motor, and send the target position command to motor, while monitoring the actual positions by reading the encoder information.

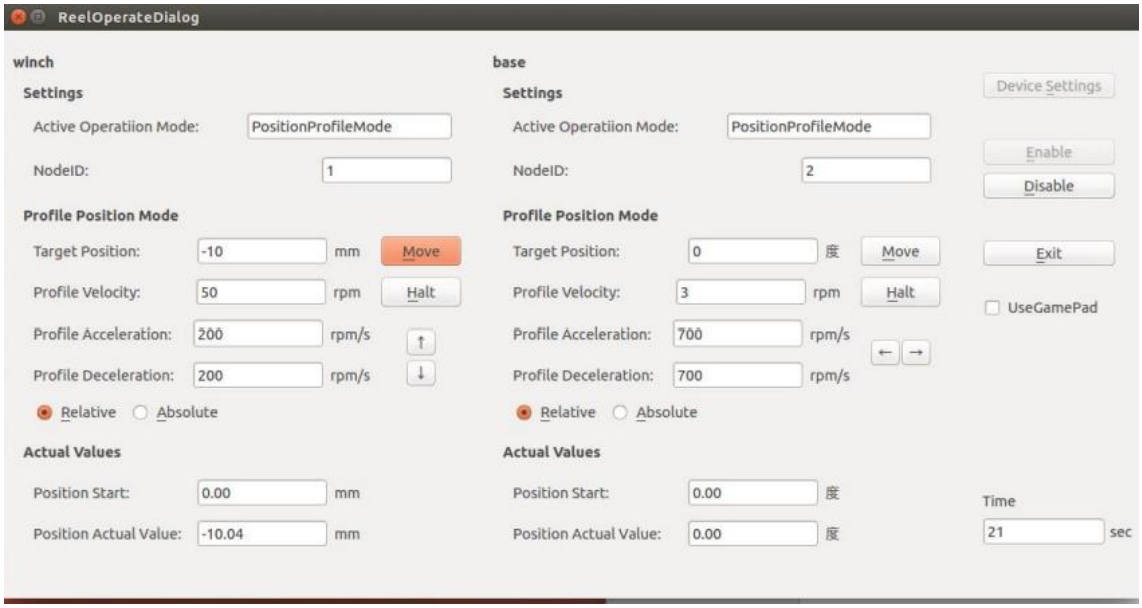


Figure 2.16 ROS node /reel_operate GUI.

At last, along with the refined visual odometry method, I also used the LRF sensor to perform SLAM, as a side comparison method. I utilized the hokuyo_node driver provided in ROS to acquire laser data, also known as /urg_node in Figure 2.14. Then, gmapping algorism was selected as the SLAM method, which could generate a topological map with the trajectory of robot attached to it, based on the data gathered by laser range finder and the wheels' motor encoder. As the /slam_gmapping node shows, it gathered the information from /tf as robot odometry information, and also the basic kinematic chain information from /tf_static. As the result, the generated topological map and the relative robot position could be observed with the visualization tool Rviz. Figure 2.17 shows one of the reconstructed maps by the gmapping algorism viewed from Rviz. It's a map in top view which could be used to compare with the refined visual odometry method in terms of intuitive view.

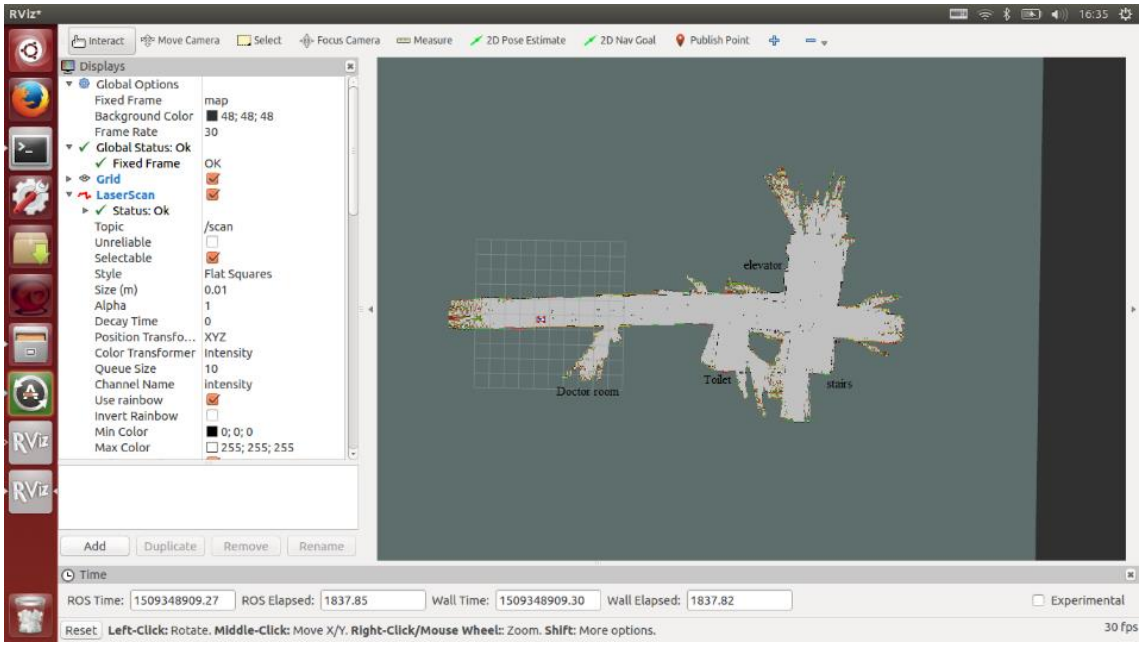


Figure 2.17 One of the reconstructed topological maps by gmapping.

2.6 Summary

This chapter presented the proposal of the robotic investigation system RhinoUS-II, as an ultrasonic sensor carrier. Compared with the conventional robotic investigation system PMORPH2 which used underwater camera to gather images of PCV's pedestal, the proposed robot selected 2D phased ultrasonic sensor as new underwater investigation sensor. Besides, an extra rotating DoF along the center axis was added to adjust sensor orientation. The difference between the proposed robotic system and the conventional one PMORPH2 is shown in Table 2.2 and Table 2.3.

Table 2.2 Comparison of the robot structure between PMORPH2 and RhinoUS-II

Systems	Robot Type	Winch Mechanism
PMORPH2	Self-propelled robot	1 Degree of Freedom
RhinoUS-II	Wheeled mobile robot	2 Degree of Freedom; Slip ring

Table 2.3 Comparison of the underwater sensing devices and methods between PMORPH2 and RhinoUS-II.

Systems	Sensing Devices	Sensing Methods	Feature
PMORPH2	Tip camera & dosimeter	Image synthesis	Low visibility and radiation resistance problem
RhinoUS-II	Ultrasonic sensor	2D flow mapping & aperture synthesis	More robust to radiation and turbidity

Through preliminary experiments, the proposed robotic investigation system has been proved feasible to adjust the ultrasonic sensor's orientation and put it into water by passing through the narrow grating lattice. Through experiments conducted by Kikura lab [33][35][36], the feasibility of using this 2D phased array ultrasonic sensor to provide better awareness of fuel debris and water leakages has been confirmed.

Chapter.3 Utilization of Grating Texture and Motion Patterns of a Mobile Robot for 2D Map Building

In this chapter, two robustness improvement methods for VO will be proposed, in order to enable it to utilize characteristics from the specific environment (grating texture) and the robot (proposed wheeled robot "RhinoUS-II"). This chapter will introduce the basic working principles of the conventional 2D map building method VO (Section 3.1), and then propose the two improvement methods. Within these two methods, the first one is a perspective transformation method to adjust the original inclined view to more intuitive top view (Section 3.2). The second method is a feature tracking failure compensation method to utilize the wheeled robot's motion patterns to improve the mapping correctness (Section 3.3). Then, the experimental verification of these two methods will be presented with Section 3.4.

3.1 Basic Working Principles of VO

Visual Odometry (VO) is the process of incrementally estimating the ego-motion of an agent (e.g., vehicle, human, and robot) using only the input of a single or multiple cameras attached to it. In the situation of the wheeled robot RhinoUS-II, the robot itself will be the agent, while the grating floor will be the environment. Classical VO can be divided into two types, direct VO and feature-based methods. The main difference lies on the processing method to estimate the relative motion from the input images. Direct VO can also be categorized into three directions, including dense VO, semi-dense VO, and sparse VO. Concretely speaking, dense VO uses all the pixels from the input images, while semi-dense VO selects pixels with specific strategy like intensity gradient, and sparse VO uses sparsely selected pixels [37]. Therefore, based on the concerned pixels, direct VO method performs an estimation by minimizing the photometric error with non-linear optimization algorithms[38]. Therefore, direct VO usually has more computational burden compared with feature-based VO. On the other hand, feature-based VO matches the extracted features from image sequence into pairs, and estimates the relative motion by minimizing reprojection errors. Thus, feature-based VO usually has better speed but relies on the available features in the scene. In this research, taking account of the calculation burden, feature-based VO was selected as the basis.

The general working procedure of a conventional feature-based VO is as follows:

- Conduct the initial camera calibration to acquire the necessary camera parameters like intrinsic matrix. Zhang's checkerboard calibration method [39] is a typical method by using the known size and structure of checkerboard. Besides, it's possible to simplify the calibration process by setting the world coordinate system to the checkerboard's corner.
- Feature extraction and matching. Usually in the case of feature-based VO, the selection of suitable correspondence features and feature detectors is of great importance. Mainstream pipelines consist of Scale Invariant Feature Transform (SIFT), Speeded-up Robust Features (SURF), and Oriented FAST and Rotated BRIEF (ORB), etc. Here, I used ORB [40] feature

considering its good balance of processing speed and accuracy. When matching the extracted feature pairs, I used the conventional sequential order by matching the adjacent image pairs following the image sequence order.

- Estimate the relative camera motion based on the matched feature pairs. Here, I used the pipeline named EstimateRigidTransformation provided in OpenCV library [41]. The input of this function is the matched points pair in two subsequent frames, and I denoted the matched point group in the n-th frame as P_n. $P_n[i]$ indicates the i-th point in the matched points pair. Such point pairs contain the coordinates of the matched points in the camera coordinate system. The output of this function is a 3x3 matrix denoted as $\boldsymbol{C_n}$, which also can be divided into a rotation matrix $\boldsymbol{A}$ and a translation vector $\boldsymbol{b}$, shown as below. Here, I denoted the element $\theta_{c,n}$ indicates the instant rotation angle of the camera in the n-th frame, $x_{c,n}$ and $y_{c,n}$ indicate the translation in x and y-direction. It's a typical planar motion matrix without scale factor. The definition can also be written in form of Equation 3.1 and 3.2.

$$\boldsymbol{C_n} = \begin{bmatrix} \cos(\theta_{c,n}) & -\sin(\theta_{c,n}) & x_{c,n} \\ \sin(\theta_{c,n}) & \cos(\theta_{c,n}) & y_{c,n} \\ 0 & 0 & 1 \end{bmatrix} \tag{3.1}$$

$$\boldsymbol{C_n} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{0} & \boldsymbol{1} \end{bmatrix} \tag{3.2}$$

Therefore, the estimation of the relative camera motion (the transformation matrix) will be conducted by minimizing the reprojection errors between the original frame and the reprojected frame. The reprojection errors can be written in form of Equation 3.3:

$$[\boldsymbol{A}^*|\boldsymbol{b}^*] = arg \min_{[\boldsymbol{A}|\boldsymbol{b}]} \sum_i |\boldsymbol{P_{n+1}}[i] - (\boldsymbol{A}\boldsymbol{P_n}[i] + \boldsymbol{b})|^2 \tag{3.3}$$

- Integrate the estimated instant camera motion to localize the camera under the global coordinate system. Usually, this process will give rise to accumulation of pose error in each estimation, therefore how to deal with the accumulated error is the optimization direction of this process.

On the other hand, the previous on-site investigation of PMORPH2 also utilized vision-based localization method in "pedestal grating". As has been mentioned, PMORPH2 utilized two stereo cameras installed on the crawlers of the robotic body to capture visual images to localize the position of the robot. The localization method can be regarded as a combination of two localization methods as:

- The stereo localization method for the positional relationship between the camera and the target object to calculate the initial position of the robot based on the known structure. This process is similar with the initial calibration process in VO, but the object with known structure changed from the regular checkerboard to a prism.

- The repetitive tracking method using repetitive structures to remove accumulating error in displacement estimation from the initial position. This process aims to remove the accumulating error by using repetitive structures like grating, which is similar to the bundled adjustment idea of recognizing the locations that robot has visited.

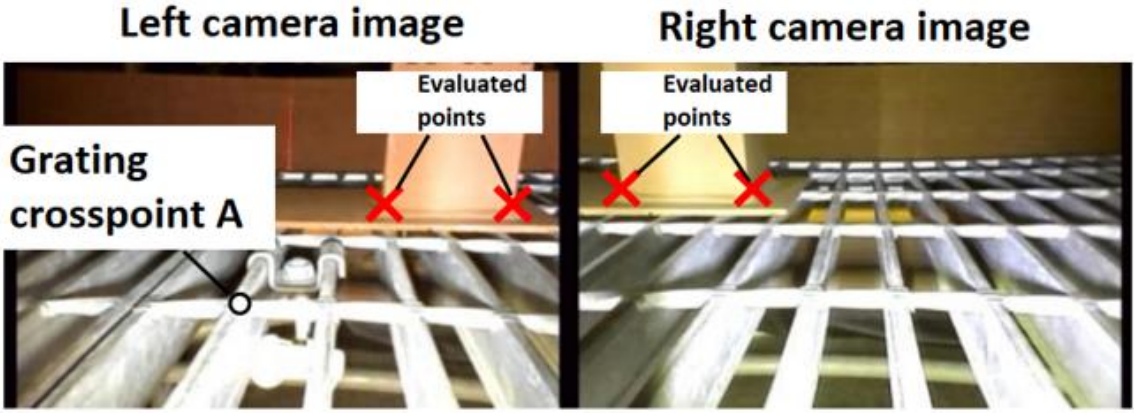


Figure 3.1 The inclined view by the stereo cameras on crawlers. [8]

Figure 3.1 shows the visual images taken by the stereo cameras by PMORPH2 when performing the initialization process in the preliminary experiment. I can notice that there are two kinds of structures applied in the localization method. The “Evaluated points” from the prism served as the known structure to initialize the position of the robot. And the grating cross point served as the repetitive structure to remove the accumulated error. Compared with the methods utilized by PMORPH2, feature-based VO with a monocular camera does not rely on known structure like prism to initialize.

While, one of the key problems in the conventional feature-based VO method and the repetitive tracking method of PMORPH2 are that both of them could not fully utilize the available characteristics in the environment, as the inclined view made it hard to take advantage of the grating lattices with known size and regular shape. Besides, both of them were highly reliable on the available features in images, thus might suffer from robustness degradation when the process of feature tracking failed without necessary countermeasures. While, the regular motion patterns of the specific robot could have been a valuable reference data, but these methods could not fully utilize the motion patterns.

Therefore, in order to improve the performance and robustness of VO by fully utilizing the grating lattices and robot’s motion pattern, I proposed a perspective transformation method to generate a top view map (Section 3.2), and proposed a feature tracking failure compensation method on the basis of RhinoUS-II’s motion patterns (Section 3.3).

3.2 Perspective Transformation Method to Utilize Grating Lattices' Size and Shape as Prior Knowledge

The core problem, as has been repeated multiple times previously, is that the original inclined view made it hard to fully utilize the grating lattices' information. Therefore, the perspective transformation method is proposed to generate a top view.

The concept of perspective transformation is shown in Figure 3.2. The upper left part of the figure is the robot test platform with a monocular camera inclined with the grating floor surface. Therefore, the original view from the camera will also be inclined, as the lower-left part shows. After the perspective transformation, the grey camera indicates the virtual transformed camera, and its view has been transformed to an intuitive top view shown in the upper right part.

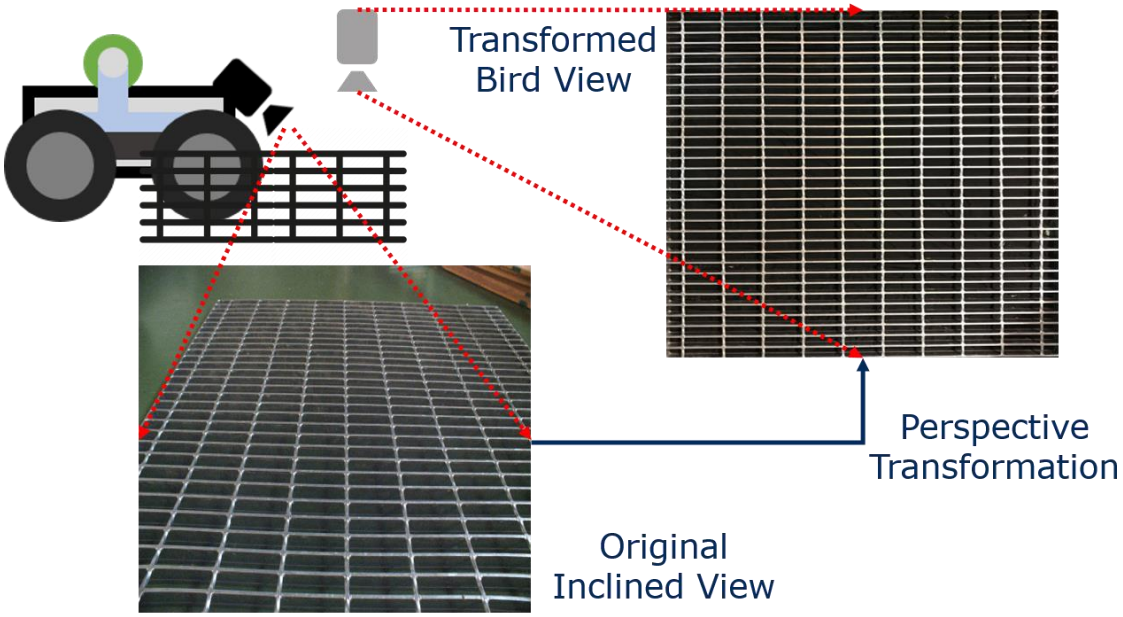


Figure 3.2 The concept of perspective transformation.

In terms of these two views, both views were focusing on the grating floor, but the poses were different. One pose was the original inclined view, and the other pose was the vertical view. To realize the perspective transformation, a homography matrix which relates the transformation between the original camera surface and the grating floor surface would be necessary. Homography matrix is a planar matrix that relates the transformation between two planes. I placed the checkerboard on the ground floor and took pictures to acquire the extrinsic matrix of the camera and the homography matrix, as Figure 3.3 shows. The concerned coordinate systems can be defined as below:

- The global coordinate system. Its origin is set at one point on the grating floor, with the x-y plane parallel with the plane of the grating floor. Conveniently I use the side of the grating lattice to define the x and y-axis. Denoted as (O_g, X_g, Y_g).
- The robot coordinate system. Its origin is set as the center of the robot, with the x-y plane parallel with the surface plane of the robot. I use the side of the robot to define the x and y-axis. Denoted as (O_r, X_r, Y_r).
- The perspective camera coordinate system. Its origin is set at the corner point of the checkerboard for the camera initialization process. Its x-y plane is supposed to be the plane of the checkerboard. And the x and y-axis are defined with the sides of the checkerboard. Denoted as (O_c, X_c, Y_c).

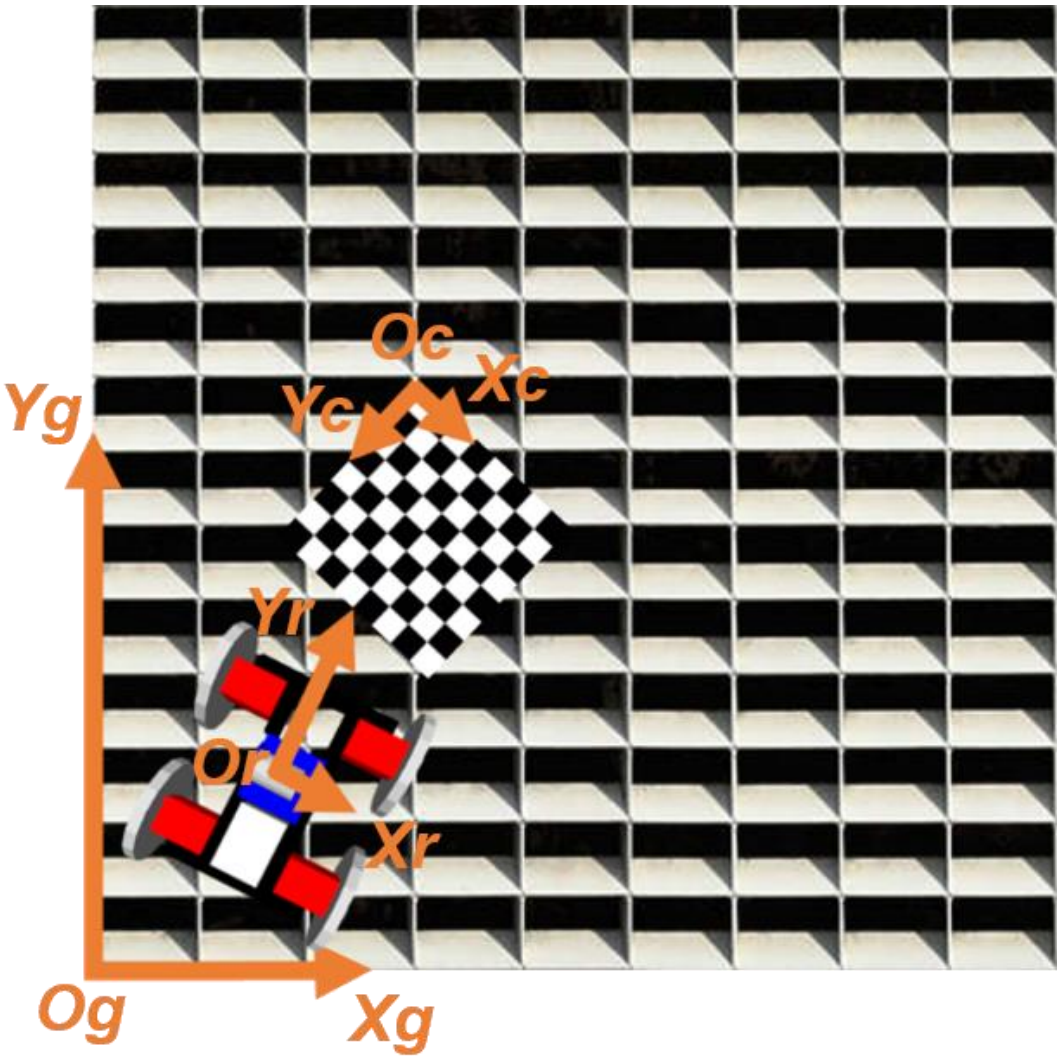


Figure 3.3 The coordinate systems concerned in perspective transformation.

When estimating the homography matrix between (Or, Xr, Yr) and (Oc, Xc, Yc), at least four common points in these coordinate systems will be necessary. Here, the checkerboard with known

structure and size can provide prior knowledge of the corresponding relationships between the grid points in these coordinate systems. The grid points are denoted as (x_i, y_i) and $(x`_i, y`_i)$, and the homography matrix is denoted as $\boldsymbol{H}$, its element at i-th row and j-th column being h_{ij}. Then, the homography matrix can be estimated by minimizing the reprojection error, as Equation 3.4 shows:

$$\boldsymbol{H}^* = \min_{\boldsymbol{H}} \sum_i \left(x_i` - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left( y_i` - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 \quad (3.4)$$

By implementing the perspective transformation in the calibration process of VO, I successfully transformed the original inclined view into the top view, as Figure 3.2 shows. This refinement on the viewpoint transformation provides more intuitive view, and allowing the further usage of regular and repetitive grating structure in removing accumulated error.

3.3 Feature Tracking Failure Compensation Method based on RhinoUS-II's Motion Patterns

In this section, the proposed compensation method designed to improve the robustness of VO against "Feature Tracking Failure" by using motion patterns will presented. This section will start with the definition of feature tracking failure and discuss its possible reasons (Section 3.3.1). The compensation method can be divided into three main steps. The first step is to recognize feature tracking failure from the dataset and judge its reason (Section 3.3.2). Then, based on the motion patterns of the robot, a trajectory clustering step is utilized to extract the neighboring images to the failed frame (Section 3.3.3). At last, an approximate estimation of the instant motion at the failed frame will be provided as the compensation step (Section 3.3.4).

3.3.1 Definitions and Sources of Feature Tracking Failures

Feature Tracking Failure refers to the situation where the correspondence matching processes of VO failed to provide a reliable matched correspondence between two frames. Since the instant motion estimation in VO is highly reliable on the correspondence extracted, the failure will lead to degraded performance or complete failure of the system. The reasons of feature tracking failure may can be listed below:

- The overlapping degree between two frames is not enough. This can happen for a variety of reasons, such as narrow field of view, fast motion, or occlusion. Figure 3.4 shows the one of the failure situations that the subsequent frames have relatively low overlapping degrees. The left image is one frame, the right image is the next frame, and the overlapping areas are circled in red. Obviously, the overlapping areas in these two frames are quite small compared with the total areas of images, which will lead to feature tracking failure. The possible

solution can be limiting the camera motion, increasing the frame rate of camera, or adjusting the down-sampling rate from the video.

- The available features in the two frames are not enough. The possible scenarios in which VO cannot extract enough features are like: The scene is lacking distinctive features like sufficient textures, for example, the white walls. Figure 3.5 shows the situation where the scenario in the subsequent images cannot provide enough features, even the overlapping degree between the two images are quite high. The possible solution can be keeping distances with the areas with few textures.

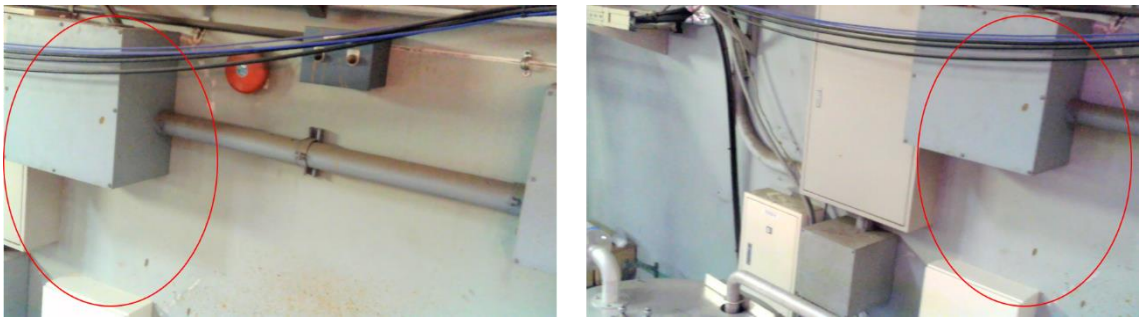


Figure 3.4 One of reasons of feature tracking failure: Overlapping degree is low. Left: The last frame; Right: The current frame.

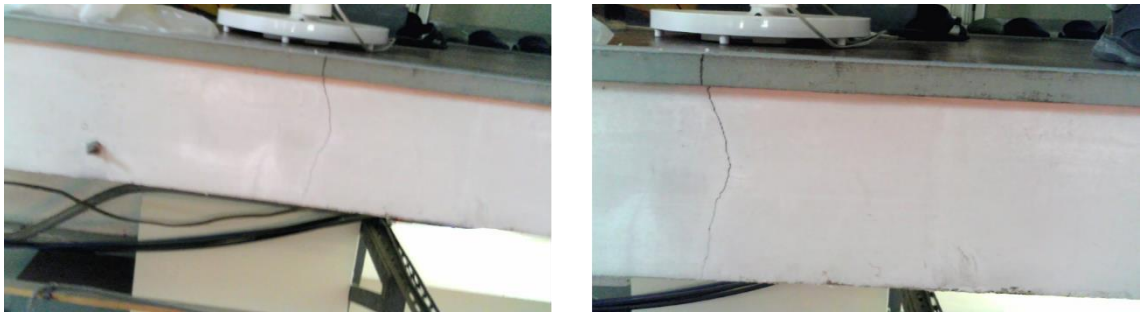


Figure 3.5 Another reason of feature tracking failure: The scene is lacking available features. Left: The last frame; Right: The current frame.

3.3.2 Step 1: Recognition of Feature Tracking Failure Source

On the basis of the discussed reasons of feature tracking failures, I defined several quantities to evaluate the key parameters mentioned above, and used them to recognize the feature tracking failure, as the first step of the proposed compensation method. Figure 3.6 shows a typical image frame after the perspective transformation with a normal condition. The small points in green color indicate the extracted feature points, while the green lines connect the matched feature point pairs in the adjacent frames. The small points in blue color indicate the feature points which fail to have a matched pair with the adjacent frames. On the basis of the transformation relationship between the matched feature point pairs, the instant camera motion between the adjacent images can be estimated through Equation 3.3. Then, the estimated camera motion is represented by the

red circle with a red vector. This figure is a typical representation of a normal condition for VO, since the total number of the extracted feature points is quite abundant, and the number of the successfully matched correspondences between two frames is also very high.

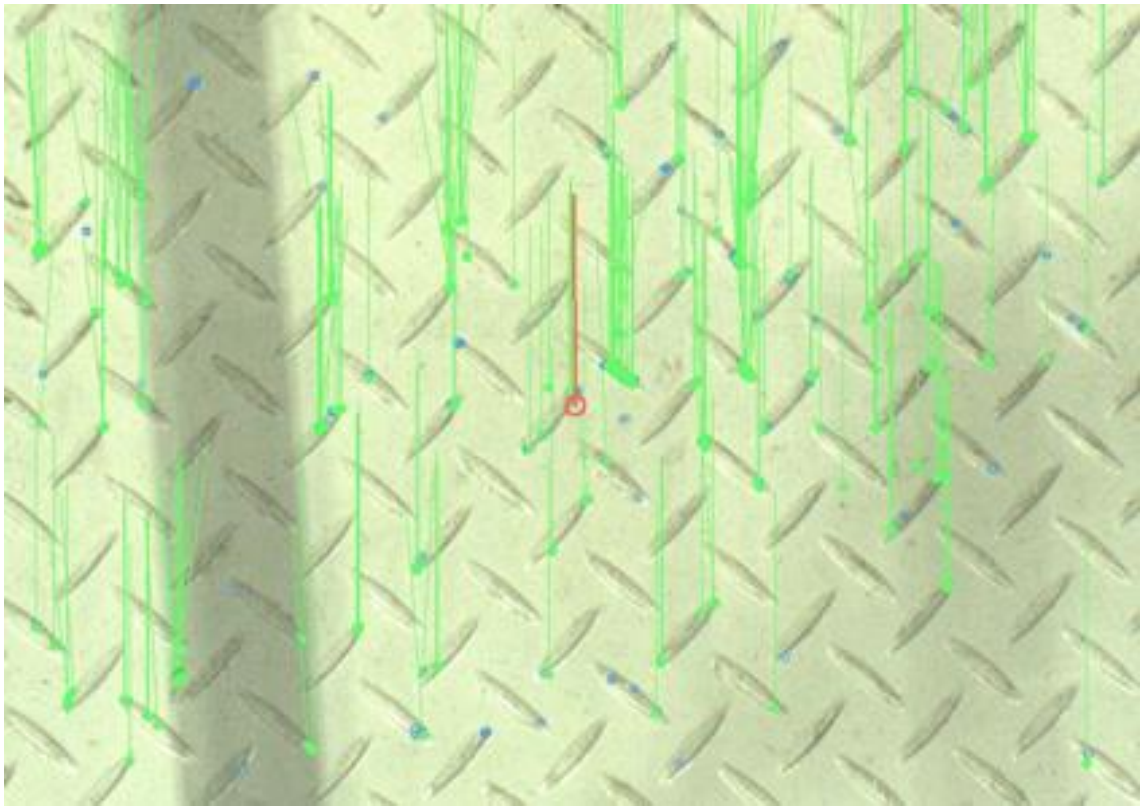


Figure 3.6 The transformed image in top view with a normal condition.

Here, as has been defined in Section 3.1, the successfully matched correspondence with the subsequent frame in n-th frame are denoted as P_n, the green point with green optical flow in Figure 3.6. Besides, I defined the whole extracted feature points including the green points and blue points (points without matched correspondence) in n-th frame as F_n. Therefore, the matching rate of the extracted feature points in n-th frame can be denoted as R_n, which is calculated by Equation 3.5:

$$R_n = \frac{num\ (P_n)}{num\ (F_n)} \tag{3.5}$$

Therefore, the matching rate R_n can be utilized as an approximate estimation of the overlapping degree between the adjacent images. When the robot is moving at a steady speed without occlusion, the subsequent image frames captured by the camera should share a lot of overlapping areas, which is also the basic assumption of visual odometry. However, when the robot is moving at a relatively high speed compared with low frame rate, the overlapping areas between the adjacent frames may drop, which will lead to low matching rate. While, as long as

the environment can provide abundant visual features, the total number of the available feature points F_n can be still high.

On the other hand, when the environment does not consist of enough visual features, like white walls without textures, the number of extracted feature points F_n will decrease, while as long as the overlapping degree between the adjacent frames is decent, the matching rate may not be influenced by the feature lacking. Therefore, in terms of the normal condition of a single image frame, both the matching rate R_n and the number of extracted feature points F_n should meet the threshold requirements. When the matching rate R_n is below the threshold, the situation can be categorized into low overlapping degree between the adjacent image frames, which may result from a sharp turn or high-speed motion. While, the low value of the number of extracted feature points F_n indicates the available feature lacking. Therefore, a preliminary step of recognizing the situations of feature tracking failures is to monitor the defined R_n and F_n, and then use the threshold to judge whether the overlapping degree is enough or the available features are abundant.

3.3.3 Step 2: Motion Pattern-Based Trajectory Clustering

As the main target of the improvement method, utilization of the specific motion patterns can help the feature-based VO method to understand the specific features in the frames which inherit from the robot itself. The inheriting relationship between the camera image and the robot motion patterns can be constructed on the basis of the general smooth camera model of MonoSLAM in terms of the estimated camera motion.

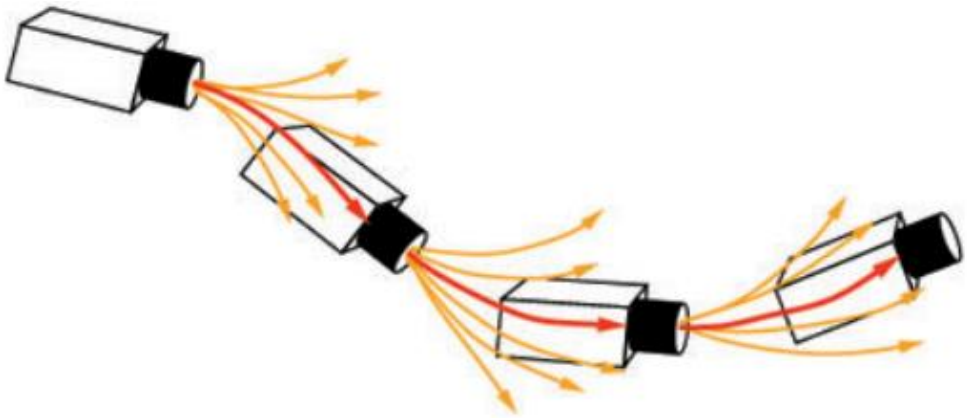


Figure 3.7 The concept of general smooth camera model by MonoSLAM. [42]

The model assumes that the observed statistical model of the camera motion in a time step should be on average we expect. Figure 3.7 shows the basic concept of the general smooth camera model, while the orange lines indicate the expected camera motions at the short time step, and the observed model of the camera motion can be approximated by the average expectation. Especially, in the situation of the planar motion on the grating floor, the camera motion in the normal condition should only include the planar components (the camera should always be on the plane). In other words, the camera motion in the real-world should be spatially continuous (the camera

should not leap) and can be approximated based on the subsequent motions (the camera motion should be smooth).

Therefore, the general smooth camera model combined with the situation of the wheeled robot provided three assumptions listed as below:

- The camera should always be on the plane. Although the estimation method concerned in this study (Equation 3.3) is also capable to perform estimation on spatial motion with the "fullAffine" parameter, and the results should consist of 6 degrees in total. In the specific situation where the camera is supposed to move on the same plane, the existence of component out of the grating plane can be used as flag of estimation failure, while the possible reason can be unexpected collision or vibration.
- The camera should not leap. In the environment with repetitive textures or similar objects, the visual odometry process may mismatch the extracted features when it failed to distinguish the differences. Such a problem may occur more frequently when the visible range is limited. In the specific situation of grating floor, the grating lattices in each frame is highly similar, which increased the danger of mismatch. While still, the matching process is conducted by following the capture order, which is able to avoid the camera position leap due to feature mismatch.
- The camera motion should be smooth. In other words, the velocity in a short time step should not change severely. Therefore, the assumption can be further refined that the velocity vectors in the same clusters should share the similar features, according to the specific motion pattern of the robot itself.

As the precondition, such investigation mobile robots are usually manually-controlled by operator. Besides, the rescue robot usually uses the game pad as the controller, which only contain simple commands like moving forward, backward, turning left and turning right. Therefore, the motion patterns of a wheeled robot on a regular grating plane are relatively predictable. A natural assumption can be obtained that the available motion of the robot, and the robot-mounted camera is restricted on the horizontal floor plane. Therefore, the robot position on the plane can be described with respect to an arbitrary inertial frame by a posture vector $\boldsymbol{\xi_r}$ in Equation 3.7:

$$\boldsymbol{\xi_r} = (x_r, y_r, \theta_r)^{\mathrm{T}} \tag{3.7}$$

In this posture vector, x_r and y_r stand by the coordinates of the center of mass point of the robot chassis, and θ_r describes the orientation of the robot chassis with respect to the inertial frame. Since the camera is not fixed on the center of the robot, a vector $\boldsymbol{R_c}$ connecting the robot center to the camera center is shown in Figure 3.8. The camera position on the plane can be described by the posture vector $\boldsymbol{\xi_c}$ in Equation 3.8.

$$\boldsymbol{\xi_c} = (x_c, y_c, \theta_c)^{\mathrm{T}} \tag{3.8}$$

Figure 3.8 shows the simplified relationship between the camera and the robot, while I denoted the vector connecting the robot center and the camera center as vector $\boldsymbol{R_c}$. I define the direction that the robot faces as the Y_r axis, while the side of the robot as the X_r axis, and similar for the camera coordinate system. I defined that the angular velocity of the robot as ω_r. As has been stated in Section 3.1, I set the camera coordinate system along the corners of the checkerboard, and put the checkerboard parallel with the robot direction for convenience. Under the robot coordinate system, I denoted the robot velocity as ${}^{R}\boldsymbol{V_r}$, and the camera velocity as ${}^{R}\boldsymbol{V_c}$. ω_r is used to present the angular velocity of the robot center. Then, the planar velocity vector of the camera and the robot can be calculated by Equation 3.9:

$$ {}^{R}\boldsymbol{V_c} = {}^{R}\boldsymbol{V_r} + \boldsymbol{\omega_r} \times \boldsymbol{R_c} \tag{3.9}$$

Then, the camera velocity represented with the base vectors of the camera coordinate system can be converted with a transformation matrix based on the orientation deviation between the robot center and the camera center, as Equation 3.10 shows.

$$ {}^{C}\boldsymbol{V_c} = \begin{bmatrix} \boldsymbol{\cos(\theta_c - \theta_r)} & \boldsymbol{-\sin(\theta_c - \theta_r)} \\ \boldsymbol{\sin(\theta_c - \theta_r)} & \boldsymbol{\cos(\theta_c - \theta_r)} \end{bmatrix} {}^{R}\boldsymbol{V_c} \tag{3.10}$$

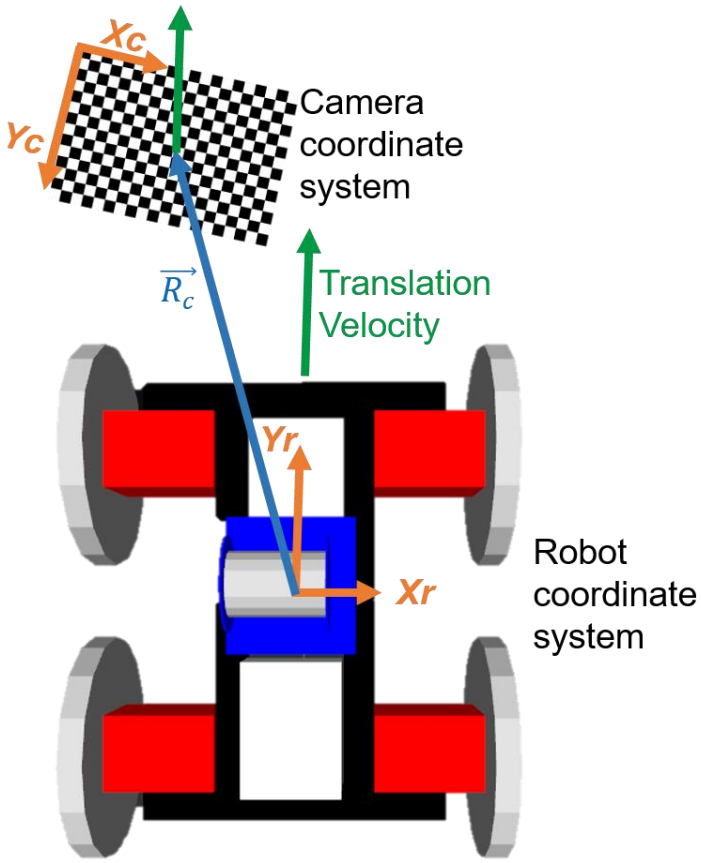


Figure 3.8 The coordinate system relationship between the transformed camera view and the robot.

Taking account of the remote operation processes, it's obvious that the orientation deviation between the robot and the camera may lead to inconvenience, especially when the robot needs to rotate to change the forwarding direction. Therefore, in the initial calibration processes to decide the spatial relationship between the robot and the camera, I intentionally installed the robot camera at positions where $\boldsymbol{R_c}$ is parallel with the longitudinal direction of robot, to avoid the orientation deviation between θ_c and θ_r. A direct influence is that the transformation between robot and camera can be approximately assumed as a simple translation. Such a simplified geometric relationship forms the basis of the following motion patterns.

Among the various four-wheel robots, RhinoUS-II has a differential drive mobile robot structure, which means that the ipsilateral two wheels are synchronously steered to keep the same instantaneous center of rotation [43]. Starting from the motion of the robot center, I assumed that most motions of a wheeled robot performing planar motion can be decomposed into two patterns, listed below:

- Pure translation along the longitudinal direction. Usually, this component only contains the velocity component along the longitudinal direction, and has negligible component along the horizontal direction due to robot chassis shake.
- Pure rotation around the robot center. Usually, this component only contains an angular velocity around the robot center.

Based on the representation of the robot center velocity ${}^{R}\boldsymbol{V_r}$ with these components, I can transform it into the camera center velocity under the camera coordinate, ${}^{C}\boldsymbol{V_c}$ by using Equation 3.9 and Equation 3.10. Therefore, when the robot motion presents the features of such patterns of pure translation or pure rotation, the camera motion will also present the corresponding component relationship features, listed below.

- Pure translation along the longitudinal direction. Since the orientation deviation between the robot center and the camera center has been intendedly set to approach zero, the translation component will keep the original feature after the transformation.
- Pure rotation around the robot center. According to Equation 3.9, the camera motion can be simplified into the motion of a rotating link end, which means that its instant velocity vector should be vertical to the rotating link, vector $\boldsymbol{R_c}$.

Figure 3.9 shows two typical instant camera velocities under such motion patterns. Similar with Figure 3.6, the green points represent the matched extracted feature points; the green lines indicate the matching flow connecting the feature points pair of two neighbor frames; the red circle and red line indicate the camera center and estimated camera velocity vector. The left part of this figure shows the instant velocity vector when robot performs a pure translation along the longitudinal direction. Basically, the camera velocity vector only contains the component along the longitudinal direction, similar with the robot velocity vector. In this figure, the estimated

velocity vector is (0.03, 20.15), with the horizontal direction as X-axis, and the longitudinal direction as Y-axis. While, in the right figure, the rotating angular velocity around the robot center has been converted to a horizontal velocity component. The estimated velocity vector is (16.09, -0.13).

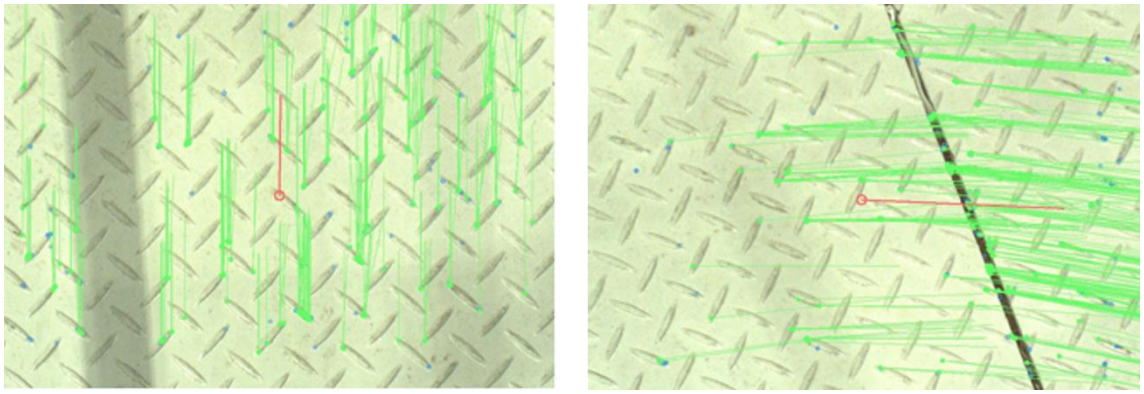


Figure 3.9 The camera velocity vectors under two motion patterns. Left: Pure translation along the longitudinal direction; Right: Pure rotation around the robot center.

Therefore, on the basis of the camera velocity vectors' features under the typical patterns, I defined a flag quantity F_t which concerned with velocity vector component in Y-axis (V_y) and in X-axis (V_x) in each image frame, as Equation 3.11 shows:

$$\boldsymbol{F_t} = \begin{cases} \mathbf{1} - \dfrac{|\boldsymbol{v_x}|}{|\boldsymbol{v_y}|} & \left(|\boldsymbol{v_y}| \geq |\boldsymbol{v_x}|\right) \\ \mathbf{1} - \dfrac{|\boldsymbol{v_y}|}{|\boldsymbol{v_x}|} & \left(|\boldsymbol{v_x}| > |\boldsymbol{v_y}|\right) \end{cases} \tag{3.11}$$

By using F_t as the flag quantity, when it exceeds 0.90, I considered that the magnitude difference between the velocity components were huge enough to be categorized into the motion patterns of "Pure Translation" and "Pure Rotation". Then, the camera trajectory can be divided into clusters where the motion should remain the similar motion pattern. Figure 3.10 shows the clustering processes of a camera trajectory, in purple color. The trajectory can be divided into two clusters: A cluster with pattern of "pure translation", where the velocity vector in the cluster shared the similar pattern feature that the component in Y-axis is much larger than that in X-axis. The right part of the trajectory is a cluster with mixed pattern of translation and rotation.

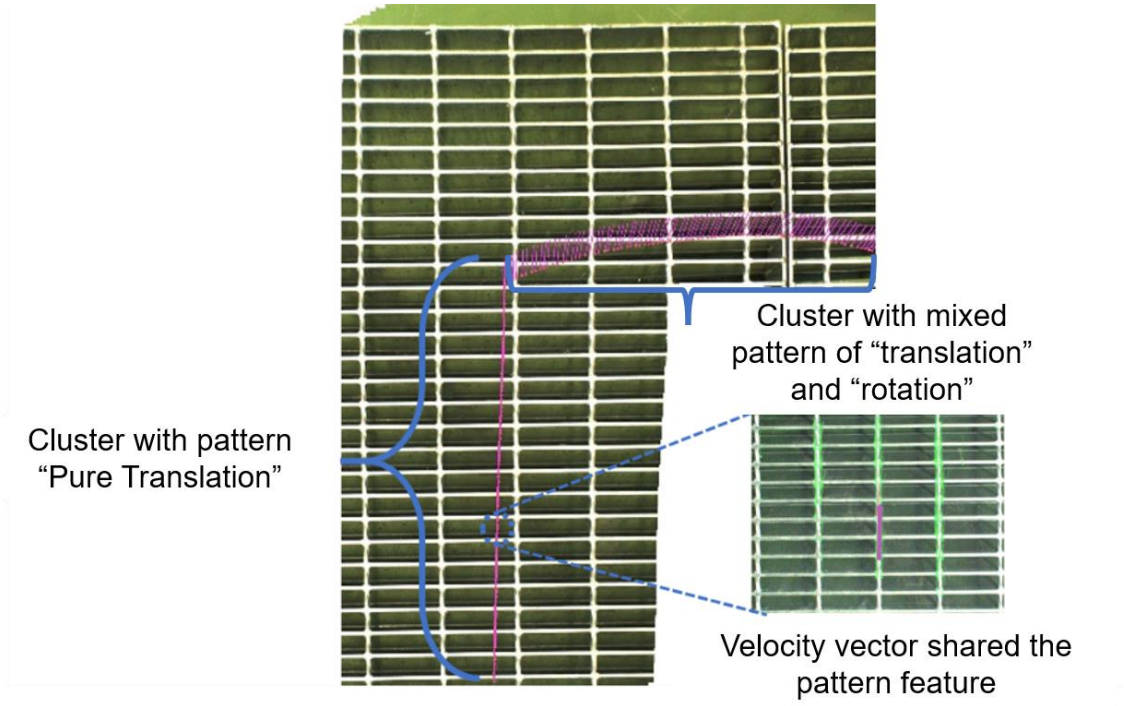


Figure 3.10 The concept of trajectory clustering according to motion patterns.

3.3.4 Step 3: Approximate Compensation with Average Expectation

At last, through the step 1, the image frames with feature tracking failure can be recognized; Then, the step 2 clustered the trajectory into small clusters according to motion patterns of RhinoUS-II robot. At last, the step 3 will provide an approximate compensation of the failed image frame by using average expectation.

The basis of using average expectation to replace the mis-estimated motion of the failed image frame can track back to the general smooth camera model mentioned in Section 3.3.3. The model assumes that the observed statistical model of the camera motion in a time step should be on average we expect. Therefore, after the clustering, the mis-estimated velocity vector when the feature tracking failure problem occurred may have different velocity features compared with the cluster. However, the cluster served as a short time step in the general smooth camera model, which indicates the possibility of approximating the misestimated camera motion with the average expectation. In the early stage of development, I directly used the estimated motion matrix of the last frame of the same cluster to replace the misestimated motion matrix. Such a simplified method would lead to larger errors compared to the Gaussian profile utilized by MonoSLAM. Instead, the misestimated velocity can be approximated by using the average expectation in the cluster, as the compensation method. Therefore, as the result, the compensation method should contribute to improving the correctness of mapping even when the robot meets with low overlapping degrees or temporary areas with few features.

3.4 Verification Experiments

Till now, I have presented the refinement on the conventional VO method, including implementing a perspective transformation method, which allows to transform the original inclined view into intuitive top view thus to utilize grating information as prior knowledge (Section 3.2). Besides, in order to improve the robustness of feature-based VO processes, I proposed a compensation method, including a recognition step, a trajectory clustering step, and a compensation step.

Therefore, in this section, the effectiveness of the proposed two improvement methods will be experimentally conducted. The concerned benchmarks for performance evaluation will be presented in Section 3.4.1; Section 3.4.2 will show the environment setups and the designed trajectory for the experiment. The experimental verification of the compensation method will be presented in Section 3.4.3. Then, the performance of the refined VO will be presented in Section 3.4.4. At last, a preliminary experiment on the available range of the camera parameters to keep required accuracy is presented with Section 3.4.5.

3.4.1 Benchmarks for Performance of Refined VO

As was pointed out in Section 3.1, the application environment of the concerned VO can be concluded as “Pedestal grating”. Concretely speaking, the robot will utilize the VO method to localize itself while approaching the target position on the grating floor. Therefore, the localization accuracy of the VO method will be of great importance. As for the situation of PMORPH2, the accuracy requirement was set as 100mm/1500mm, while 100mm is the length of the grating lattice. Therefore, the localization accuracy served as one of the benchmarks in the performance experiment of refined VO, while the accuracy can be calculated by comparing the estimated posture $\boldsymbol{\xi_r}$ with the ground truth posture.

$$\boldsymbol{\xi_r} = (x_r, y_r, \theta_r)^{\mathrm{T}} \tag{3.7}$$

Apart from the localization accuracy, the performance of map generation served as another benchmark. An ideal map should reflect the realistic structure and positional relationship, and keep consistency with the ground truth environment. Especially, because of the integration of the estimated instantaneous motion in one frame, the errors in each estimation would accumulate and influence the whole localization and map generation. The refinement allowed to transform the inclined view to a top view, so the grating lattices are displayed vertically with the view direction. Then, I implemented the automatic recognition of lattices with known size and regular shape on the basis of YOLO v7 [56], and counted the number of traversed lattices to assist the accumulated error removal processes. In order to quantitively study the smoothness of generated map, I defined a specific quantity named “Regular Grating Rate” for the grating floor situation, as an approximate estimation on the relative consistency of the reconstructed grating lattices.

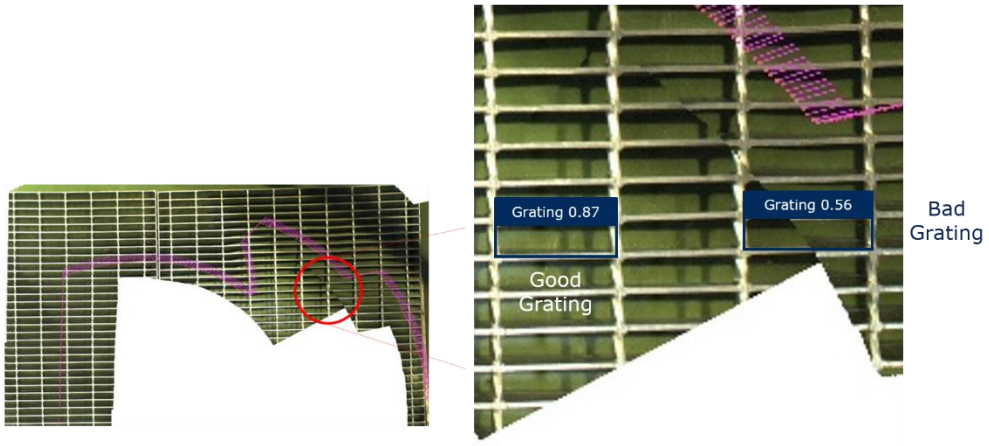


Figure 3.11 A typical map in top view generated by refined VO, and grating counted by YOLO v7. The value in frame indicates the regularity of the detected grating lattice.

Figure 3.11 shows a typical map in top view generated by the refined VO. The purple lines in this figure indicate the camera route concerned. In this figure, most of the grating lattices are quite regular, and can be automatically recognized on the basis of the known size and shape. But in the red circle part, there are slightly distorted grating lattices, which cannot be correctly recognized due to the distortion. Then, I denoted the number of the visible grating lattices as Ng_v, while the number of regular grating lattices recognized was denoted as Ng_r. Then the map smoothness benchmark regular grating rate GR can be calculated as the ratio:

$$GR = \frac{Ng_r}{Ng_v} \tag{3.12}$$

Along with the processes of the refined VO, the localization accuracy and map generation performance of the LRF-based method gmapping (Section 2.4) was also used as compare target.

3.4.2 Environment Setups for VO Performance Experiment

Figure 3.12 shows the experiment environment for the VO performance experiment. The environment is the combination of a water tank and the external 2 floors, so the robot can move on the second floor as the grating floor, while the investigation areas are set at the bottom of the water tank as the bottom pedestal. Basically, it is a simplified simulation of the spatial relationship of the grating floor and the bottom pedestal for the further developments, built by Naraha Remote Technology Development Center and JAEA.

Figure 3.12 The experiment environment in Naraha Remote Technology Development Center.

Now, the focus came to the second floor as the experiment room. The room can be divided into two parts, the external floor and the tank room, as Figure 3.13 shows. The overall size is 7.5m x 6.68m. At the center of the tank room, there is a cylindrical water tank with diameter being 4.5m. Beyond the water, there is a wooden bridge with grating floor across the tank, as the simulation of the opening to the bottom pedestal (Figure 3.14). The boundary of this room is covered with 3 acrylic walls and sparse guardrail in the external floor direction. The deep blue line indicates the designed route of the experiment. The robot is supposed to start from the corner of the external floor, passing through the floor and move to the front end of the wooden bridge.

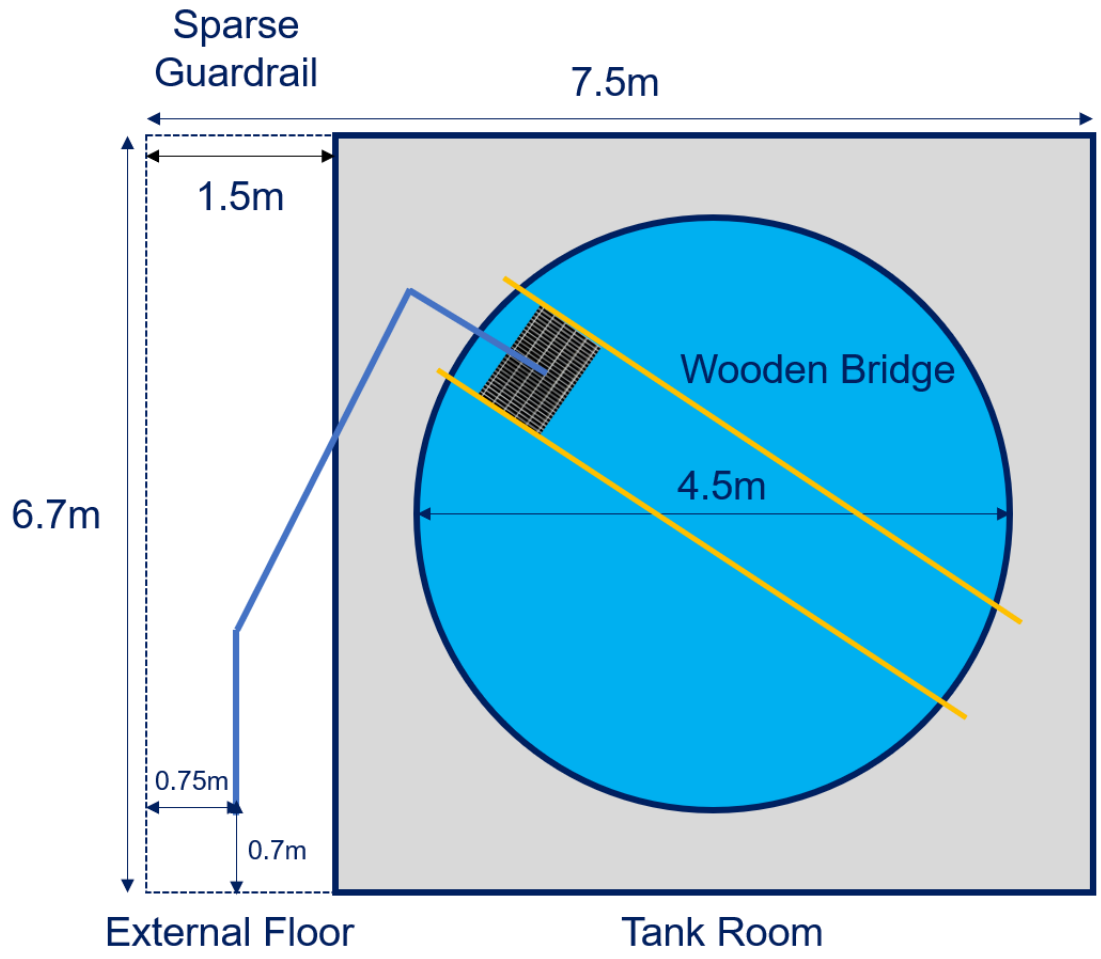


Figure 3.13 The experiment room in top view.

As Figure 3.15 shows, I made some modifications on the environment, as a variation environment setup. The original environment floor only has a small area of grating floor on the wooden bridge since setting up the whole areas with grating will be quite difficult and time-consuming, while the assumed working situation of the method was set as whole grating floor. Therefore, I covered the ground floor with printed textures of grating lattices. These two environment setups are denoted as Setup 1 (Original floor) and Setup 2 (Floor covered with printed textures). Then, the displacement between the starting point and the ending point will be estimated by the refined VO method and LRF-based SLAM method, and compared with the measured ground truth by the laser scanner. The specification of the camera setup and the robot information is shown in Table 3.1.

Figure 3.14 The wooden bridge across the water tank (As the route end point).

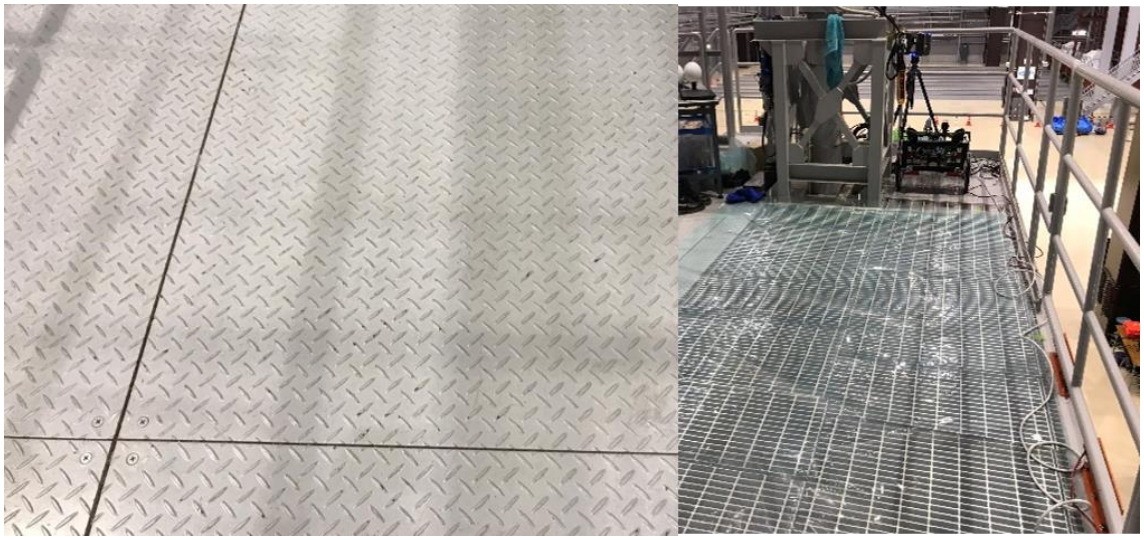


Figure 3.15 The environment setup variation in the external floor. Left: Original ground floor; Right: Floor covered with printed textures of grating lattices.

Table 3.1 The specification of robot setup.

Specification		Value
Camera	Frame Rate	30
	Resolution	1280x720
	Height (mm)	400
	Inclination (degree)	50
Robot	Average Speed (mm/s)	200

3.4.3 Experiment on the Effects of Compensation method

As has been presented in Section 3.3, the proposed compensation method included a recognition step, a trajectory clustering step, and a compensation step. Therefore, in order to verify the effects of the compensation method, the experiment was conducted by following these three steps.

The experimental method went as: I performed the designed route from the initial point to the end point. While, I used the camera installed on RhinoUS-II to capture images during the movement. The captured image sequence would be utilized as the input of the VO method without the proposed compensation method. A map of the past path would be generated with conventional VO method. Then, the extracted features and the match correspondences in the images would be calculated, and transferred to the recognition step. Therefore, the recognizing step should be able to detect time periods where there might be a feature tracking failure. The corresponding areas of the detected time periods would be focused. Then, the VO with compensation method would generate the map again, and comparing with the map before compensation according to the detected areas.

Through a test run of the designed trajectory under the environment setup without printed textures, the captured images were utilized as the input of the VO method. The total number of the images was around 1200. The frame rate was set as 30. A map was generated by the original VO method without compensation method. Then, the extracted features and the match correspondences in the images were transferred to the recognition step. Here, Figure 3.16 shows the relationship between F_n, the number of extracted feature points and the time stamp. Obviously, the time period around 34s to 38s had a sharp drop. According to the recognizing step, F_n is an approximate estimation of the available features, which indicated the past areas at 34s to 38s may have feature tracking failures, and the reason should be lacking available features.

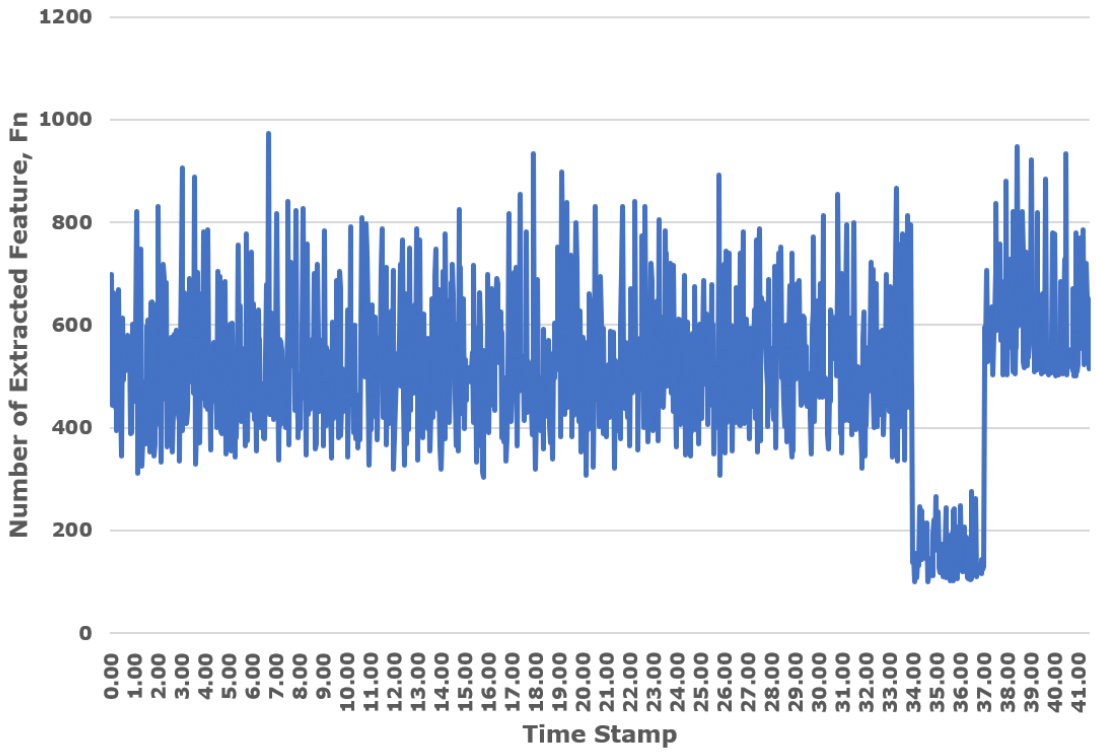


Figure 3.16 The number of extracted feature points F_n according to the time stamp in the test run image sequence.

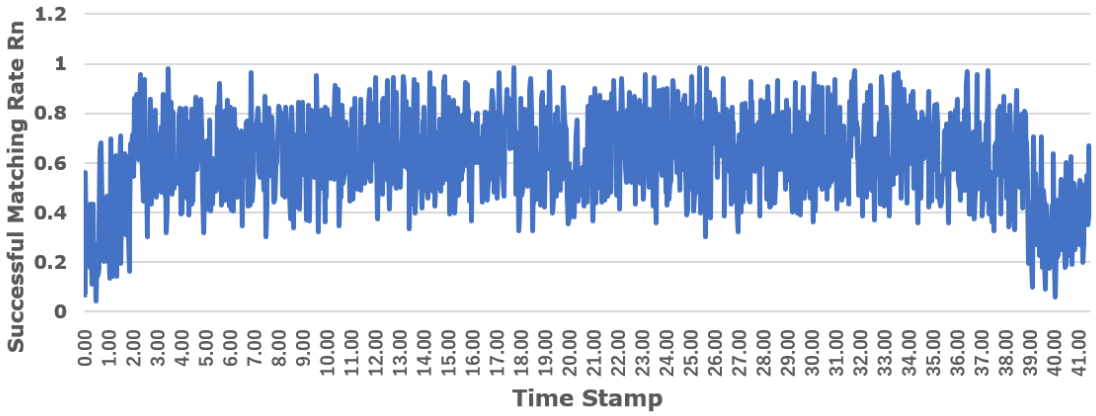


Figure 3.17 The successful matching rate R_n according to the time stamp in the test run image sequence.

Similarly, the monitored quantities, R_n, the successfully matching rate and, according to the time stamp is shown in Figure 3.17. Since R_n served as an approximate estimation of the overlapping degree between the adjacent images, the drop of R_n at the period of 0s~2s and that of 39s~40s indicated that these periods might suffer from low overlapping degree. Here, the detected

areas with possible feature tracking failure with time periods are listed in Table 3.2, denoted as Area #1, 2, 3 from left to right.

Table 3.2 The detected periods with feature tracking failure and the possible reason.

Areas	Time Periods	Possible Reason
1	0.1s~1.9s	Low overlapping degree
2	33.8s~37.1s	Lacking available features
3	39.1s~41.3s	Low overlapping degree

Based on the recognition results of the time periods where there might be a feature tracking failure, the corresponding areas of these periods are marked in the generated map without the proposed compensation method. Figure 3.18 shows the generated map by VO without compensation method, with detected areas marked in red circle. The purple line indicates the estimated camera trajectory.

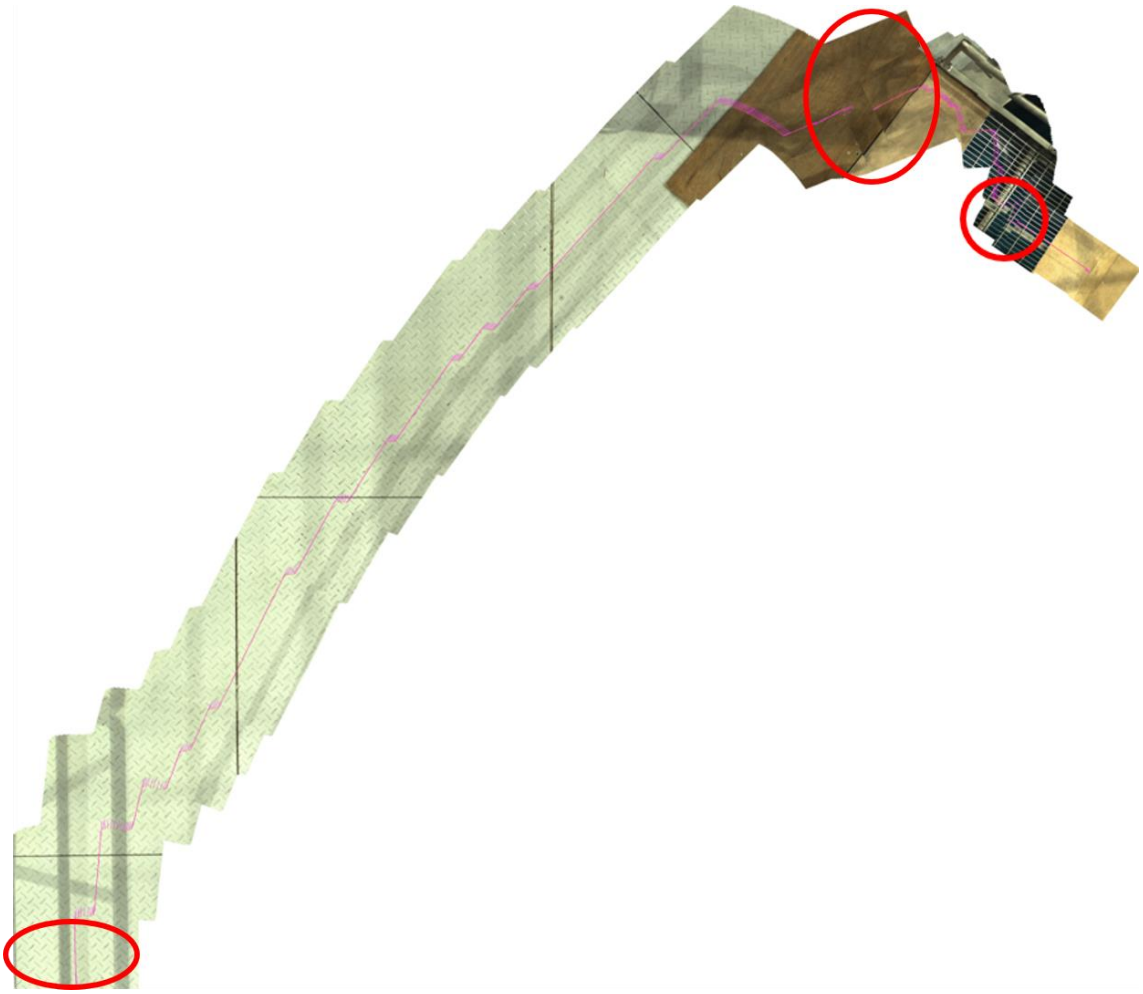


Figure 3.18 The generated map without the proposed compensation method. The red circle: Detected feature tracking failure position by the recognition step.

To confirm the correctness of the proposed recognition step, the areas of the detected periods will be focused. Besides, in order to confirm the effect of the proposed compensation method upon the influences of the possible feature tracking failure, the image sequence was reconstructed again with the compensation method. The comparison between the focused three areas before and after compensation method is shown in Figure 3.19, Figure 3.20 and Figure 3.21.

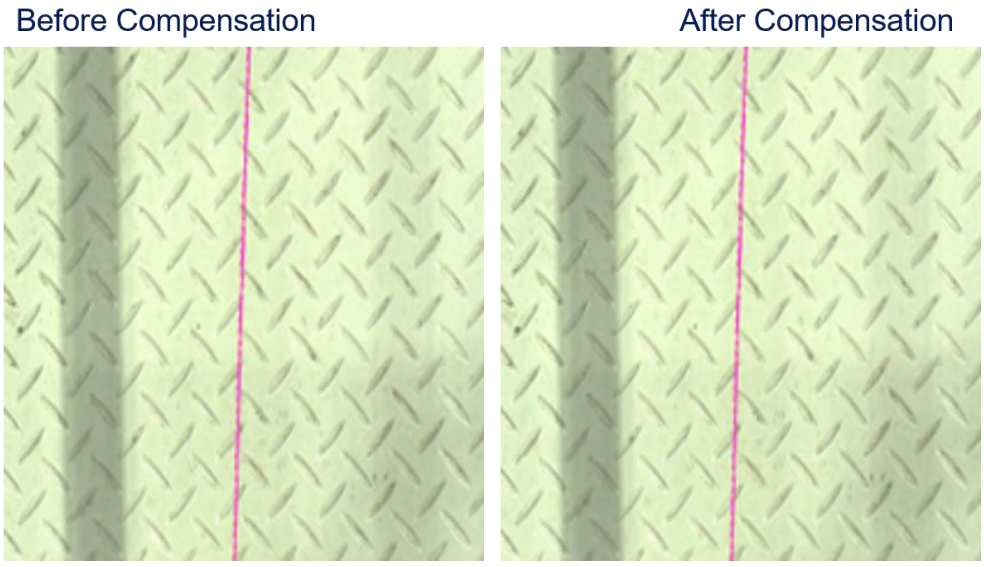


Figure 3.19 The comparison between the generated map before and after the compensation method of Area #1.

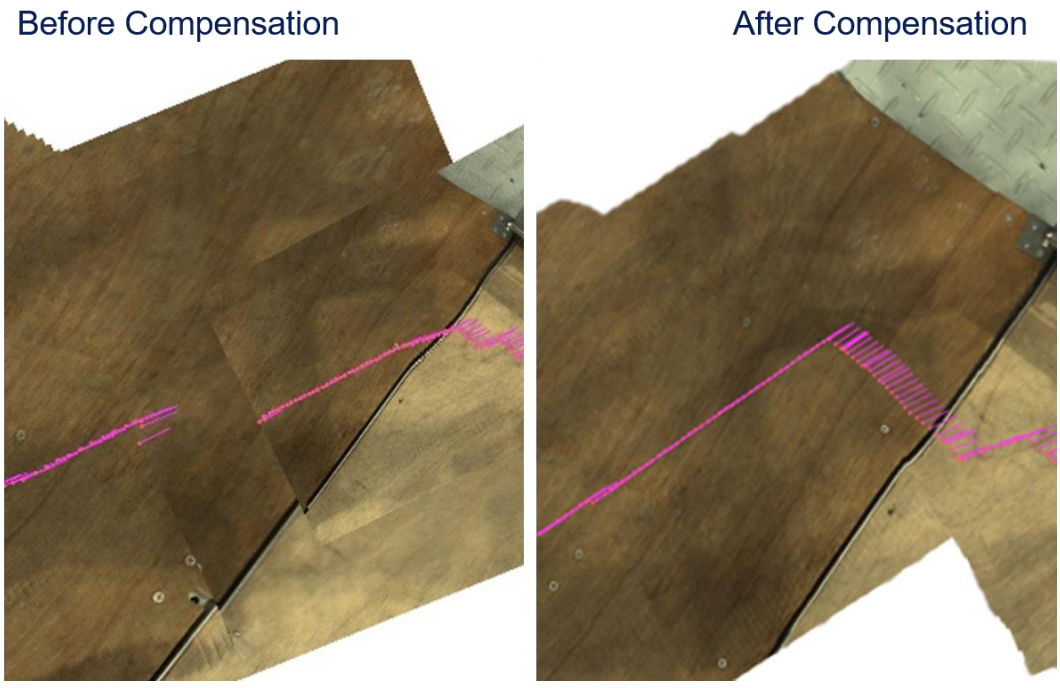


Figure 3.20 The comparison between the generated map before and after the compensation method of Area #2.

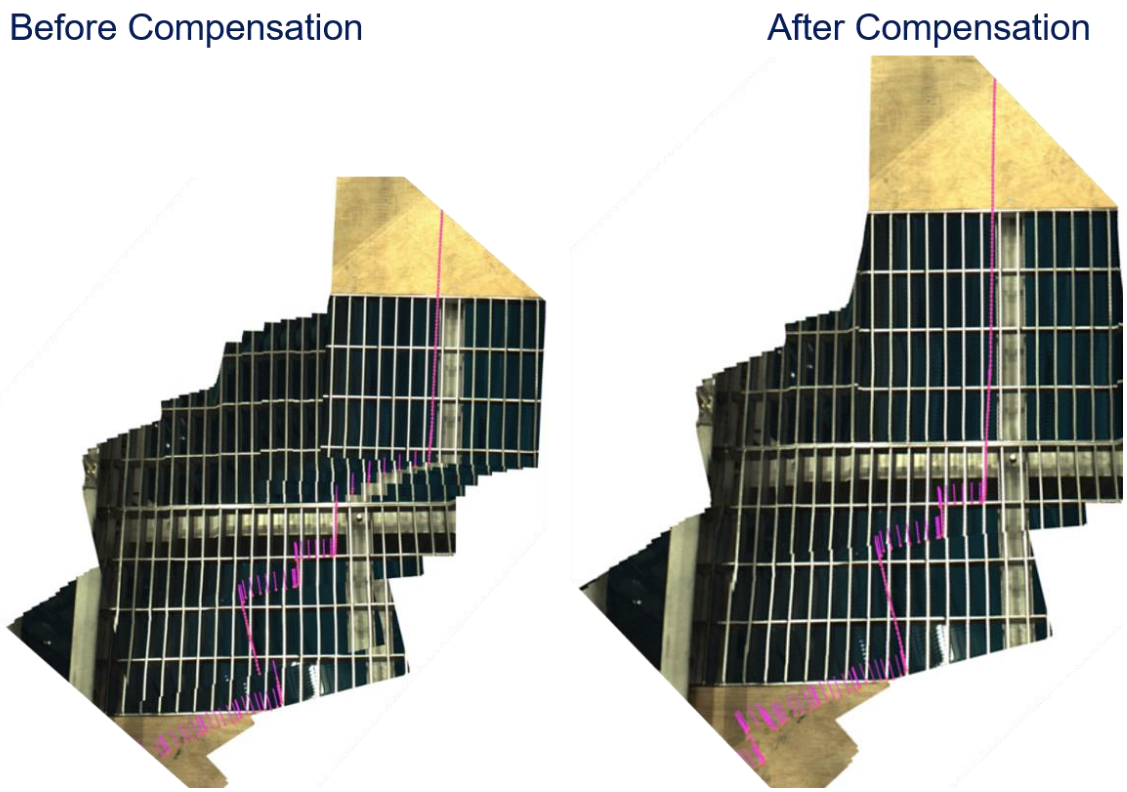


Figure 3.21 The comparison between the generated map before and after the compensation method of Area #3.

Through the visual comparison, there is no obvious distortion or inconsistency in the generated maps of Area #1, and the difference of the compensation method could not be confirmed. Therefore, the Area #1 could be considered as a misrecognition. In this area of time period from 0.1s~1.9s, the wheeled robot should be starting to move, so the starting vibration may reduce the overlapping degrees between the images. While, in terms of the map generation, there have been many steady images of this areas before the robot started to move, so the vibration didn't influence the map.

For the Area #2, it's near the boundary of the floor and the wooden bridge. In the left part of Figure 3.20, there is obvious broken trajectory and the generated map before compensation, and the reconstructed wooden bridge is distorted due to the unreliable instant motion estimation, which should result from a feature tracking failure. The recognition step successfully detected the area with feature tracking failure with the drop of number of available feature points. Reflecting the wooden bridge, during this period, the camera view only contained the wooden bridge, which could not provide enough feature points. Then, after utilizing the proposed compensation method, the generated map got much smoother and consistent with the actual wooden bridge. Therefore, the effect of the compensation method got visually confirmed.

At last, the final Area #3 is on the wooden bridge with grating lattice and supporting links. The left part is the generated result without the proposed compensation method, and I can observe obvious inconsistency and distortion around the grating lattices and the cross-shaped supporting links. In order to numerically confirm the effect of the proposed compensation method, I utilized the defined regular grating rate in Section 3.4.1, and the GR in the left part before compensation was measured as 62%. After applying the compensation method, the GR increased to 90%, indicating that the consistency and regularity of the generated grating got improved by the compensation method. Therefore, the effect of the proposed compensation method got confirmed experimentally, including increasing the consistency, smoothness, and regularity of the generated map.

3.4.4 Experiment Results of Refined VO Performances

In the last section, the effect of the compensation method has been confirmed. In this section, the performance of the refined VO with the compensation method will be evaluated based on the benchmarks defined in Section 3.4.1.

For both the environment setups, I performed the designed route from the initial point to the end point. While, I used the camera installed on RhinoUS-II to capture images during the movement, and conducted the refined VO method to localize the robot displacement relative to the initial point. The ground truth posture was measured with an external laser scanner Faro Focus 3Dx130, which is capable to provide decent and accurate measurement of the surrounding environment with error rate under 2mm, while the maximum measurement range is 130m. Along the experiment, I also utilized laser range finder to perform the LRF-based SLAM method gmapping as a comparison target. While, in the actual experiments, I did not follow the strict straight route, instead, I added some random rotations in the routes. Each experiment was repeated three times.

As the results on the localization accuracy, Table 3.3 shows the robot posture comparison between that estimated by the refined VO method and the ground truth measured by laser scanner under the environment setup 1. Therefore, by comparing with the measured robot posture, the localization accuracy of the refined VO method was confirmed. Still, as long as there are abundant visual features, VO is able to provide decent localization accuracy with maximum error rate in X-direction being 13mm/1500mm, 40mm/1500mm in Y-direction, both within the requirement range of 100mm/1500mm. Figure 3.22 shows one of the generated maps under the environment setup 1, while Figure 3.23 shows a topological map generated by gmapping. The topological map by gmapping reconstructed the shape of the experiment room by receiving the reflected laser signal, but cannot provide information of the floor. While, the refined VO method successfully reconstructed the grating lattices visible to the camera, as an intuitive top view map. On the other hand, the visible range in the top view map got reduced, as a problem. For the environment setup

1, only the natural grating areas on the wooden floor were selected. Then, the number of the recognizable grating lattices were counted and compared with all the visible grating lattices. Therefore, the regular grating rate was also calculated as the benchmark of map smoothness.

Table 3.3 The benchmarks result of refined VO in environment setup 1 (Original Floor).

Method	X-direction displacement (mm)	Y-direction displacement (mm)	Orientation (°)	Regular Grating Rate (%)
VO-1	2908	3451	-118.2	89.2
3D Scanner-1	2883	3528	-128.1	/
VO-2	3000	3216	-132.1	95.1
3D Scanner-2	2981	3299	-129.3	/
VO-3	2987	3311	-142.0	91.6
3D Scanner-3	3012	3403	-140.9	/

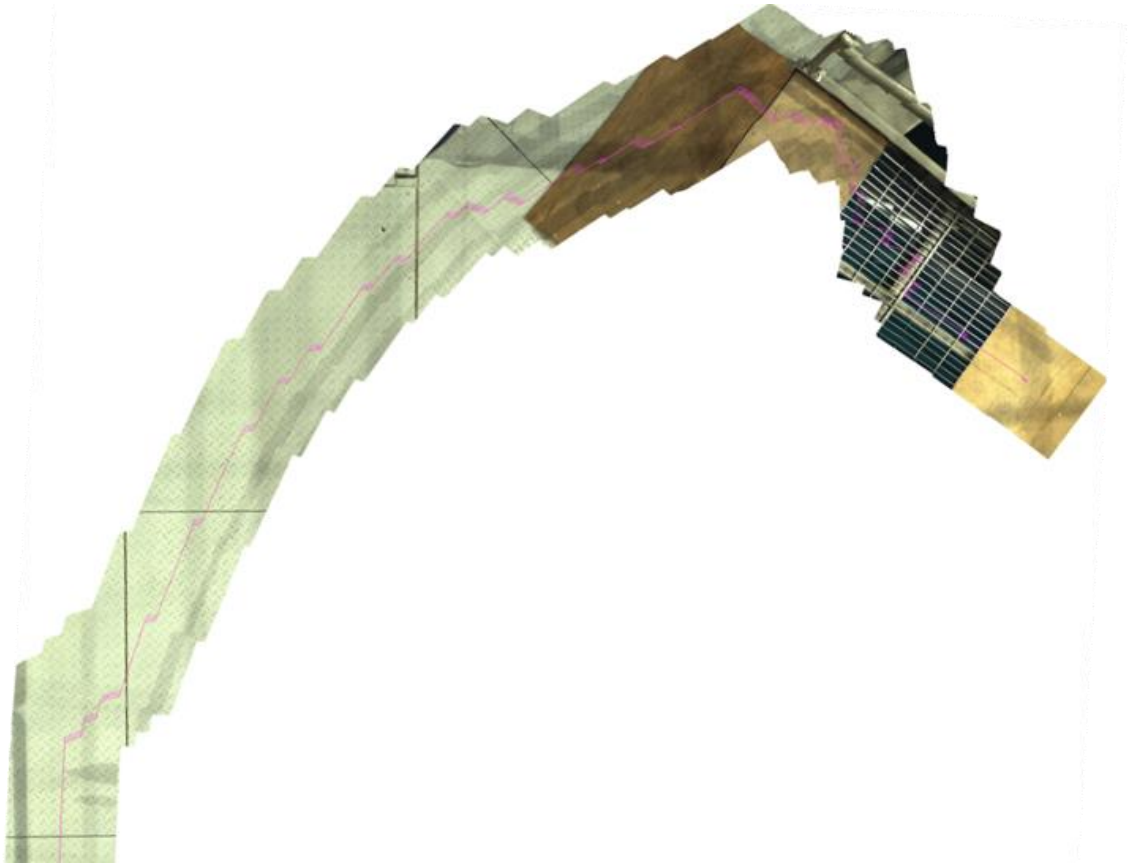


Figure 3.22 One of the generated maps by refined VO under environment setup 1 (original floor).

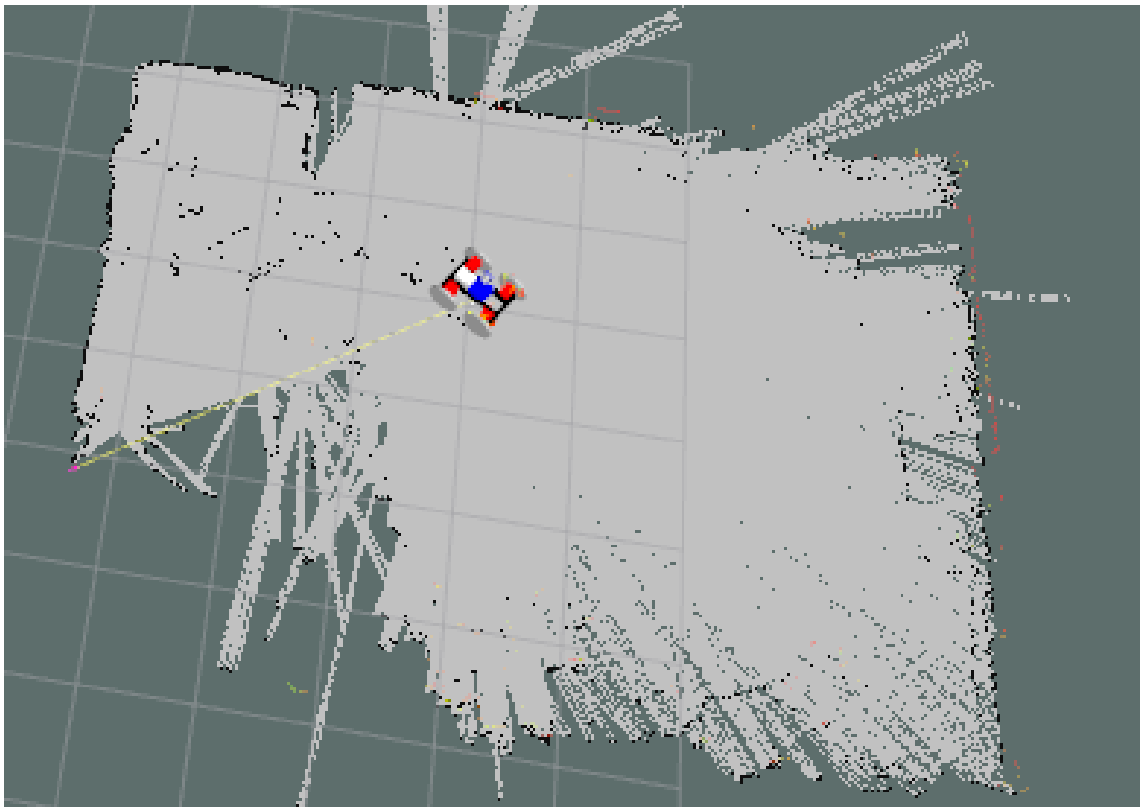


Figure 3.23 A topological map generated by the LRF-based SLAM method gmapping.

Table 3.4 The benchmarks result of refined VO in environment in environment setup 2 (Floor covered with printed textures).

Method	X-direction displacement (mm)	Y-direction displacement (mm)	Orientation (°)	Regular Grating Rate (%)
VO-1	3043	3022	-117.2	97.2
3D Scanner-1	3025	3180	-126.9	/
VO-2	3134	3122	-127.2	94.1
3D Scanner-2	3146	3226	-129.1	/
VO-3	3249	3130	-133.9	98.4
3D Scanner-3	3192	3041	-134.2	/

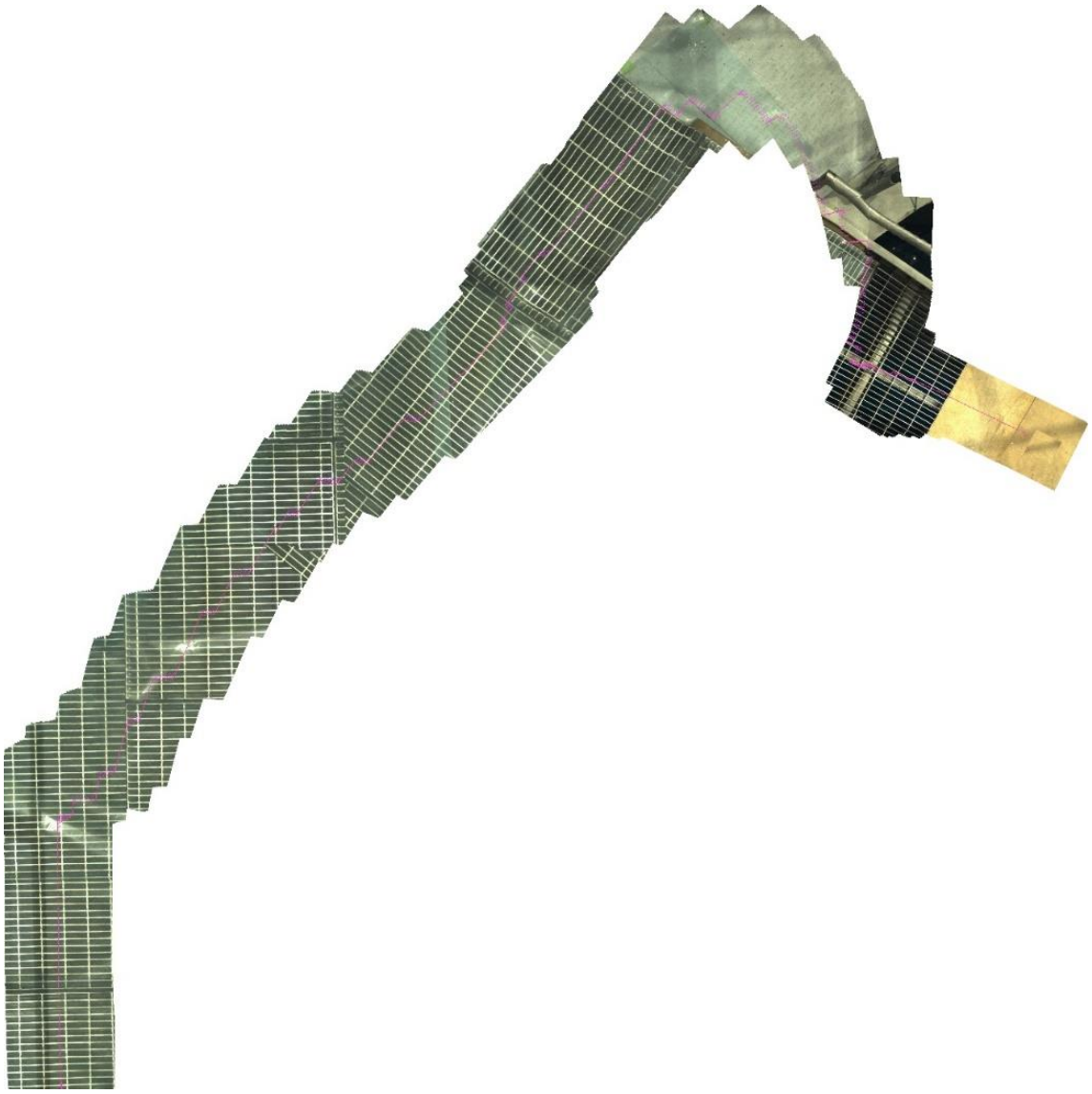


Figure 3.24 One of the generated maps under environment setup 2 (floor covered with printed textures).

Similarly, the results of localization accuracy of robot posture, and generated maps under the environment setup 2 with printed textures are shown below. The maximum error rate in X-direction was 26mm/1500mm, and 74mm/1500mm in Y-direction, both slightly increased compared with environment setup 1. Figure 3.24 shows the generated map in top view. Owing to the covered textures of grating lattices, the regular grating rate got improved compared with setup 1, which indicates that the overall map smoothness got enhanced.

3.4.5 Camera Parameter Ranges to Keep Required Accuracy

Through the previous experiments, the basic performance of the refined VO has been confirmed under the default camera parameters denoted in Table 3.1. While, in the on-site investigations, more restrictions could apply depending on the environmental conditions, like the

size limitations to pass through access penetration, and the camera performance limitations in extreme situations. To provide useful information for the further development of the practical model, in the following sections, I focus on the parameters of the camera pose and the threshold designed for recognizing the feature tracking failure, and use an experimental method to investigate the available ranges to keep the error rate within 100 mm/1500 mm.

The concerned parameters are listed below:

- Camera height. The default value was set as 400mm, and this experiment will study with range of 250mm~600mm.
- Camera inclined angle. The default value was set as 50 deg, and this experiment will study with range of 30 deg ~ 70 deg.
- Frame per second. The default value was set as 30, and this experiment will study with range of 12~60.
- Minimum number of extracted feature points. The default value was set as 250, and this experiment will study with range of 50~350.

The experimental method will be investigating the changing tendency of the localization accuracy corresponding to each parameter individually within the adjustable range, by keeping other parameters constant. Similarly, the laser scanner served as the ground truth reference device. As for the trajectory, two types of simplified motion patterns were used, including pure translation along +Y direction for 1000mm, and pure rotation around the robot center for 45 deg in a clockwise manner. For these two motion patterns, the corresponding quantities to evaluate the localization accuracy are defined as follows.

For the translation pattern, the error rate of translation displacement ER_t was defined as the difference between the estimated displacement in Y-direction and the ground truth. A shift error quantity SE_t was defined as the displacement error in X-direction. For the rotation pattern, the error rate ER_r was defined as the proportion of the orientation difference over the ground truth orientation, SE_r was defined as the absolute trajectory root-mean-square-error rate. The threshold for these error rates was set as 100 mm/1500 mm, as well as 3 deg/45 deg. The changing tendency of these indexes with changing parameters are presented with Figure 3.25 and Figure 3.26. The vertical axis indicates the error rates with the line chart, and the horizontal axis shows the changed parameter. I also included the shift error value in the figure with the bottom bar chart as an alternative index of the accuracy.

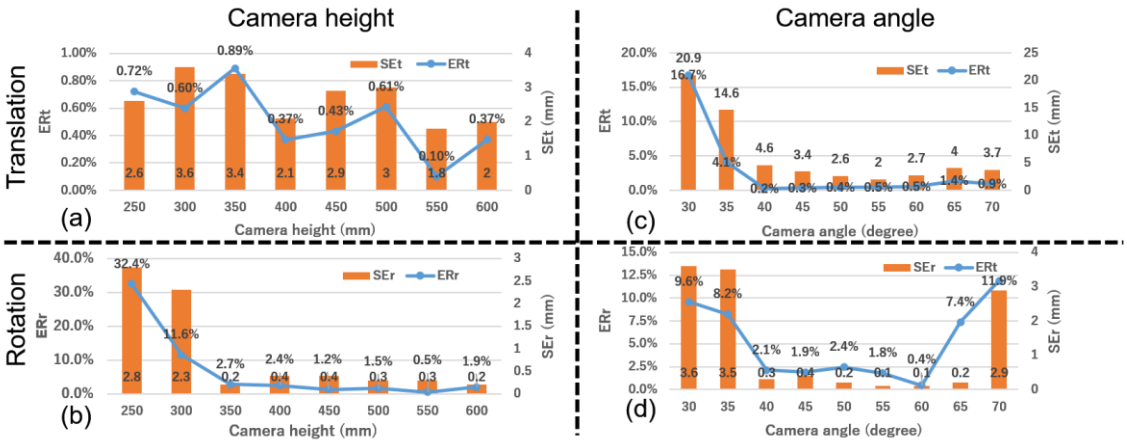


Figure 3.25 Error rate and shift error with changing camera height and angle. (a), (c): Translation motion; (b), (d): rotation motion. (a) and (b) show the effect of the camera height, while (c) and (d) show the effect of the camera angle.

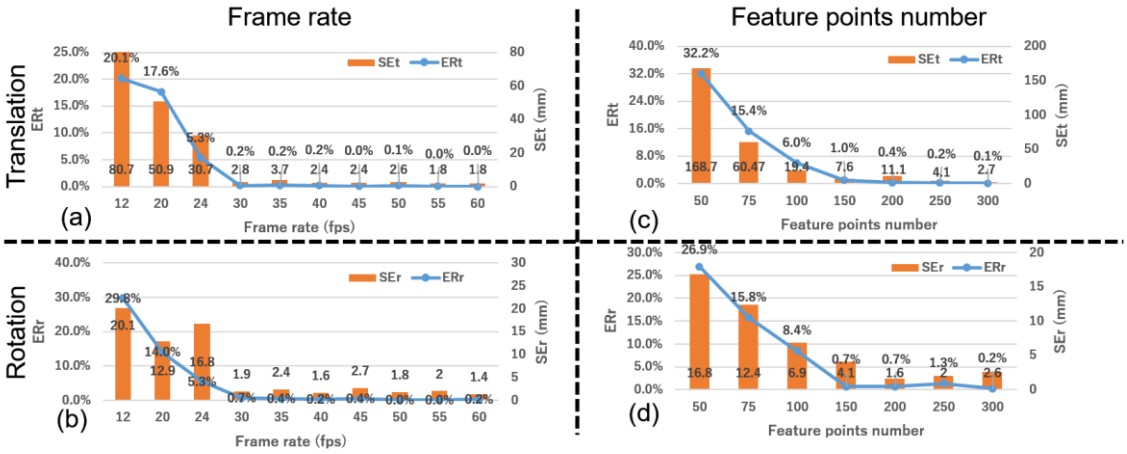


Figure 3.26 Error rate and shift error with changing frame rate and feature point number. (a), (c): Translation motion, (b), (d): rotation motion. (a) and (b) show the effect of the frame rate, while (c) and (d) show the effect of the number of feature points.

As the left part of Figure 3.25 shows, under changing camera heights, both error rates and shift errors in the translation motion were at approximately the same level. I concluded that the influence of the changing camera heights on the localizing performance was relatively small in translation motion. I attributed this tendency to the transformation of the originally inclined camera view to the top view; the homography matrix was not sensitive to the height itself when there were no rotating components in the instant motion matrix. However, when the robot performs primarily rotating motion, both the error rates and the shift error appear quite significant when the camera heights are less than 300 mm. I supposed that when the camera height was relatively low, the available view range of the grating would correspondingly reduce. Furthermore,

the perspective transforming process acquired the bird's-eye view at the expense of a reduced view range. However, the available view range was closely related to the available number of feature points. When trying to track features of relatively large objects, a sufficient view range would become relevant. Thus, to ensure a sufficient view range after the perspective transformation, the camera height should be set to more than 300 mm.

Under changing camera angles, the variation in the translation motion was similar to that of changing the camera height in the rotation motion. Error rates and shift errors appeared significant when the camera angle was less than 40°, while the rotation results limited the camera angles from 40° to 60°. In the perspective transformation, the homography matrix contained mappings in both the vertical and horizontal directions. The mapping component in the horizontal direction was similar to the radius around the original focus in the original view in the rotating motion, which would influence the localization accuracy significantly. Thus, I considered that a suitable range of camera angles should be set to limit the mapping component in the horizontal direction. Because I aimed to transform the original view to the top view, the homography matrix may not be robust or accurate when the original view approaches the horizontal plane, for example, when the angle is set to over 70°. However, the results also showed that the horizontal camera configuration of PMORPH may not be directly applied in the visual odometry method.

As Figure 3.26 shows, the tendency appeared similar to the error rates, and shift errors remained at a relatively low level with the changed parameter over a specific value. We considered the parameters of the programming setting to be similar because the frame rate directly influenced the available frames when passing the objects, and the number of feature points was relevant to the reliability of instantaneous motion estimation. However, in an environment where there are more types of textures, the requirements on the frame rate and number of feature points may also increase. This is because at the boundaries of different textures, the feature points from different textures might be difficult to match in time without sufficient frames in the feature tracking process. However, although setting the frame rate and feature point number might also increase the computational burden of the program, I did not observe any influence on the current experimental settings.

The result of the available range of camera parameters can be referred in Table 3.5:

Table 3.5 Range of parameters meeting the error rate requirement.

Parameters	Available ranges
Camera height (mm)	>300mm
Camera inclined angle (degree)	40°–60°
Frame per second (frames/s)	>24 fps

Minimum number of feature points	>150

3.5 Summary

This chapter proposed two improvement methods for Visual Odometry to utilize grating information and wheeled robot motion patterns as prior knowledge. The first improvement method was a perspective transformation method, which transformed the original inclined view to intuitive top view, thus providing the possibility of effectively counting the number of grating lattices. The second improvement method was a compensation method, which used the specific motion patterns of RhinoUS-II to cluster the trajectory and then correct the misestimation of motion vector.

Through experiments, the performance of the refined VO method was quantitively studied, from the aspect of localization accuracy and map generation. Two variations of the environment setups were conducted, including using original ground floor and covering the floor with printed textures of grating lattices. In both the environment setups, the refined VO achieved decent localization accuracy, as the maximum positional error rate turned out to be within 12.5mm/1500mm and 40.5mm/1500mm in X, and Y direction, which met the requirement of under 100mm/1500mm.

Besides, based on the proposed perspective transformation method, the refined VO successfully generated intuitive maps in top view, thus allowing to count grating lattices as prior knowledge. Furthermore, the effect of the compensation method was confirmed by comparing the uncompensated map with the compensated one, which turned out to be much smoother and more consistent with the ground truth, by improving the regular grating rate from 62% to 90%.

Chapter.4 Proposal of a Light-weighted and Flexible Wire Drive Manipulator

As has been introduced in Section 1.3, for the future Unit 2 reactor investigation, the conventional tough manipulator “Veolia-MHI Arm” had met a lot of problems, including heavy weight, installing actuators on joints, hard to retrieve from joint failure, and vulnerable against collisions. These problems limited the feasibility of using such tough manipulator in unknown environment investigation taking account of possible unexpected collisions. In order to help solve these problems, this chapter proposes a light-weighted and flexible wire drive manipulator, as a preliminary exploration of using flexible manipulator in unknown environment investigation. Therefore, the exploration will also follow the flow of system introduction and existing problems discussion. This chapter consists of 5 sections. The discussion will start with preliminary introduction of the concerned mechanism and the hardware prototype in Section 4.1, and introduction of the remote controlling system in Section 4.2. Then, the common existing problem of a flexible manipulator in terms of motion controllability degradation will be discussed in Section 4.3.

4.1 BWD Mechanism and BWD Robot

Bundled Wire Drive mechanism, is a variation of wire-drive mechanism from tendon-drive mechanism. Compared with the conventional transmission mechanism, tendon-drive mechanism can provide higher freedom in terms of deploying mechanical component at distances, like the tendon-muscle relationship in human hands. The typical rope material is steel due to its excellent strength in various situations, and the mechanism with steel wire is also called wire-drive mechanism. However, the wire made of steel is usually quite heavy and lacking flexibility, which limited the feasibility of designing tendon-drive robot with steel rope transmission mechanism. While, the development of synthetic fiber rope gave rise to new possibility of developing flexible and light-weighted long-reach arm with tendon-drive mechanism.

Currently, the new synthetic fiber rope has achieved 1.5~2 times tensile strength compared with the conventional steel wire, while keeping 1/5~1/8 density. Furthermore, synthetic fiber rope showed excellent performance in softness and flexibility. In 2016, our research group studied the strength degradation feature of synthetic fiber rope, and proposed a practical design guidelines of a mechanical drive system with synthetic fiber ropes [15]. The results showed that the synthetic fiber rope would have similar degree of strength degradation with the conventional steel wire, but much more flexible and light-weighted at the same time. Compared with the conventional drive mechanism with wire made of steel due to its excellent strength, new synthetic fiber ropes have exceeded the strength of steel, besides, light-weighted and flexible at the same time. Furthermore, in 2018, the bending durability of synthetic fiber rope was also studied based on the bending durability evaluation method in Japan Industrial Standards [16]. Even in the case of applying quite severe bending conditions, the synthetic fiber rope made of ultra-high molecular weight

polyethylene showed decent strength and durability, compared with the other synthetic fiber rope and steel wires. However, the average friction coefficients of fiber ropes are around 1/3~1/5 of the steel wire, which made it hard to continue to use the conventional fixation method to fiber ropes. In 2018, a terminal fixation method using a grooved pulley and a figure-eight knot was proposed, as the basic study on the fixation of fiber rope [18]. On the basis of these researches as guidelines, development of long-reach articulated arm with synthetic fiber rope have been progressing rapidly. Multiple variations of wire drive mechanism using synthetic fiber ropes have been proposed, including BWD mechanism.

BWD mechanism, allowed to use multiple synthetic fiber ropes to drive multiple DoF, while the multiple ropes are bundled in the same transmitting rails. A similar concept of using bundled wires for manipulator driving can be found in the works of a remotely-operated surgery manipulator proposed by Ikuta in 2002 [19]. Figure 4.1 shows the concept of BWD mechanism, where the lines in different colors indicate the bundled ropes sharing the same pathway. Figure 4.2 reveals the joint of the surgery manipulator using decoupled wire drive mechanism.

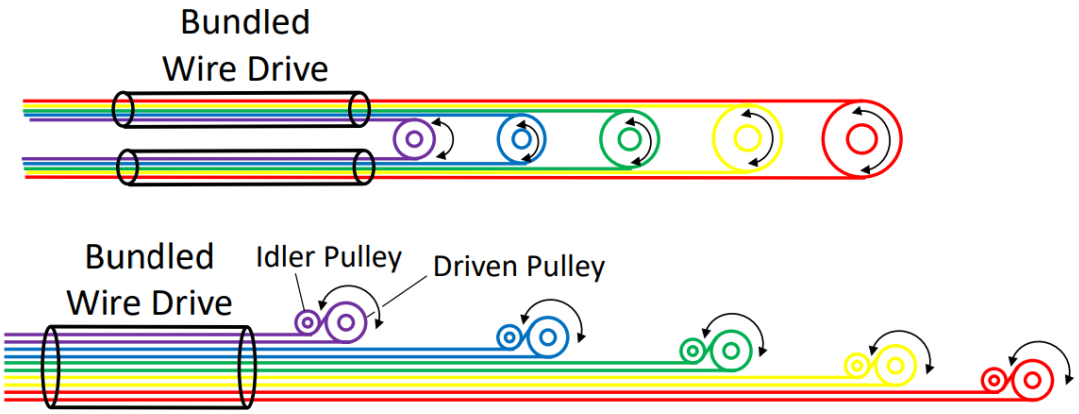


Figure 4.1 Concept of BWD mechanism. [18]

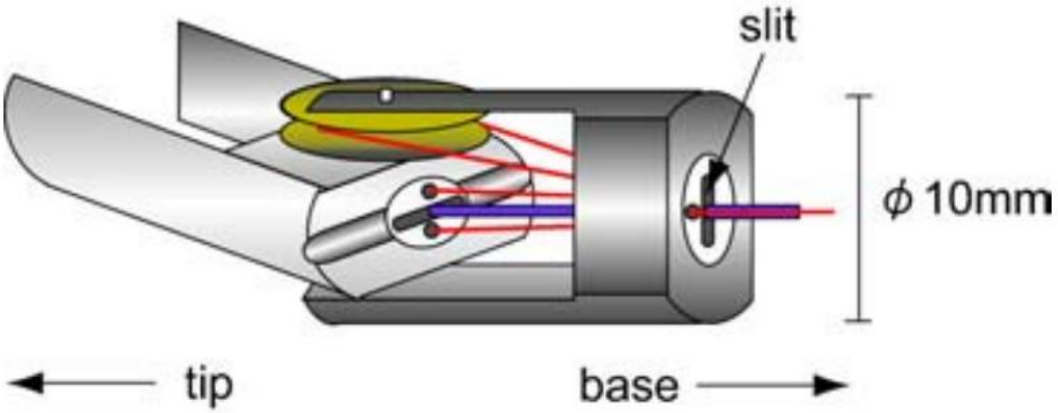


Figure 4.2 Structure of a remotely-operated surgery manipulator joint with decoupled wire driving mechanism [19].

In order to allow the bundled ropes to slide along each other, a relay mechanism was proposed by our research group in 2018, as Figure 4.3 shows. The design of the relay mechanism can be traced back to the human hand structure, which also has multiple bundled tendons to drive each finger. The left part of Figure 4.3 shows the relay mechanism structure with an offset in the axial direction. The right part of Figure 4.3 shows the top view of the relay mechanism, so the relayed wires pass through the same pathway and slide to each other.

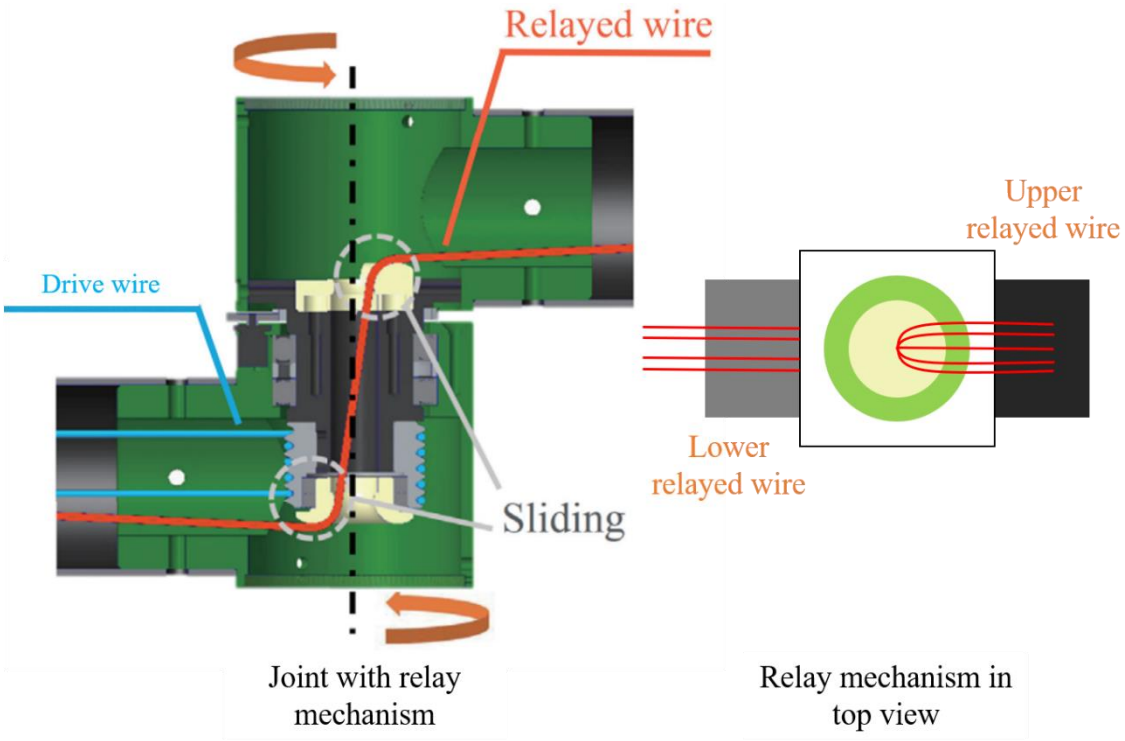


Figure 4.3 The relay mechanism which allows axial sliding of ropes. [18]

In 2019, BWD 3 DoF robot was implemented as a prototype to study the feasibility of BWD mechanism in robot arm design. Figure 4.4 shows the arrangements of tendons, holes and arm pulleys, which allows all the actuators to be placed on the base. Although the current prototype has only 3 DoFs, it has the potential to implement more DoFs to 15 joints. Joint 1 is a simple rotatable joint actuated by a single motor through a timing belt and pulley; Joint 2 and Joint 3 are actuated by two motors through antagonistic tendons, while one motor winds up the tendon and the other motor feeds the tendon, forming an antagonistic relationship. Therefore, the drive wires of J2 will pass through J1 with the proposed relay mechanism; Similarly, the drive wires of J3 will pass through both J1 and J2. The implementation of wire drive mechanism allowed to transmit actuation power from the actuator to the driven joint located at a long distance, as all the motors are installed in the base structure, leaving quite a long distance with the corresponding joints. Then, the experimental results may serve a reference for the future development of long-

reach articulated arm with wire drive mechanism, in regard of the situation of the No.2 reactor situation.

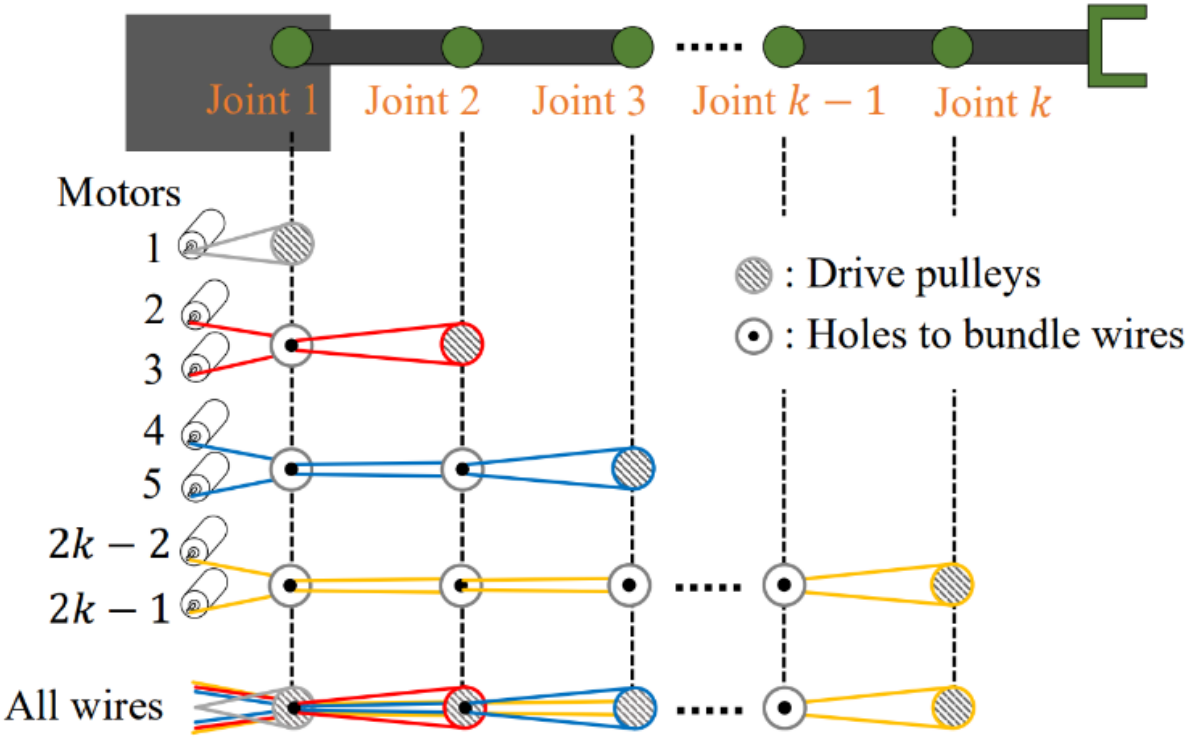


Figure 4.4 Arrangements of tendons, holes, and pulleys of the BWD arm.

As for the investigation sensor, a RealSense D435i camera was selected as the test sensing device, that consists of RGB sensor, stereo infrared sensors for depth estimation, and an IMU to assist the ego-localization. Especially, the integration of IMU may provide a side odometry information apart from the unreliable robot odometry. Table 4.1 shows the specification of RealSense D435i.

Table 4.1 Specification of RealSense D435i

Parameter	**Value**
Operation Range (min-max)	0.2m~10m
Depth Resolution and FPS	1280x720 30fps
Depth Field of View	H:87 V:58
Depth Technology	Active Stereoscopic
RGB Sensor	Yes
Inertial Motion Unit	Yes

Table 4.2 Specification of the prototype bundled wire drive arm. [18]

Metrics	Value
Length	1.8 m
Diameter	50 mm
Weight	33.3 kg (Arm:1.8 kg)
Number of Axis	3
Pose Repeatability	0.4 mm
Position Stabilization Time	0.99 s
Position Overshoot	4.45 mm
Minimum Resolution	0.059 deg

In this thesis, the prototype robot will be utilized as a preliminary experimental platform in regard of the possible problems of bundled wire drive mechanism corresponding to its compactness and light-weighted features. Therefore, the available DoF and the length of the prototype robot are relatively limited. Still, the BWD mechanism has the potential to implement with more DoFs and expand the rotation axis to other direction. In 2020, a 6 DoF articulated robot arm with BWD mechanism was developed, with mechanism for changing the rotation axis of joint to pitch or yaw [21].

4.2 Remote Controlling

In the case of controlling RhinoUS-II, ROS [34]has been introduced as the middleware framework to build the remote controlling system. Similarly, the remote controlling system of BWD robot was also developed under the ROS framework. Compared with the conventional commercial robot supported by the ROS community, BWD robot is an original robot with a unique transmission mechanism. Therefore, to utilize BWD robot into ROS frameworks, several customized control interfaces were implemented in this thesis, and the corresponding functions are listed below:

- Describe the original BWD mechanism in the ROS style.
- Simultaneously display the current robot pose.
- Provide controllers of the robot pose or joint pose.

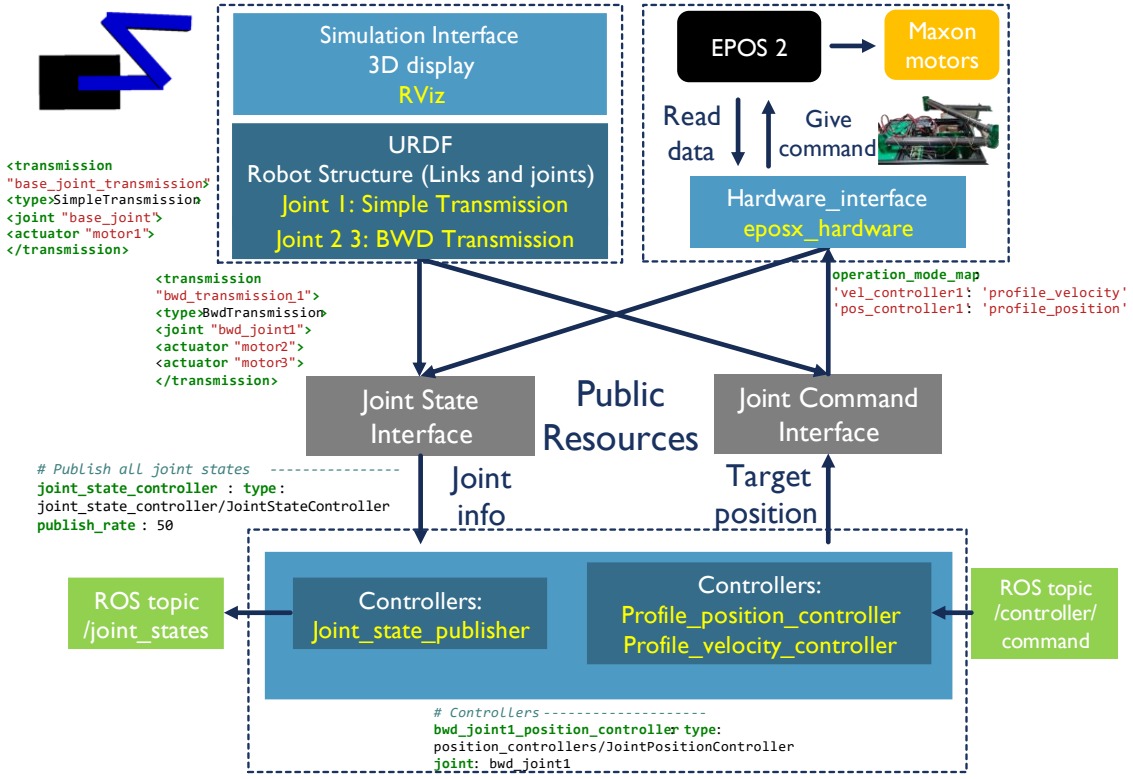


Figure 4.5 The structure of the general control interfaces for an articulated robot in ROS.

Figure 4.5 expands the basic structure of the universal control system for the articulated robot. Generally, it can be divided into three components: simulation interface, robot hardware interface, and controller manager. The development and details of the components are included in the Appendix B.

As Figure 4.5 shows, the joint command interface keeps the communication with the bottom controller EPOS2 through the controller hardware interface. Therefore, the encoders of the motors read the amount of wound rope, and the position, velocity and torque of the motors. Then, the readings of the motors will be transferred to the joint state interface and publish as joint state. When controlling the angle of each joint, the profile position/velocity modes of the EPOS2 controller are utilized, which are corresponding to the profile position/velocity controllers in controller manager. For the profile velocity controllers, the profile velocity, profile acceleration and deacceleration can be set. The profile position controllers tend to conduct a simple Point to Point control.

4.3 Source of Motion Controllability Degradations

The advantages of the BWD mechanism in terms of developing light-weighted arm have been discussed, while the problems of motion controllability degradation are also unavoidable.

Especially, because the conventional environment mapping method SfM could not fully utilize the robot's motion pattern as prior knowledge, it requires high image quality for the available features. Therefore, the compatibility between SfM and flexible manipulators was quite limited.

As has been discussed in the previous sections, the realization of compact joint design is based on the low friction coefficients of the new synthetic fiber rope, thus allowing the bundled wires to slide to each other. Since the joints are driven by ropes with the low friction coefficients, the stiffness of the joint is also relatively low due to the low antagonistic tension. On the one hand, the low joint stiffness makes the joint quite flexible, thus having good tolerance against collisions. On the other hand, the positional controllability will also be limited by the low antagonistic tension, leading to overshoot.

While, when increasing the antagonistic tension, the wires in the relay pulley will have bigger friction, thus leading to stationary deviation. Then, in the worst cases, the danger of wire breaking may occur when the wires are stuck by the friction. Therefore, the positional controllability degradation can be regarded as a problem of the low joint stiffness, as a common problem of tendon-driven mechanism. Especially, the BWD mechanism deploys multiple wires on the same pathway, while the bundled wires slide to each other. The interference between the ropes due to their twisting is also the possible source of controllability degradation.

4.4 Summary

In this chapter, the light-weighted long-reach manipulator using wire drive mechanism is proposed. The developed BWD prototype robot had 3 horizontal DoFs and be extremely light-weighted in the arm part, which also indicated the potential to develop longer arm with more DoFs. Compared with the conventional manipulator "Veolia-MHI Arm", the proposed manipulator is much lighter and more compact, easier to recover from joint failure, and stronger against collisions. Therefore, it may serve as a light-weighted option for the tough Veolia-MHI arm. While, as a flexible wire drive manipulator, it also has the common problem of motion controllability degradation, which will be the topic in next chapter.

Chapter.5 Utilization of Motion Patterns of a Flexible Wire Drive Manipulator for 3D Reconstruction

In the last chapter, the light-weighted and flexible manipulator BWD robot has been proposed to solve the existing problems of the conventional tough manipulator Veolia-MHI Arm. However, a new problem has occurred that the feasibility of using such a flexible wire drive manipulator to gather images for environment mapping process would be limited. The reasons of the feasibility limitation can be considered from two aspects: The first aspect is that the unavoidable motion controllability degradation of wire-drive mechanisms may damage quality of images captured with camera mounted on the flexible manipulator, which is also the common problem of such flexible manipulators. The second aspect is that the conventional environment mapping method SfM is highly relying on the available features from images, thus usually requiring images with high quality. These two aspects have given rise to the compatibility problem between SfM and flexible manipulators.

In order to solve this problem, this chapter will present a compatibility improvement method designed for SfM to improve the image quality based on the specific motion patterns of the flexible manipulator BWD robot as prior information.

5.1 Image Quality Degradation Problems from Motion Controllability Degradation

The previous experiments on the positional deviation evaluation have shown that the sources of the controllability degradation are low joint stiffness and interference between multiple joints. In terms of concrete effects, the interference will increase the positional deviation when multiple joints moved at the same time, and overshoot may occur when starting to move and approaching the ending point. Therefore, the image quality degradation due to wire-drive mechanism's motion controllability degradation needs to be studied and evaluated through experiment.

5.1.1 Positional Deviation

Robot Odometry is known as an intuitive and quick method to estimate the end effector trajectory based on the joint angles and forward kinematics. In the case of installing the investigation camera on the end effector, the problem of tracking the camera trajectory could be solved by referring to robot odometry. Especially when the investigation sensor cannot localize itself with its data. However, problems like link deflections and joint flexibility often occurred when it came to a long-reach articulated manipulator, which may lead to robot odometry deviation. Especially, in the case of BWD mechanism, the interference between each DoF may lead to robot odometry deviation, too.

In terms of SfM processes, the deviation usually does not influence the image registration processes since it does not rely on prior odometry input. In the case of SLAM processes, some of SLAM algorisms can accept external odometry information from the robot as an assisting

information. Besides, as has been indicated in last section, the usage of camera motion features as prior knowledge, like tendencies of robot odometry, or continuity will require relatively accurate robot odometry. Concretely speaking, since the image sequences are from the camera installed on end effector of robot, the trajectory should maintain the continuity feature, like the general motion camera model. However, the image registration processes may not maintain the continuity feature because the extracted features were weakly-localized.

As the conclusion, the existence of positional deviation will damage the reliability of robot's motion pattern as prior knowledge.

5.1.2 Potential Motion Blur

When discussing the specific feature in dataset which may influence the method's performance, the concept "Common Corruption" has been commonly used in aspect of neural network [49]. In the case of using image sequences from camera mounted on a flexible manipulator, Motion Blur is the most likely common corruption. Motion blur, usually occurred when the speed of camera is changing severely, while the quick speed change often happens when the robot overcomes the dead zone. For the BWD mechanism, the joints could not be actuated until the tension reached the required level to overcome the static friction, which usually formed a dead zone when the robot started moving. Besides, the previous experiment showed the increasing tendency of interfere between DoF when the rotation angle increased. Besides, due to the low joint stiffness, BWD robot usually needed more stabilization time to stop moving, leading to positional overshoot.

As for the influences of motion blur, it's natural that the motion blur in images will damage the general dataset image quality, like low available features. On the other hand, in the camera pose estimation process of SfM, it is highly relying on the available features. Usually, the image sequences from the recorded video are extracted through a fixed frame rate, but the frames with motion blur may not contain enough features or overlap for image registration. Then, a direct influence is the ratio of number of registered images over the total number should drop, leading to fewer triangulated points in the reconstruction.

After the sparse reconstruction, the dense reconstruction process with Multi View Stereo (MVS), fewer registered images will lead to fewer available images for reconstruction reference. When some blurry images with motion blur were registered, the inconsistent blur part may influence the dense pixelwise correspondence search processes. While, the typical MVS software colmap_rob introduced a multi-view geometric consistency term for depth/normal estimation, which has been proved to be quite strong in recognizing and refining the inconsistence [26].

On the other hand, the brand-new technique Novel View Synthesis (NVS) which uses a fully-connected deep network NeRF (Neural Radiance Fields for View Synthesis) has been gaining great popularity in 3D modeling. The successfully registered images will be utilized to train the

NeRF network, then the network will render the specific viewpoint in color and density based on the training results. Therefore, the existence of blurry images will form a corruption in the network training process, resulting in inconsistent noise in the rendered model. In the following section, the degrees of motion blur from the BWD robot will be experimentally studied.

5.1.3 Experimental Evaluation of Positional Deviation

To evaluate the positional deviation of the BWD robot, the robot arm is commanded to move to a series of positions in space, while the actual positions of the robot joints are measured, and compared to the desired position. To obtain the ground truth trajectory, an optical motion capture system Motion Analysis Kestrel mocap was used as the reference, which is able to achieve an accuracy within 1 mm in a 3- × 3- × 3-m cube space and a response speed within 3 ms. On each joint of the BWD robot, a specialized marker for motion capture was installed. Therefore, the motion capture system can recognize and track the positions of the markers, thus to form a skeleton of the robot arm. A L-frame was placed on the base of the BWD robot to define the coordinate system. Here, the forward direction of the arm is defined as positive Y-direction, while the horizontal direction is defined as positive X-direction.

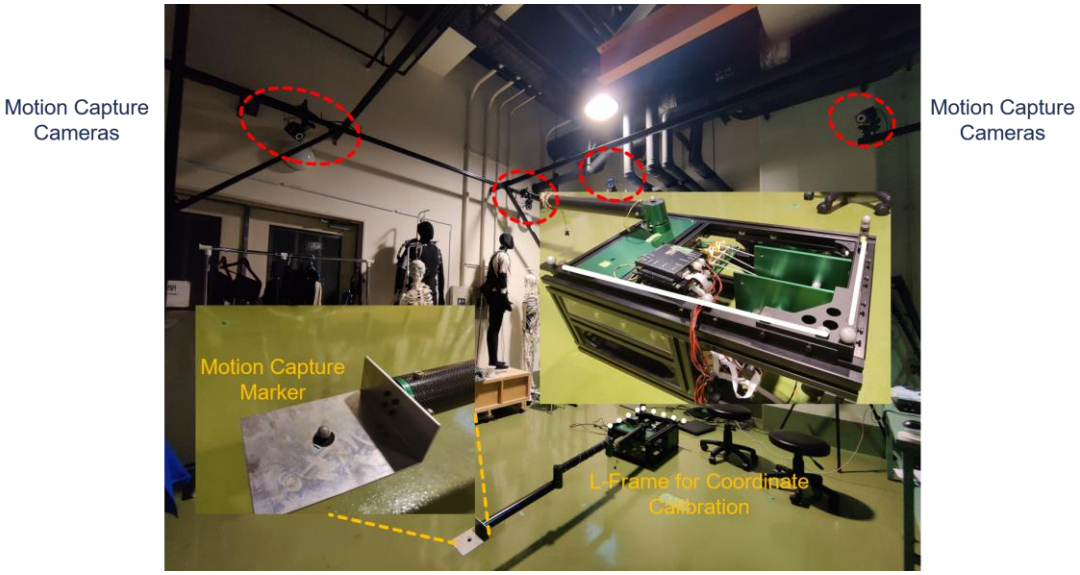


Figure 5.1 The environment setup of motion capture measurements.

Considering the possible interactions between bundled ropes, four motion patterns were employed in the experiment with an angular speed of 20 degree/s, listed below.

- Pattern 1: Pure rotation of J3 within its maximum range from −160° clockwise to +160°.
- Pattern 2: Rotating J2 to −90°, then rotating J3 from −140° clockwise to +140°.
- Pattern 3: Moving the end effector along the negative Y-direction for 600 mm, which is the length of a single link.
- Pattern 4: Similar to Pattern 3, a Cartesian translation along the positive X-direction for 1200 mm.

Obviously, the patterns 1 and 2 only moved one joint at the same time. While, the pattern 2 rotated the J2 first, and kept the wires of J2 stretched, and rotated the J3 again. Therefore, as an internal comparison, the difference between Pattern 1 and 2 can be assumed as the influences of wires of J2. While, the patterns 3 and 4 moved multiple joints at the same time, as an inverse kinematics movement. Therefore, the influences of the interference between multiple joints can be studied by inspecting the results of pattern 3 and 4.

The comparison between the desired trajectory and the measured trajectory is shown in Figure 5.2 and Figure 5.3, while the left part of these figures is the defined coordinate system, the direction from the base to end effector J3 is defined as the positive Y-direction. The robot arm (solid line) indicates the initial pose of the motion pattern. The dotted line indicates the final pose.

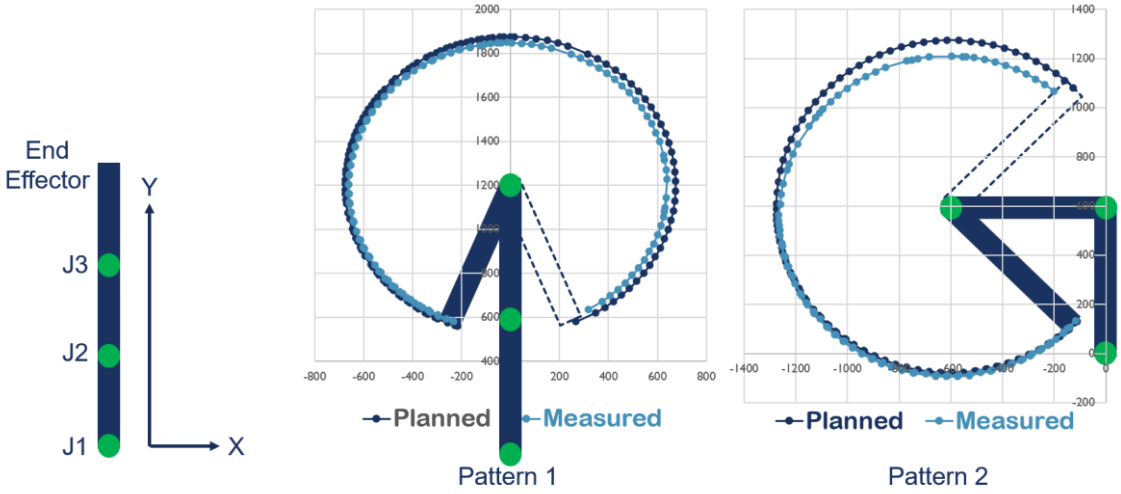


Figure 5.2 End effector trajectory comparison results of patterns 1 and 2 (deep blue: planned trajectory by robot odometry; light blue: measured trajectory by motion capture)

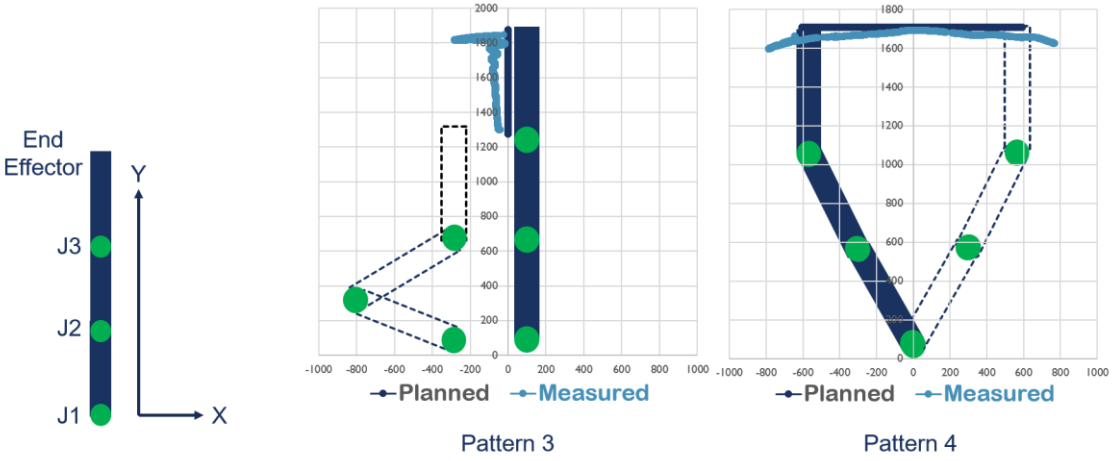


Figure 5.3 End effector trajectory comparison results of patterns 3 and 4 (deep blue: planned trajectory by robot odometry; light blue: measured trajectory by motion capture)

Table 5.1 Average positional deviation of robot odometry in each pattern

Pattern	Average Positional Deviation
1	45.1 mm
2	57.6 mm
3	76.8 mm
4	101.1 mm

The results of the comparison of end effector trajectories are shown in Table 5.1. As Figure 5.2 and Figure 5.3 show, patterns 1 and 2 only include the pure rotation of a single joint, and the comparison show that there was a steady rise in the deviation with joint rotation. This indicates that the deviation may originate from low joint stiffness. In addition, pattern 2 was conducted to keep the ropes of J2 stretched, which introduced the static interference and led to more deviation. The figure shows the phenomenal growth of deviation in patterns 3 and 4, where Cartesian translations were conducted with multiple joints, thus introducing dynamic interference between multiple joints. In particular, the deviations were mainly distributed in the beginning phase (pattern 3) or formed an overshoot in the stopping phase (pattern 4). The concentrated distribution of deviations suggests that they may result from overcoming the static friction in the beginning phase or overshooting owing to the relatively long stabilization time.

5.1.4 Experimental Evaluation of Motion Blur

In order to study the degrees of motion blur from BWD robot, it's necessary to define the degrees of motion blur quantitively.

The experimental evaluation of motion blur was conducted on the basis of a benchmark D proposed by Su et al in 2011 [50]. Su proposed a blur detection method using Singular Value Decomposition (SVD), which is also region available. The benchmark of D is defined as below:

$$D = k\beta + (1-k)\left(\frac{\Omega_b}{\Omega}\right) \tag{5.1}$$

Where k is the weight to the concerned two aspects. One of the aspects, β indicated singular value feature of the image, which may increase when the relative camera speed increased; The other one was the ratio of the blur region area (Ω_b) over the whole image area (Ω). The basic work principle of the SVD decomposition of the images can be referred in Appendix C.

As the experimental method, the images concerned are borrowed from the images captured by equipped camera in the experiment of Section 5.1.3. The average angular speed of the robot arm was 20 degree/s, and the pattern 1 rotated the J3 from -160 degree to +160 degree. The frame rate was 30 frames per second, then down-sampled to 10 fps for the experiment. Therefore, the number of images in this pattern was 186 frames. Figure 5.4 shows the estimated results of D in the image sequence dataset of pattern 1.

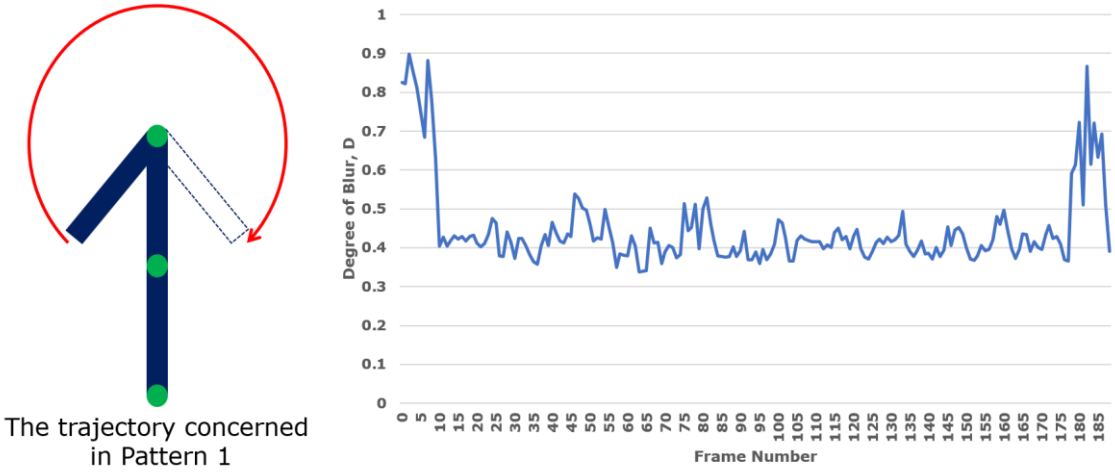


Figure 5.4 The estimated degree of blur in Pattern 1.

As Figure 5.4 shows, the average degree of blur was around 0.4 when the angular speed of the camera was set 20 degree per second. While, in the beginning 10 frames and the ending 10 frames, obvious peaks of the degree of blur can be observed, which is also consistent with the previous experiment results that the overshoot may easily occur due to low joint stiffness. The rest patterns had the similar results and the average period of the overshoot was around 1s, close to the average stabilization time of BWD robot. Besides, in the Pattern 3 and 4 where there was multiple joints movement concerned, the peak values of the observed motion blur and the maximum length of the overshoot period were slightly bigger than those in Pattern 1 and 2.

5.2 Proposal of Improvement Method: A Dataset Creation Method for Flexible Wire Drive Manipulator

The experimental results of Section 5.1.3 and Section 5.1.4 have revealed that the motion controllability degradation problem from flexible manipulator will introduce two main problems: Positional deviation, and motion blur. Especially, the positional deviation between the actual camera poses and the poses calculated from forward kinematics will damage the corresponding relationship between the captured images and the specific motion patterns, thus reducing the reliability motion patterns as prior knowledge. And, the existence of blurry images will reduce the image quality of the dataset, thus introducing inconsistent factors to the environment mapping. Therefore, in order to improve the reliability of motion patterns of flexible wire-drive manipulator as prior knowledge, influences from positional deviation should be avoided. Besides, to improve the available image quality for further SfM method, influences from motion blur should be mitigated, too.

Therefore, this section proposes a dataset creation method, which tries to mitigate the existence of positional deviation and motion blur, by designing overlapping trajectories and selecting images with low blurry degree, on the basis of the flexible manipulator's motion patterns as prior knowledge.

5.2.1 Motion Patterns Features of the Flexible Wire Drive Manipulator

In order to utilize the specific motion patterns of the Flexible Wire Drive manipulator, it's necessary to clarify the patterns first. The previous experiments utilized with 4 motion patterns, including two rotation trajectories of a single joint, and two translation trajectories with multiple joints' simultaneous movement.

According to the evaluation results of position deviation of these four patterns, there was a phenomenal growth of positional deviation when conducting Cartesian translations with multiple joints, thus introducing dynamic interference between multiple joints (Table 5.1). In particular, the deviations were mainly distributed in the beginning phase or formed an overshoot in the stopping phase. The concentrated distribution of deviations suggests that they may result from overcoming the static friction in the beginning phase or overshooting owing to the relatively long stabilization time. On the other hand, the motion blur evaluation results also showed the similar tendency that the beginning phase and stopping phase tend to have more motion blur (Figure 5.4). Therefore, a preliminary conclusion of the flexible wire drive manipulator motion patterns can be made:

- The translation trajectory with multiple joints' simultaneous movement will introduce more positional deviation, which may result from the interference between multiple joints. While, the rotation trajectory with single joint movement tends to have lower positional deviation.
- In both rotation and translation trajectories, the blurry images tend to concentrate in the beginning phase and the stopping phase of a motion loop, which may result from overcoming static friction and low joint stiffness.

5.2.2 Dataset Creation Method Step 1: Trajectory Design

The first step of the proposed dataset creation method is the trajectory design step which aims to avoid trajectories with high positional deviation. For an investigation aimed to obtain a reconstruction of the environment or a specific object, taking account of the intuitive representation, the textured 3D point cloud from various perspectives and viewpoints is most suitable. Therefore, with regard to the trajectory of the sensor, it should also contain images from as rich as perspectives and viewpoints. Figure 5.5 shows a typical ideal viewpoints distribution for a specific object reconstruction, in other words, rotation around the target object is preferred.

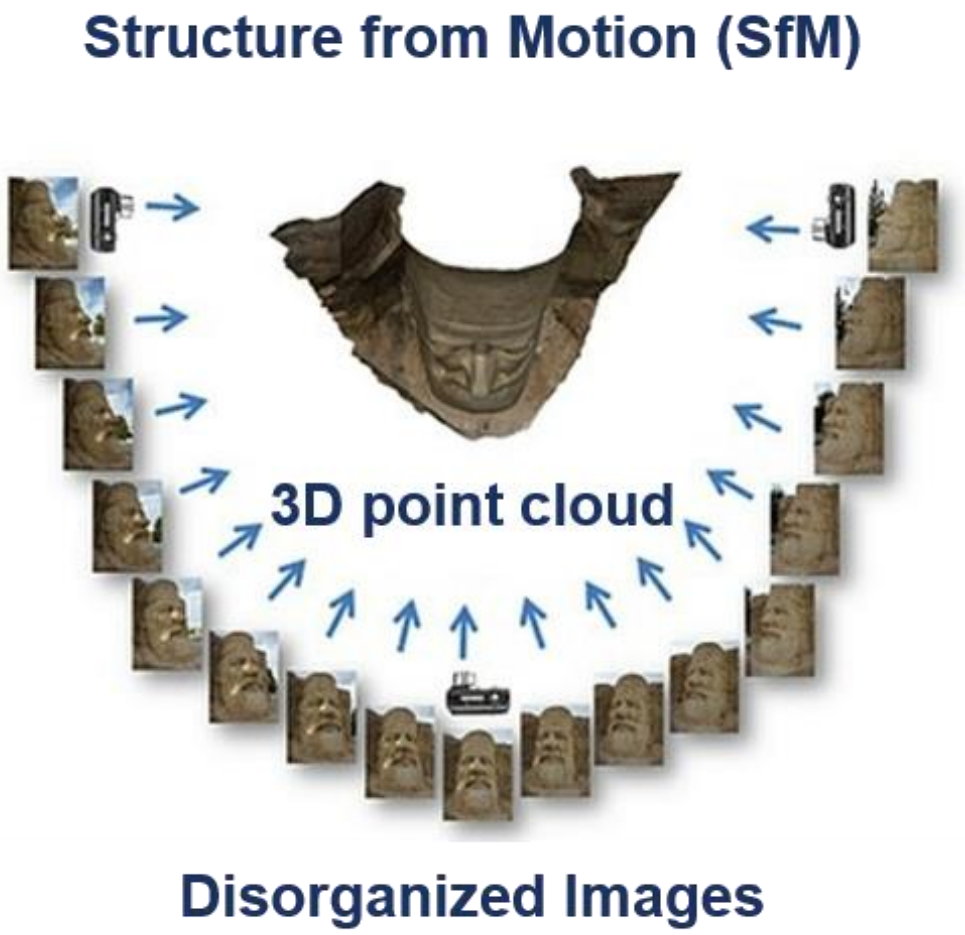


Figure 5.5 The ideal viewpoints for a specific object in reconstruction.

However, for the surrounding environment with boundaries, the total surrounding viewpoints will be hard to achieve. By using the feature of flexible manipulator motion pattern in Section 5.2.1, when creating an image dataset, the total trajectory should avoid using trajectories with multiple joints' movement. However, using only pure rotation trajectories may result in a limited range of motion and may be time-consuming due to the need of individually moving each joint. A natural compromising way of designing trajectory will insert rapid translation trajectory to adjust the viewpoint position, which can be denoted as a transition trajectory, as Figure 5.6 shows. Black camera icons represent the camera poses and colorful rings indicate visible ranges of the corresponding trajectories. For trajectories #1 and #2, only pure rotation of a single joint is used, resulting in relatively good image quality with low positional deviation. However, the transition movement (orange) uses rapid translation through matched joint movement to adjust the viewpoint position, which may introduce more deviation. Directly removing the images captured during the transition movement, however, would disrupt the continuity of the trajectory and result in partial reconstruction.

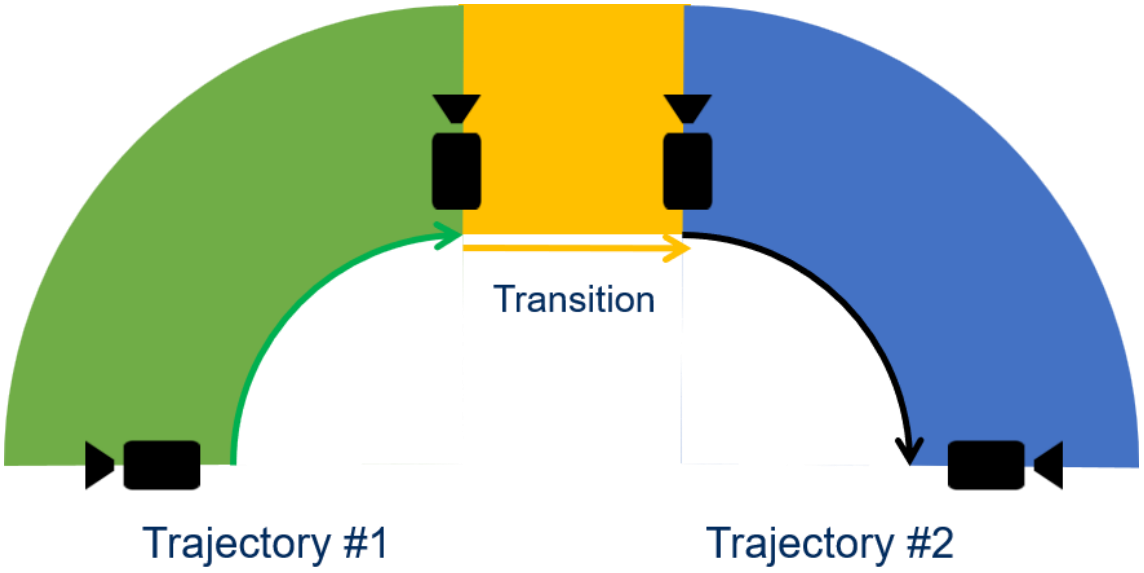


Figure 5.6 The concept of conventional trajectory design with rapid transition trajectory. where the images of transition are necessary for continuity. Green: Trajectory #1; Orange: Transition; Blue: Trajectory #2.

Therefore, in order to discard images captured during the transition trajectory while maintaining a continuous trajectory with a variety of viewpoints, the trajectory design step creates overlapping pure rotation trajectories by orienting the beginning pose of each subsequent trajectory towards the beginning pose the previous trajectory. The concept of the trajectory design step is shown in Figure 5.7.

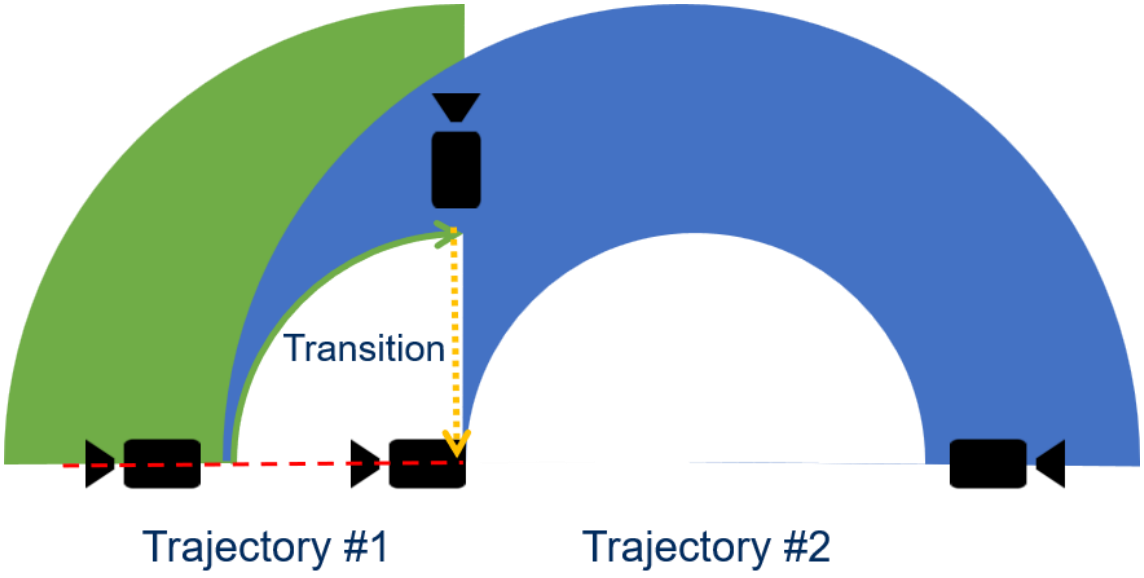


Figure 5.7 The concept of the trajectory design step. Green: Trajectory #1; Orange: Transition; Blue: Trajectory #2.

As the result, the trajectory design step makes it feasible for the dataset to discard images from matched joints' movement with high positional deviation while maintaining a continuous viewpoint trajectory for correspondence searching processes of SfM. Therefore, the high positional deviation from matched joints' movement can be skillfully avoided without obvious negative influences on reconstruction processes.

5.2.3 Dataset Creation Method Step 2: Images Selection

The second step of the proposed dataset creation method, is to select the images with high image quality while removing images influenced by high positional deviation and motion blur. Thanks to the trajectory design step, the total dataset will consist of overlapping pure rotation trajectories with low positional deviation, and rapid transition trajectories with high positional deviation. Therefore, the image selection step will first cluster the dataset according to overlapping pure rotation trajectories and rapid transition trajectories, and remove images from transition trajectories.

Therefore, as the preliminary effect, the trajectories with high positional deviation can be removed. To mitigate the influences of the motion blur, the image selection step will then utilize the concluded motion pattern feature in Section 5.2.1, which indicate that highly blurry periods usually concentrate in the beginning phase and the stopping phase of a motion loop. Therefore, the image selection step will use the quantity D to detect the blurry degree of the remaining trajectories, especially the beginning period and stopping period. The preset period length is set as the average stabilization time 1s. Then, the images with high blurry degrees will be removed, the remaining images form the created dataset.

Eventually, after the two steps of the proposed dataset creation method, the positional deviation and motion blur from wire drive motion controllability should have been removed from remaining images. So, the image quality of the dataset can be enhanced, when used for the subsequent environment mapping process, the compatibility between such flexible manipulator and SfM should be improved.

5.3 Experimental Verification of the Proposed Improvement Method with SfM and MVS/NVS

In this section, the effectiveness of the proposed dataset creation method will be studied through an experiment with SfM and MVS/NVS methods. The proposed dataset creation method is designed to improve image quality of created dataset by mitigating influences of the motion controllability degradation according to BWD robot's motion pattern features. Therefore, to confirm its effectiveness, a dataset created without the proposed dataset creation method will be prepared as the control group, and a dataset created with the proposed method will be the

experimental group. The environment mapping performance of these two datasets will be compared to confirm the effectiveness.

In this Section, Section 5.3.1 presents the utilized benchmarks to evaluate the performance of the vision-based mapping methods, including accuracy, completeness and runtime; Then, Section 5.3.2 shows the setups of the environment and the concerned trajectories. The effectiveness of the dataset creation method will be confirmed with the following two sections with 2 steps of the proposed dataset creation method (Section 5.3.3 for trajectory design, and Section 5.3.4 for image selection). Section 5.3.5 discusses the existing robustness problem against more realistic and harsher environment through experiments with lower illumination.

5.3.1 Benchmarks for Vision-based Mapping Methods Performance

When evaluating the environment reconstruction performance of mapping methods, the commonly acknowledged benchmarks of accuracy, completeness, and runtime proposed by Schops et al. in 2017 were utilized [51]. They are based on a range of distance thresholds that work with the distance between the ground truth 3D points and their closest reconstruction points. In this study, a tolerance threshold of 20 mm we utilized, so the points within 20 mm of the ground truth points are defined as "accurate points." The number of accurate points is N_{acc}, and the number of points of the reconstructed cloud is defined as N_{rec}. The accuracy benchmark, Acc, is defined as the number fraction using Equation 5.2:

$$Acc = \frac{N_{acc}}{N_{rec}} \tag{5.2}$$

For each point in the ground truth, if the distance between itself and its closest point from the reconstructed result is within the distance threshold, then this point is defined as "completed." The number of completed points is N_{com}, and the number of points of the ground truth cloud is defined as N_{truth}. The completeness benchmark, Com, is defined as the number fraction using Equation 5.3:

$$Com = \frac{N_{com}}{N_{truth}} \tag{5.3}$$

The concept of these two benchmarks is shown in Figure 5.8.

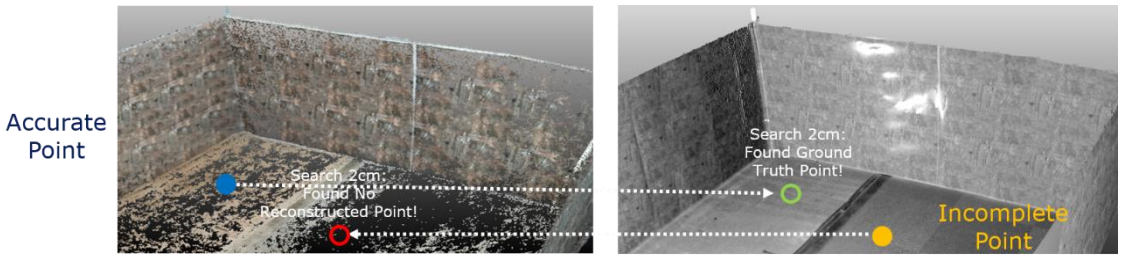


Figure 5.8 The concept of Accurate and Completeness. Left: the reconstructed point cloud; Right: the ground truth by laser scanner.

5.3.2 Environment Setup and Designed Trajectory

In the experimental method, the SfM pipeline of colmap was used to acquire the estimated camera parameters and trajectories. Then, processing was performed with colmap_rob as the MVS method and instant-ngp as the NVS method. The reconstructed point clouds were evaluated with the benchmarks explained above through a comparison with the ground truth provided by an external laser scanner. A FARO Focus 3Dx130 scanner was chosen. Its maximum distance reaches 130 m, and the accuracy is within 2 mm. A simplified environmental setup was used to eliminate unnecessary factors.

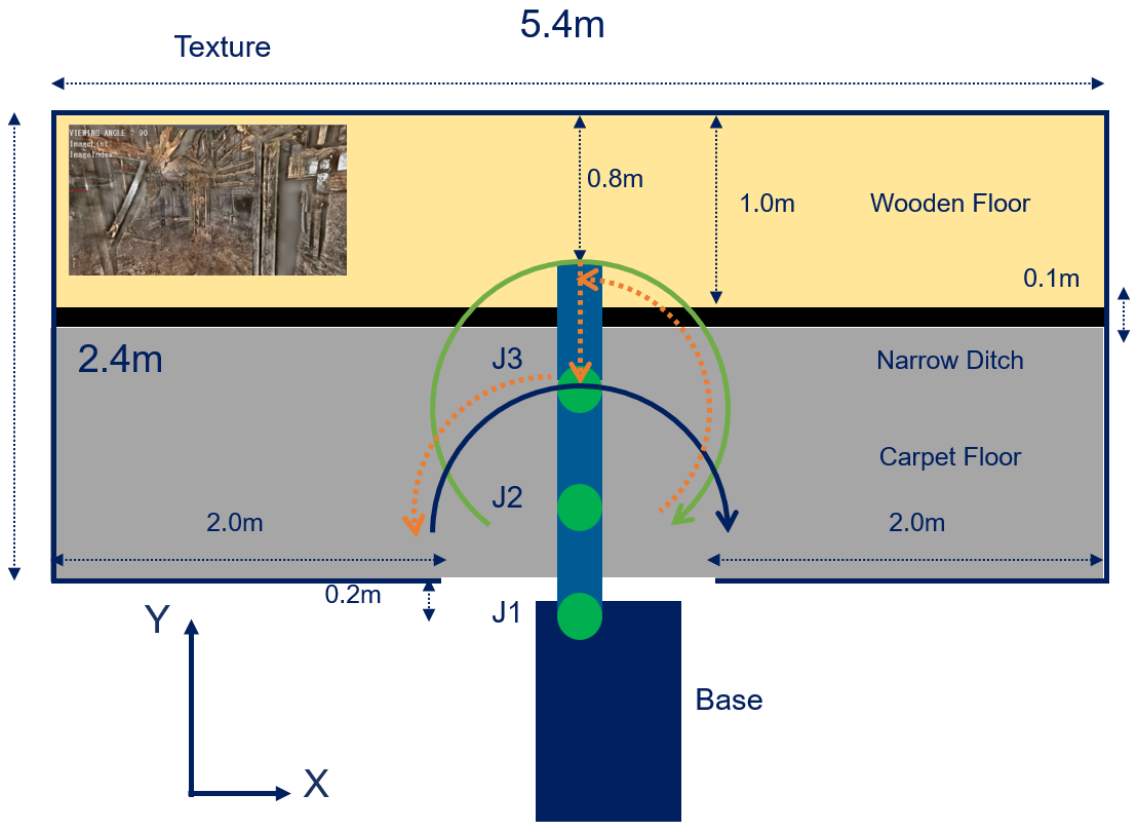


Figure 5.9 The environment setup for the reconstruction experiment. Green: Trajectory #1; Orange: Transition; Blue: Trajectory #2.

As Figure 5.9 shows, the environment was a rectangular bounding box of 5.4 × 2.4 × 1 m, and the robot arm entered the box through a preserved entry at its bottom center. The upper part of this rectangular box was a wooden floor with a width of 1.0 m, and the rest was covered with a gray carpet. Between these two floors, there was a narrow ditch with a width of 0.1 m. When the robot arm was placed, an 0.2-m offset was intentionally added along the negative Y-axis, maintaining a 0.8-m distance between the end effector and wall. On all the boundary walls of the environment, printed photographs from previous investigations conducted in the No. 2 reactor were attached [52]. These photographs provided the necessary feature points for the SfM

processes and also concentrated on the reconstructed points with regular shapes. Considering the preliminary results on the robot odometry deviation, a trajectory composed of pure rotation and Cartesian translation was intentionally used. The trajectory could be simplified into three phases, including a pure rotation around J3 from −160° to +160°, a Cartesian translation along the −Y-direction of 600 mm, and another pure rotation around J3 from −90° to +90°. This setup was defined as setup pattern 1.

Considering the preliminary results on the robot odometry deviation, the trajectory was designed by following the proposed dataset creation method, which composed of pure rotations and rapid Cartesian translation as transition. As Figure 5.9 shows, the end point trajectory can be divided into trajectory #1 (green line), a pure rotation around J3 from -160 deg to +160 deg; three rapid transition movements (orange line), including J3 rotating back to center, a cartesian translation along -Y direction for 600mm, and J3 rotation to adjust the starting pose of subsequent trajectory; At last, the trajectory #2 with J3 rotating from -90 to +90 deg. The parameters, including angular velocity, are presented in Table 3.

Based on setup pattern 1, two setup patterns were added by placing a box and cylinder in the environment to increase the spatial complexity. The size of the box was 0.6 × 0.6 × 0.6 m, and that of cylinder was 150 mm (radius) × 1000 mm. For pattern 2, the box and cylinder were placed close to the far left and right corners of the environment to modify the boundary shape slightly. In pattern 3, they were placed at specific positions and orientations to evaluate the object localization accuracy. When the origin was set as the center of the entry, the center position of the box was (−2.0 m, 1.7 m) and rotated along the Z-axis clockwise by 45°. The position of the cylinder was (1.5 m, 1.2 m). The three setup patterns are shown in Figure 5.10.

In each environmental setup, the planned trajectory described above was used while capturing videos with a RealSense D435i camera at a 1280 × 720 resolution. Then, images were extracted at 2 fps and used to conduct SfM processes with colmap, and the subsequent reconstruction processes were performed with colmap_rob and instant-ngp. The run configurations are presented in Table 5.2.

Table 5.2 Run configuration of reconstruction processes

Component	Configuration
CPU	Intel Core i7-9700K @ 3.60 GHz
Memory	32 GB
Graphic Port	NVIDIA GeForce RTX 3070
CUDA	11.6
colmap	3.8
OptiX	7.3

Vulkan	Yes

Figure 5.10 Three setup patterns for the environment.

5.3.3 Experimental Verification of the Trajectory Design Step

In the experiment setup, the image dataset has been created according to the proposed dataset creation method. While, the original dataset still includes images of poor quality, which consists of rapid transition trajectories and periods with high blurry images.

On the other hand, the proposed image selection method involves two steps: the removal of images taken during transition movements, and the filtering out of blurry images. As such, the following experiment will compare the original dataset and the dataset which removes images from rapid transition trajectories with high positional deviation. Then, the effect of the trajectory design step in the proposed dataset creation method can be confirmed by comparing the reconstruction performance of these two datasets. For convenience, these two datasets are denoted below and shown with Figure 5.11.

- Original Dataset: images from all trajectories. (Including high positional deviation)
- Transition Removed Dataset: images from only overlapping rotation trajectories. (Relatively low positional deviation)

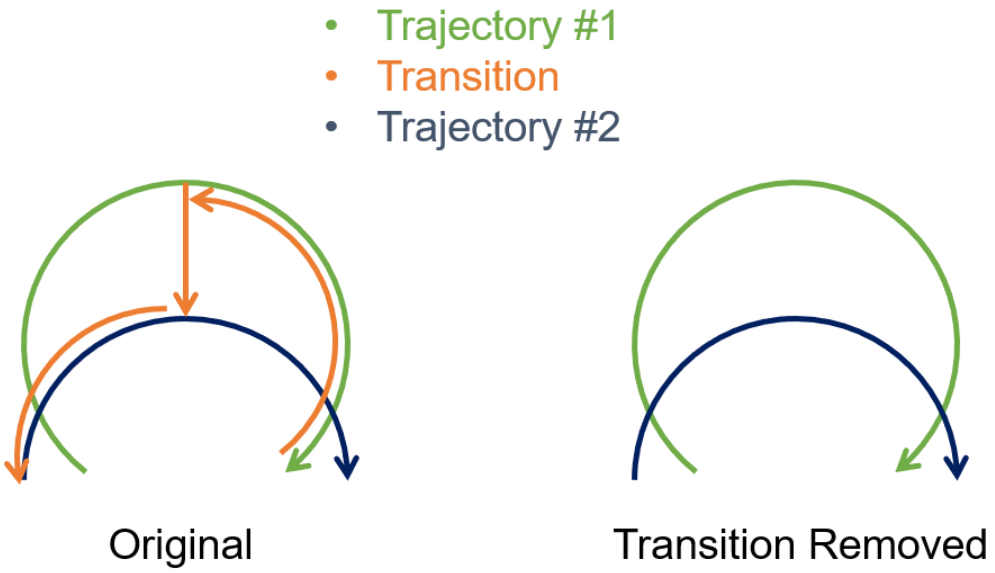


Figure 5.11 The difference between the two datasets in this experiment.

Then, the reconstruction performance of the mapping methods with these two datasets will be compared, to confirm the influences of the trajectory design step. Table 5.3 and Table 5.4 presented the reconstruction performance of colmap_rob and instant-ngp with different trajectory dataset input. In this section, the target is to compare the reconstruction performance before and after applying the proposed trajectory design strategy. Therefore, the cross comparison on the performance of colmap_rob and instant-ngp will not be conducted. As Table 5.3 shows, removing the images captured during transition results in a significant reduction in the number of input images and a corresponding reduction in average runtime of around 17.6%. While, although the images from the phase 2 got removed, the general accuracy and completeness did not reduce, instead, slightly increased.

Table 5.3 Reconstruction results of all three setup patterns by colmap_rob before and after removing the transition images.

colmap_rob	Metric	Original	Transition Removed
Setup Pattern 1	Accuracy (%)	86.7	88.0
	Completeness (%)	60.8	62.1
	Runtime (s)	8062	6213
Setup Pattern 2	Accuracy (%)	90.4	88.2
	Completeness (%)	68.0	67.7
	Runtime (s)	7264	6596
Setup Pattern 3	Accuracy (%)	81.7	85.2
	Completeness (%)	59.7	59.2
	Runtime (s)	7445	5999

As Table 5.4 shows, similarly, the accuracy got improved after removing the transition images, which indicated that the higher positional deviation brought more influence to the processes. Besides, the completeness didn't have large change, which indicated that the Phase 1 and Phase 3 had successfully covered the visible areas of phase 2. While, although the input number of images reduced, the runtime of instant-ngp didn't have obvious difference.

Table 5.4 Reconstruction results of all three setup patterns by instant-ngp before and after removing the transition images.

instant-ngp	Metric	Original	Transition Removed
Setup Pattern 1	Accuracy (%)	61.3	67.4
	Completeness (%)	94.5	94.1
	Runtime (s)	7.6	6.2
Setup Pattern 2	Accuracy (%)	71.4	75.6

	Completeness (%)	89.1	87.2
	Runtime (s)	6.3	6.3
Setup Pattern 3	Accuracy (%)	62.5	69.9
	Completeness (%)	92.1	91.2
	Runtime (s)	5.1	5.0

5.3.4 Experimental Verification of the Image Selection Step

Furthermore, in order to study the effect of the image selection step in the proposed dataset creation method, this section will conduct the comparison between transition removed dataset and blur filtered dataset, and the definition is listed below:

- Transition Removed Dataset: images from only overlapping rotation trajectories. (Relatively low positional deviation, but with blurry periods)
- Blur Filtered Dataset: the final dataset created by the proposed method which also removed the highly blurry images.

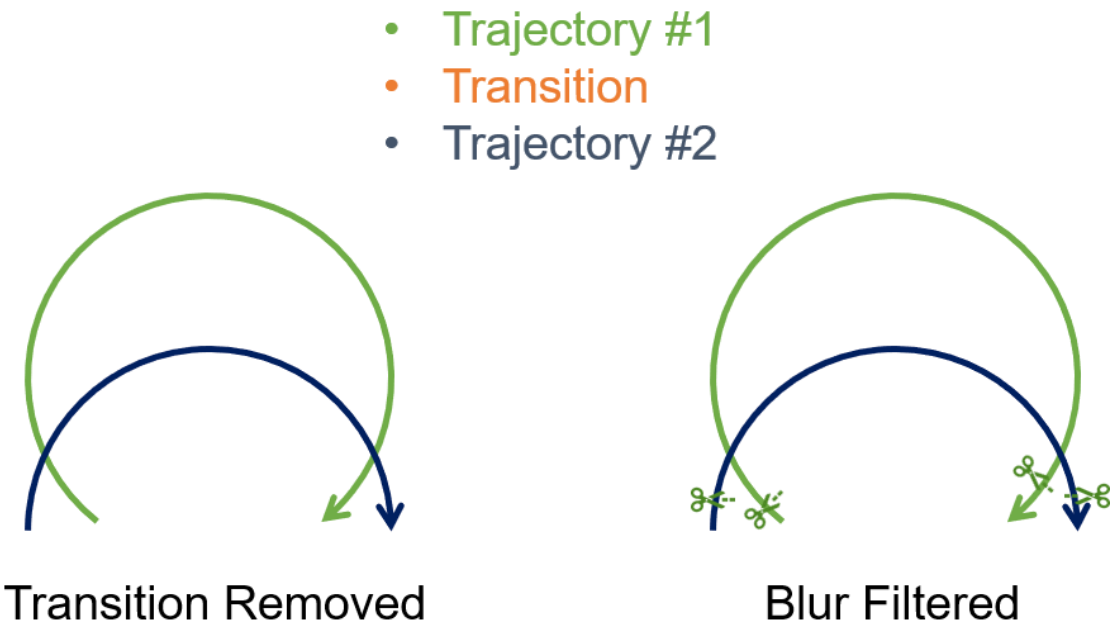


Figure 5.12 The concept of the two datasets for image selection step experiment.

Figure 5.13 presents a comparison of the reconstructed point clouds produced by colmap_rob under three different setup patterns, with the left clouds based on the transition removed dataset and the right clouds based on the blur-filtered dataset. There is no notable difference between the reconstruction results. In contrast, when using instant-ngp for reconstruction, the presence of motion blur results in fuzzy noise in the reconstructed results. The proposed image selection method is effective in reducing this noise, particularly around key objects in the environment setup. Figure 5.14 and Figure 5.15 show the comparison of reconstruction results before and after

conducting the blur filtering, both in terms of an overview of the point cloud and close-up views of key objects in the setup. The corresponding numerical differences in reconstruction performance are presented in Table 5.6.

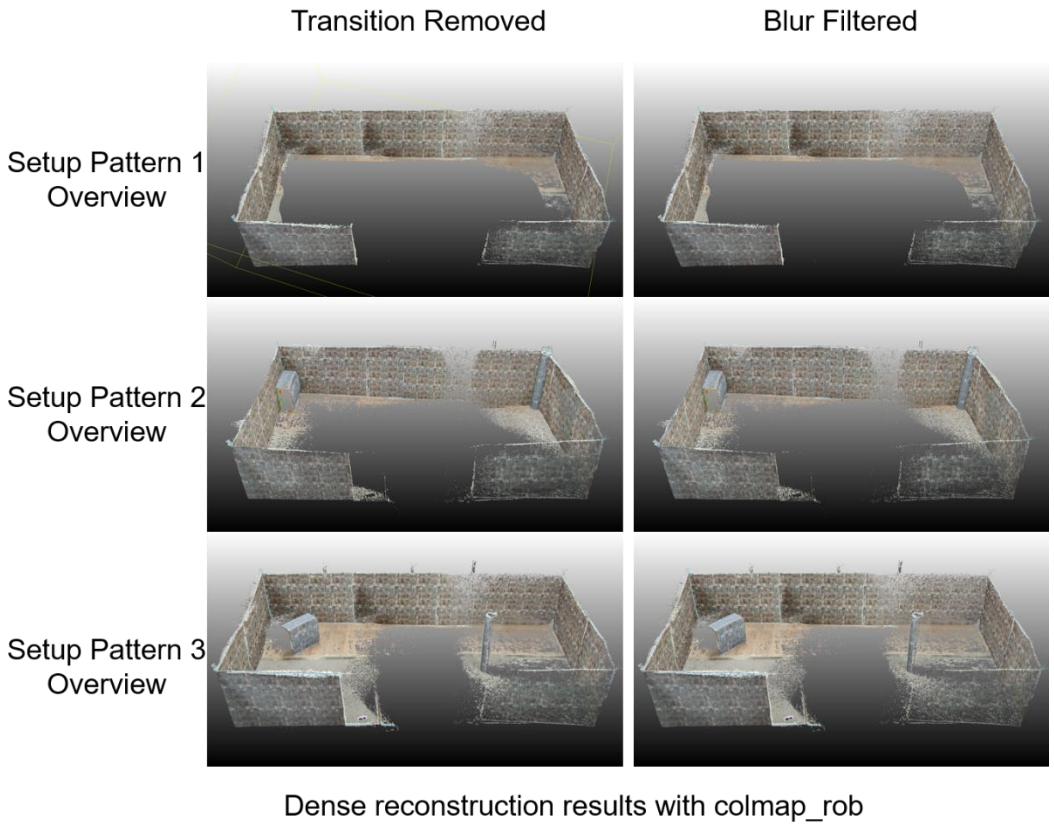


Figure 5.13 Reconstruction results by colmap_rob with transition removed dataset (left) and blur filtered dataset (right) under three setup patterns.

Table 5.5 Reconstruction performance comparison results by colmap_rob before and after image selection under three setup patterns.

colmap_rob	Metric	Transition Removed	Blur Filtered
Setup Pattern 1	Accuracy (%)	88.0	88.6
	Completeness (%)	62.1	60.9
	Runtime (s)	6213	5962
Setup Pattern 2	Accuracy (%)	88.2	88.7
	Completeness (%)	67.7	60.1
	Runtime (s)	6596	6007
Setup Pattern 3	Accuracy (%)	85.2	90.7

	Completeness (%)	59.2	64.1
	Runtime (s)	5999	5762

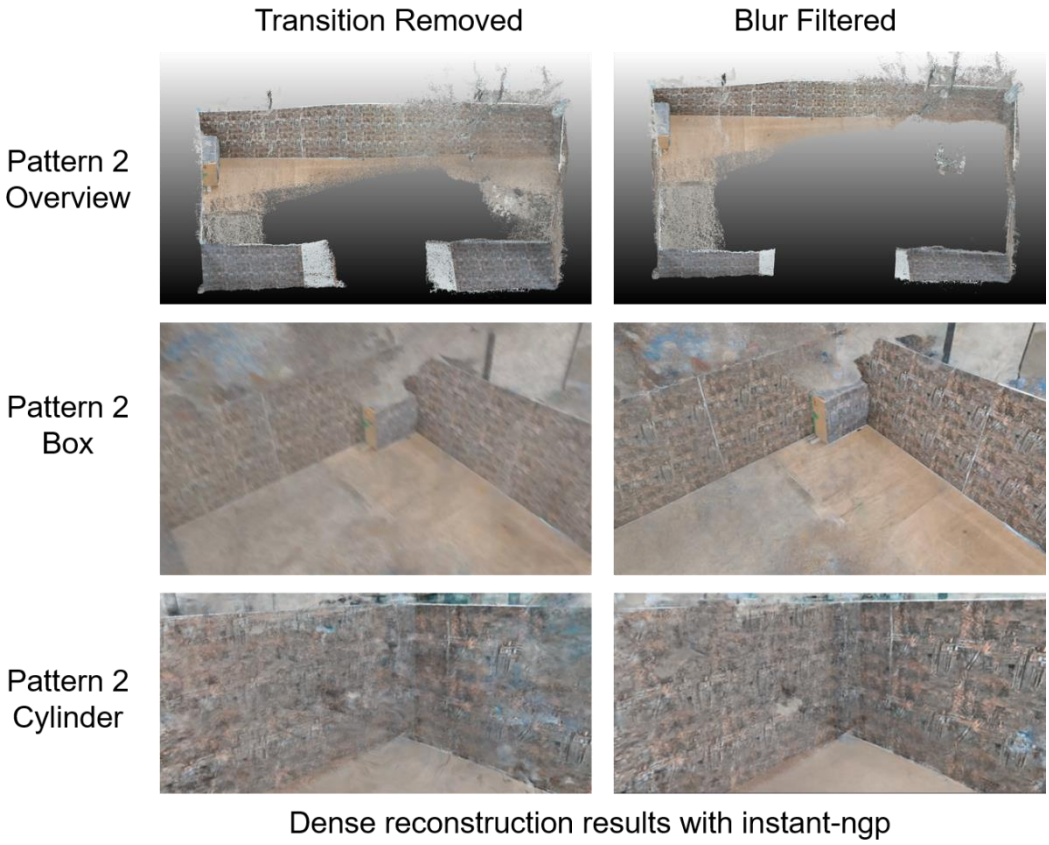


Figure 5.14 Reconstruction results by instant-ngp with transition removed dataset (left) and blur filtered dataset (right) under setup pattern 2

Table 5.6 Reconstruction performance comparison results by instant-ngp before and after image selection under three setup patterns.

instant-ngp	**Metric**	**Trajectory Adjusted**	**Blur Filtered**
Setup Pattern 1	Accuracy (%)	67.4	71.1
	Completeness (%)	94.1	91.9
	Runtime (s)	6.2	5.8
Setup Pattern 2	Accuracy (%)	75.6	79.4
	Completeness (%)	87.2	86.3
	Runtime (s)	6.3	6.2
Setup Pattern 3	Accuracy (%)	69.9	73.1

	Completeness (%)	91.2	90.0
	Runtime (s)	5.0	5.0

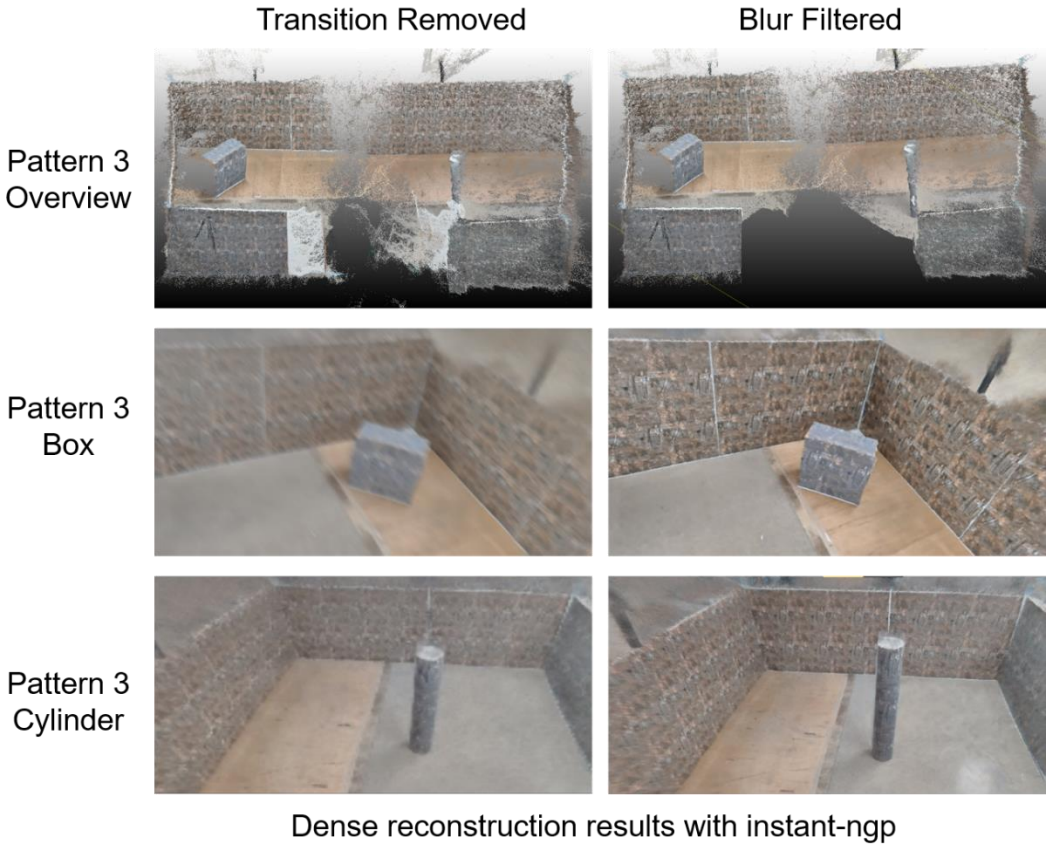


Figure 5.15 Reconstruction results by instant-ngp with transition removed dataset (left) and blur filtered dataset (right) under setup pattern 3

By comparing the reconstructed point clouds, we can find that the fuzzy noise gets reduced, which may be the representation of the blurry image when utilized in the training of NeRF. Since the fuzzy noise should not exist in the ground truth, its existence will generate inconsistency between the ground truth, thus leading to reconstruction accuracy degradation. While, the step of blur filtering in the proposed image selection method recovers relative average accuracy for instant-ngp at 3.5%, while at the cost of about 1.5% relative completeness drop. The drop of completeness drop should result from removing the images in the blurry periods, like the starting and stopping periods, and the transition phase trajectory. Therefore, the completeness of areas captured in these periods might be dropping, like the white edges loss in the reconstruction results.

5.3.5 Robustness Study against Limited Illumination

Till now, the feasibility and effectiveness of the proposed dataset creation method have been verified through experiments. The results revealed that the proposed method could improve the image quality of dataset gathered from flexible manipulators, thus improving the compatibility

between SfM and flexible manipulators. While, apart from the image quality degradation from robot's characteristics, the degradation from harsh environment like limited illumination is also an important topic remaining to be solved. As a preliminary robustness study of mainstream vision-based environment mapping methods, with the proposed dataset creation method, a reconstruction performance comparison experiment was conducted under different illumination conditions. The experimental introduction and results can be referred in Appendix D.

5.4 Summary

In this chapter, the influences of the common motion controllability degradation were experimentally discussed, including positional deviation and motion blur. The positional deviation will damage the reliability of manipulator's motion patterns as prior knowledge; The motion blur will damage the image quality in the total dataset, thus reducing the feasibility of conducting SfM with the dataset. In order to mitigate the influences of these two problems, this chapter proposed one improvement method for Structure from Motion to utilize flexible manipulator's motion patterns features as prior knowledge. The dataset creation method includes two steps: a trajectory design step to avoid high positional deviation trajectories, and an image selection step to remove blurry images. Both of these two steps were conducted on the basis of the specific motion pattern of the manipulator.

The effectiveness of the proposed dataset creation method was also experimentally confirmed by comparing the reconstruction performance of SfM with MVS/NVS before and after applying the dataset creation method. The dataset creation method was proved to be effective at avoiding trajectories with the adverse effects of bundled wire interference while preserving the reconstruction performance with continuous viewpoints, by creating overlapping trajectories with high image quality. The total dataset creation method was effective at improving the average reconstruction accuracy of colmap_rob by 3.7% at the cost of 1.3% completeness drop, while reducing 22% runtime; The result for instant-ngp was 14.7% accuracy improvement at the cost of 2.7% completeness drop, and 9.7% runtime reduction.

Chapter.6 Conclusion

This thesis focuses on three challenges of conventional robotic investigation systems and environment mapping methods in terms of preliminary investigation stage of fuel debris retrieval phase for future Fukushima decommissioning works.

One of the main challenges encountered is the limited visibility and high radiation levels in Unit 1, which hindered the ability of the underwater camera of PMORPH2 to provide adequate awareness of fuel debris or water leaks. To address this problem, I proposed a mobile robotic investigation system called RhinoUS-II, which is equipped with a 2D phased array ultrasonic sensor. Through experiments with a hardware prototype, I confirmed the feasibility of using the newly-implemented 2DoF winch mechanism to navigate through narrow grating lattices (100mm x 30mm). Additionally, the implementation of the 2D phased array sensor demonstrated improved robustness against underwater turbidity and radiation in experiments. This suggests that the proposed mobile robotic platform RhinoUS-II can serve as a suitable carrier for the newly-implemented 2D phased array ultrasonic sensor, which can provide better spatial awareness of turbid and radioactive underwater environments by using shape & flow mapping methods. As the result, the utility of this proposed system may enable operators to locate fuel debris and water leakages more effectively, thus to minimize necessary investigation time.

Another challenge is that the feasibility of using the tough manipulator "Veolia-MHI Arm" in Unit 2 is hindered by its weight, weak collision resistance, and difficulties in recovering from joint failure. As an alternative option for the investigation stage of Unit 2, I proposed a lightweight long-reach wire-drive manipulator called the BWD robot, which can install actuators at the base and easily retrieve from joint failure and has a high tolerance for collisions. The unique bundled wire drive mechanism of the BWD robot reduces the weight of the arm part to only 1.8 kg. The low joint stiffness of the BWD robot makes it more compliant and adaptable to the extremely complex environment, and also allows for better tolerance against possible collisions, making it safer and more forgiving if it comes into contact with obstacles, which reduces the risk of damage to both the arm and other equipment. Therefore, the implementation of the proposed BWD robot offers insight into using flexible manipulators with high tolerance for collisions in such unknown environment investigations, which may help save expenses on repairs or re-development by enhancing forgiveness in case of emergency collisions.

The third challenge is that the compatibility of Visual Odometry (VO) and Structure from Motion (SfM) with the specific environment and robot is limited due to their inability to fully utilize characteristics as prior knowledge, such as grating lattices with known size and specific robot motion patterns. To address this, three improvement methods are proposed that allow VO and SfM to utilize characteristics as prior knowledge.

By implementing a perspective transformation method, VO can provide an intuitive top-view map, thus allowing for the counting of past grating lattices. The availability of a clear top-

view map offers a more intuitive understanding of the layout and spatial relationship of the grating floor for remote operators, enabling them to reach the target investigation points more effectively. Especially, the top view map introduces the number of past grating lattices as an additional information source for feature-based VO method, which improves both the robustness and spatial awareness of remote investigations.

The proposed motion pattern-based compensation method allows feature-based Visual Odometry (VO) methods to utilize the robot's motion patterns as prior information to correct maps generated during feature tracking failures. These failures are more likely to occur in realistic and complex environments. The implementation of this method has the potential to enhance the robustness of visual inspections in such environments, and expand the areas that can be inspected.

Lastly, the proposed dataset creation method utilizes the motion patterns of the BWD robot, thus skillfully mitigating the image quality degradation problem caused by the BWD mechanism's motion controllability degradation. The total accuracy improved. Back to the context of the Fukushima reactor investigation, this method allows for the use of flexible manipulators while preserving the image quality of the captured image dataset. Thus, the proposed method improves the compatibility between flexible manipulators and environment reconstruction method SfM, which values motion controllability. As such, it enhances the feasibility of using flexible manipulators while conducting decent environment mapping in unknown environments.

In general, this thesis discusses the utilization of specific environmental conditions and robot’s motion patterns to improve performance of conventional environment mapping methods, especially for the grating textures in Fukushima reactors, and the specific patterns of two robots. Although, the overall developing direction of environment mapping method is to conduct the mapping process with only images input, therefore avoiding necessary dependence on other specific conditions. While, apart from the specific Fukushima situation and the specific robots, the idea of utilizing such information as prior knowledge should be generic, since they are available for performance improvement and don’t work as a necessary input. The similar idea of using specific environment condition can be picked as “Manhattan World Assumption”, which assumes that most artifacts like buildings are built parallel to the ground cartesian coordinate system. This assumption is often utilized in SLAM optimization designed for the city scenes with many buildings. When applying to other robots, the core problem will be extracting the motion patterns’ features and adjust the improvement method correspondingly, which will require preliminary experiments on the concerned specific motion patterns in terms of motion performance as well as image quality like Section 5.1.3 and Section 5.1.4.

The future work directions are listed below: (1) Consider the necessary robustness enhancement measures against harsher condition close to actual Fukushima investigation, like

low illumination, radiation, foggy. (2) Conduct effectiveness verification of the proposed methods on other vision-based environment mapping methods.

Reference

[1] Roadmap on the Way to Decommissioning (2021). Japan: Tokyo Electric Power Company Holdings. https://www.tepco.co.jp/en/hd/decommission/project/roadmap/index-e.html. 2021 May, accessed Jan 2023.

[2] Tokyo Electric Power Company Holdings (2013). The progress and lesson in the Fukushima Nuclear Power Plant accident [in Japanese]. http://www.tepco.co.jp/nu/fukushima-np/outline/pdf/outline01.pdf. March 2013, accessed Mar 2019.

[3] Information Portal for the Fukushima Daiichi Accident Analysis and Decommissioning Activities (2012). Japan: Tokyo Electric Power Company Holdings. https://fdada.info/home/vis_3d/unit1_2_3d/. February 2012, accessed Mar 2019.

[4] Tokyo Electric Power Company Holdings (2022). Development of a technology to investigate inside the Reactor Primary Containment Vessel (PCV). https://www.tepco.co.jp/en/nu/fukushima-np/handouts/2015/images/handouts_150430_04-e.pdf. April 2015, accessed Jan 2020.

[5] International Research Institute for Nuclear Decommissioning and Tokyo Electric Power Company Holdings (2017). About the internal investigation of Unit 1 reactor [in Japanese]. https://www.tepco.co.jp/nu/fukushima-np/handouts/2017/images1/handouts_170327_14-j.pdf. March 2017, accessed June 2018.

[6] Unit 1 Fuel Debris Retrieval (2022). Japan: Tokyo Electric Power Company Holdings. https://www.tepco.co.jp/en/hd/decommission/progress/retrieval/unit1/index-e.html. 2022 May, accessed Oct 2022.

[7] Hitachi GE (2017). Development of PMORPH, the internal PCV investigation robot [in Japanese]. https://irid.or.jp/wp-content/uploads/2017/02/20170203_21.pdf. Feb 2017, accessed July 2018.

[8] Konishi T et al. Development of position estimation technology for robot in extreme environments. Proceedings of the 35th Annual Conference of the Robotics Society of Japan, 2017 Sep 10-14, Saitama, Japan. [in Japanese]

[9] International Research Institute for Nuclear Decommissioning and Tokyo Electric Power Company Holdings (2017). The investigation results by the self-propelled investigation device [in Japanese]. https://www.tepco.co.jp/nu/fukushima-np/handouts/2017/images1/handouts_170216_11-j.pdf. Feb 2017, accessed Nov 2022.

[10] International Research Institute for Nuclear Decommissioning (2018). The current situation of technique development concerned in Fukushima Daiichi decommissioning [in Japanese]. https://irid.or.jp/_pdf/20180921.pdf. Sep 2018, accessed 12 Dec 2022.

[11] Tokyo Electric Power Company Holdings (2018). The results of the internal investigation of the No.2 reactor containment vessel [in Japanese]. https://irid.or.jp/wp-content/uploads/2018/01/20180119.pdf. Jan 2018, accessed 22 Sep 2022.

[12] International Research Institute of Decommissioning (2022). The access device (Robot Arm) for further investigation and trial fuel debris extraction of the No.2 reactor [in Japanese]. https://irid.or.jp/wp-content/uploads/2022/03/20220309_robot_arm.pdf. March 2022, Accessed March 2022.

[13] International Research Institute of Decommissioning (2020). Development of investigation technique of reactor containment vessel [in Japanese]. https://irid.or.jp/wp-content/uploads/2020/10/2019001x6penepcvtyousa.pdf. July 2020, Accessed March 2022.

[14] International Research Institute of Decommissioning (2021). Development of investigation technique of reactor containment vessel [in Japanese]. https://irid.or.jp/wp-content/uploads/2021/07/2020009X6penePCVnaibusyousaityousa202107.pdf. May 2021, Accessed March 2022.

[15] Horigome A, Endo G (2016). Basic study for drive mechanism with synthetic fiber rope - investigation of strength reduction by bending and terminal fixation method. Advanced Robotics 30(3):206-217.

[16] Horigome A, Endo G (2018). Investigation of Repetitive Bending Durability of Synthetic Fiber Ropes. IEEE/International Conference on Robotics and Automation (ICRA2018), IEEE Robotics and Automation Letters, 3-3:1779-1786.

[17] Endo G, Horigome A, Wakabayashi Y, Takata A (2018). Basic study for drive mechanism with synthetic fiber rope (Terminal fixation method using a grooved pulley and a figure-eight knot). Transactions of the JSME (in Japanese), Japan Society of Mechanical Engineers, Vol. 84, No. 864, pp. 18-00067, June 2018.

[18] Endo G, Wakabayashi Y, Nabae H and Suzumori K (2019). Bundled Wire Drive: Proposal and Feasibility Study of a Novel Tendon-Driven Mechanism Using Synthetic Fiber Ropes. IEEE Robotics and Automation Letters 4(2) 966-972. doi: 10.1109/LRA.2019.2893429.

[19] Ikuta K, Daifu S, Hasegawa T, Higashikawa H (2002). Hyper-finger for Remote Minimally Invasive Surgery in Deep Area. In: Dohi, T., Kikinis, R. (eds) Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002. MICCAI 2002. Lecture Notes in Computer Science, vol 2488. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45786-0_22

[20] Intel RealSense Depth Camera D435i (2022). United States: Intel. https://www.intel.co.jp/content/www/jp/ja/products/sku/190004/intel-realsense-depth-camera-d435i/specifications.html. Accessed Dec 2022.

[21] Wakabayashi Y, Endo G, Nabae H and Suzumori K (2020). A development of long-reach and lightweight robot arm using bundled wire drive. Second report: A reconfigurable wire drive mechanism. 2020 JSME Conference on Robotics and Mechatronics, Proceedings of the 2020 JSME Conference on Robotics and Mechatronics, The Japan Society of Mechanical Engineers, May 2020.

[22] Agarwal S, Snavely N, Simon I, Seitz SM and Szeliski R (2009). Building Rome in a day. 2009 IEEE 12th International Conference on Computer Vision, 2009, pp. 72-79, doi: 10.1109/ICCV.2009.5459148.

[23] Schönberger JL and Frahm J (2016). Structure-from-Motion Revisited. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 4104-4113, doi: 10.1109/CVPR.2016.445.

[24] Agarwal S, Mierle K, et al (2015). Ceres Solver. http://ceres-solver.org.

[25] Lindenberger P, Sarlin P, Larsson V and Pollefeys M (2021). Pixel-Perfect Structure-from-Motion with Featuremetric Refinement. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 5967-5977. doi: 10.1109/ICCV48922.2021.00593

[26] Schönberger, Johannes Lutz and Zheng, Enliang and Pollefeys, Marc and Frahm, Jan-Michael (2016). Pixelwise View Selection for Unstructured Multi-View Stereo. European Conference on Computer Vision, 2016.

[27] Mildenhall, Ben, et al (2020). Nerf: Representing scenes as neural radiance fields for view synthesis. European conference on computer vision. Springer, Cham, 2020.

[28] MÜLLER, Thomas, et al (2022). Instant neural graphics primitives with a multiresolution hash encoding. arXiv preprint arXiv:2201.05989, 2022.

[29] Engel J., Schöps T., Cremers D. (2014). LSD-SLAM: Large-Scale Direct Monocular SLAM. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8690. Springer, Cham. https://doi.org/10.1007/978-3-319-10605-2_54

[30] Labbé M, Michaud F, et al (2019) RTAB-Map as an Open-Source Lidar and Visual SLAM Library for Large-Scale and Long-Term Online Operation. Journal of Field Robotics 36(2) 416–446.

[31] Endo G, Kikura H, et al (2016). Proposal of a Robot System Transporting an Ultrasonic Sensor for Measuring Flows of Contaminated Water in Reacter. Proceedings of the 10th International Symposium on Ultrasonic Doppler Methods for Fluid Mechanics and Fluid Engineering, 2016 Sep 28-30, Tokyo, Japan.

[32] Beller SL, Brown LH (1984). Design and Operation of the Core Topography Data Acquisition System for TMI-2. Idaho (USA): EG & G Idaho, Inc; 1984, (GEND-INF 012)

[33] Kikura H, Kawachi T, Ihara T (2015). Study on Ultrasonic Measurement for Determination of Leakage from Reactor Vessel and Debris Inspection. Proceedings of the 11th National Conference on Nuclear Science and Technology, 2015 Aug 6-7, Da Nang, Vietnam.

[34] Stanford Artificial Intelligence Laboratory et al (2018). Robotic Operating System. Retrieved from https://www.ros.org.

[35] Kawachi T, Kimura S, Ihara T, Kikura H, Kimoto K (2015). Development of Simultaneous Measurement System of Object Surface Shape and Two-dimensional Flow Mapping by using Ultrasonic Array Sensor [in Japanese]. IEICE Technical Report 115:39-44.

[36] Nishiwaki R, Takahashi H, Endo G, Kikura H (2018). Study on Flow Mapping Method of Ultrasonic Flow Measurement by Robot Remote Control [in Japanese]. Proceedings of the 96th Japan Society of Mechanical Engineers Fluids Engineering Conference, 2018 Nov 29-30, Muroran, Japan.

[37] Kim S, Kim I, Vecchietti LF, Har D (2020). Pose Estimation Utilizing a Gated Recurrent Unit Network for Visual Localization. Applied Sciences. 2020; 10(24):8876. doi: 10.3390/app10248876.

[38] Li R, Wang S, Gu D (2018). Ongoing evolution of visual slam from geometry to deep learning: Challenges and Opportunities. Cognit. Comput. 2018, 10, 875–889.

[39] Zhang ZY (2000). A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22:1330-1334. doi: 10.1109/34.888718.

[40] Rublee E, Rabaud V, Konolige K, Bradski G (2011), ORB: An Efficient Alternative to SIFT or SURF. Proceedings of 2011 International Conference on Computer Vision, Barcelona, Spain, 6-13 Nov 2011. doi: 10.1109/iccv.2011.6126544.

[41] Bradski G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools.

[42] Davison JA, Reid DI, Molton DN, Stasse O (2007). MonoSLAM: Real-Time Single Camera SLAM. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007 April; 29:1052-1067.

[43] Campion G., Chung W. (2008) Wheeled Robots. In: Siciliano B., Khatib O. (eds) Springer Handbook of Robotics. Springer, Berlin, Heidelberg, doi: 10.1007/978-3-540-30301-5_18

[44] Durrant-Whyte H, Bailey T (2016). Simultaneous localization and mapping: part I. IEEE Robotics & Automation Magazine. 2016 June; 13:99-110.

[45] Okada Y, Kojima S, Mizuno N (2018). eposx_hardware. https://github.com/yoshito-n-students/eposx_hardware.

[46] Quattrini LA, Coskun A, et al (2016). Experimental Comparison of Open-Source Vision-Based State Estimation Algorithms. In: Kulić. D, Nakamura Y, Khatib O, Venture G. (eds) 2016 International Symposium on Experimental Robotics. Springer Proceedings in Advanced Robotics, vol 1. Springer, Cham. doi:10.1007/978-3-319-50115-4_67.

[47] Li, Haoang, et al. A monocular SLAM system leveraging structural regularity in Manhattan world. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018. p. 2518-2525.

[48] Jensen R, Dahl A, Vogiatzis G, Tola E and Aanæs H (2014). Large Scale Multi-View Stereopsis Evaluation. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 406-413, doi: 10.1109/CVPR.2014.59.

[49] Hendrycks D, Dietterich T (2019) Benchmarking neural network robustness to common corruptions and perturbations. In: 2019 International Conference on Learning Representations, New Orleans, USA, May 6–9, 2019.

[50] Su B, Lu S, Tan CL (2011). Blurred Image Region Detection and Classification. Proceedings of the 19th ACM International Conference on Multimedia. Scottsdale, Arizona, USA, Nov 28-Dec 1. 2011.

[51] Schops T, Schönberger JL, et al (2017) A multi-view stereo benchmark with high-resolution images and multi-camera videos. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, July 22–25, 2017.

[52] Tokyo Electric Power Company Holdings (2018). About the results of the internal investigation of the No. 2 reactor containment vessel [in Japanese]. https://www.tepco.co.jp/nu/fukushima-np/roadmap/2018/images1/d180426_08-j.pdf. Accessed Jan 8 2021.

[53] Orest K, Tetiana M, Junru W, Wang ZY (2019). DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better. In Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 8878-8887.

[54] Zhang K, Ren W, Luo W, Lai W, Stenger B, Yang MH, Li H (2022). Deep image deblurring: A survey. International Journal of Computer Vision, 130(9), 2103-2130.

[55] Li DJ, Deng HP, Yao G, Jiang JC, Zhang YB (2022). Gamma Radiation Image Noise Prediction Method Based on Statistical Analysis and Random Walk. Sensors 2022, 22(19):7325. https://doi.org/10.3390/s22197325.

[56] Wang CY, Bochkovskiy Am Liao MHY (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696, 2022 - arxiv.org.

Publications

International Journal

1. Wang, Z., Endo, G., Takahashi, M. et al. Study of a robotic system to detect water leakage and fuel debris-System proposal and feasibility study of visual odometry providing intuitive bird's eye view-. Robomech J 7, 34 (2020). https://doi.org/10.1186/s40648-020-00184-z (Chapter 2~3)
2. Wang, Z., Endo, G., Takahashi, H. et al. Dataset creation and selection methods with a wire drive flexible manipulator for vision-based reconstruction. Robomech J 10, 4 (2023). https://doi.org/10.1186/s40648-023-00241-3 (Chapter 4~5, Appendix D)

Presentation in the International Conference

1. Gen Endo, Toshiya Nagai, Atsushi Takata, Zhenyu Wang, Hiroshige Kikura, Hideharu Takahashi. Development using Advanced Super Dragon Articulated Robot Arm (2) Development of an Articulated Robot Arm with Telescopic Structure. International Topical Workshop on Fukushima Decommissioning Research 2022 (FDR2022), Oct. 2022.

Presentation in the Domestic Conference

1. Zhenyu Wang, Masashi Takahashi, Gen Endo, Koichi Suzumori, Hiroyuki Nabae, Hideharu Takahashi, Hiroshige Kikura. Empirical Evaluation on the Localization Accuracy of Visual Odometry on the Gratings for PCV Internal investigation. Proceedings of the Robotics and Mechatronics Conference 2019 June 5-8, Hiroshima, Japan.

Publications not Related to this Thesis

1. Zhenyu Wang, Liang Gong, Qianli Chen, Yanming Li, Chengliang Liu, Yixiang Huang. Rapid Developing the Simulation and Control Systems for a Multifunctional Autonomous Agricultural Robot with ROS. International Conference on Intelligent Robotics and Applications.2016 Aug. 22-24, Hachioji, Tokyo, Japan.
2. Liang Gong, Changyang Gong, Zhao Ma, Lujie Zhao, Zhenyu Wang, Xudong Li, Xiaolong Jing, Haozhe Yang, Chengliang Liu. Real-time human-in-the-loop remote control for a life-size traffic police robot with multiple augmented reality aided display terminals. International Conference on Advanced Robotics and Mechatronics (ICARM). 2017 Aug 27-31, Hefei & Taian, China.

Acknowledgement

I would like to express my heartfelt gratitude to my advisor, Prof. Endo, for his invaluable guidance, support, and encouragement throughout my research journey. Besides, during the writing process of this thesis, Prof. Takayama, Prof. Matsunaga, Prof. Kikura, and Prof. Suzumori have provided a lot of pertinent comments, which greatly contributed to the refinement of this thesis. I would like to deeply thank them for it. Besides, I want to thank cooperators from Kikura lab, Takeshi Moriya for valuable help during the cooperating experiments. I want to thank my colleagues from Suzumori lab, Endo lab for their helpful insights, constructive feedback, and discussion throughout the research journey. And my parents, it has been over 3 years since I met them last time. I miss them. I want to thank you all for lifting me up every time when I fell.

And, I'm grateful to MEXT for their scholarship financial support. This work is supported by the Nuclear Energy Science & Technology and Human Resource Development Project (through concentrating wisdom) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. This work was also supported by the JAEA Nuclear Energy S&T and Human Resource Development Project under Grant Number JPJA19P 19210348. This research was partially conducted by the TEPCO & Tokyo Tech Collaborative Research on Frontier Decommissioning Technologies.

When I reflect the period as a PhD student, I feel long and short at the same time. It definitely is the most painful period till now, which has brought much depression to me. Therefore, I feel way too long, and hope it could end sooner. But the period has been basically overlapping with the Corona pandemic, which has significantly changed the lives of numerous people. The dramatic three years made everything look extremely fast. Now, I have come to nearly the end of the PhD period, I still find myself a green hand in academics. Therefore, I feel so short, as time waits no one.

During my 5 to 6 years as a student, I have often had nightmares about not being able to submit my exam papers on time. Although I never had such an experience, it's simply a lingering sense of anxiety. However, the greatest fear I have faced is the specter of "Originality," which is also the driving force behind human progress. If academics did not value originality, and if we did not strive to go beyond the limits of our existing knowledge, human progress would come to a standstill. It is definitely essential to explore the unknown, and due to its difficulty, the academic world is often referred to as "Standing on the shoulders of Giants." This phrase acknowledges the importance of building upon the knowledge and achievements of those who have come before us. The dwarfs could have seen further by standing on the shoulders of giants. But the dwarfs also undertake the responsibility of building the giant higher, or becoming the new giant. At times, a strange question comes to me that, will some of dwarfs be afraid of the danger of dropping from

that height, so unable to climb then? Or, the height of giant has been way too high, making it extremely hard for a “normal” dwarf to continue building? I might have been one of the scared dwarfs. But as I near the end of my doctoral journey, I can finally say that I have made some small contributions to the shoulders of the giants, and perhaps helped increasing their height a little bit. In fact, the title of “Doctor” might be the best gift for those dwarfs who did it.

Wang Zhenyu

2023/01/29

Appendix A

Working Principles of the 2D Phased Array Sensor

In terms of the phased array technique, Figure 0.1 presents the basic principle of array ultrasonic sensor. Here a indicates the width of one element, and d indicates the inter-element spacing. The steering angle θ and the time delay Δt is related to the speed of sound c and the inter-element spacing d, as Equation A.1 shows.

$$\theta = \arcsin\left(c\frac{\Delta t}{d}\right) \tag{A.1}$$

Therefore, based on a different predetermined time delay, we used an electronic pulse to excite the element in the sensor. After being excited, the element would emit a spherical ultrasonic wave forming interference with the wave from neighboring elements. The pattern of the interference is dependent on the time delay; While, we could change the time delay to adjust the steering angle and focal point of the ultrasonic beam by using the beam steering principle.

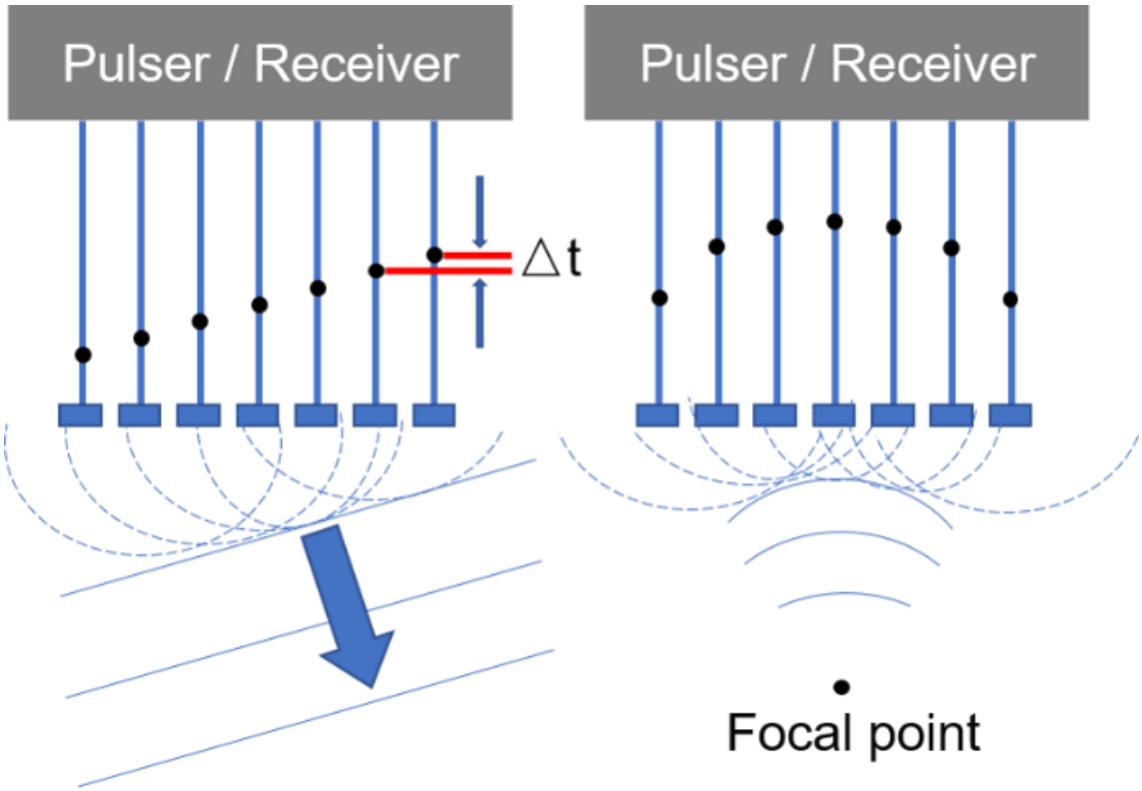


Figure 0.1 The principle of phased array technique.

Working Principles for "Water under pedestal grating"

The processes of underwater localization included the localization of water leakage to avoid further outflow of polluted water, and that of fuel debris for further tasks.

For the localization of water leakages, I used the UVP method to reconstruct the velocity field, and the working principle is shown in Figure 0.2. The upper grey part indicated a linear array ultrasonic sensor with 8 horizontally distributed elements, denoted as 1-ch to 8-ch. An emitted ultrasonic pulse and element in the array sensor would receive the echo reflected from the surface of a particle in the measurement line. The Doppler frequency observed at each element is described in Equation A.2 as below,

$$f_{D_i} = \frac{f_0}{c}(\boldsymbol{e_e} + \boldsymbol{e_l}) \cdot \boldsymbol{V} \qquad (A.2)$$

where f_{Di} is the Doppler frequency which is observed at the i_{th}-channel element, f_0 is the frequency of the emitted pulse echoed by the particle, $\boldsymbol{e_e}$ is the unit vector in the direction of the measurement line, $\boldsymbol{e_l}$ is the unit vector in the direction from the particle to the i_{th}-channel element and $\boldsymbol{V}$ is the particle velocity. From Equation A.2, the doppler-shift frequency observed at each element is different because of the difference of element positions.

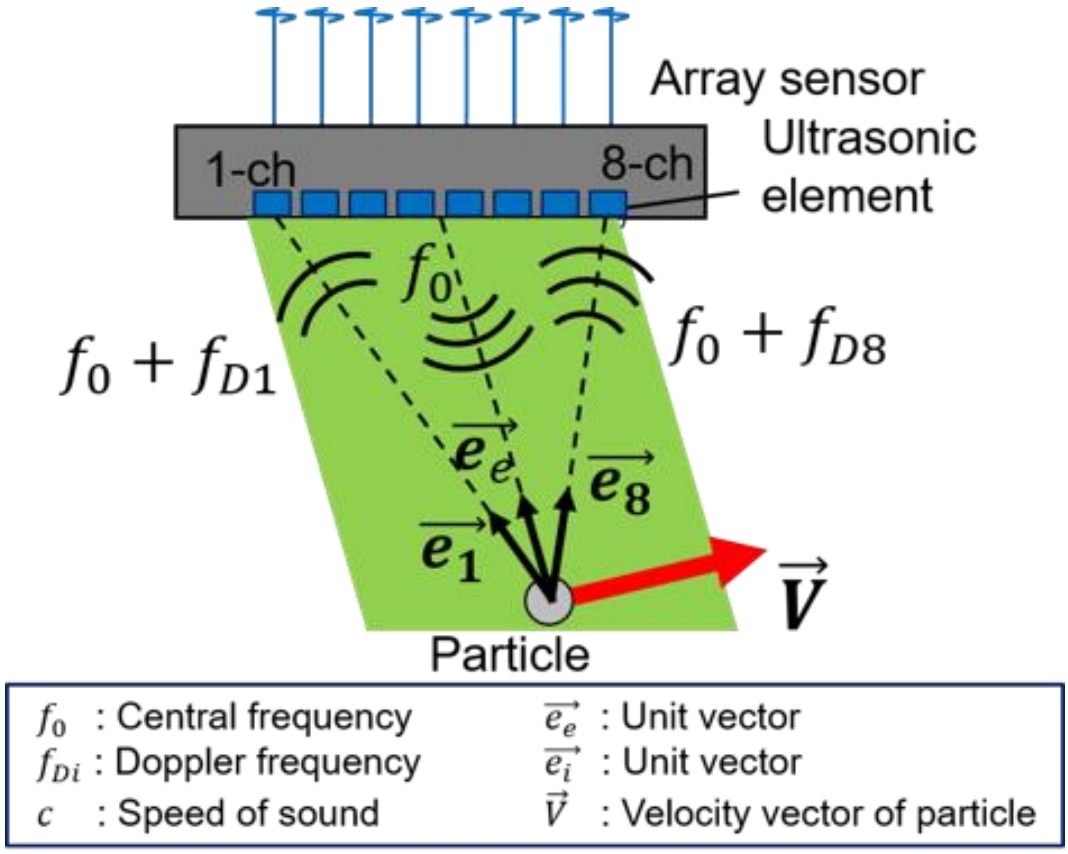


Figure 0.2 The principle of velocity vector reconstruction with a linear array sensor.

Therefore, the velocity of the detected particle can be reconstructed based on the doppler frequency from each channel element by Equation A.3 and A.4:

$$\boldsymbol{V} = \frac{c}{f_0}\begin{bmatrix}\boldsymbol{e_e} + \boldsymbol{e_1} \\ \boldsymbol{e_e} + \boldsymbol{e_8}\end{bmatrix} \cdot \begin{bmatrix} f_{D_1} \\ f_{D_8}\end{bmatrix} \tag{A.3}$$

$$f_{D_i} = \frac{f_0}{c}\,(\boldsymbol{e_e} + \boldsymbol{e_1}) \cdot \boldsymbol{V} \tag{A.4}$$

Thus, a velocity vector in a measurement point can be calculated by analyzing the echoes received by different elements in an array sensor. Moreover, the velocity vector distribution in the measurement line can be obtained by this processing at each measurement point. Applying the phased array technique, flow mapping can be conducted with only one array sensor [33]. On the basis of the flow mapping measured at different positions, it's possible to localize the position of unknown outlet by combining the flows.

On the other hand, when using ultrasonic sensors to detect the shape of objects, conventionally, a single ultrasonic transducer emits ultrasonic pulses and receives the reflected echo from the object's surface. A common problem of this method is that the strength of the reflected echo may be quite weak when the direction of the emitted pulse forms an angle with the surface of the object that is to be inspected. Therefore, even if the transducer performs a surface inspection while moving parallel to the object, part of the object still cannot be inspected by this method [35]. Evidently, this problem is likely to occur when inspecting objects with rugged surfaces, such as fuel debris. Our research group applied the aperture synthesis method to receive the reflected echo with multiple arrayed elements. The principle of aperture synthesis method is shown in Figure 0.3.

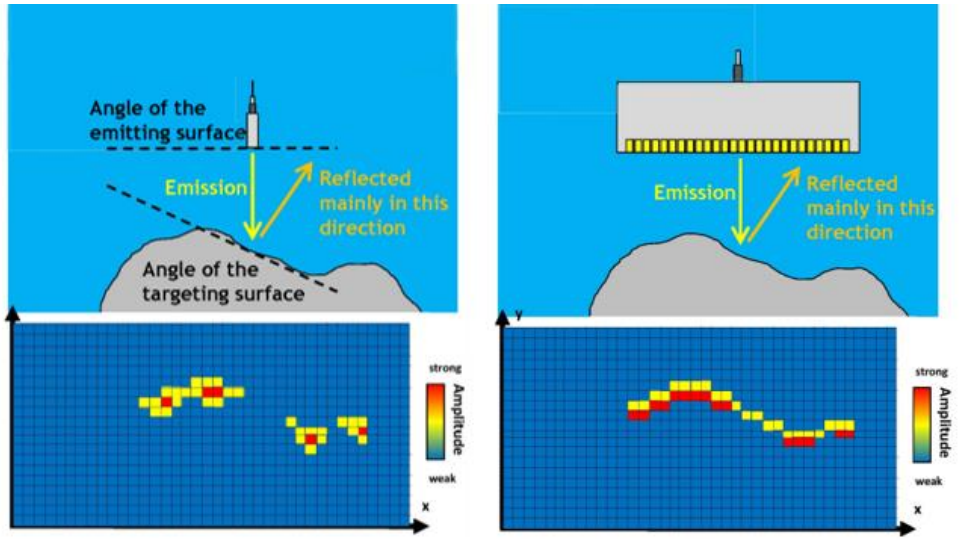


Figure 0.3 Schematics of the aperture synthesis method. Left: Measurement with conventional ultrasonic sensor; Right: Measurement with aperture synthesis method and array sensor. [35]

The left part of Figure 0.3 showed the situation of using only one conventional ultrasonic sensor to detect the targeting surface which may form angle with the emitting surface. Therefore, the reflected signals will be hard to analyze due to low signal-to-noise ratio. While, usage of linear array sensor with multiple elements will help receive the reflected signals to reconstruct the targeting surface, also known as aperture synthesis method.

Experiments with Ultrasonic Sensors

The verification experiments for the array ultrasonic sensor in terms of underwater localization was also conducted according to two aspects: water leakages localization with UVP method, and shape reconstruction of fuel debris with aperture synthesis method.

For the water leakages localization experiment, the sensor was put into a water tank where an outlet pipe was set at the center of the bottom floor. The water tank size is 1160mm x 720mm, and a magnet water pump was placed at one side to circulate the water at 14L/min. The ultrasonic sensor was carried with a wire pulley driven robot arm installed on the wheeled test platform robot. Here, we denoted the center of the water tank, also the position of the mock-up water leakage as the coordinate system center. Then, the planned trajectory in green line, started with a horizontal movement from about (-80, 250) to (80, 250); then sinking downward about 150mm, turning left to (-80, 100). Then, the reconstructed flow mapping of three points on the final phase is shown in Figure 0.5.

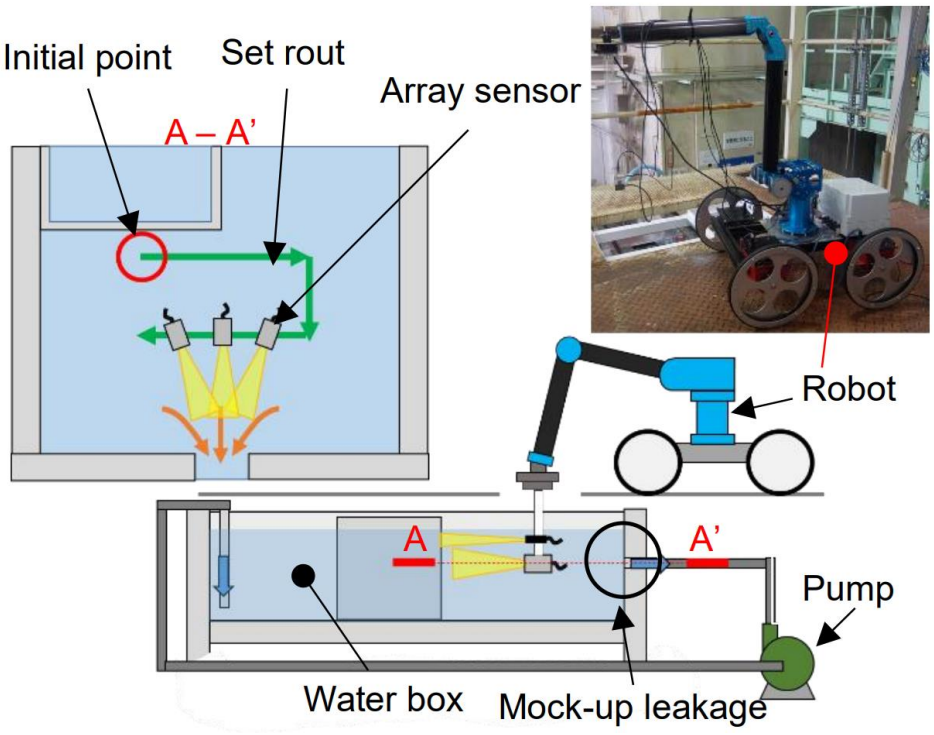


Figure 0.4 The experiment setup for water leakages localization. [36]

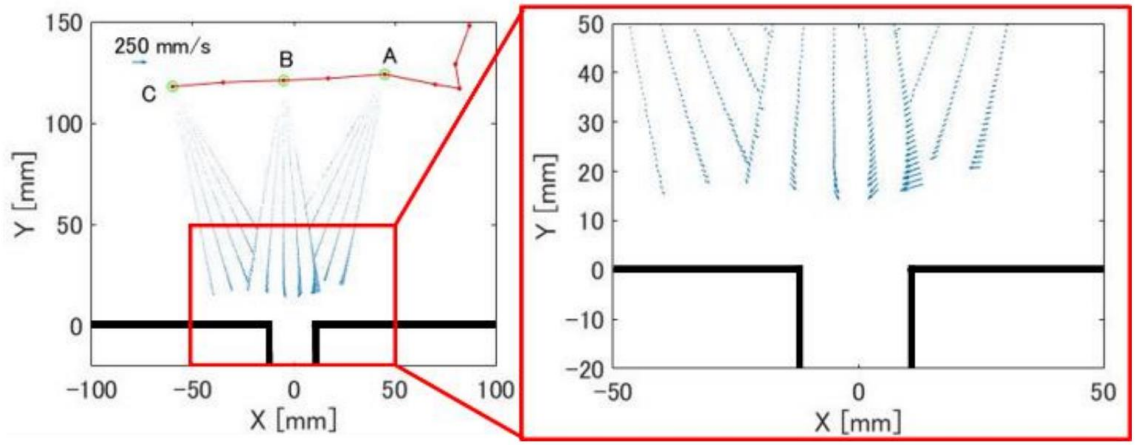


Figure 0.5 The measured results of remote flow mapping. [36]

Three green points were picked from the trajectory as A B C. The blue lines indicated the reconstructed flow mapping at each point. Therefore, we could observe that by combining the flow mapping measured from different positions, it's feasible to localize the water leakage in the water tank.

For the feasibility experiment of shape reconstruction on fuel debris, the assumption on the fuel debris characteristics should be made. In the experiment, we used a slug in rugged shape (about 100mm x 40mm) as the presumed fuel debris, as Figure 0.6 shows. The slug was placed in a water box, and a linear array ultrasonic sensor was utilized to detect its shape. The ultrasonic basic frequency was set as 8MHz, ultrasonic pulse was reflected to various direction on the surface of the rugged shape. Then, we picked the lower shape of the slug, emphasized in red line, as the target shape. We compared the reconstructed shape results before (Figure 0.7) and after using aperture synthesis method (Figure 0.8).

Figure 0.6 The slug for fuel debris shape reconstruction experiment. [33]

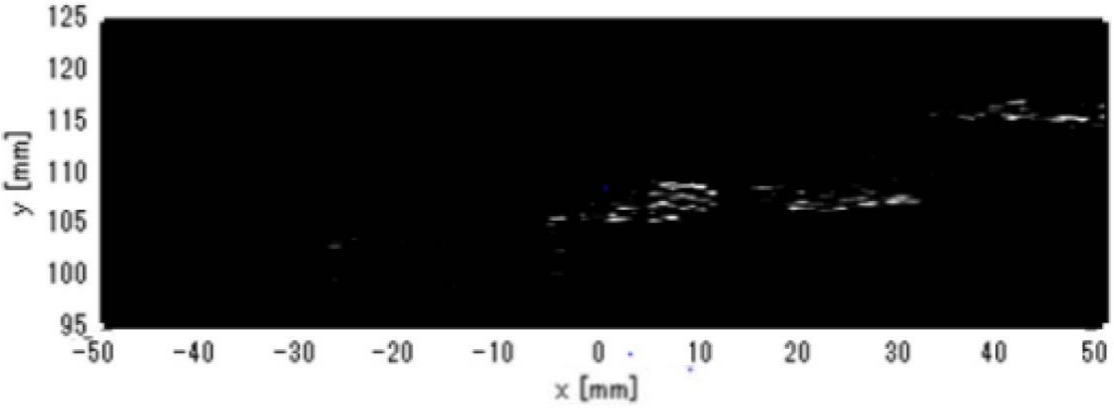


Figure 0.7 The reconstructed shape before using aperture synthesis.

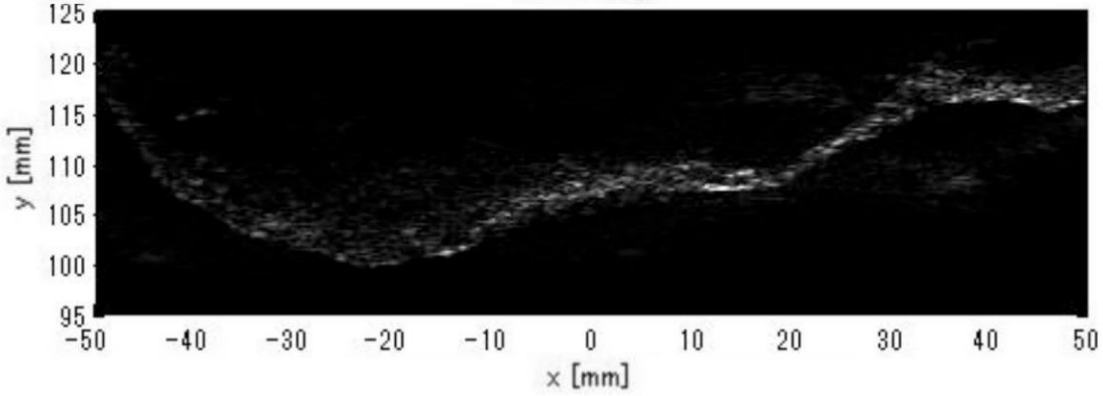


Figure 0.8 The reconstructed shape after using aperture synthesis.

Through the comparison, we could confirm that the obtained shape without aperture synthesis was not continuous, while the aperture synthesis method did help receive the reflected signals with the array elements to reconstruct clear and continuous shape [33].

Appendix B

Simulation Interface for Robot Structure Description

In order to provide intuitive representation of the robot, a simulation interface was implemented to describe the structure of the robot. The popular visualization toolkit Rviz was selected as the primary 3D display tool. The simulation interface uses Unified Robot Description Format (URDF) to describe the robot structure, which usually defines an articulated robot with elements of rigid links and joints. In the case of the BWD robot, the configuration can be simplified into a chain of base link, and the subsequent three links, as Figure 0.1 shows. In terms of the visual form, the 3D model files of each links will be remapped to the URDF links. While, Joint 1 is a simple rotatable joint actuated by a single motor through a timing belt and pulley; Joint 2 and Joint 3 are actuated by two motors through antagonistic tendons with a new relay mechanism, whose working principle needed to be defined in the interface.

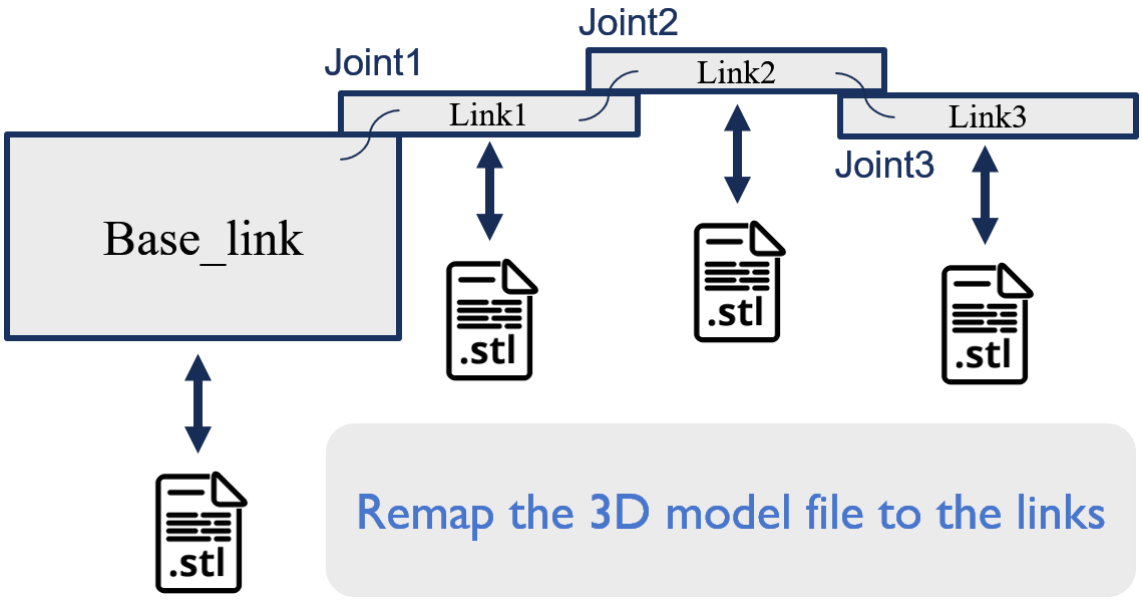


Figure 0.1 The simplified configuration of BWD robot.

Before defining the newly proposed BWD joints with antagonistic tendons, I referred to the available transmission mechanisms in ROS. Currently, in the framework of ROS, the conventional transmission mechanisms can be divided into 3 types, listed below:

- Simple Transmission Type. The simplest transmission types that one actuator controls one joint with timing belt or gears, as Figure 0.2 shows. In the situation of the 3 DoF BWD robot arm, the transmission relationship between the joint 1 and the motor 1 can be exactly described with this transmission.

- Differential Transmission Type. This transmission relates two actuators and two joints through a differential mechanism.
- Four-bar-linkage Transmission Type. This transmission relates two actuators and two joints through a mechanism in which the state of the first joint only depends on the first actuator, while the second joint depends on both actuators.

Obviously, the BWD joint does not belong to any of them. While, to define a new transmission type with URDF description, the most important points lie on the corresponding relationships of the positions, velocities, and torques of the paired joint and the actuator. When driving one joint with two antagonistic ropes, one motor winds up the rope and the other one feeds it. In order to avoid ropes slack, the winding/feeding speeds of two antagonistic ropes are set equal. Here, θ_j stands for the rotation angle of the joint, θ_{mw} indicates the rotation angle of the winding up motor, and θ_{mf} indicates the rotation angle of the feeding motor. D_{motor} and D_{joint} indicate the diameters of the pulley of the motor and joint. The relationship between them can be simplified as Equation A.5 shows.

$$\theta_j = \frac{D_{motor}}{D_{joint}}\theta_{mw} + \frac{D_{motor}}{D_{joint}}\theta_{mf} \qquad (A.5)$$

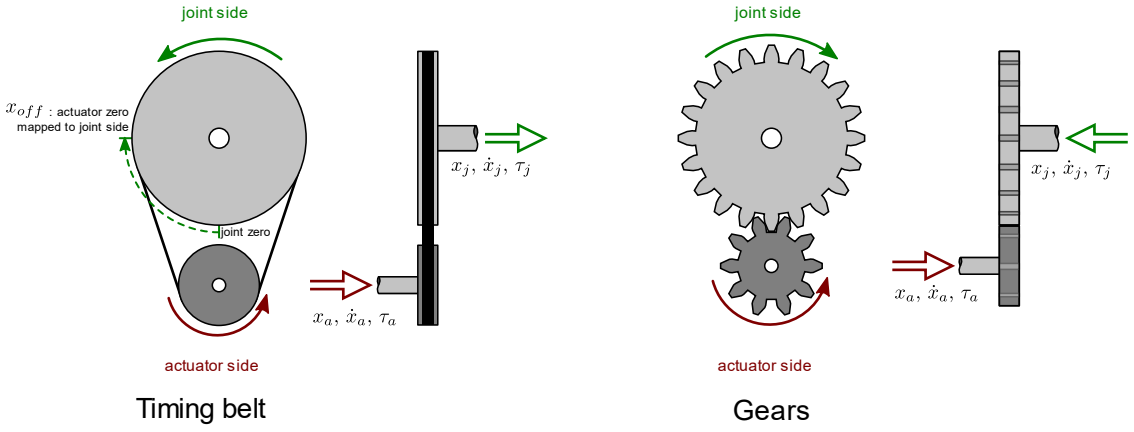


Figure 0.2 The simple transmission type in ROS.

Therefore, the transmission relationship of the two motors and one joint can be described by modifying the simple transmission type, from one motor to one joint relationship. In terms of the new transmission type, the corresponding functions which deal with the transformation between the position, velocity, and torque between joint and motors were implemented. In order to ensure the corresponding relationship between two motors and one joint, I referred to the constructor of differential transmission and implemented the number checking function. At last, the implemented new transmission type works as a new transmission plugin in ROS, and the initialization process is listed below. The necessary values included the corresponding actuator name, and its role in the antagonistic relationship, and the mechanical reduction (D_{motor} and D_{joint}).

```
<transmission name="bwd_transmission_1">
 <type>transmission_interface/BwdTransmission</type>
    <actuator name="motor2">
      <role>actuator1</role>
      <mechanicalReduction>-270.4</mechanicalReduction>
    </actuator>
    <actuator name="motor3">
      <role>actuator2</role>
      <mechanicalReduction>-270.4</mechanicalReduction>
    </actuator>
    <joint name="bwd_joint1">
      <offset>0</offset>
      <mechanicalReduction>1.0</mechanicalReduction>
    </joint>
</transmission>
```

Hardware Interface for Communication

On the basis of the newly implemented transmission type of BWD joint, I successfully defined the transmission relationship between the joint and the motor. Then, the following step came to the implementation of hardware interface to communicate with the motors. Usually, the communication should include the readings from encoder, and the commands from controller. The encoder readings will be transferred to the controller manager, which publishes the joint state into a ROS topic named /joint_states. Then, by subscribing this topic, the visualization tool Rviz can display the correct robot pose simultaneously. On the other hand, when remotely controlling the robot, the hardware interface receives the command from controller manager, and sends to the lower controllers. For the special case of BWD robot, the bottom-layer controllers are EPOS2 provided by Maxon. Since EPOS is supported by ROS community, here I utilized the ROS package named eposx_hardware developed by Okada [45]. The package provides the drivers for EPOS under ROS framework, and supports extensions of ros_control, which provides the controllers managers indicated in Figure 4.5. Table 0.1 shows the specification of the hardware interface of BWD robot.

Table 0.1 The specification of hardware interface of BWD robot.

Parameters	Value
Device	EPOS2

Protocol_stack	MAXON SERIAL V2
Interface	USB
Timeout (ms)	500

Controller Manager for Joint Controlling

Thanks to the interfaces provided by eposx_hardware, I can easily get touch with the powerful toolkit ros_control. The controller manager mainly exchanges information with the public resources, joint state interface and joint command interface. It manages both the pose-publishing controller and the command-giving controllers. In the situation of EPOS2, the available controlling methods in EPOS include profile position mode, homing mode, profile velocity mode, position mode, velocity mode, and current mode. Here I defined a profile position controller and a profile velocity controller in eposx_hardware, which corresponds to the profile position mode, and profile velocity mode respectively.

When defining the controllers to control the joint pose, I first referred to the available controlling modes in EPOS. For the profile position mode, it only required the target position in unit of qc. For the profile velocity mode, the main input is the commanded rotational speed in unit of rpm. Therefore, I checked the provided controllers in ROS and chose JointPositionController to fit with the profile position mode. Similarly, I chose JointVelocityController to fit with the profile velocity mode. Although it's possible to define the original controllers or set the PID values for the controllers, these controllers could provide the basic bang-bang control.

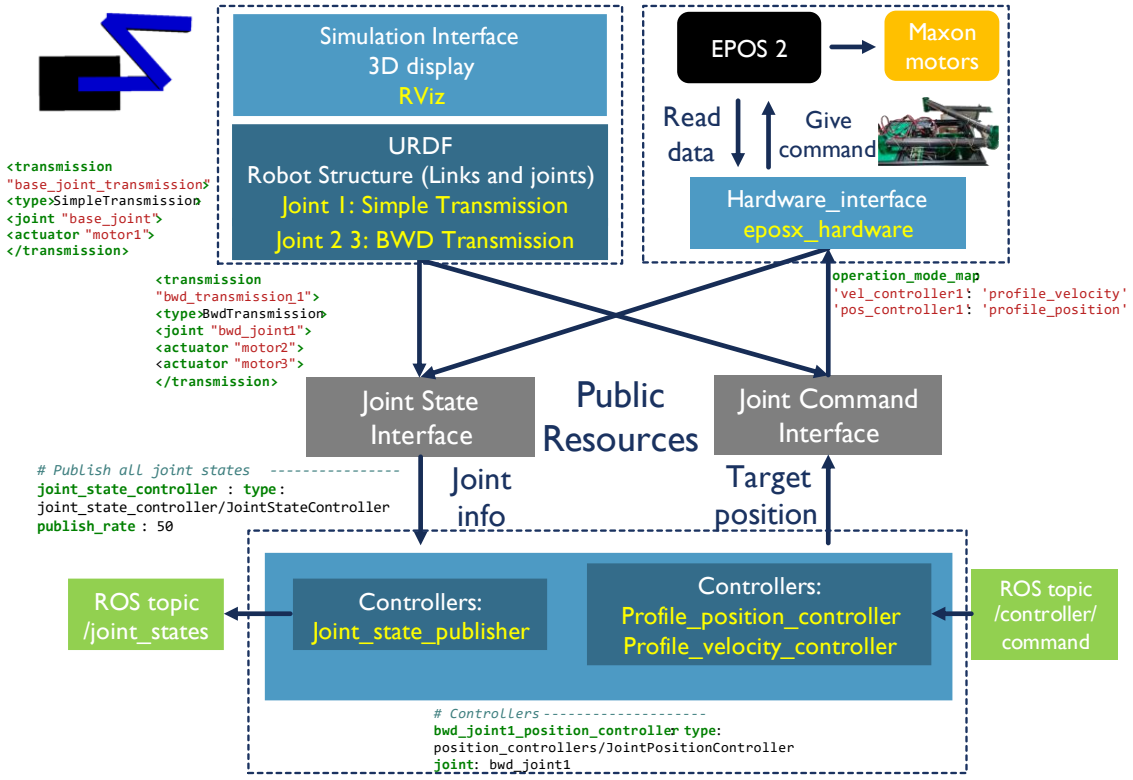


Figure 0.3 The final interfaces for BWD robot.

Figure 0.3 shows the final structure of the customized interfaces for BWD robot in ROS. I utilized RViz as the 3D display tool, and used URDF format to describe the robot structure. By defining the BWD transmission plugin, I could relate the 2 actuators and 1 joint in the URDF file; thus, the instant pose of the joints could be calculated and displayed in RViz. Besides, I utilized the templates of position/velocity controllers provided in ROS and related them to the profile position/velocity modes provided in EPOS. Therefore, all the controllers could be loaded and send commands on the position or velocity of each joint through the command topic.

Appendix C

Singular Value Decomposition (SVD) in Blur Detection

Singular value decomposition (SVD) is one of the most useful techniques in linear algebra, and has been applied to different areas of computer science. Given an image *I*, its SVD can be represented by

$$I = U\Lambda V^T \tag{C.1}$$

where *U, V* are orthogonal matrices and Λ is a diagonal matrix that is composed of multiple singular values arranged in decreasing order. The image can therefore be decomposed into multiple rank 1 matrices (which are also called eigen-images) as follows:

$$I = \sum_{i=1}^{n} \lambda_i u_i {v_i}^T \tag{C.2}$$

Therefore, the decomposed eigen images will have scale degradation when i increases, while the eigen images themselves can reflect the details of the image. The bigger i the eigen image has, the more details in the image. Thus, for the blurry images, the eigen images with huge scale (thus fewer details) will have bigger weight λ_i. Therefore, the weight distribution of the eigen images can be utilized to estimate the degree of blur (degrees of details).

Appendix D

Robustness Degradation Evaluation Experiments under Limited Illumination Conditions

In the main text of this thesis, the core idea of this thesis can be concluded as utilizing environmental and robotic characteristics to improve the performance of environment mapping methods. In other words, the concerned characteristics served as favorable conditions like grating texture with known size and regular shape, or robot's specific motion patterns as prior knowledge. While, apart from favorable conditions, realistic environments also give rise to unfavorable conditions like radiation, low illumination condition, etc. Not only the robustness of the concerned environment mapping method, but also that of the concerned camera will be seriously influenced. While, available knowledge of the robustness degradation of mainstream environment mapping methods and commercially-available cameras is still lacking. As a preliminary quantity study on this existing robustness degradation problem, a comparison experiment to evaluate the robustness degradation of mainstream environment mapping methods and commercially-available cameras was conducted, in order to provide (1) reference robustness degradation data for future countermeasure development, and (2) prior knowledge of commercially-available camera performance degradation for specialized camera development. The concerned environment mapping methods are concluded with Table 0.1. The camera utilized in this experiment is the same camera RealSense D435i in the main text of this thesis.

Table 0.1 Concerned environment mapping methods

Method	Software	Mode
SfM+MVS	Colmap + colmap_rob	Dataset creation method implemented
SLAM	RTAB-Map	RGB-D
		Stereo IR + IMU

When conducting under the dark environment, the SLAM method RTAB-Map has been selected as a competitor for SfM with proposed dataset creation, which is able to enhance its robustness by using depth sensor and Inertial Motion Unit (IMU) input. Besides, progresses on dealing with illumination-invariant environment based on the framework of RTAB-Map have been published.

As for the experimental setup, the same experimental setup in Section 5.3 was used. On its basis, two more illumination condition variations were introduced, so the illumination conditions included three variations.

- Variation 1: The first variation is bright.
- Variation 2: The second is dark.
- Variation 3: A 15-W light-emitting diode (LED) illumination device with a maximum illumination of 1200 lm was mounted on J3 of the arm. Therefore, it could follow the rotation of J3 to provide tracking illumination, as Figure 0.1 shows. A local light spot was observed on the illuminated plane.

Along the boundary walls of the environment, 18 points with 0.6-m intervals were selected, and an illumination meter was used to measure the average illumination levels in each variation (Figure 0.2). The measured average illumination at selected points along the boundary walls are listed in Table 0.2. Similar to the environment setup and the planned camera trajectory defined in Section 5.3, a RealSense D435i camera was used to capture videos during movement. Then, the image sequences will be processed by three different methods' setup as comparison.

Figure 0.1 Illumination condition variation 3 with LED illumination device mounted on J3.

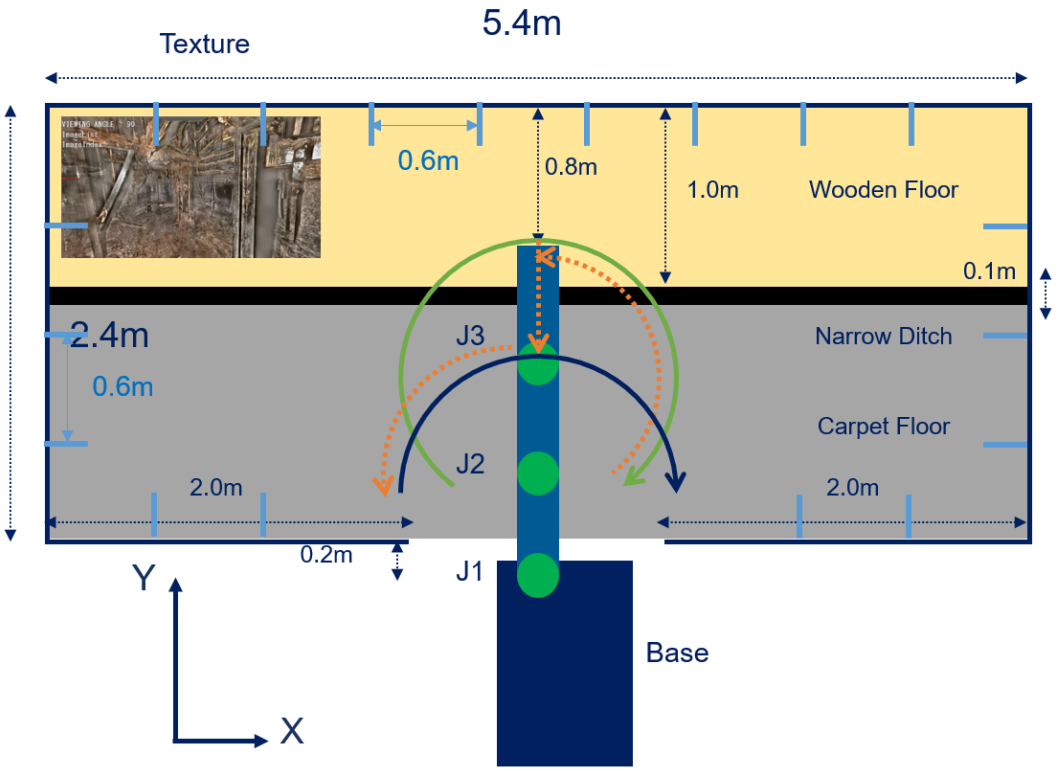


Figure 0.2 Distribution of illumination measurement points.

Table 0.2 Average illumination measured by 18 points along the boundary walls in each illumination condition variation.

Illumination Condition	Average Illumination
V1: Bright	423.3lx
V2: Dark	4.5lx
V3: Mounted	246.3lx

Similarly, the reconstruction performances will be evaluated by using the same benchmarks and ground truth laser scanner. As Table 0.3 shows, the generated point clouds of the concerned method setups under three different illumination condition and three environment setup patterns are listed. The corresponding reconstruction performances evaluated by the benchmarks of *Accuracy* and *Completeness* are listed in Table 0.4, Table 0.5 and Table 0.6.

In bright conditions, SfM algorithms are able to reconstruct environments with high levels of detail and relatively dense point clouds, although the completeness of these reconstructions is often limited, particularly on the ground floor where there may be a lack of feature points for triangulation. In contrast, when mounted illumination is used, SfM suffers from a significant drop in completeness due to limited illumination. This is reflected in the loss of details and degradation of point cloud density, as well as the appearance of double vision problems in some areas (Pattern

2, the cylinder in corner). The inexistent corners led to more accuracy drop in this pattern (88.7% to 33.1%). The relative average accuracy drop of SfM was 43%, and 42% for completeness drop. In contrast, SLAM algorithms are able to generate point clouds with more information from the ground floor, but the overall density of these point clouds limits completeness. In dark conditions, only the use of stereo infrared (IR) sensors with an inertial measurement unit (IMU) is able to provide reconstruction results, thanks to the inner structural light provided by the stereo IR sensors. However, the resulting point clouds are significantly less dense and detailed, leading to a significant drop in completeness (43.6% to 13.4%). When a limited illumination condition is created using a mounted LED device, RGB-D sensors tend to produce maps with less noise, particularly around boundary walls, compared to those produced using stereo IR+IMU. This is likely due to the relatively accurate depth estimation provided by the depth sensor and the additional color detail provided by the RGB camera. However, RGB-D sensors are unable to work in completely dark conditions due to the lack of visual features, while stereo IR cameras are able to work in these conditions due to the internal structural light they create. RGB-D combination suffered from relative average accuracy degradation at 30%, and 15% in completeness. For stereo IR+IMU, the relative average accuracy drop was 30%, and 20% in completeness.

Table 0.3 Point clouds generated by three methods under different illumination conditions and environment setup patterns.

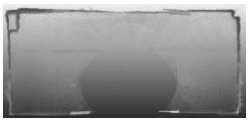
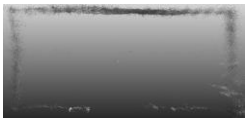
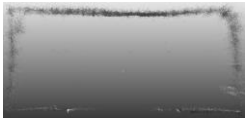
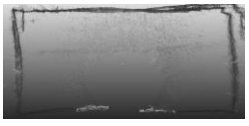
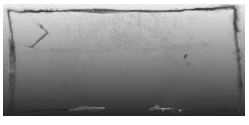
Condition	**Setup**	**Colmap + colmap_rob**	**RTAB-Map (RGB-D)**	**RTAB-Map (Stereo IR + IMU)**
Bright	Setup Pattern 1			
	Setup Pattern 2			
	Setup Pattern 3			
Dark	Setup Pattern 1	Failed	Failed	
	Setup Pattern 2	Failed	Failed	
	Setup Pattern 3	Failed	Failed	
Mounted	Setup Pattern 1			
	Setup Pattern 2			
	Setup Pattern 3			

Table 0.4 Performance comparison of SfM+MVS under different illumination conditions and setup patterns.

SfM + MVS	Metric	Bright	Dark	Mounted
Setup Pattern 1	Accuracy (%)	88.6	-	58.7
	Completeness (%)	60.9	-	38.9
Setup Pattern 2	Accuracy (%)	88.7	-	33.1
	Completeness (%)	60.1	-	36.1
Setup Pattern 3	Accuracy (%)	90.7	-	60.1
	Completeness (%)	64.1	-	30.7

Table 0.5 Performance comparison of RTAB-Map (RGB-D) under different illumination conditions and setup patterns.

RGB-D	Metric	Bright	Dark	Mounted
Setup Pattern 1	Accuracy (%)	76.7	-	62.1
	Completeness (%)	50.4	-	40.6
Setup Pattern 2	Accuracy (%)	74.9	-	50.7
	Completeness (%)	44.2	-	42.4
Setup Pattern 3	Accuracy (%)	74.1	-	44.6
	Completeness (%)	49.1	-	38.9

Table 0.6 Performance comparison of RTAB-Map (Stereo IR + IMU) under different illumination conditions and setup patterns.

Stereo IR + IMU	Metric	Bright	Dark	Mounted
Setup Pattern 1	Accuracy (%)	70.1	45.1	50.1
	Completeness (%)	43.6	13.4	37.1
Setup Pattern 2	Accuracy (%)	64.9	42.6	46.7
	Completeness (%)	38.4	18.2	30.2
Setup Pattern 3	Accuracy (%)	64.7	47.4	41.6
	Completeness (%)	49.8	9.1	37.9

As the conclusion of this experiment, these comparison results under more realistic environmental conditions may provide insights for improving the safety and robustness of vision-based reconstruction methods in future decommissioning development, and serve as a reference data of the performance of commercially available camera for specialized camera development.