RESEARCH ARTICLE

Dataset creation and selection methods with a wire drive flexible manipulator for vision-based reconstruction

Zhenyu Wang[*] , Gen Endo, Hideharu Takahashi and Hiroshige Kikura

Abstract

In order to improve spatial awareness for future investigations of reactor No. 2 at the Fukushima nuclear power plant, it is necessary first to acquire the environment model through reconstruction. To gather images for this task, we have developed a flexible, compact, and lightweight manipulator called the Bundled Wire Drive robot. However, due to mechanism' shortcomings, the feasibility of using this robot is limited by potential degradations of image quality, including odometry deviation and motion blur. Based on the motion characteristics of the robot, we have proposed a dataset creation and selection method to mitigate the impact of these degradations. The effectiveness of this method was verified through experiments with a hardware prototype robot, which demonstrates that it is possible to avoid the influence of matched joint movement deviation by using overlapping simple trajectories; and pre-filtering out blurry images, which usually concentrate on the beginning and stopping periods. Additionally, we conducted a robustness study of mainstream reconstruction methods under limited illumination conditions to quantitively study the performance degradation in a more realistic environment.

Keywords Fukushima Decommissioning, Wire drive mechanism, Image processing, Robustness comparison, SfM, SLAM

Introduction

The decommissioning tasks upon the follow-up treatment of the Fukushima Nuclear Power Plants are of great significance. Over a decade has passed since the great earthquake. However, the decommissioning tasks are still at an early stage due to the extremely complex environments inside the reactors. Especially, the situation of each reactor, according to the degree of damage in the earthquake, varies from each other greatly. Therefore, the core problems lie on lacking prior information and experience in such unique environments. To gather these information, preliminary internal investigations conducted with remotely-operated robotic platforms have been widely studied and carried out.

The decommissioning tasks of the follow-up treatment of the Fukushima nuclear power plants are of great significance. However, three reactors were damaged, and the internal situations of these reactors are different because of the degree of damage, making the decommissioning tasks extremely complex. Therefore, the core problems lie in the lack of prior information on and investigation experience of the unique environments. To contribute to the preliminary information-gathering processes, studies on remotely operated robotic platforms and the necessary environment reconstruction methods have been gradually gaining attention.

For the reactor No. 2, Tokyo Electric Power Company Holdings (TEPCO) and the International Research Institute for Nuclear Decommissioning (IRID) have been planning the further investigations. Similar to previous

*Correspondence:
Zhenyu Wang
wang.z.ak@m.titech.ac.jp
Department of Mechanical Engineering, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-Ku, Tokyo 152-8550, Japan

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 2 of 18

investigations, the follow-up investigations will use one of the through holes in the containment vessel (X-6 penetration) as the robot entrance. Therefore, passing through the X-6 penetration can be regarded as the preliminary condition for the robotic platform, which gives rise to new requirements according to the access strategies.

To pass through the X-6 penetration, IRID utilizes a long-reach articulated arm with precise operational controllability considering the narrow size of the penetration. In this article, we will call this arm as "Veolia-MHI Arm" [2]. The arm structure includes a carriage part, several expandable links, and a telescopic front end (Fig. 1). The specifications of the prototype access device are listed in Table 1 [1].

Most of the actuators are installed on the joints of the arm and link materials with stainless steel and aluminum, making the arm quite heavy. Currently, experiments on passing through the mock-up X-6 penetration and measurements of the deflections of the access device are in progress [3]. Experiments have shown that weight can lead to deflections, component deformations, and positional misalignments. The high joint stiffness guaranteed precise position control while making the arm relatively vulnerable to collisions. In addition, the size requirements owing to X-6 penetration limit the available size of actuators on the joints, which may not provide sufficient torque in some cases.

Compared with the conventional manipulator structure, which installs actuators on joints and focuses on precise motion controllability, the manipulator structure with a wire drive mechanism has all the actuators installed on the base, realizing a compact structure with extremely low weight.

Therefore, our research group have proposed the concept of a Bundled Wire Drive (BWD) mechanism using low friction synthetic fiber ropes [4]. Compared with conventional metal wire ropes, which usually avoid sliding contact and collisions, the BWD mechanism uses multiple bundled synthetic fiber ropes allowing sliding contact. Figure 2 shows the prototype model of the

Table 1 Specification of the prototype "Veolia-MHI Arm"

Metrics	Value
Length	18 m
Weight	4.6 ton
Number of axis	18

horizontally expandable manipulator with the BWD mechanism.

To enhance the spatial awareness, vision-based reconstructions of an unknown environment like Fukushima reactor usually aim to acquire a comprehensive model, such as a 2D grid map or 3D point cloud of the surrounding environment based on the images captured by cameras installed on the investigation manipulator. As a common problem of wire-drive mechanism, friction and elasticity of rope limits the motion controllability of a wire-drive manipulator, possibly resulting in image quality degradations when creating image datasets for vision-based reconstruction. The image dataset quality degradation problem formed restrictions on the feasibility of using such soft wire-drive arms in this context.

Therefore, in order to mitigate the detrimental effects of image quality degradations like robot odometry deviation and motion blur, while preserving the soft feature of the arm, we proposed two image dataset improvement methods based on the motion features of the wire-drive manipulator. (1) A dataset creation method by trajectory design. (2) A data image selection method by transition removal and blur filtering. Both of them work as the preprocessing of Structure from Motion (SfM), which is a photogrammetry technique to generate 3D point cloud reconstruction with other post-processing techniques.

- The dataset creation method works by utilizing overlapping pure rotation of single joint, and using rapid matched joints' movements (transition) to adjust the viewpoint and workspace. The method makes it feasible for the dataset to discard images
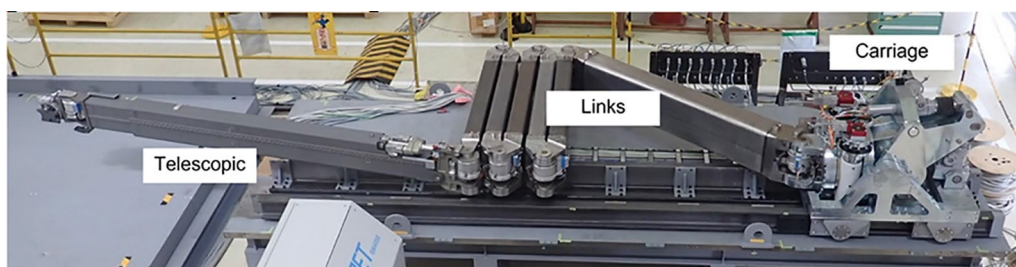
Fig. 1 Structure of the prototype "Veolia-MHI Arm" [1]

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 3 of 18
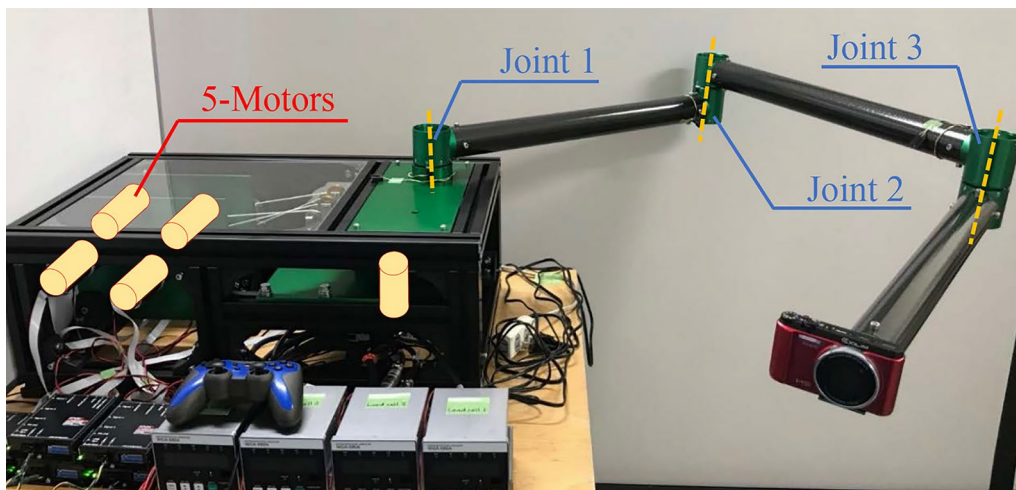
Fig. 2 The prototype bundled wire drive robot with three horizontal expandable Degree of Freedom [4]

from matched joints' movement with low image quality while maintaining a continuous viewpoint trajectory.

- The data image selection method discards the images from matched joints' movement as transition and filters out the images period with high blurry degrees automatically by using a specific quantity *D*. This approach is particularly well-suited to wire-drive arm movements, as the blurry peaks tend to concentrate at the beginning and stopping periods of a motion pattern.

The aim of this article is to verify the effectiveness of the proposed dataset creation and selection methods through experiments with the hardware prototype; And, provide quantitively available reference data for decommissioning researchers on the reconstruction performance degradation of mainstream methods like SfM and SLAM under a more realistic environment, by conducting robustness experiment with limited illumination condition.

This article is organized as follows. In Section II, we presented the core features of BWD robot platform, and discussed the existing problems of the mechanism. In Section III, we quantitively evaluated the degree of the possible influences due to BWD problems, including robot odometry deviation, and potential motion blur, and proposed the corresponding dataset improvement methods. Section IV presented the feasibility and effectiveness verification experiments of the proposed dataset improvement methods and the preliminary robustness study under low illumination condition. Section V concluded the article and discussed the future works.

Bundled wire drive robot prototype

Figure 3 shows the arrangements of tendons, holes, and arm pulleys, which enable all the actuators to be placed on the base, and also shows the potential to implement an arm with more degrees of freedom. Joint 1 is a simple rotatable joint actuated by a single motor through a timing belt and pulley. Joints 2 and 3 are actuated by two motors through antagonistic tendons. One motor wind up the tendon, and the other motor feeds the tendon, forming an antagonistic relationship.

Table 2 lists the specifications of the prototype robot. Owing to the wire drive mechanism, all the motors were installed in the base structure, making the arm part only 1.8 kg. The key feature of the BWD mechanism is that multiple wires are bundled and deployed on the same pathway, which may introduce interference between bundled wires (Additional file 1). Besides, the available antagonistic tension is relatively low, therefore providing relatively low joint stiffness. Therefore, the motion controllability will get degraded, leading to more stabilization
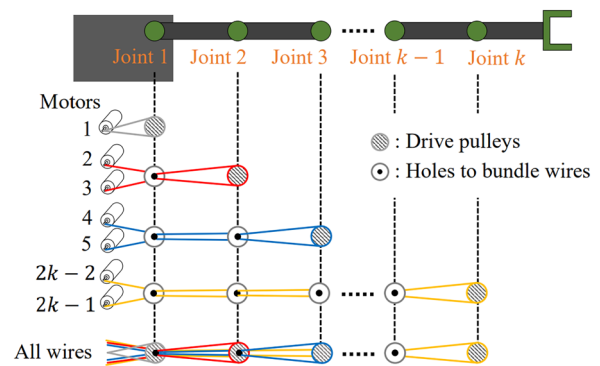
Fig. 3 Arrangements of tendons, holes, and pulleys of the arm [4]

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 4 of 18

Table 2 Specification of the prototype bundled wire drive robot [4]

Metrics	Value
Length	1.8 m
Diameter	50 mm
Weight	33.3 kg (Arm:1.8 kg)
Number of Axis	3

time or overshoot. Considering the collision vulnerability of the tough Veolia-MHI arm with high joint stiffness, we aimed to cover the influences while keeping the soft feature. In the next section, we would like to discuss the influences on vision-based investigations and propose dataset improvement methods from aspects of creation and selection.

Image quality degradation and improvement methods proposal

The performance of vision-based investigations is influenced by various factors. In this section, the discussion proceeds according to the dual problems logic proposed by Quattrini [5]. The vision-based investigations can be divided into the dual state estimation problems of tracking the camera trajectory and reconstructing the environment. The degree of the influences on image quality will be quantitively studied through evaluation experiments. Then, the dataset creation and selection methods will be proposed on the basis of both the evaluation results and the unique motion patterns of wire-drive manipulator.

Robot odometry deviation evaluation

Robot odometry is an intuitive and quick method for estimating the end effector trajectory based on joint angles and forward kinematics. When installing the investigation camera on the end effector, the problem of tracking the camera trajectory can be solved by referring to robot odometry. This is particularly true when the investigation sensor cannot localize itself with its data. However, such problems as link deflections and joint flexibility often occur with a long-reach articulated manipulator, possibly leading to deviations in robot odometry. The deviation usually does not influence the SfM image registration processes because it does not rely on prior odometry input. While, there is prior information available from robot motion patterns, such as trajectory continuity, which can be utilized to refine the SfM processes.

Because the image sequences are from the camera installed on the end effector of the robot, the trajectory should maintain the continuity feature, as in the general motion camera model [6]. However, image registration processes might not maintain the continuity feature

because the extracted features are weakly localized [7]. Before implementing the refinement method with robot odometry, it is necessary to evaluate the degree of deviation.

To evaluate the robot odometry deviation of the BWD robot, the end effector trajectory obtained by robot odometry was compared with the ground truth trajectory. An optical motion capture system with motion analysis markers was used to measure the trajectory. It could achieve accuracy within 1 mm in a 3- × 3- × 3-m cube space and a response speed within 3 ms. Considering the possible interactions between bundled ropes, four motion patterns were employed in the experiment. The results of the comparison of end effector trajectories are shown in Fig. 4. The figure defines the direction from the base to end effector J3 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. Based on the coordinate system defined above, four patterns were identified (Additional file 4):

- Pattern 1: Pure rotation of J3 within its maximum range from −160° clockwise to +160°. The average angular velocity is 20 deg/s.
- Pattern 2: Rotating J2 to −90°, then rotating J3 from −140° clockwise to +140°. The same angular velocity is used.
- Pattern 3: Moving the end effector along the negative Y-direction for 600 mm, which is the length of a single link. The average translation velocity is 33 mm/s.
- Pattern 4: Similar to Pattern 3, a Cartesian translation along the positive X-direction for 1200 mm. The average translation velocity is 80 mm/s.

As Figs. 4 and 5 show, the deep blue line shows the planned trajectory by robot odometry, while the light blue shows the measured trajectory by motion capture. The arm pose in the solid line indicates the starting pose of the motion pattern, while the pose in the dotted line shows the ending pose. Patterns 1 and 2 only include the pure rotation of a single joint, and the comparison shows that there was a steady rise in the deviation with joint rotation. This indicates that the deviation may originate from bundled wire interference. In addition, pattern 2 suffers from more deviation because the ropes of J2 are kept stretched. The figure shows the phenomenal growth of deviation in patterns 3 and 4, where Cartesian translations were conducted with 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
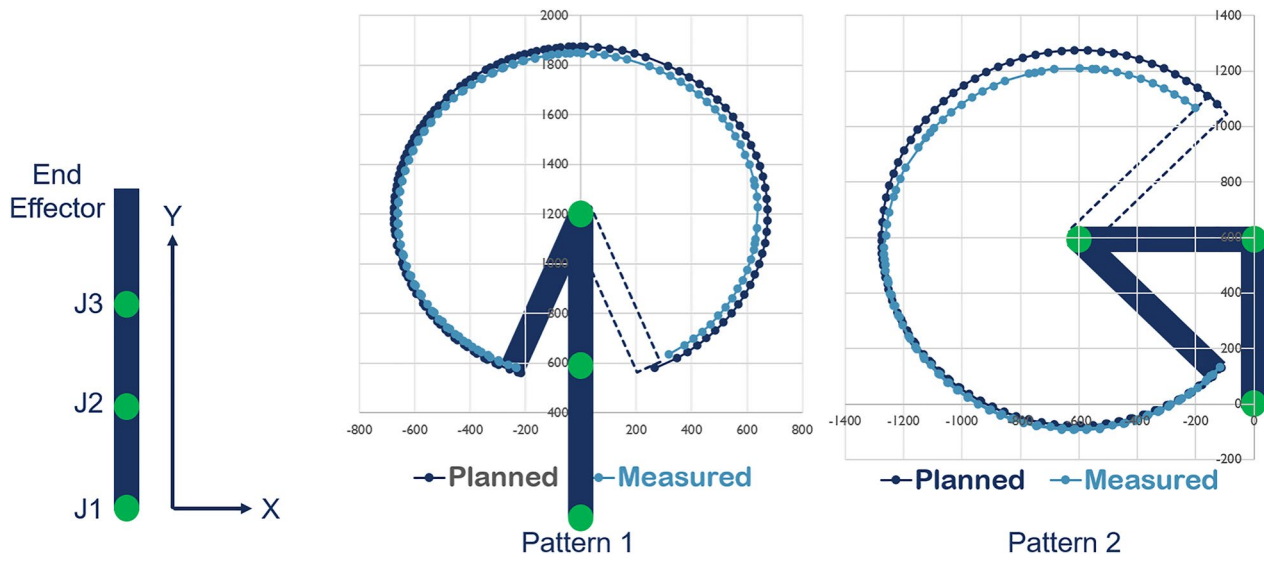
Fig. 4 End effector trajectory comparison results of Pattern 1 and 2. **A** The coordinate system in the image. **B** The result of Pattern 1. **C** The result of Pattern 2
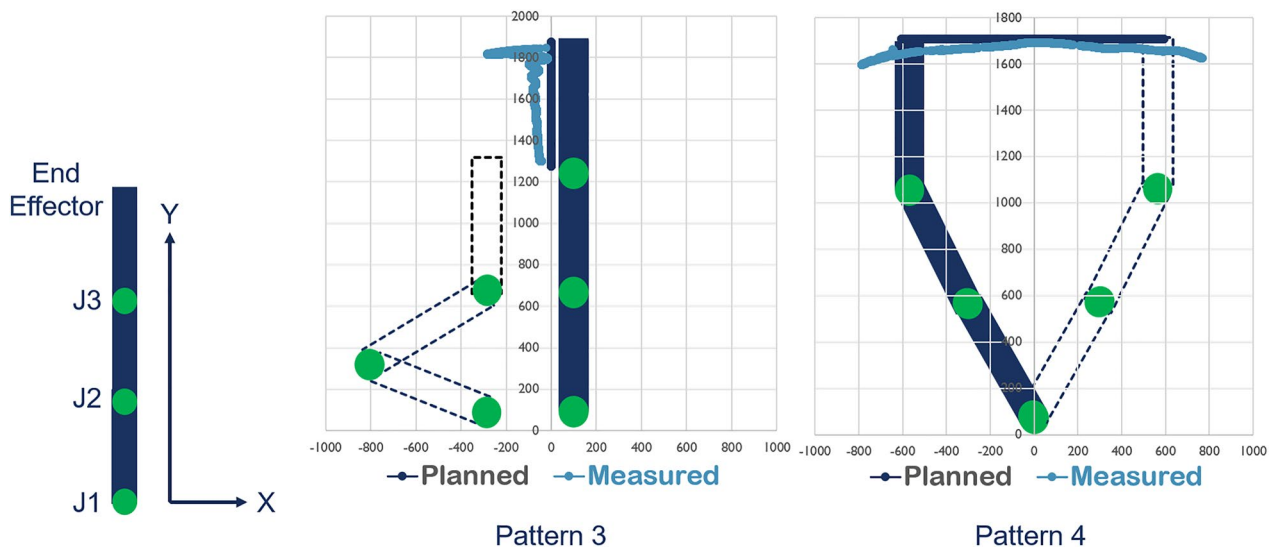
Fig. 5 End effector trajectory comparison results of Pattern 3 and 4. **A** The coordinate system in the image. **B** The result of Pattern 3. The position of the base joint is shifted in the X-direction to increase visibility. **C** The result of Pattern 4

phase or overshooting owing to the relatively long stabilization time. Compared with patterns 1 and 2, the trajectories with multiple joints' simultaneous movement may introduce interference between joints, thus increasing the deviation. Therefore, it is believed that performing pure rotation with a single joint can avoid introducing extra deviation.

Potential motion blur evaluation

As has been discussed, the position stabilization time and overshoot caused by low joint stiffness can influence vision-based investigations. When image sequences from a camera mounted on a flexible manipulator are used, motion blur is the most likely common influence.

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 6 of 18

To quantitively evaluate the degree of motion blur, the quantity D proposed by Su, et al. in 2011 [8] was utilized in this article. This quantity is based on the blurry region detection method by using Singular Value Decomposition (SVD). The quantity D is defined as below:

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

where k is the weight to the concerned two factors. One of the factors, β indicates singular value feature of the image, which has positive correlation with the blurry degree of image; The other one is the ratio of the blur region area (Ω_b) over the whole image area (Ω).

As the experimental method, we mounted a RealSense D435i camera on the end joint of the BWD arm; Then, we gathered images while conducting the same motion patterns utilized in "Introduction" section; At last, we down-sampled the gather images and used the quantity D to evaluate the degree of motion blur according to the motion patterns. The concerned experimental parameters' setup is listed in Table 3.

Figure 6 shows the evaluation result of blurry degree of Pattern 1, and two obvious peaks of high blurriness can be confirmed at the beginning and stopping periods of the designed trajectory. The situation of Pattern 2 has the similar tendency, which indicates that in these motion patterns with pure rotation of a single joint, the motion blur may concentrate in the beginning period and the stopping period. Figure 7 shows the result of Pattern 3, while the blurry period at the beginning of the motion is much longer than Pattern 1, which indicates that multiple joint movements introduce more vibration. Pattern 4 also shows a similar tendency. Reflecting on the features of the BWD mechanism, we can assume that the joints cannot be actuated until the tension reaches the required level to overcome the static friction, which forms a dead zone. And the processing of overcoming the dead zone may introduce an unstable period with high blurry degrees, especially multiple joint movements, bringing more interference between bundled wires. Besides, the low joint stiffness may require more stabilization time, leading to an overshoot at the stopping period.

Table 3 Parameters' setup for following experiments

Metrics	Value
Camera	
Resolution	1280×720
Frame Rate	30
Down-sampled Frame Rate	10
Average Angular Velocity	20 deg/s
Robot Arm	
Average Translation Velocity (X-axis)	80 mm/s
Average Translation Velocity (Y-axis)	33 mm/s

Dataset improvement methods proposal

Through the evaluation experiments on the degrees of odometry deviation and motion blur degree, we can conclude that:
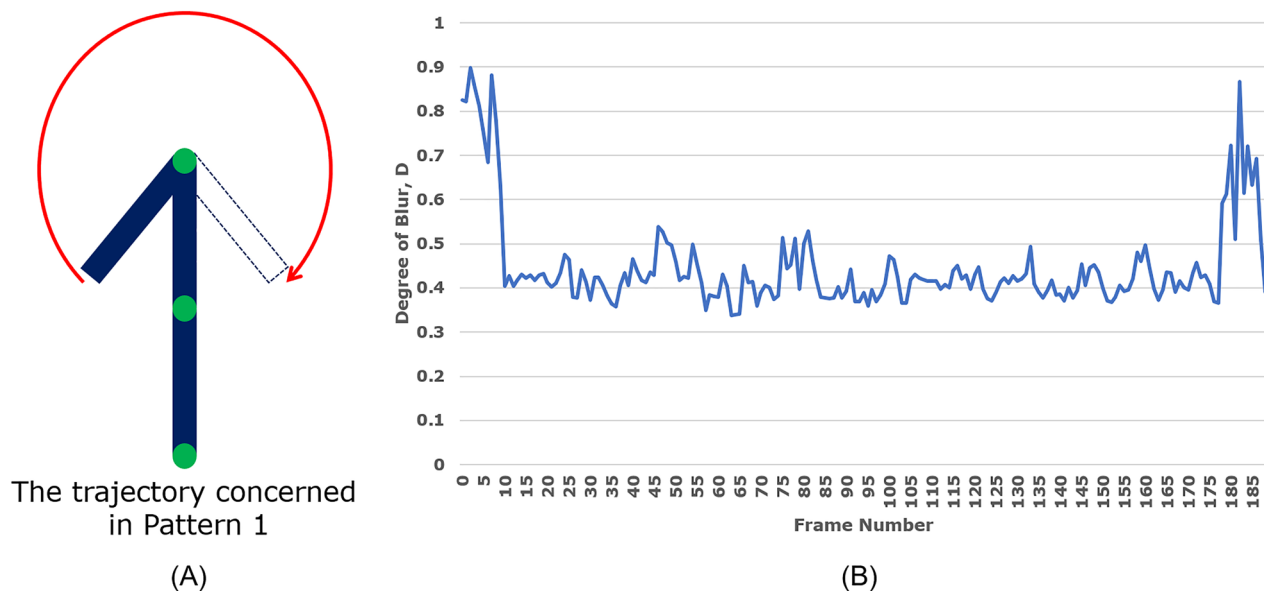
Fig. 6 The degree of blur evaluation result of Pattern 1. **A** The trajectory of Pattern 1. **B** The tendency of blurry degree D according to the frame number

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 7 of 18
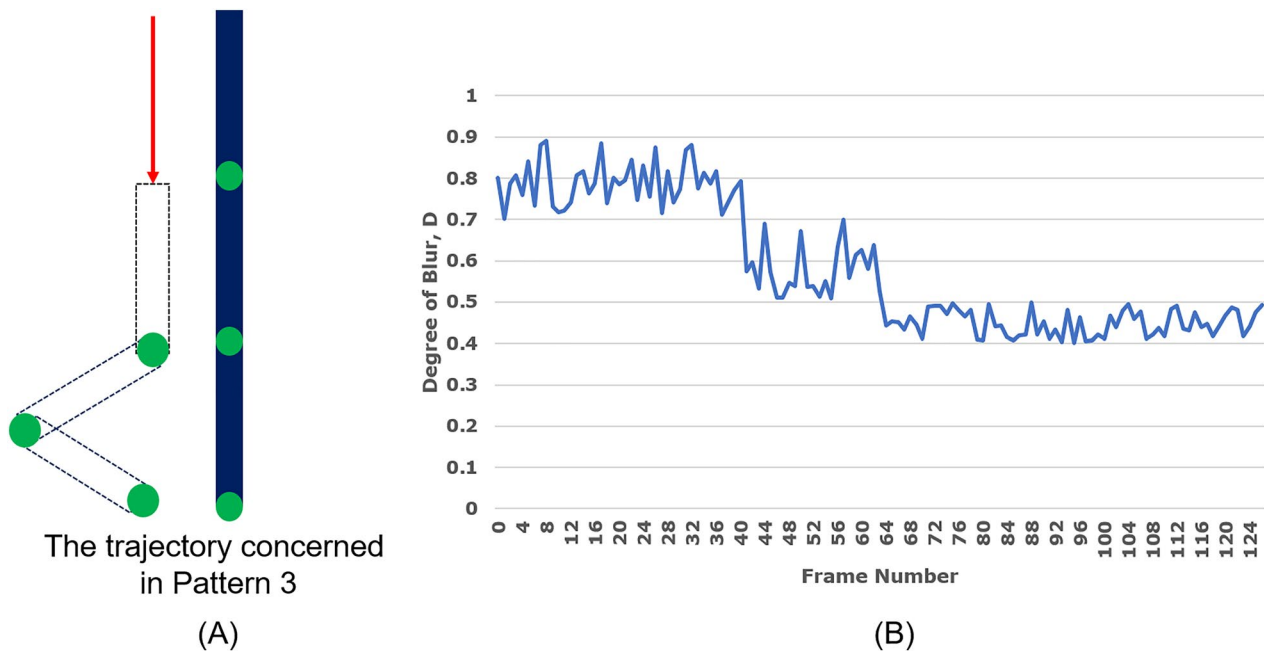
Fig. 7 The degree of blur evaluation result of Pattern 3. **A** The trajectory of Pattern 3. **B** The tendency of blurry degree D according to the frame number


Fig. 8 The reconstruction flowchart with the proposed dataset improvement methods (in orange color)

- The positional odometry deviation may increase when conducting simultaneous matched joint movement, possibly due to the interference between bundled wires.
- The motion blur is usually concentrated on the beginning and stopping periods, possibly due to dead zone and low joint stiffness overshoot. Especially multiple joint movements may extend the blurry periods.

Therefore, the key direction of mitigating the detrimental effects should be avoiding matched joint movement since it introduces more deviation and blurs to image quality.

Thus, two dataset improvement methods are proposed to reduce image quality degradation due to multiple joint movements while keeping continuous viewpoints and compact trajectories. The flowchart of Structure from Motion with the proposed strategies as pre-processing is shown in Fig. 8.

The goal of creating a dataset is often to capture images of an environment from a variety of perspectives, which requires a wide range of viewpoints. Using
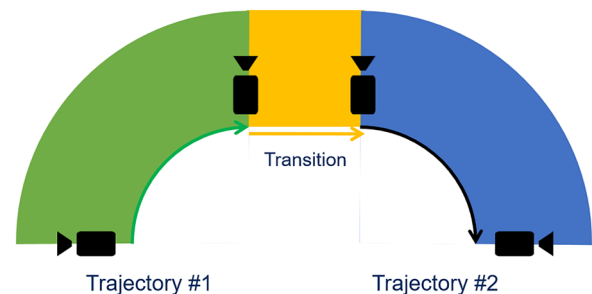
Fig. 9 The conventional trajectory design idea, where the images of transition are necessary for continuity. Green: Trajectory #1; Orange: Transition; Blue: Trajectory #2

pure rotation trajectories alone may result in a limited range of motion and may be time-consuming due to the need to individually move each joint. Figure 9 shows a conventional trajectory design concept, with black camera icons representing the camera poses and colorful rings indicating the visible ranges of the corresponding trajectories. For trajectories #1 and #2, only pure rotation of a single joint is used, resulting in relatively

Wang *et al. ROBOMECH Journal* (2023) 10:4

Page 8 of 18

good image quality. However, the transition movement (orange) uses rapid translation through matched joint movement to adjust the viewpoint position, which may introduce more deviation and blur. Removing the images captured during the transition movement, however, would disrupt the continuity of the trajectory and result in partial reconstruction.

The dataset creation method seeks to discard images captured during unstable motion, such as that produced by matched joint movement, while maintaining a continuous trajectory with a variety of viewpoints. To achieve this, the method designs overlapping pure rotation trajectories by orienting the beginning pose of each subsequent trajectory towards the beginning pose of the previous trajectory (as shown by the red dotted line). This allows for the creation of overlapping ranges of viewpoints among the remaining trajectories, even if the images taken during the transition movement are removed (Fig. 10).

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

After gathering image sequences by conducting the designed trajectory, the image selection method aims to remove the images from the transition movement and mitigate the influences of motion blur. The previous experiments showed that the blur peaks concentrated on the beginning and stopping periods of a motion loop. Therefore, the image selection method takes advantage of the concentration feature to detect blurry periods and conduct fast filtering. The actual process of the image selection method is described below.
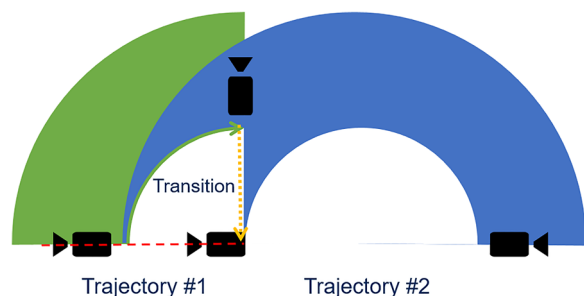
- Clustering the image sequence into remaining trajectories and transitioning according to the data creation method.
- Remove the images from the transition movement.
- Using quantity D to detect the blurry degree in the clustered trajectories, especially the beginning period and stopping period. (Preset period length is the average stabilization time 1 s)
- Finding the image with the lowest blurry degree in the detected blurry periods and filtering the other images.
- Using the rest images as SfM input.

Therefore, the images of the blurry periods which may result from starting vibration and stopping overshoot can be automatically removed, thus mitigating the influences of the image quality degradation on the image sequence.

Experiments

In this section, the experimental targets are:

- Verify the effectiveness of the proposed dataset creation and selection methods on the reconstruction performance of SfM.
- Study the robustness of SfM method with the proposed strategies under a dark environment.

Therefore, the following sections will discuss the related methods and reconstruction performance benchmarks in "Related methods and benchmarks" section, the environment setup, and the concerned trajectory in "Environment setup and trajectory design" section. Then, the verifications will be presented in "Effectiveness verification of proposed image selection method on transition image removal" and "Effectiveness verification of proposed image selection strategy on blur filtering" sections. The robustness study under dark illumination will be clarified in "Robustness study against limited illumination" section.

Related methods and benchmarks

For the proposed strategies designed for SfM, colmap [9] has been selected as the SfM pipeline, as a general-purpose pipeline for camera poses estimation and sparse reconstruction. Then, for the dense reconstruction pipelines, colmap_rob [12] was selected as the Multiple View Stereo (MVS) pipelines, as the conventional combination with SfM. Then, instant-ngp [11] has been selected as the Novel View Synthesis (NVS) pipeline, which is the new technique for view synthesis with Neural Radiance Fields for View Synthesis (NeRF) [10]. Especially it has shown excellent rendering performance and the potential for extremely fast training speed compared

Fig. 10 The concept of the dataset creation method in trajectory design. Green: Trajectory #1; Orange: Transition; Blue: Trajectory #2

Wang *et al. ROBOMECH Journal* (2023) 10:4

Page 9 of 18

with conventional methods. When conducting under the dark environment, the SLAM method RTAB-Map [13] has been selected as a competitor for SfM with proposed strategies, 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 [14] have been published.

To evaluate the reconstruction performance, the commonly acknowledged benchmarks of accuracy, completeness, and runtime proposed by Schops et al. in 2017 are utilized [15]. 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 was 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 Eq. 2:

$$Acc = \frac{N_{acc}}{N_{rec}} \tag{2}$$

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 Eq. 3:

$$Com = \frac{N_{com}}{N_{truth}} \tag{3}$$

The reconstructed point clouds would be evaluated with the benchmarks clarified above, by comparing with the ground truth provided by an external laser scanner. We chose FARO(R) Focus 3Dx130 scanner, whose maximum distance reaches 130 m while the accuracy is within 2 mm.

Environment setup and trajectory design
As Fig. 11 shows, the environment was a rectangular bounding box of $5.4 \times 2.4 \times 1$ m, and the robot arm entered the box through a preserved entry at its bottom center. The upper part of this rectangular box is a wooden floor with a width of 1.0 m, and the rest is covered with a gray carpet. Between these two floors, there is 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
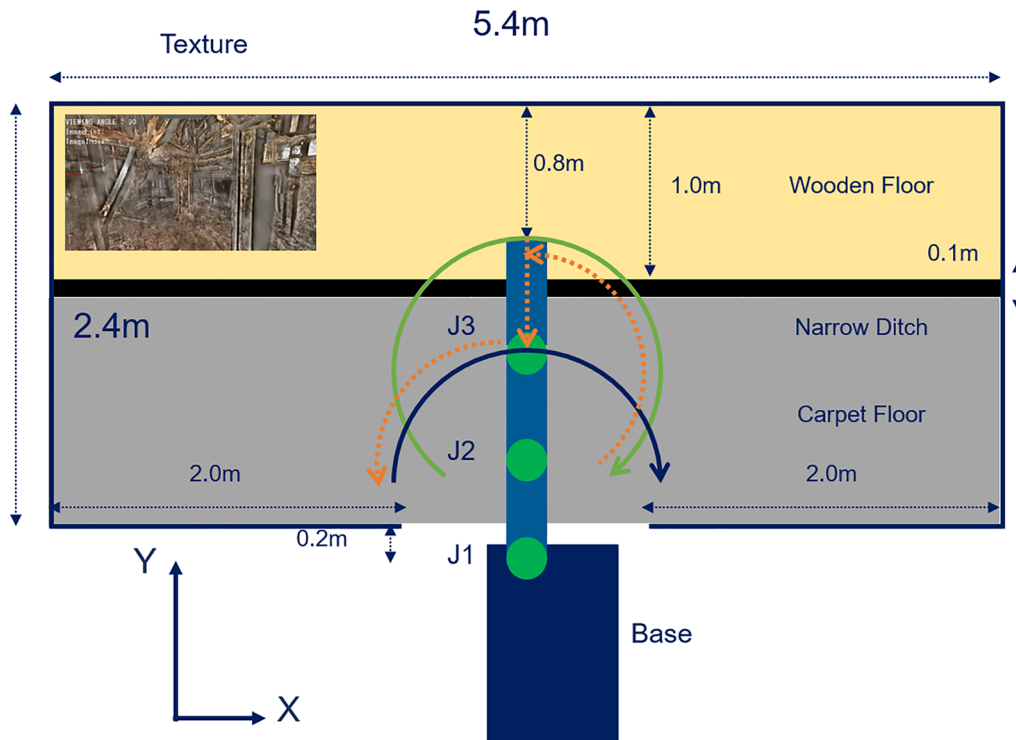
Fig. 11 The environment setup for the reconstruction experiment. Green: Trajectory #1; Orange: Transition; Blue: Trajectory #2

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 10 of 18

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 [15, 16]. These photographs provided the necessary feature points for the SfM processes and also concentrated on the reconstructed points with regular shapes. 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 (Additional file 2). As Fig. 11 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 600 mm, 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. The size of the box was $0.6 \times 0.6 \times 0.6$ m, and that of cylinder was 150 mm (radius) $\times 1000$ 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 Fig. 12.

In each environmental setup, the planned trajectory described above was used. 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 4.

Effectiveness verification of proposed image selection method on transition image removal

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. 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 experiments will compare the reconstruction results obtained using three different datasets, according to the extent to which the image selection method is applied. Figure 13 shows the difference between the concerned three datasets.

As the feasibility verification of the dataset creation method, we will confirm the influences of removing images from transition phase by comparing the reconstruction performance between original dataset and transition removed dataset first.

Then, the reconstruction performance of the sensing methods with these two datasets will be compared, to

Table 4 Run configuration of reconstruction processes

Component	Configuration
CPU	Intel(R) Core (TM) i7-9700 K @ 3.60 GHz
Memory	32 GB
Graphic Port	NVIDIA GeForce RTX 3070
CUDA	11.6
colmap	3.8
OptiX	7.3
Vulkan	Yes
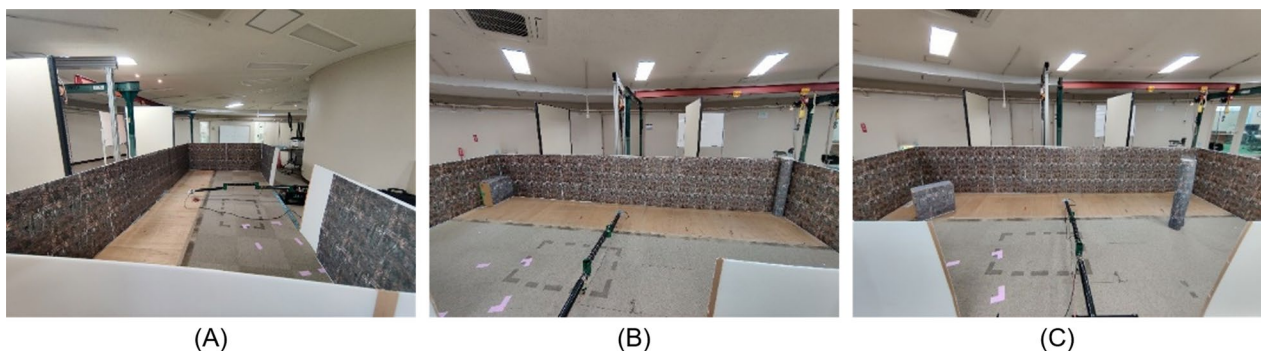
Fig. 12 Three setup patterns for the environment. **A** The default setup pattern 1 without extra objects. **B** Setup pattern 2 with box and cylinder placed in the corners. **C** Setup pattern 3 with box and cylinder placed in specific positions

Wang *et al. ROBOMECH Journal* (2023) 10:4

Page 11 of 18
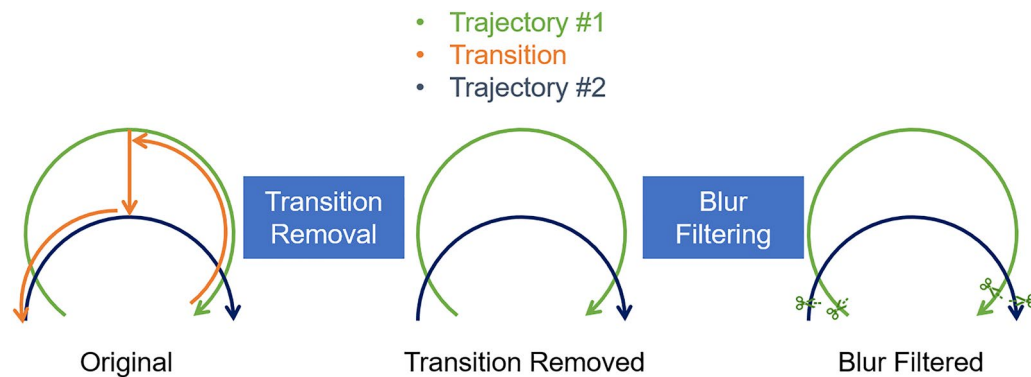
Fig. 13 The concerned datasets comparison

Table 5 Reconstruction performance comparison results of all three setup patterns by colmap_rob before and after removing the transition images

colmap _rob	Metric	Original dataset	Trajectory adjusted
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

Table 6 Reconstruction performance comparison results of all three setup patterns by instant-ngp before and after removing the transition images

instant-ngp	Metric	Original dataset	Trajectory adjusted
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

confirm the influences of the proposed dataset creation method.

Tables 5 and 6 presents the reconstruction performance of colmap_rob and instant-ngp using different dataset as input. In this section, the focus is on confirming the

effects of the proposed dataset processing methods, so, a cross comparison of colmap_rob and instant-ngp is not conducted. As shown in Table 5, 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%. However, the general accuracy and completeness do not decrease, but rather slightly increase, despite the removal of images from the transition phase.

As shown in Table 6, similar results occurred in the case of instant-ngp except for runtime reduction.

Therefore, through the comparison experiments, the results indicate that it's feasible to remove the images from transition movement while having no obvious negative influences on the reconstruction performance by using the overlapping dataset creation method. No obvious difference in the visual reconstruction results is observed after the removal.

Effectiveness verification of proposed image selection strategy on blur filtering

Furthermore, in order to study the effect of the proposed image selection method on blur filtering, this section will conduct the comparison between trajectory adjusted dataset and blur filtered dataset. As Fig. 13 shows, the difference is that the blurry images in the beginning and stopping periods have been detected and filtered out.

Figure 14 presents a comparison of the reconstructed point clouds produced by colmap_rob under three different setup patterns, with the left clouds based on the trajectory-adjusted dataset and the right clouds based on the blur-filtered dataset. There is no notable difference between the reconstruction results. Table 7 indicates that, in general, the impact of the blur filtering step in the image selection method on the reconstruction performance of colmap_rob is relatively small.

In contrast, when using instant-ngp for reconstruction, the presence of motion blur results in fuzzy noise in

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 12 of 18
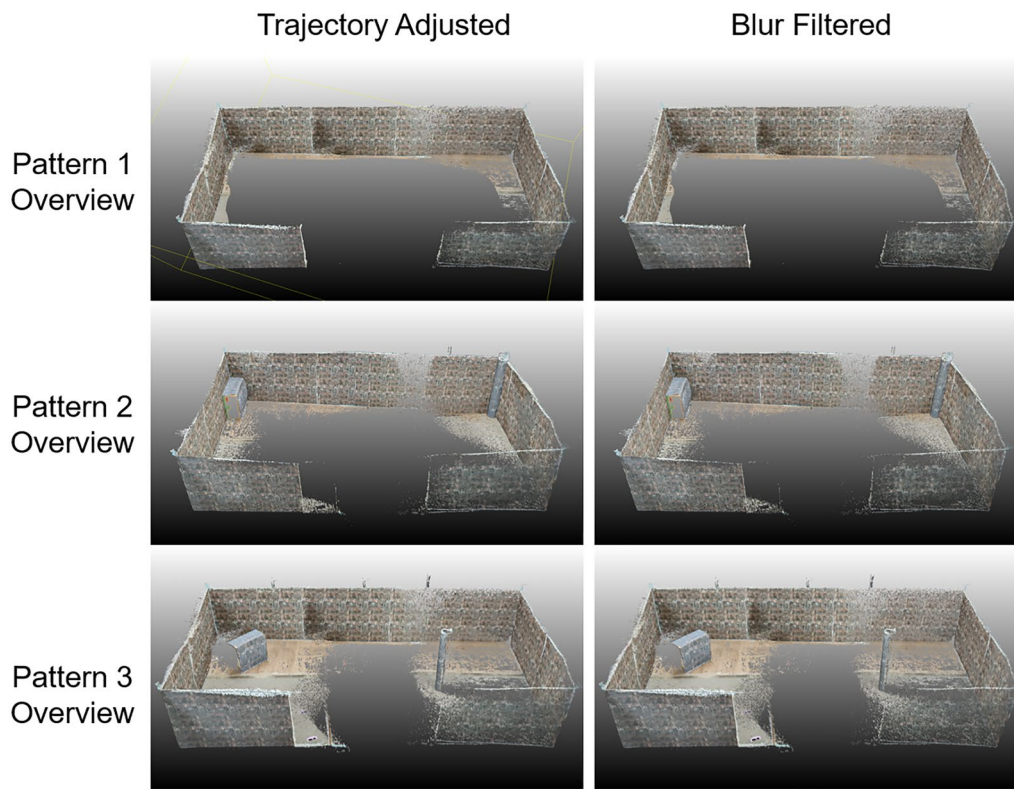
Fig. 14 Reconstruction results by colmap_rob with trajectory adjusted dataset (left) and blur filtered dataset (right) under three setup patterns

Table 7 Reconstruction performance comparison results by colmap_rob before and after blur filtering under three setup patterns

colmap_rob	Metric	Trajectory adjusted	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

the reconstructed results. The proposed image selection method is effective in reducing this noise, particularly around key objects in the environment setup. Figures 15 and 16 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 8.

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.

Robustness study against limited illumination
Till now, the feasibility and effectiveness of the proposed dataset improvement methods have been verified through experiments. In the following part, we would

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 13 of 18
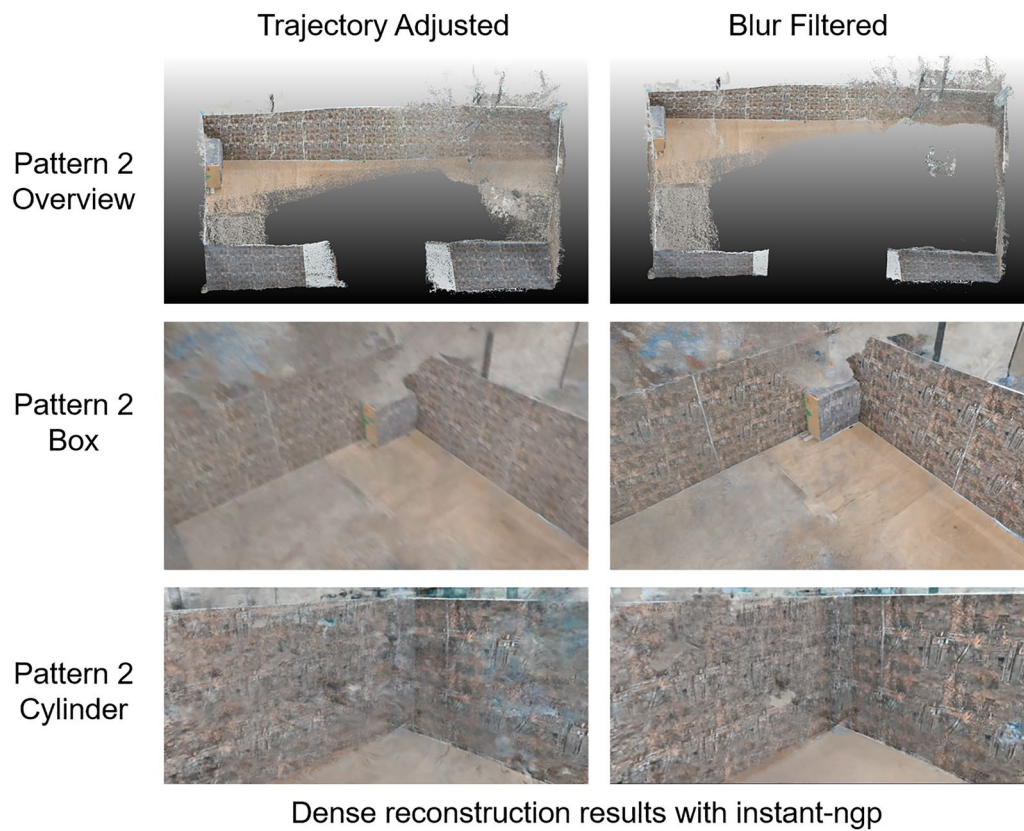
Fig. 15 Reconstruction results by instant-ngp with trajectory-adjusted dataset (left) and blur filtered dataset (right) under setup pattern 2

like to study the robustness of the refined SfM with proposed methods against limited illumination condition, as a simulation of more reaslitic environment to the decommissioning investigation. In the experiment, the same setup patterns were used for the environment discussed in "Experiments" section. On this 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 Fig. 17 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 (Fig. 18). Table 9 lists the average illumination measured at selected points along the boundary walls (Additional file 3).

Similar to the environment setup and the planned camera trajectory defined in "Experiments" section, 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, listed below (Table 10):

Similarly, the reconstruction performances will be evaluated by using the same benchmarks and ground truth laser scanner. As Table 11 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 Tables 12, 13 and 14 (Additional file 2).

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,
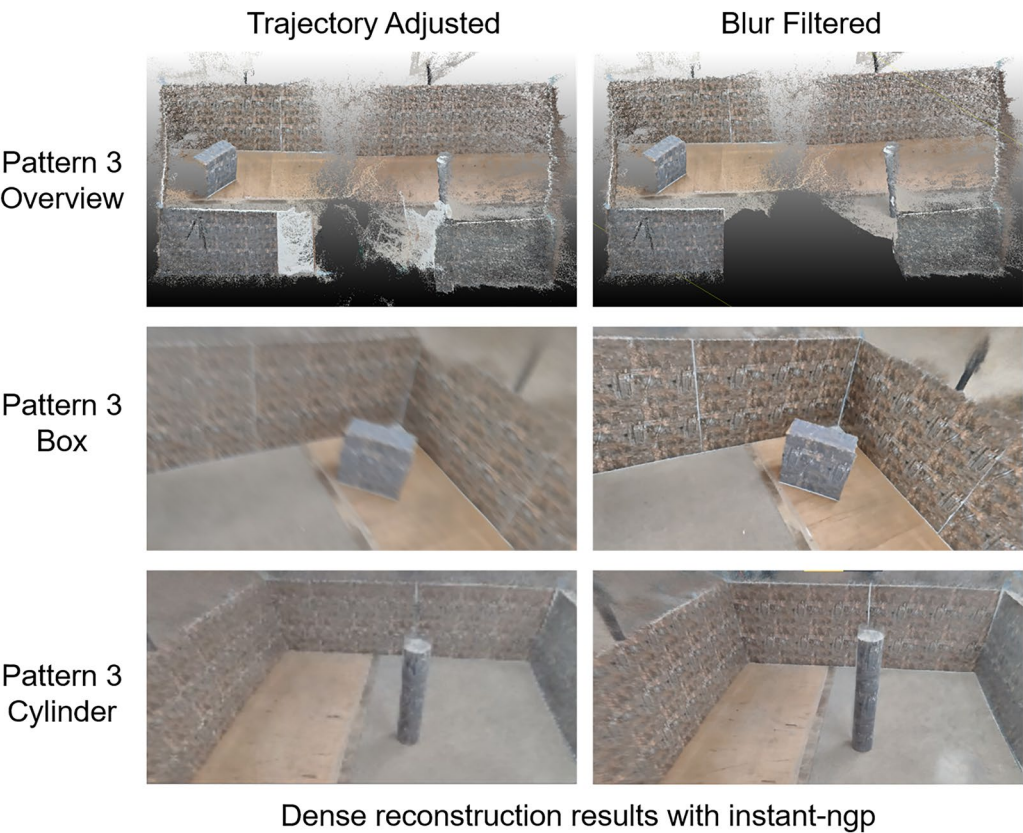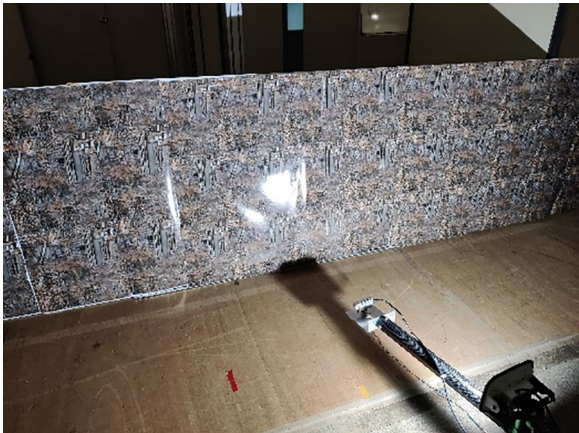
Fig. 16 Reconstruction results by instant-ngp with trajectory-adjusted dataset (left) and blur filtered dataset (right) under setup pattern 3

Table 8 Reconstruction performance comparison results by instant-ngp before and after blur filtering 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

Fig. 17 Illumination condition variation 3 with LED illumination device mounted on J3

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 (Additional file 3).

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

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 15 of 18
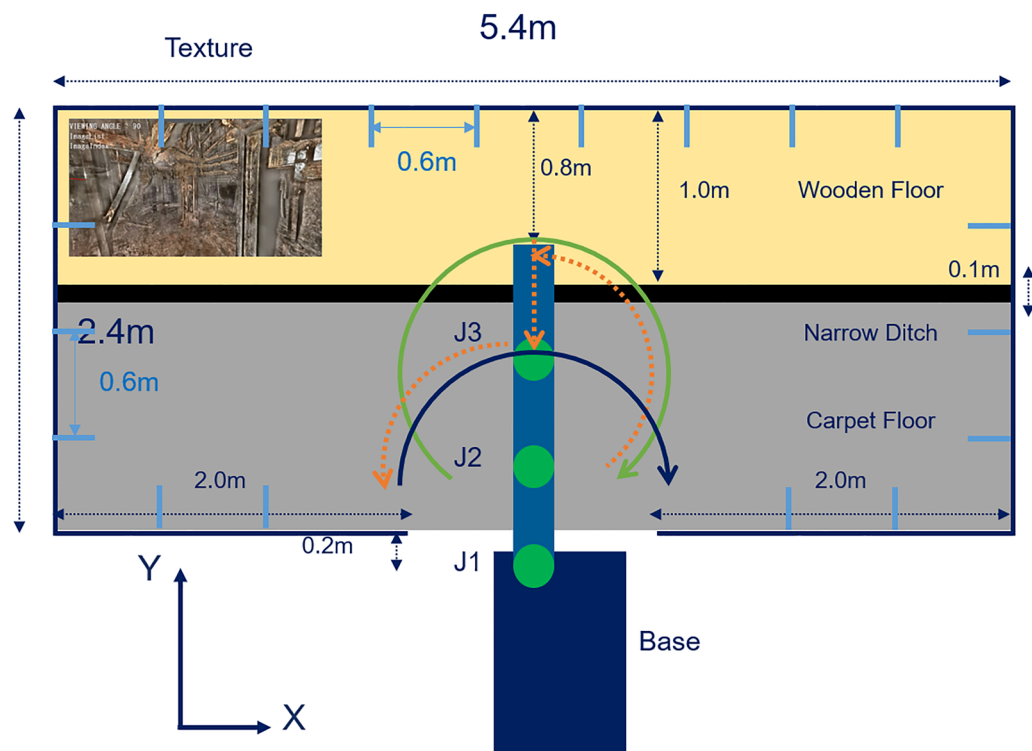
Fig. 18 Distribution of illumination measurement points

Table 9 Average illumination measured by 18 points along the boundary walls in each illumination condition variation

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

Table 10 Three concerned methods' setup for robustness study under limited illumination condition

Method	Software	Mode
SfM + MVS	Colmap + colmap_rob	Image selection method used
SLAM	RTAB-Map	RGB-D
		Stereo IR + IMU

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
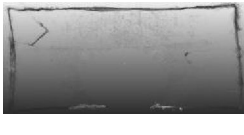
average accuracy degradation at 30%, and 15% in completeness. For stereo IR + IMU, the relative average accuracy drop was 30%, and 20% in completeness (Additional file 4).

Conclusion

In this article, the problems of image quality degradation due to limited motion controllability of wire-drive mechanism has been discussed, including odometry deviation and motion blur. We quantitively evaluated the degrees of these degradation through experiments, and proposed a dataset creation and image selection method on the basis of the experimental results. By taking advantage of the robot's motion features, these two dataset improvement

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 16 of 18

Table 11 Point clouds generated by three methods under different illumination conditions and environment setup patterns

Condition	Method/environment 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			

methods aim to mitigate the detrimental effects of the image quality degradation.

Through verification experiments with colmap (SfM) and colmap_rob (MVS)/instant-ngp (NVS), the effectiveness of the proposed dataset creation and selection methods was evaluated. 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 blur filtering step in the selection method was effective at improving the reconstruction accuracy of instant-ngp by 3.5% at the cost of a 1.5% completeness drop and reducing fuzzy noise in

Wang *et al. ROBOMECH Journal* (2023) 10:4

Page 17 of 18

Table 12 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 13 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 14 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

the reconstructed point clouds. Therefore, the proposed methods may contribute to expanding the feasibility of using a soft arm with better tolerance against possible collisions in the environmental investigation by skillfully reducing image quality degradation.

Furthermore, the reconstruction performances of mainstream methods of SfM and SLAM are quantitively compared under the limited illumination condition. The feature-based SfM method provides best performance in bright condition, while it suffers from significant performance degradation when the illumination condition reduced from 423 to 246 lx, including average relative accuracy degradation at 43%, and 42% average relative completeness drop. Especially,

the performance degradation reflects on more camera poses estimation failures, and details loss due to lacking points for triangulation. In contrast, even in the bright condition, completeness of SLAM may be limited by the density of the reconstructed point cloud (38%~50%), but have more potential to enhance the robustness against limited illumination conditions by using other sensors like infrared sensors. In the dark condition, only the Stereo IR and IMU combination provided reconstruction under a dark environment, while the completeness was extremely low due to the density problem (9%~18%). In the mounted illumination condition, both the RGB-D and Stereo IR + IMU combinations suffered from similar performance degradation to some degrees. Both the average relative accuracy degradations of RGB-D and Stereo IR + IMU were around 30%. The completeness degradation of RGB-D was 15% and 20% for Stereo IR + IMU. 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.

The following three directions will be part of our future works: (1) Autonomous robot odometry deviation recognition and removal, (2) Refining image selection method to achieve more accurate motion deblurring pre-processes, (3) Conduct robustness study of SfM in more realistic environment and develop the robustness enhancement refinement methods on the basis of the experimental results under low illumination.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40648-023-00241-3.

Additional file 1. Demo video of BWD robot, including the sliding ropes deployed on the same pathway, and the narrow-field exploration with a camera.

Additional file 2. Designed trajectory in the verification experiment.

Additional file 3. Illumination variation in the robustness study experiment.

Additional file 4. Concerned 4 motion patterns in the evaluation experiment.

Author contributions
ZW and GE conducted the design of experiments. HT and HK contributed to the completion of the manuscript. All authors read and approved the final manuscript.

Wang *et al. ROBOMECH Journal*   (2023) 10:4

Page 18 of 18

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

References

1. 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. Accessed 22 March 2022
2. Veolia Nuclear Solutions & MHI robotic arm system. https://www.youtube.com/watch?v=UeJmJzX9Feo. Accessed 7 Jan 2023
3. International Research Institute of Decommissioning, Tokyo Electric Power Company Holdings (2022) The preparation progress of the PCV internal investigation and trial extraction operation for the No.2 reactor [in Japanese]. https://irid.or.jp/wp-content/uploads/2022/09/202208252goukiPCVnaibutyousa.pdf. Accessed 22 Sep 2022
4. Endo G, Wakabayashi Y, Nabae H, Suzumori K (2019) Bundled wire drive: proposal and feasibility study of a novel tendon-driven mechanism using synthetic fiber ropes. IEEE Robot Autom Lett 4(2):966–972. https://doi.org/10.1109/LRA.2019.2893429
5. Quattrini L.A, 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:https://doi.org/10.1007/978-3-319-50115-4_67.
6. Davison AJ, Reid ID, Molton ND, Stasse O (2007) MonoSLAM: real-time single camera SLAM. IEEE Trans Pattern Anal Mach Intell 29(6):1052–1067. https://doi.org/10.1109/TPAMI.2007.1049
7. Lindenberger P, Sarlin P, Larsson V and Pollefeys M (2021) Pixel-perfect structure-from-motion with featuremetric refinement. In: 2021 IEEE/CVF international conference on computer vision, 12–15 Oct 2021. doi: https://doi.org/10.1109/ICCV48922.2021.00593
8. Su B, Lu S, Tan CL (2011) Blurred image region detection and classification. In: Proceedings of the 19th ACM international conference on multimedia. Scottsdale, Arizona, USA, Nov 28-Dec 1. 2011
9. Schönberger J L, Frahm J (2016) Structure-from-motion revisited. In: 2016 IEEE conference on computer vision and pattern recognition, Las Vegas, USA, 27–30 June 2016
10. Mildenhall B, Srinvasan P, et al (2020) Nerf: representing scenes as neural radiance fields for view synthesis. In: 2020 European conference on computer vision, Springer, Cham, 23–28 Aug 2020
11. Müller T, Evans A, et al (2022) Instant neural graphics primitives with a multiresolution hash encoding. arXiv preprint arXiv:2201.05989, 2022
12. Schönberger J.L, Lutz J, et al (2016) Pixelwise view selection for unstructured multi-view stereo. In: 2016 European conference on computer vision, Amsterdam, Netherlands, October 8–16, 2016
13. 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. J Field Robot 36(2):416–446
14. Labbé M, Michaud F (2022) Multi-session visual SLAM for illumination-invariant re-localization in indoor environments. Front Robot AI. https://doi.org/10.48550/arXiv.2103.03827
15. Schops T, 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, 22–25 July 2017
16. 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 8 Jan 2021

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.