

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications in Construction
Engineering & Management

Durham School of Architectural Engineering
and Construction

Spring 6-2022

Single-pass inline pipeline 3D reconstruction using depth camera array

Zhexiong Shang

University of Nebraska-Lincoln, szx0112@huskers.unl.edu

Zhigang Shen

University of Nebraska - Lincoln, shen@unl.edu

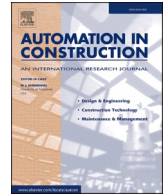
Follow this and additional works at: <https://digitalcommons.unl.edu/constructionmgmt>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computational Engineering Commons](#), and the [Construction Engineering and Management Commons](#)

Shang, Zhexiong and Shen, Zhigang, "Single-pass inline pipeline 3D reconstruction using depth camera array" (2022). *Faculty Publications in Construction Engineering & Management*. 25.
<https://digitalcommons.unl.edu/constructionmgmt/25>

This Article is brought to you for free and open access by the Durham School of Architectural Engineering and Construction at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications in Construction Engineering & Management by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



Single-pass inline pipeline 3D reconstruction using depth camera array

Zhexiong Shang^a, Zhigang Shen^{b,*}

^a Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, 122 NH, Lincoln, NE 68588, United States of America

^b Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, 113 NH, Lincoln, NE 68588, United States of America

ARTICLE INFO

Keywords:

ILI
Visual inspection
Pipeline
Depth camera array
Camera calibration
Robotic
Motion estimate
Dense point-cloud model

ABSTRACT

A novel inline inspection (ILI) approach using depth cameras array (DCA) is introduced to create high-fidelity, dense 3D pipeline models. A new camera calibration method is introduced to register the color and the depth information of the cameras into a unified pipe model. By incorporating the calibration outcomes into a robust camera motion estimation approach, dense and complete 3D pipe surface reconstruction is achieved by using only the inline image data collected by a self-powered ILI rover in a single pass through a straight pipeline. The outcomes of the laboratory experiments demonstrate one-millimeter geometrical accuracy and 0.1-pixel photometric accuracy. In the reconstructed model of a longer pipeline, the proposed method generates the dense 3D surface reconstruction model at the millimeter level accuracy with less than 0.5% distance error. The achieved performance highlights its potential as a useful tool for efficient in-line, non-destructive evaluation of pipeline assets.

1. Introduction

Pipelines are essential infrastructures for transporting energies throughout the nation. Per the U.S. Energy Information Administration (EIA), there are more than 2.6 million miles of transmission and gathering pipelines in the United States, supporting about 34% of the annual energy consumption of the country [1]. Although pipelines are considered as the safest and the most cost-effective way to deliver the energy (as opposed to the tanker trucks and freight trains), integrity failures of pipelines often cause catastrophic losses. Per the statistics from the Pipeline & Hazardous Materials Safety Administration (PHMSA), there were over 12,500 reported pipeline incidents between 2001 and 2020, resulting in over 280 fatalities at a cost of approximately \$US 10 billion [2]. In practice, pipeline integrity can be negatively affected by many factors, including excavation damage, structural or material failures caused by aging, and/or by the static or dynamic stresses from the surrounding environment, severe temperature fluctuations, soil movements, floods, etc. Thus, periodical inspection is needed to maintain the integrity and service life of the pipelines.

Pipeline inspection involves the techniques of detecting and locating the pipe's inherent defects, such as deformation, corrosion, pitting, cracks etc. Modern pipe inline inspection (ILI) and nondestructive evaluation (NDE) were often carried out by pipeline inspection gauges, also known as 'smart pig'. During the inspection, the pigs move slowly

inside pipelines and measure the irregularity on the surface using on-board sensors [3]. The most common sensors installed on the pigs for ILI purposes are magnetic flux leakages (MFL) [4], ultrasonic transducers (UT) [5] and eddy current testing (ECT) [6]. However, most of these NDE sensors are configured for detecting one specific type of defects in each inspection [7]. In addition, the pigs are often large and heavy and can only be used in the later generation pipelines that were specifically designed to allow the pig's operation. Older or legacy pipelines still account for a large portion of the operating pipelines, whose service conditions often need more frequent assessments. Currently, there is a lack of effective and efficient ILI tools to acquire the pipeline condition data for comprehensive assessments.

Following the advancements in optical sensors, visual inspection has gained significant attention in the last decade. By mounting a rotatable camera system on a robotic rover, video data of the pipe's internal surfaces can be obtained while the rover moves across the pipe [8]. The collected videos are then processed with the image processing and machine learning algorithms to identify the pipe's defects [9]. While the RGB camera has the advantage of being low-cost and capturing rich information compared to other NDE sensors, the RGB image data is largely limited to capturing 2D textures of the pipe internal surface [10]. For example, the small bulges and deflections of the buried pipes are often not detectable from single images [11]. Recently, image-based 3D reconstruction has been recognized as a viable tool to better assess the

* Corresponding author.

E-mail addresses: ssx0112@huskers.unl.edu (Z. Shang), shen@unl.edu (Z. Shen).

<https://doi.org/10.1016/j.autcon.2022.104231>

Received 13 October 2021; Received in revised form 24 February 2022; Accepted 24 March 2022

Available online 31 March 2022

0926-5805/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

pipeline defects such as visible cracks, corrosion, pitting, bending, and holes [12]. It involves the techniques of detecting and matching features between the adjacent images, finding the correspondences for image triangulation, optimizing the camera poses as well as the 3D points through bundle adjustment (BA), and recovering the model geometry through surface reconstruction [13]. To reconstruct a full pipeline model using small number of images, front-viewed omnidirectional vision sensors (e.g., fisheye camera, catadioptric camera) were often used [14,15]. By sequentially registering the obtained images, the internal surfaces of the pipe walls can be recovered in high-level details [16]. Despite the progress, the monocular sensor-based techniques often suffer from challenging illuminating conditions, the sparsity of the surface features, and the scale ambiguity of the structure-from-motion (SfM) [17]. To address the issues, some researchers used a strategy to assume that the pipeline geometries are known based on the original pipeline designs [18]–[20]. However, the strategy is likely to fail, since the assumed prior knowledge of the pipeline geometrics in the design stage often do not match the as-builts, and often do not match the actual conditions after the pipeline's decades of exposure to underground geotechnical forces.

The proliferation of depth cameras in robotic applications creates opportunities to address the limitations of the monocular camera in ILI. In addition to RGB image data, depth cameras can acquire direct measurements of the object's distance at the pixel level. Typical depth cameras include the time-of-flight (ToF), structured light, and stereo triangulation. In recent years, the reduced cost and improved performance of depth camera sensors make them applicable in ILI. While existing studies have developed varied in-pipe inspection systems using depth cameras (such as the front-viewed stereo system [21], the verged stereo system [22], and the trinocular system [23]), the reconstructed models either were too sparse and noisy, or only covered a partial of the pipe, or required heterogeneous textures being painted across the pipe wall. Up to now, complete, dense, and accurate 3D reconstruction of pipelines through a single-pass inspection remain a challenging task.

To overcome the identified challenges, the authors introduce a uniquely designed depth camera array (DCA) for ILI 3D pipeline reconstruction. Three primary contributions are expected from this work: First, the developed oblique-view DCA overcomes the major limitations identified in [21,22]. This new camera array layout ensures the full surface coverages with sufficient point density while keeping a minimal number of cameras. Second, the developed DCA calibration method allows the registration of the four oblique depth cameras into a unified pipe model at approximately one millimeter level accuracy. Compared to the existing depth cameras calibration methods [24–26] that used a moving checkerboard to capture a sequence of overlapped images, this new calibration method does not require overlapped coverage among cameras. In addition, only one shot is required from each camera, which significantly reduces the overall computational cost. Third, a fused RGBD video generated from the four oblique RGBD cameras is used for DCA motion estimation. Unlike the state-of-the-art monocular SfM techniques [18,19], the new method does not require prior geometric knowledge of the inspected pipelines for motion estimate. This method also outperforms the recently developed laser profiling technique [27] because it allows the image data being acquired at much higher traveling speed (~ 1 m/s) along the pipelines.

The rest of the paper is structured as follows. Section 2 describes the related work. Section 3 presents the hardware of the ILI system and the design choice of the onboard cameras. Section 4 introduces the proposed calibration method for the DCA. Section 5 discusses the automated approach for 3D pipeline reconstruction. Section 6 illustrates the experimental setup and the evaluation results. Section 7 concludes the article with the key findings and suggestions for future studies.

2. Related work

2.1. Inline pipeline 3D reconstruction

There have been extensive studies about inline pipe 3D reconstruction. Among the different onboard sensing systems, the omnidirectional camera (e.g., catadioptric cameras, fisheye cameras) is the most popular design choice due to the full pipe surface coverage at a single shot [28–30]. The collected images are then passed into the SfM workflow to obtain the 3D pipe model. To name a few, Kannala et al. [15] divided the image-based in-pipe reconstruction into three major steps: First, the method split the collected fisheye images into triplets and computed the local pipe structures by tracking the feature points within each triplet. Second, the locally reconstructed models were merged into a longer sequence with hierarchy bundle adjustment. And finally, model fitting algorithm was employed to remove the outliers and convert the sparse point cloud into a pipe surface model. Instead of directly processing on the raw image data, Esquivel et al. [31] unrolled every input image using spherical projection which facilitates the feature tracking and the image triangulation process. After the sparse reconstruction, the method employed the known camera poses and the pipe geometry to reduce the model distortion. To correct the scale ambiguity directly from the monocular image data, Hansen et al. [18] merged the known pipe geometry (i.e., diameter) with the camera calibration data. Sliding window sparse bundle adjustment (SBA) is employed to compute the inline robot poses as well as the pipe surface map. The method also detected the pipelines as straight and T-shape, enabled the 3D reconstruction of a network of pipelines. A similar pipe reconstruction strategy was presented in Kagami et al. [19] where the geometry of a small pipe network was reconstructed using an endoscopic camera. The method integrated the conic shape detection into the objective function of BA optimization, which makes it robust to the scale drifting errors. However, similar to the above-mentioned approaches, the method requires the dimension (e.g., diameter) and/or the shape (e.g., straight, elbow, cylinder, etc.) of the pipe to be surveyed as the input (for model fitting operation), which may fail to recover the in-situ geometry of the pipelines.

Compared to the usage of the omnidirectional cameras, the studies on pipeline reconstruction using depth cameras are limited because of the limited camera FOV and the depth inhomogeneity. For example, Hansen et al. [22] proposed a verged stereo system for the 3D pipe surface mapping. The system pre-tuned the baseline/parallax angle of the stereo system for the selected pipe such that the depth map of the pipe surfaces can be recovered through multi-view geometry. However, due to the limited FOV of the stereo system, multiple passes are required to cover the entire pipe wall. More recently, Bahnsen et al. [21] investigated the performance of 3D pipe mapping using a single front-viewed RGBD camera (i.e. RealSense D435 [32]). While the 3D accuracy can be improved with the proper camera configuration (i.e., enable active IR projector or under sufficient illumination condition), the study only achieved the centimeter-level accuracy which is insufficient to accurately recover the pipe geometry. Haertel et al. [23] proposed a trinocular stereo system to reconstruct the 3D geometry of straight pipes. Each 3D point was estimated by correlating the homogeneous points between the cameras. While the system reached the high geometrical accuracy, the presented method requires heterogeneous textures for the accurate 3D points estimation which might not be applicable for operating pipelines where the in-pipe textures are often sparse and repeated. In [27,33], a real-time in-pipe inspection robot was developed that combines RGBD mapping with laser profiling for 3D reconstructions at millimeter level accuracy. The system generates the point cloud model by triangulating the readings of a stereo IR system at the projected laser ring. The color readings of a RGB camera are then embedded into the point cloud through extrinsic calibration. Because the depth values are only computed at the laser rings, the robot has to move extremely slow (0.2 m per minutes) in order to obtain a dense map, which limits its efficiency for inspecting long pipelines. In addition, wheel encoder was

used to localize the robot within the pipe and that may result in the drifting problem when the pipe surface is not even. In contrast to the existing studies, our ILI system used multiple depth cameras that generates a dense, complete, and high-fidelity 3D pipe reconstruction with a single pass. In addition, our method tracks the in-pipe motions based solely on the image data and enables the robot to travel across pipelines at moderate speed (i.e., 1 m/s).

2.2. 3D reconstruction using multiple depth cameras

In recent years, dense 3D reconstruction via multiple low-cost, commodity depth cameras have gained increased popularity in the visual computing and computer vision communities. Conventional multi-camera calibration involves the techniques of estimating the camera extrinsic parameters. This can be achieved by capturing a planar checkerboard placed at several positions. The camera extrinsic transformations can then be estimated by minimize the reprojection errors of the checkerboard's crossing points detected in the RGB images [34]. However, such optical approach often results in the poor registration result due to the missed consideration of the depth data. An alternative approach is to use the robust Simultaneously Localization and Mapping (SLAM) and SfM techniques that aligns the both the color and depth images either between the cameras [35,36] or through a specifically designed 3D reference target [37,38]. While these approaches can generate the improved results, the methods are laborious to apply, have the certain requirements of the camera movements (e.g., for loop closure), and might be only applicable to certain camera layouts and/or in the controlled environment where there are sufficient textural and the geometrical features (e.g., 1D line, 2D planes, 3D corners) to be extracted from.

In practice, the irregular geometric distortion of the depth cameras negatively affects 3D reconstruction quality [39]. One way to address this issue is to allow the depth data to deform non-rigidly while reconstructing the scene [40]. Recent works [24,26,41] jointly calibrated the cameras intrinsic/extrinsic parameters and registered the depth measurements into an artificial-free reconstruction through a field of space varying transformations. Specifically, Deng et al. [26] separated the 3D space captured by the depth cameras into a set of 3D voxels. The rigid transformations that align the overlapped 3D data within each voxel are individually estimated. While this method yielded visually appealing results, it only morphs the scene geometry rather than corrects the actual depth distortion. To overcome this limitation, Beck and Froehlich [41] developed a sweeping-based volumetric calibration method that corrects the depth distortion in tandem with the extrinsic calibration. The method took the video while moving a checkerboard across the calibration space. A motion-capturing system was used as the ground truth data within each voxel of the space. However, this technique requires the operators to carefully move the checkerboard and requires an external motion capturing system to track the movements. To address these limitations, in this study, the authors proposed a new calibration approach for the onboard DCA that eliminates the need of the moving cameras/checkboard and the external tracking system. The new approach only requires a single shot image from each camera, which significantly reduces the manual efforts needed to calibrate the ILI system. We achieve this by constructing a pipe-shaped 3D marker map, formatting the in-pipe calibration space based on pipe geometry, and integrating the cylinder fitting operations with the 3D data registration. Moreover, the proposed calibration method only needs to perform on a standard pipeline once and can be applied to a pipeline with varied surface texture or geometrical dimensions within a pre-defined tolerance.

3. The system hardware briefing

The designed robotic ILI system consists of three components: the computing module, the actuator module, and the sensing module. The

computing module includes a Micro PC with Intel Core i7 CPU and Samsung 980 PRO SSD connected via PCIe for onboard image processing and data retrieving. The actuator module is a four wheel-drive rover powered by a 36-V battery pack. This module enables the rover to move inside the pipe without tether and supports the long-distance inspection purposes. Currently, the motion of rover is remotely controlled by an operator, but the speed controller can be programmed at fixed traveling speeds along the pipelines. The DCA sensing module is the core of the proposed system, which is composed of four Intel RealSense D435 cameras (abbreviated as RS camera) [32]. We selected the RS cameras due to its small size, light weight, and high resolution when compared to other available products. The RS camera relies on the stereo triangulation of two IR sensors to estimate the depth values at each pixel. This strategy would not cause the IR interference noise when multiple cameras are utilized, as opposed to the time-of-flight (ToF) cameras (e.g., Microsoft Kinect). However, stereo cameras have the minimal resolvable depth (Min-Z), which is hardware constrained by the baseline of the two sensors. This Min-Z constraint makes it infeasible to place the cameras perpendicular to pipe walls when the pipeline's diameters are close to or less than Min-Z. To overcome this limitation, instead of facing the cameras outward, we tilt each camera inwards to cover the opposite pipe walls (as shown in Fig. 1).

For a standard 14" (355.6 mm inside diameter) pipeline, we experimentally found that setting the tilted angle around 35 degrees fulfills both the constraints of the Min-Z and the density of the depth readings. This setup also helps reduce the number of RS cameras needed to cover the entire pipe surface. A reduced number of cameras is beneficial for real-time image processing and data storage. A forward-looking LED light and circular LED lights are attached around the cameras, to provide necessary illuminations inside the pipeline (Fig. 1 (b)).

The next two sections describe the 3D reconstruction techniques based on the DCA module: Section 4 covers the method that calibrates and registers the RS cameras into a unified 3D pipe model, and Section 5 covers the method on how to incorporate the calibration results into a camera motion estimation method for the automated 3D pipeline reconstruction.

4. DCA calibration

Fig. 2 is an overview of the proposed DCA calibration method. The method starts with defining the calibration space given the cameras setup and the pipe geometry (Section 4.1). We construct a pipe-shaped 3D marker map as the reference for the visual and the depth correction across the calibration space (Section 4.2). The proposed calibration is performed in a coarse-to-fine fashion: We initially estimate the extrinsic transformations of each RS camera by formulating it as a Perspective-n-Point (PnP) problem [42]. We achieve this by matching the markers detected in each color image to the correspondent locations in the 3D map (Section 4.3). Using the estimated camera poses as the initial transformations, we then correct the depth distortions by volumetric registering the camera projected marker point clouds to the map. We decompose the map based on the distributed marker locations and find the rigid body transformation within each decomposed space by locally aligning the camera extracted markers to the map (Section 4.4). And finally, we smooth and generalize the marker map computed transformations throughout the calibration space. The output of the proposed method is a 3D lookup table consisting of a smooth field of rigid body transformations that fuses the RS cameras into an accurate and unified pipe model.

4.1. Calibration space

Fig. 3 demonstrates the calibration spaces of the onboard RS cameras. Because the subject is the internal pipe wall surface, only the spaces around the pipe wall need to be calibrated. For a straight pipe segment, the calibration space is formed as a circular tube with the length

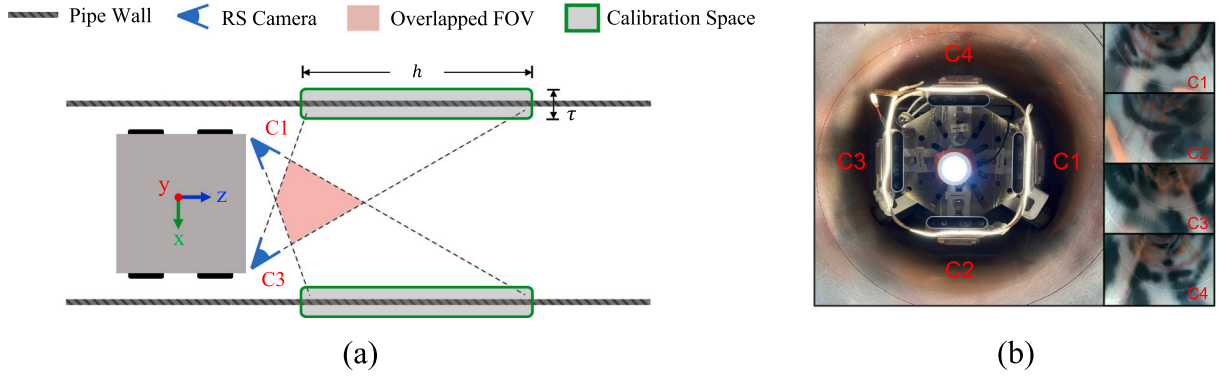


Fig. 1. (a) Conceptual diagram of the proposed ILI system and the layout of the RS cameras; (b) Front-view of the DCA module while the rover is placed in the pipe. The in-pipe color image taken from each camera is also presented.

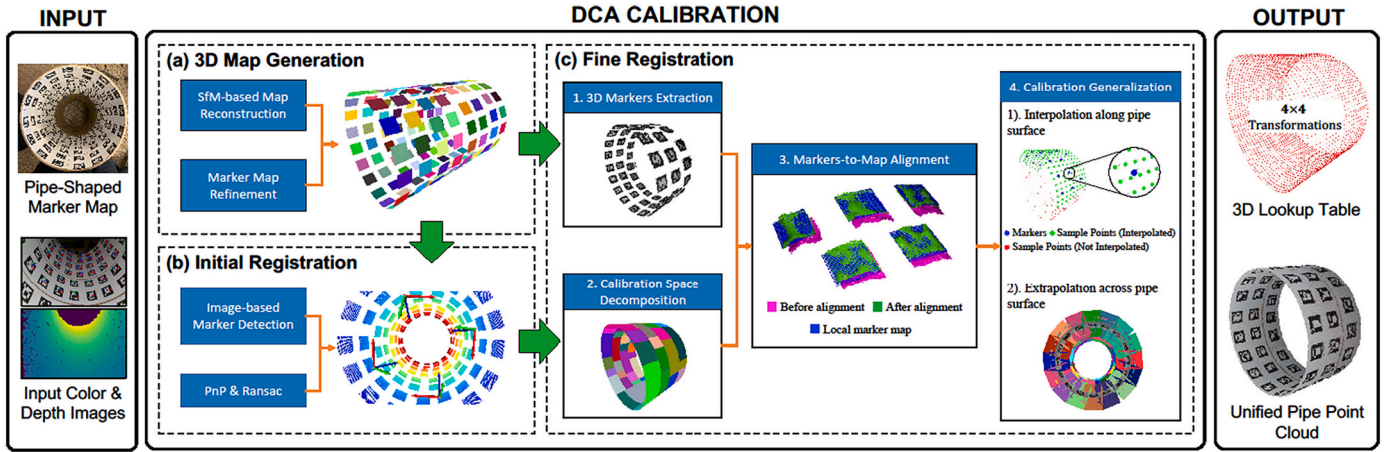


Fig. 2. Workflow of the proposed DCA calibration method

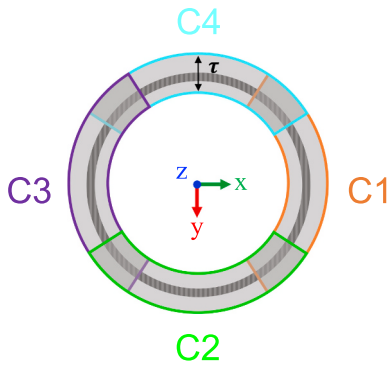


Fig. 3. Front view of the calibration space formed by the four RS cameras in the robot coordinate system. Dashed lines denote the pipe wall. Each color highlights a 3D calibration space of the correspondent RS camera.

d defined based on the projection of the camera FOV along the pipe axis and the thickness τ as the tolerance given the pipe diameters. Noted existing multi-cameras calibration methods often relied on the overlapped spaces to conduct the calibration. In the proposed ILI system, there are no common spaces shared by all four cameras (as shown Fig. 1 (a)). This condition, combined with the tubular shape of the calibration space, makes it difficult to be handled by the exiting strategies [24,25,38].

Several coordinate systems are used in the calibration. We define C_i ($i = 1..4$) as the coordinate system of the i th camera where (u, v) is the

coordinate on the image plane and d is the z-depth along the camera's principal axis. We set the 3D pipe coordinate system in the Euclidean space as P where the z-axis always passes through the pipe's centerline. The coordinate system of the in-pipe robot is defined as R with (x, y, z) denoting the pitch, roll, yaw of the robot motion. The origin is set at the geometric center of the rover with the z axis pointing along the pipe's axis. Because the rover might not always be located at the pipe's centerline, a translation is needed to transform from the robot to the pipe coordinate system. And finally, we define W as the world coordinate system that can be transformed from the robot coordinates if the GPS at the starting location of the in-pipe robot is known.

4.2. Pipe-shaped marker map

In this study, we construct a pipe-shaped marker map M as the 3D reference for calibrating the RS cameras. The map is composed of a set of unique ArUco markers [43] attached at the inner surface of a circular pipe (Fig. 2). The markers are designed at the unified size (i.e., 4×4 cm) and attached at the pipe at varied spacing (i.e., from 2.5 cm to 4.0 cm). This setup avoids the geometrical ambiguity for the correction of depth distortion (detailed in Section 4.4). The initial 3D representation of the map is generated by collecting a set of overlapped color images and processed using the marker-based structure-from-motion [44]. Because each marker is unique, the poses of the markers are estimated by combining the markers detection with the SBA optimization. The sub-millimeter accuracy reported on the small scene shows the practicability of using the reconstruction as the 3D representation of the physical map. Fig. 4 (a) shows the rendered marker map with the associated

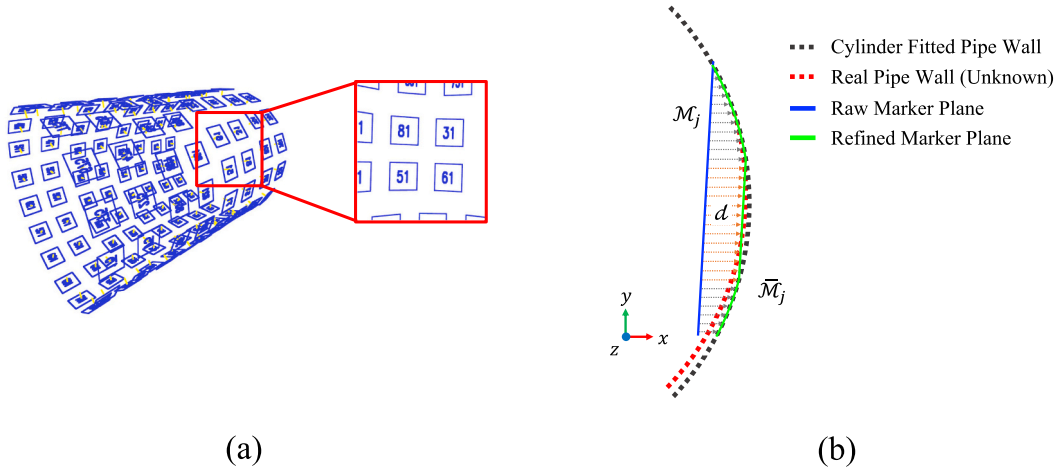


Fig. 4. (a) The initially rendered marker map and the highlighted marker ids in the map; (b) the refinement of the initially rendered marker \mathcal{M}_j (in blue) to the curved representation $\bar{\mathcal{M}}_j$ (in green) in the map.

marker ids. It is noted that the initially reconstructed markers are assumed to be planar [44]. However, in reality, the markers attached at the circular pipe wall are bent. These deformations must be recovered for the accurate 3D representation.

Due to the initially rendered markers are deformed along the pipe axis with the profile of each marker following the curvature of the pipe segment, these markers can be further refined based on the pipe geometry. In this study, we handle this problem by (1) fitting the pipe into a cylinder model; and (2) correcting the markers by projecting the points from the marker planes to the cylinder surface. In the first step, we reconstruct the circular pipe as a perfect cylinder using the normal to the marker planes (i.e., computed based on the cross product of two vectors on each marker). The cylinder axis is estimated with Principle Component Analysis (PCA) [45] where the eigenvector associated with the least eigenvalue is selected. We then project the corners of the detected markers along the estimated cylinder axis, and fit the projected 2D circle with the least square (LS) optimization [46]. The fitted center and radius of the circle, along with the cylinder axis, generates the parametric pipe surface model. In the second step, we actively bend each initially rendered marker $\mathcal{M}_j (j = 1..m)$ by projecting each point on the marker plane along its normal direction. The points with the minimal distance to the surface of the cylinder model are selected as the point on the refined marker $\bar{\mathcal{M}}_j$ (as in Eq. (1)).

$$\bar{\mathcal{M}}_j = \underset{\sigma}{\operatorname{argmin}} d(\mathcal{M}_j' + \sigma \mathcal{N}_j', P) \quad (1)$$

where \mathcal{M}_j' and \mathcal{N}_j' respectively denote the points and the normal vector to the rendered marker \mathcal{M}_j , and $\bar{\mathcal{M}}_j'$ is the projected marker point. P denotes the fitted pipe surface, d is the Euclidean distance, and σ is the coefficient that indicates the step of the projection. To avoid the imperfect condition of the pipe segment (i.e., pipe with minor deformations or local surface unevenness), we restrict σ within ± 2 cm. Fig. 4 (b) illustrates the concept of the map refinement at a single marker. Clearly, the refined marker is a closer estimation to the real pipe geometry as opposed to the initial estimation. This map will then be utilized as the 3D reference target for the DCA calibration.

4.3. Initial registration

After the marker map is reconstructed, we put the rover inside the pipe with each camera covering a portion of the map. A single RGBD image is taken from each camera (as in Fig. 2). Because the RS cameras are factory calibrated, the extrinsic transformations between the cameras can be initially estimated based on the color images. By matching

the 3D locations of the markers in the map to the correspondent pixel coordinates in the 2D color image, the relative pose of the cameras are obtained through a list of 2D-to-3D correspondence. Specifically, we extract the marker corners from the color images and find the correspondent 3D points in the refined map by matching the markers with the same ids. The Perspective-n-Point [42] is utilized to recover the 6D pose of each camera. Because more than one marker is covered by each camera, redundant corner points are extractable. Thus, random sample consensus (RANSAC) [47] is employed to handle the redundancy while increasing the robustness of the pose estimation. Fig. 2 (b) shows the 6d poses of estimated RS cameras using the PnP method. Although the transformation based on the factory calibrated images is insufficient to ensure a high-quality registration, it does provide an initial range estimate, which reduces the search space needed to find the best results in the fine registration step.

4.4. Fine registration

The objective of the fine registration is to correct the non-linear, systematic depth distortion [48] of each camera based on the initial registration and the refined marker map. The basic idea is to find a field of rigid body transformations across the tubular calibration space where the overall registration errors from the cameras' depth readings to the marker map is minimized. To achieve that, we first extract the markers from the RGBD image of each camera. Based on the initial transformations, the 3D markers can be projected into the pipe coordinate system. Next, we decompose the marker map into submaps based on the distribution of the markers in the map. Within each decomposed submap, point set registration is employed to find the single best rigid body transformation that aligns the markers extracted from the images to the correspondence in the map. Finally, we generalize the calibration from a collection of 2D submaps to the 3D calibration space. The method outputs a 3D lookup table consisting of a smooth field of rigid transformations, which automatically converts the RGBD images taken from the cameras into a unified 3D pipe point cloud. Fig. 2 (c) shows the workflow of the fine registration, and we discuss the major steps in detail in the following paragraphs.

4.4.1. 3D markers extraction

For each camera, we extract the markers from the RGBD image, and convert it into a point cloud model for the post processing. Specifically, we extract the four corners of every marker detected in the color image. We connect the four corners as a polygon and fill the pixels within the polygon using the flood-fill algorithm. Because the depth frame is synchronized with the color image, the depth readings at the same pixels are

cropped. Combining the extracted color and depth pixels, the 3D markers are extracted from each camera. Fig. 2 (c-1) shows the 3D markers point cloud that is initially extracted from the four cameras, and then projected into the pipe coordinate using the initial transformations.

4.4.2. Submap decomposition

We decompose the marker map into the local submaps such that the transformation estimated within each submap is rigid. Because the markers are already spatially distributed in the map, we can use the locations of the markers to evenly decompose the map. It is noted that ArUco markers have the symmetrical geometry (i.e., square), thus direct marker-to-marker alignment may cause the orientation ambiguity. To avoid such issue, we define that each submap must cover a neighborhood of the markers. By setting the center of the current marker as the pivot, the neighbor markers are selected if its distance to the pivot is less than a pre-defined threshold. In this study, we set the distance threshold as 7 cm and require at least two neighbor markers to be included in each submap. Fig. 2 (c-2) color-coded the decomposed submap using the proposed method.

4.4.3. Markers-to-map alignment

In this study, we select the probability-based point set registration to compute the local transformation within each submap because the 3D markers cropped in each submap might still contain the non-systematic random noises [48]. Compared to the commonly used Iterative Closest Point (ICP) [49] which is sensitive to such noises, the probabilistic-based registration, such as the coherent point drift (CPD) [50] interprets the point cloud as a probability density distribution (especially the Gaussian mixture model) which is more robust to the outliers. The transformation of the point cloud can then be found using the Expectation Maximization (EM) algorithm [51]. In this study, we employ a variant of the CPD, the FilterReg [52] that augments the E step by solving the correspondence search through the Gaussian filtering method. The method has the improved efficiency and the comparable accuracy for a small dataset that is the ideal solution for our case. With the parallel processing, the transformations within all the submaps can be computed in seconds. Fig. 2 (c-3) shows an example of the markers-to-map alignment within a submap. Clearly, the probability-based method is robust even with the existence of the random noise on the extracted 3D markers.

4.4.4. Calibration generalization

In the previous steps, we obtained a field of rigid transformations across the marker map. However, these transformations cannot be directly applied to other pipelines because (1) the computed transformations might cause the artifacts due to the non-uniform results across the pipe: the spaces close to the centers of the markers are well calibrated while the space at the boundaries between the submaps may be ill-calibrated; (2) the transformations computed from the marker map (attached on the internal pipe surface) only work if the pipeline to be inspected has identical shape (e.g., circular, size) to the calibration pipe. In practice, there are patches, bulges, and even deflections existing on the operating pipelines, and the calibration result should be tolerant to such variations. Thus, we present a method to smooth and generalize the transformations across the calibration space. Our method is performed in two sequential steps: First, we re-sample the submaps along the pipe surfaces into a dense set of 2D cells. Scatter data interpolation is employed to fill the empty cells and achieve the higher accuracy; Second, we extrapolate each cell into a 3D voxel across the pipe surface, filling the transformations in the 3D space.

4.4.4.1. Interpolation along pipe surfaces. The basic idea of interpolating the transformations at the pipe surface is to unroll the 3D pipe model onto a 2D plane such that the data interpolation can be performed in a linear fashion. Thus, we first fit a cylinder model to the pipe geometry. Then, we uniformly sample a dense set of point at fixed intervals on the

cylinder surface. We unroll the cylinder so that the newly sampled points as well as the pivots of the computed transformations are projected on the 2D plane. Next, we construct an octree [53] to perform the quick search of the neighbor pivots for each point. If more than one pivot is found, Inverse Distance Weighting (IDW) [54] is utilized to interpolate the transformations at the point. It is noted that we cannot directly perform the IDW on the 4×4 transformations due to the nonlinearity of the rotation matrix. Thus, we convert each rotation matrix into the quaternion and perform the IDW to the quaternions at the exponential space. The interpolated quaternions are then converted back to the rotation matrices and integrated with the interpolated translations. Fig. 2 (c-4-1) presents an example of the interpolated rotation matrices of a RS camera. After that, we construct a 2D cell centered at the sampled points and formed the 2D space of the interpolated transformation. We parallelize the above process to compute the interpolated transformation at each cell. And finally, we reproject the sampled 2D cells and the associated transformation in each cell back to the 3D pipe space.

4.4.4.2. Extrapolation across pipe surfaces. For each 2D cell (with the interpolated transformation) on the pipe, we compute the normal to the pipe surface and extrude the 2D surface into a 3D voxel along the normal vector. The height of the voxels is determined by the calibration space, indicating the tolerance (τ) on pipelines with different geometries. Fig. 2 (c-4-2) colored the 3D voxels across the pipe surface.

5. End-to-end pipe surface reconstruction

In this section, an end-to-end workflow (as Fig. 5 illustrated) is presented for the automated pipeline surface reconstruction by incorporating the calibration into a robust camera motion estimation approach [55].

5.1. Data preprocessing

The first step of the approach is the synchronization of the video sequences from different cameras because each RS camera use its own hardware clock for time stamping. Exact frame-to-frame synchronization is desired. However, hardware synchronization significantly increases the computational load and memory usage during the data acquisition. Therefore, due to the high frame rate (i.e., 60HZ) of the video collection and the relative slow driving speed of the in-pipe rover (i.e., ~ 1 m/s), we employ a temporal filter that continuously examines the incoming timestamps within the image headers. By using the timestamp from one camera as the reference, the synchronized images can be collected by finding the nearest header at a subsampled timestamp. A potential advantage of using this soft synchronization strategy is that we can reduce the images that need to be saved on the onboard computer for the inspection of longer pipelines (e.g., several miles). For the synchronized image sequences, we register the color and depth frames at each timestamp using the 3D lookup table. The output is a sequence of timestamped pipe point clouds collected as the rover travel along pipelines.

5.2. Central camera projection

For each pipe point cloud, we generate a virtual front-viewed and center positioned RGBD camera (i.e., color and depth frame). We define the virtual central camera been located at the intersection between the pipe axial line and the plane formed by the four RS cameras. The resolution of the central camera is set as [500, 500] with the intrinsic matrix determined by the pinhole model. Specifically, we define the camera focal length as a unit pixel, the principal point offset at the center of the image plane, and lens distortion as zero. To compute the color and depth values at each pixel, the unified point cloud is first projected to the image plane using the camera matrix. Next, the color/depth value at

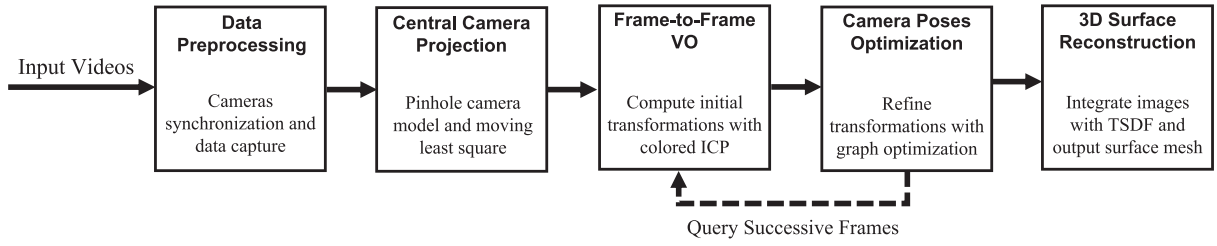


Fig. 5. Workflow of in-pipe rover pose estimation and pipeline surface reconstruction.

each pixel is calculated as the median of the neighbor points within a pre-defined distance. To minimize the color reading noises caused by varied illumination conditions of each camera, moving least square (MLS) [56] is used to interpolate and, at the same time, smooth the color readings at each pixel. Instead of constructing the polynomial fit function globally, MLS computes a locally weighted least square fit at each point. We set the neighbor distance threshold as 0.5 mm and define the error function in the quadratic form. It is noted that we only perform the MLS at the color space such that the detailed geometric features from the depth readings are preserved. Based on the proposed method, we can convert the input point clouds into a sequence of front viewed RGBD images. Fig. 6 shows the virtual camera rendered color and the depth image of the calibration pipe.

5.3. Frame-to-frame VO

In this section, the motion of the rover over timestamps is calculated based on the rendered RGBD image sequence. Compared to the unstructured point cloud, the major advantage of using the RGBD sequence is that the gradient of the local color and depth variations are parameterizable in the 2D image planes. Because the interior surface of many pipelines presents sparse textures, frame-to-frame alignment using either the 2D features (e.g., SIFT, SURF) or the pure 3D geometrical registration (i.e., ICP) might fail to achieve the accurate and robust result. Thus, in this study, we employ the colored ICP [57] that estimates the frame-to-frame transformations using the combined information of the color and the geometry. This is achieved by defining the error function as the weighted sum of the geometric and photometric disparity between the corresponding points. Because the error function is non-convex, we applied the multi-scale registration that iteratively registers the down sampled images to avoid the resulting transformation being trapped at the local optima. We empirically identified that setting the number of iterations at 4 is a good balance between the optimization convergence and the computation efficiency.

5.4. Camera poses optimization

It is worth noting that the camera trajectory computed from the frame-to-frame VO can result in the distortion of the reconstruction due to the accumulated drifting over frames. Conventionally, such a problem is handled with the loop closure detection techniques, such as the vocabulary tree [58]. However, in most cases, there are no loops existing for the inline pipe condition because the designed rover is intended to only move along the pipelines in a single pass. By assuming the depth reading from the virtual central camera is accurate, we can refine the visual odometry using a neighborhood of images. To achieve that, a pose graph is constructed with each node denoting a frame-to-frame estimated camera pose and the edges indicating the rigid transformations between the poses. Initially, the graph is linear because the edges exist only between the adjacent frames. To refine the camera poses, k successive frames at each node are queried. The same registration method as that in the visual odometry is used to compute the transformations. The edge between the nodes is constructed only if (1) there are sufficient overlaps between the selected frames; and (2) the inlier error between the correspondent points is less than a threshold χ . While k is set as 4, χ is set as 0.1 for all the test cases. We find using the RGB disparity as the only metric is sufficient to measure the inlier errors and refine the transformations. Based on the proposed strategy, the pose graph is constructed. Next, the objective function of the pose graph is formulated into the quadratic formation [55] and been optimized using g^2o [59]. The output of the optimization is a refined camera trajectory which can be used to integrate the unified point clouds at different timestamps into a pipeline surface reconstruction.

5.5. 3D surface reconstruction

Finally, a global surface fusion technique is used, which integrates the rendered RGBD sequence with the computed camera poses into a global system. The volumetric truncated signed distance function (TSDF) [60] is used to incrementally concatenate the RGBD images at each timestamp. The voxel size of the TSDF is set as 2 mm and the initial

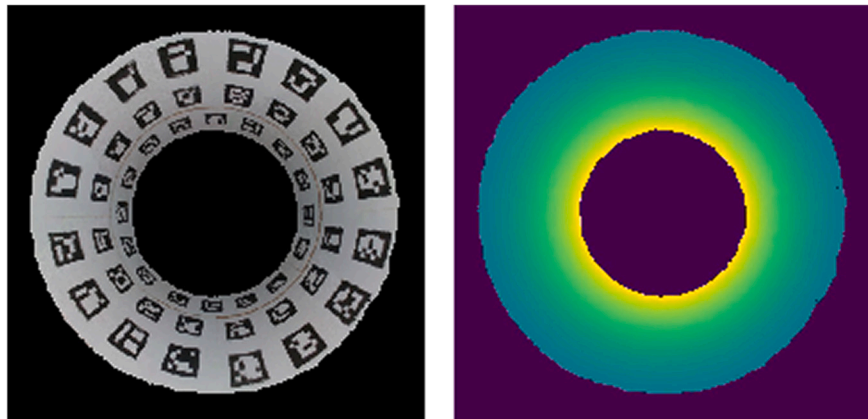


Fig. 6. Virtual camera rendered color (left) and depth (right) image of the calibration pipe.

frame is served as the reference point of the reconstructed model. In the last step, the dense pipe point cloud model is extracted from the measurement, and the model is converted into a surface reconstruction using the Poisson Surface Reconstruction [61]. The reconstructed surface model recovers both the geometry and the surface texture of the inspected pipes that can be utilized as the high-fidelity 3D inputs for further applications.

6. Experiments and results

6.1. Experimental setup

In this section, the laboratory experiments to evaluate the performance of the proposed method are presented. First, the accuracy of the DCA calibration (as in Section 4) is evaluated using three cases of the short pipe segments (Fig. 7). The pipe (a) and (b) are rigid cardboard pipes with the same dimension as the calibration pipe (i.e., the inside diameter of the pipe is 351.6 mm). We manually create the internal textures at the pipe surfaces by either attaching the thin blue adhesive tapes at the interior pipe walls or randomly spraying red/black/white paints inside the pipe. These two pipes are used to evaluate the accuracy of the calibration results when different surface patterns are presented. Pipe (c) is a steel pipe with the inner diameter of 338.6 mm and is slightly smaller than the calibration pipe. We select it to evaluate the robustness of the calibration on pipes with different dimension.

To evaluate the performance of the end-to-end 3D reconstruction approach described in Section 5, an experiment using a longer pipeline (as shown in Fig. 8) is conducted. The 2800 mm long pipeline is composed of two straight cardboard pipe segments (α) and (β). Pipe (α) is 1220 mm long with random graffiti patterns. Pipe (β) is 1580 mm long with randomly sprayed paint patterns. These patterns are created to evaluate the accuracy of the 3D reconstruction when the pipeline internal surfaces have different texture features. To provide more rigorous quantitative evaluation of the reconstruction accuracy, two exact copies of five A3-size papers marked with exact dimension (in length and width) on Cartesian coordinate system are printed with background texture of actual steel corrosion pattern (Fig. 9). One copy of the five prints is carefully (to avoid any potential stretch stress on the papers) glued onto the internal surface of the pipeline. Another copy is used as ground truth to be compared with the reconstructed dimensions. By measuring the size of the papers in the reconstructed model, the accuracy of the proposed approach can be evaluated. Fig. 9 shows the designed layouts of the five A3 prints and the detailed corrosion image in each print.

6.2. Evaluation methods

6.2.1. Short pipe segments

The performance of the proposed method is evaluated both qualitatively and quantitatively. For the short pipe segments, the accuracy is

evaluated in two aspects:

First, we evaluate the registration errors between the adjacent cameras. Two indicators are selected to quantitatively evaluate the registration accuracy, the indicators are: the geometric error ($Error_{geo}$) and the photometric error ($Error_{pho}$). The geometric error measures the geometric distance between the inlier points from each pair of the cameras; While the photometric error computes the RGB intensity variations between the inlier points. Eqs. (2) and (3) present the equation of the geometric and the photometric errors.

$$Error_{geo} = \frac{1}{|IJ|} \sum_{ij} \sqrt{\frac{\sum_{ij}^X (P(x_i) - P(x_j))^2}{|X|_{ij}}} \quad (2)$$

$$Error_{pho} = \frac{1}{|IJ|} \sum_{ij} \sqrt{\frac{\sum_{ij}^X (H(x_i) - H(x_j))^2}{|X|_{ij}}} \quad (3)$$

where IJ is the set of the point cloud pairs, X is the collection of the inlier points. The inlier correspondent points between each pair of the point cloud are calculated using the nearest neighbor technique with the radius of the neighbor size as 10 mm. For each inlier point, P describes the 3D position and H denotes the intensity value (i.e., measured as the sum of the RGB intensities).

Second, we compare the radius of the reconstructed pipes to the ground truth measurement (using laser measurement tool). To achieve that, we first fit the pipe model into a cylinder by estimating the principal axis using the PCA; Then we divide the pipe models into K segments along the pipe principal axis. For each segment, we project the points along the principal axis and estimate the radius of the projected circles using RANSAC and the LS optimization (as described in Section 4.1). The mean and standard deviation of the errors between the ground truth and the estimated radius in the segments are obtained. In this experiment, we compute the number of segments K as D/λ where λ equals to 1 mm, denoting the length of each segment. Fig. 11 (a) demonstrates the evaluation process on the calibration pipe.

6.2.2. Longer pipeline

The evaluation on the reconstruction of the longer pipeline is also divided into two aspects:

First, we qualitatively evaluate the reconstructed pipeline through the visual comparison between the 2D images generated from the reconstruction and the real-world images, including both the direct observation of the in-pipe surface patterns and the recovery of the unrolled image textures. We achieve this by first fitting the reconstructed pipeline into the cylinder using the aforementioned approach and then unroll the cylinder based on the fitted parameters. The pictures attached at the inner pipe walls can then be recovered. Noted we fitted and unrolled the pipelines α and β separately to avoid the conditions that the combined pipeline is not perfectly straight (i.e., human errors).

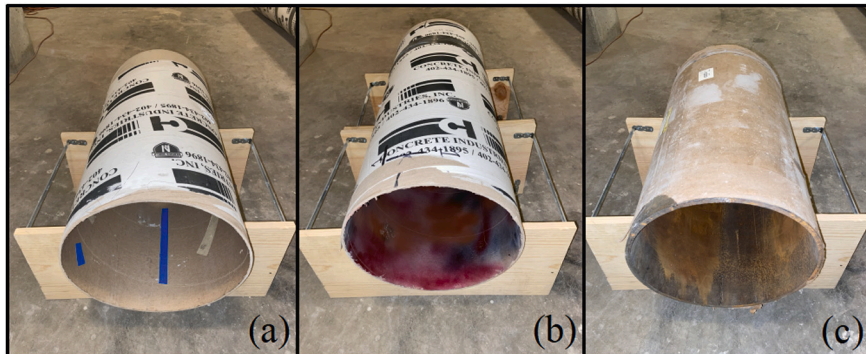


Fig. 7. The selected pipe segments for evaluating the cameras calibration method.

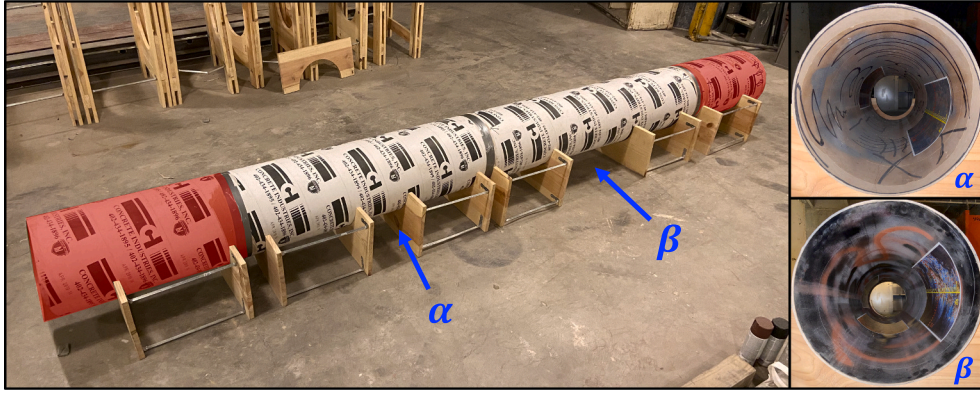


Fig. 8. The selected long pipeline in the experiment and the inner surface texture of each pipe segment.

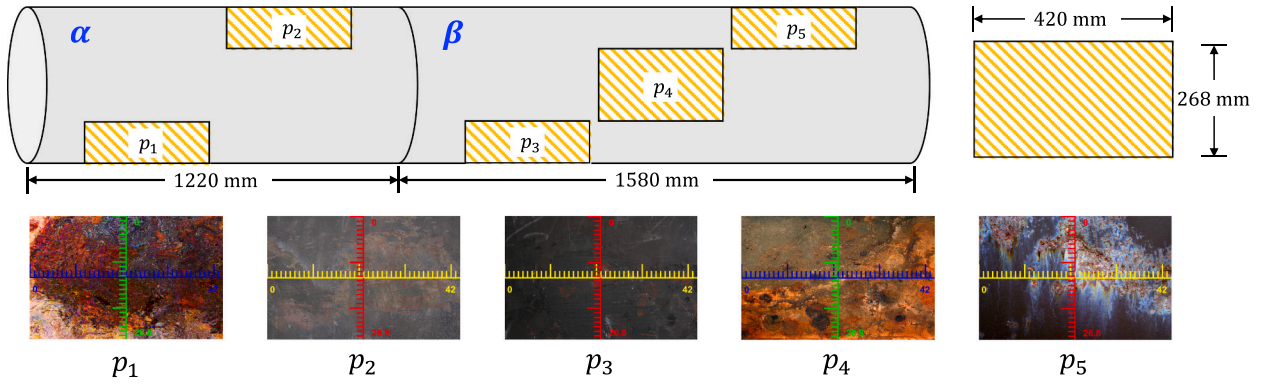


Fig. 9. The layout, the dimension, and the printed corrosion images of the designed in-pipe pictures.

Next, we perform the quantitative accuracy assessment of the reconstruction by comparing the measured dimensions of the reconstructed pipeline to the physical measurement (i.e., ground truth). We also measure the width w and height h of the recovered pictures in the unrolled model by connecting a straight line between the endpoints of both the horizontal (in blue/yellow) and the vertical (in green/red) axes. We average the results of the five measurements to minimize any human measurement errors.

6.3. Results

6.3.1. Short pipe segments

Fig. 10 demonstrates the qualitative comparison between the coarse and the fine registration results of the short pipe segments. Clearly, the fine registration shows better results when compared to the coarse registration for the three short pipes. The color-coded evaluation results (Fig. 10c vs. d) are clearly in favor of fine registration methods in all three cases. In Table 1, the quantitative evaluation of the inter-camera registration is presented. Instead of only comparing the coarse and the fine registration, we further split the fine registration results with and without the calibration generalization. In general, the results show that the fine registrations reach one millimeter-level geometric accuracy and 0.1 intensity error of the photometric accuracy that outperforms the coarse registration results. Among the fine registrations, the generalized results present an average 5% improvement at the geometric accuracy and 2.5% increase on the photometric accuracy if the geometry of the inspected pipe is identical. Without the generalization, the method fails to generate the valid result on the steel pipe (c) due to the smaller pipe diameter as opposed to the calibration pipe. It is observed that the non-generalized calibrations present the better photometric result on the calibration pipe, which may be caused by the fact that the

transformations are trained based on the marker map. To further assess the accuracy of the registration, we compare the radius of the unified pipe model (from the generalized fine registration) to the ground truth. As presented in Table 2, the error of the radius between the estimated and the ground truth reach at the one-millimeter level even for the steel pipes, which validates the accuracy of the presented method. In Fig. 11 (b), we also validate the fidelity of the projected central camera by comparing the rendered color images with the images taken by a handheld monocular camera. Even though the intrinsic and extrinsic parameters of the cameras are different, the results present a good agreement between the real and the virtual scenes. The results showed the potential of integrating the proposed method into the existing pipe NDE studies where most algorithms (10) are developed for in-pipe images taken from a front-view camera.

6.3.2. Long pipeline

Fig. 12 presents the general and the detailed views of 3D reconstruction of the long pipeline. High level matches are observed between the random patterns from the reconstructed dense point cloud model and the ground truths. Noted the color variations between the reconstruction and the ground truth in the detailed views are caused by different lighting conditions. To assess the reconstruction accuracy using the attached pictures, the 3D pipeline model is unrolled as shown in Fig. 13, where the A3 size prints attached at the internal pipe wall are flattened. Comparing to the ground truth (the five prints shown in Fig. 9), the textures in the pictures, representing the in-pipe corruptions, are recovered with high level-of-detail, validating the practicability of the proposed ILI system on pipe defects detection.

Table 3 demonstrates the accuracy of the reconstruction by comparing the point cloud measurement with the known length of the pipeline as well as the dimensions of the attached corrosion pictures. In

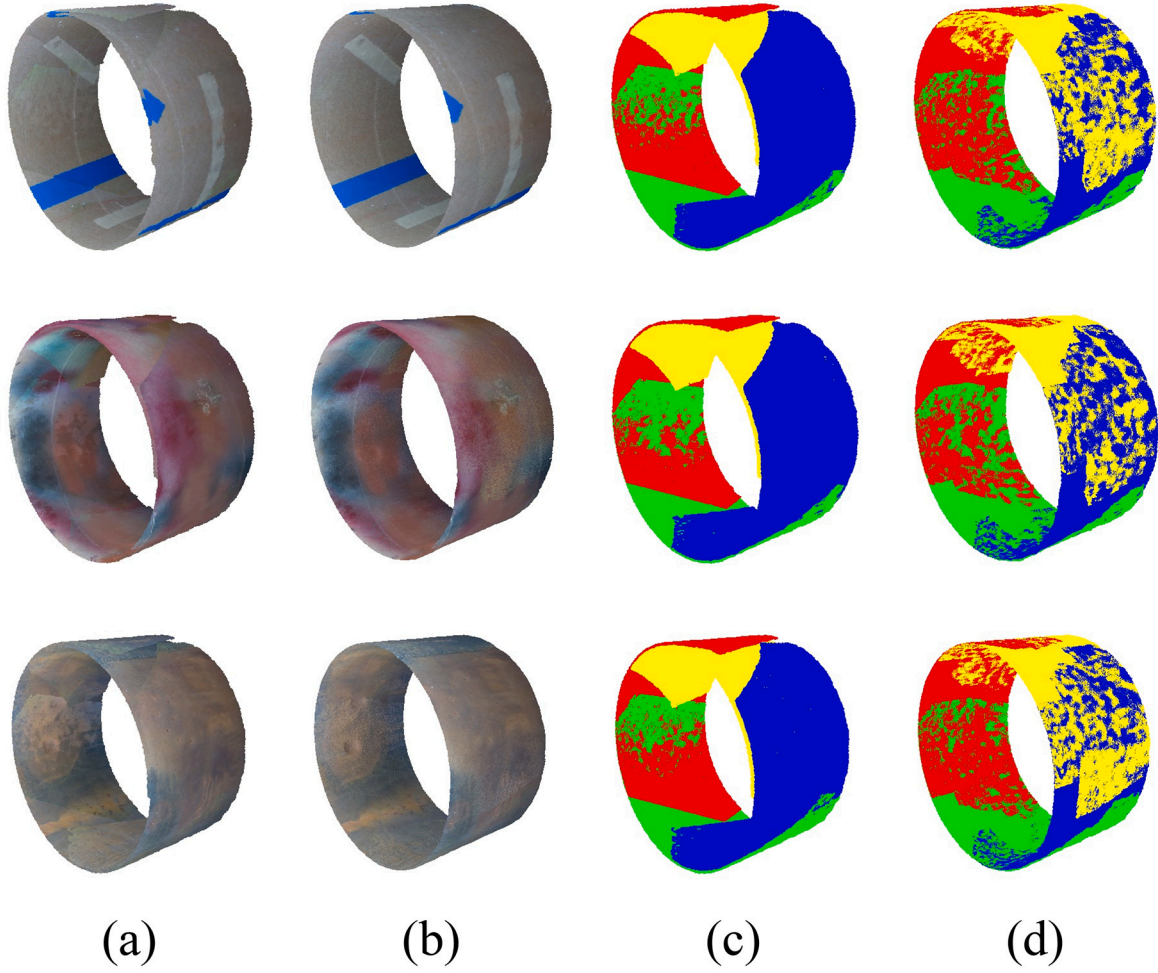


Fig. 10. Comparison of the unified pipe model of the three pipe segments using (a) the coarse registration and (b) the fine registration. (c) and (d) are color-coded registration result of (a) and (b) respectively. *Top*: pipe (a); *Middle*: pipe (b); *Bottom*: iron pipe (c).

Table 1

Evaluation of the registration errors at the selected pipe segments. Bold numbers in each column denote the smallest error values in that column.

Method	$Error_{geo}$ (mm)				$Error_{pho}$ (pixel)			
	Calibration pipe	Pipe (a)	Pipe (b)	Pipe (c)	Calibration pipe	Pipe (a)	Pipe (b)	Pipe (c)
Coarse registration	5.862	5.858	5.691	5.422	0.603	0.1590	0.1644	0.1893
Fine registration (without generalization)	1.245	1.149	1.164	×	0.306	0.1182	0.1563	×
Fine registration (with generalization)	1.192	1.091	1.101	1.112	0.357	0.1158	0.1557	0.1431

Table 2

Comparison of the registered pipe to the ground truth, unit in millimeter.

Pipe #	Ground truth radius	Measured radius (Median)	Measured error [Mean, Std]
Calibration PIPE	175.8	176.55	[0.753, 0.645]
Pipe (a)	175.8	176.61	[0.809, 0.782]
Pipe (b)	175.8	176.83	[0.987, 0.681]
Pipe (c)	169.3	168.02	[1.354, 1.108]

general, we can achieve an approximate one millimeter-level accuracy with the absolute distance error less than 0.5% on the selected pipeline. These results demonstrate comparable accuracy performance to the state-of-the-art laser profiling technique (27). However, the reconstructed models using our method are much denser, and the high-level details of the surface textures are preserved accurately. Among the

five corrosion pictures, we found that both p_2 and p_4 , where the average pixel intensities are larger (lighter-color images), contain relatively larger distance errors. We observe that these increased errors are primarily affected by the reflections from the onboard LEDs caused by the smooth surface of the printing papers. Because the robot motions are obtained purely from the RGBD images, the reconstruction accuracy is highly affected by the input images' quality. Improved performance can be expected from increased roughness to the internal wall surface of the pipelines and from an improved onboard illuminating device.

7. Conclusion and limitations

In this study, we develop an ILI visual inspection method using DCA for single-pass, full-coverage, dense 3D pipeline reconstruction. The developed camera calibration method allows reconstruction of the pipeline inline structure at millimeter or sub-millimeter geometric accuracy and at 0.1 pixel photometric accuracy. The achieved

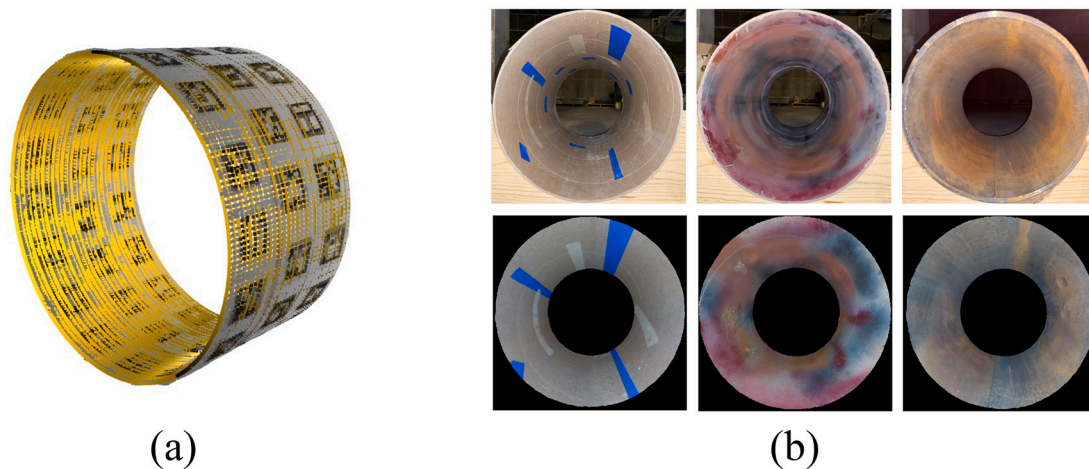


Fig. 11. (a). Circles fitting and radius estimation of the calibration pipe model, each ring is a fitted pipe segment; (b). Comparison between (Top) the ground truth image and (Bottom) the rendered images of the short pipe segments.

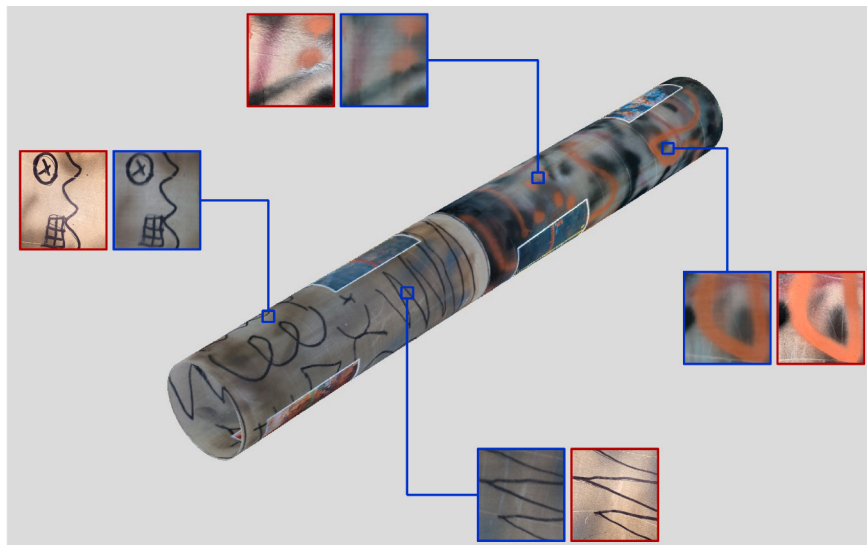


Fig. 12. 3D reconstruction of the pipeline using the proposed 3D reconstruction approach. The detailed views compare the ground truth scenes (in red: images taken with a handheld camera with flashlight on) with the reconstructed scenes (in blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

performance of 0.5% distance error in reconstruction outperform many inline NDE applications [62]. Even though the performance evaluation is reported only on the 14" diameter pipelines. Comparable results are found on 20" diameter pipeline after recalibration inside the 20" pipe. We can expect similar performance inside the pipelines of similar dimensions (e.g., 12", 16" etc) if the DCA is re-calibrated accordingly using similar illuminating condition, which is essential for any RGB camera to acquire good quality images. The performance of the DCA system inside the much-larger pipes (e.g. 30", 36, 48") is expected to be case-dependent because of many other internal/external factors (camera sensing range, resolution, illuminating intensity etc.) will come into play. Although the DCA's performance inside these much-larger pipes were not fully investigated, the proposed framework should still be applicable. However, due to the physical size limitation of the DCA and the minimal detection range of the RealSense depth cameras, it will be challenging if not impossible to use the developed DCA inside the much-smaller pipes such as 6" or 8" pipes.

The developed new visual inspection technology can potentially pave the way for low-cost and efficient ILI NDE for legacy pipelines as well as new pipelines, which is different from traditional ILI smart pigs

that relies on specially designed and expensive pig-retrieval facilities/equipment. The high-fidelity and high-density reconstructed 3D models can potentially enable simultaneous visual detections of many types of pipelines defects such serious corrosion spots, pitting, cracks, deformations, etc., through a single inspection passage using the traveling rover.

There are several areas that can be further improved in future studies. First, as stated in Section 6.3.2, the accuracy of the pipeline reconstruction is affected by the illumination conditions inside the pipelines. Further studies are needed to identify the optimal level of illumination for improved performance in different type of pipelines. Second, the performance of the proposed system is based on the test results from the straight pipe segments. Curved pipelines of various curvature values pose extra challenges to our method and need be addressed in future studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

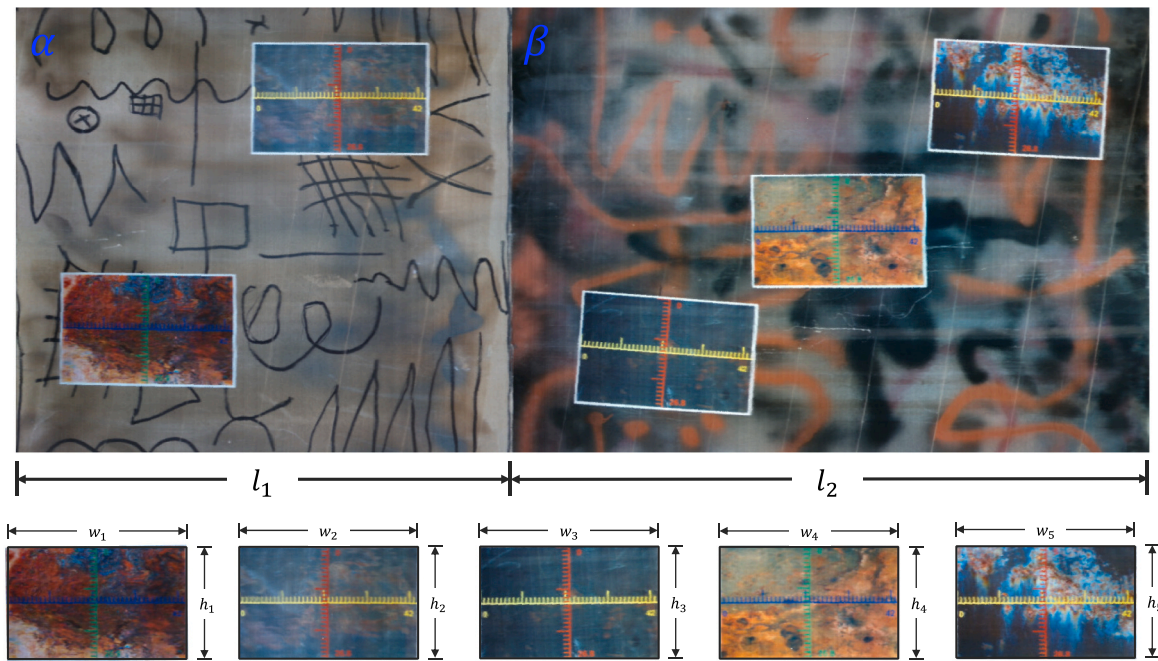


Fig. 13. Unrolled pipeline and the recovered pictures attached at the inner pipe wall.

Table 3

Evaluation of the in-pipe measurement as opposed to the ground truth.

Pipeline segments	Location	Ground truth (mm)	Measurement (mm)	Error (mm)	Abs error (%)
α	h_1	268	268.4	+0.4	0.15
	h_2		267.5	-0.5	0.19
	w_1	420	419.7	-0.3	0.07
	w_2		421.8	+1.8	0.43
	l_1	1220	1222.7	+2.7	0.22
β	h_3	268	267.3	-0.7	0.26
	h_4		269.2	+1.2	0.45
	h_5		268.8	+0.8	0.30
	w_3	420	420.5	+0.5	0.12
	w_4		418.9	-1.1	0.26
	w_5		419.6	-0.4	0.10
	l_2	1580	1578.1	-1.9	0.12

the work reported in this paper.

Acknowledgement

This work was supported by the Competitive Academic Agreement Program (CAAP) of Pipeline and Hazardous Materials Safety Administration (PHMSA) of U.S. Department of Transportation under contract 693JK31850013CAAP.

References

- [1] U.S. Energy Information Administration, U.S. Energy Facts Explained. <https://www.eia.gov/energyexplained/us-energy-facts/>, 2020 (Data accessed: 02-10-2022).
- [2] Pipeline & Hazardous Materials Safety Administration, Pipeline Incident 20 Years Trends. <https://www.phmsa.dot.gov/data-and-statistics/pipeline/pipeline-incident-20-year-trends>, 2021 (Data accessed: 02-10-2022).
- [3] J. Okamoto, J.C. Adamowski, M.S.G. Tsuzuki, F. Buiochi, C.S. Camerini, Autonomous system for oil pipelines inspection, *Mechatronics* 9 (7) (1999) 731–743, [https://doi.org/10.1016/S0957-4158\(99\)00031-8](https://doi.org/10.1016/S0957-4158(99)00031-8).
- [4] T.A. Bubenik, J.B. Nestroth, R.J. Eiber, B.F. Saffell, Magnetic flux leakage (MFL) technology for natural gas pipeline inspection, *NDT & E Int.* 30 (1) (1997) 36 (ISBN: 0963-8695).
- [5] B. Dai, H. Zhang, S. Sheng, J. Tang, Research on ultrasonic inspection of pipeline corrosion, in: 2007 IEEE International Conference on Control and Automation, ICCA, 2007, pp. 2572–2575, <https://doi.org/10.1109/ICCA.2007.4376826>.
- [6] D. Kim, L. Udupa, S. Udupa, Remote field eddy current testing for detection of stress corrosion cracks in gas transmission pipelines, *Mater. Lett.* 58 (15) (2004), <https://doi.org/10.1016/j.matlet.2004.01.006>.
- [7] Robert Bickerstaff, Mark Vaughn, Gerald Stoker, Michael Hassard, Mark Garrett, Review of Sensor Technologies for In-line Inspection of Natural Gas Pipelines, Albuquerque, NW, 2002 (Data accessed: 02-10-2022).
- [8] M.N. Mohammed, et al., Design and development of pipeline inspection robot for crack and corrosion detection, in: 2018 IEEE Conference on Systems, Process and Control (ICSPC), 2018, pp. 29–32, <https://doi.org/10.1109/SPC.2018.8704127>.
- [9] C. Piciarelli, D. Avola, D. Pannone, G.L. Foresti, A vision-based system for internal pipeline inspection, *IEEE Trans. Ind. Inform.* 15 (6) (2019) 3289–3299, <https://doi.org/10.1109/TII.2018.2873237>.
- [10] J.B. Haurum, T.B. Moeslund, A survey on image-based automation of CCTV and SSET sewer inspections, *Automat. Construct.* 111 (Mar. 2020), 103061, <https://doi.org/10.1016/j.autcon.2019.103061>.
- [11] S.K. Sinha, P.W. Fieguth, M.A. Polak, Computer vision techniques for automatic structural assessment of underground pipes, *Comput. Aid. Civil Infrastruct. Eng.* 18 (2) (2003) 95–112, <https://doi.org/10.1111/1467-8667.00302>.
- [12] T. Zhang, J. Liu, S. Liu, C. Tang, P. Jin, A 3D reconstruction method for pipeline inspection based on multi-vision, *Measurement* 98 (Feb. 2017) 35–48, <https://doi.org/10.1016/j.measurement.2016.11.004>.
- [13] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd edition, Cambridge University Press, 2004, <https://doi.org/10.1017/cbo9780511811685>.
- [14] Y. Zhang, R. Hartley, J. Mashford, L. Wang, S. Burn, Pipeline reconstruction from fisheye images, *J. WSCG* 19 (1–3) (2011) 49–58 (ISSN: 12136972).
- [15] J. Kannala, S.S. Brandt, J. Heikkilä, Measuring and modelling sewer pipes from video, *Mach. Vis. Appl.* 19 (2) (Mar. 2008) 73–83, <https://doi.org/10.1007/s00138-007-0083-1>.
- [16] A.V. Reyes-Acosta, I. Lopez-Juarez, R. Osorio-Companan, G. Lefranc, 3D pipe reconstruction employing video information from mobile robots, *Appl. Soft Comput.* J. 75 (Feb. 2019) 562–574, <https://doi.org/10.1016/j.asoc.2018.11.016>.
- [17] J.L. Schonberger, J.M. Frahm, Structure-from-motion revisited, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016, 2016 December, pp. 5157–5164, <https://doi.org/10.1109/CVPR.2016.445>.
- [18] P. Hansen, H. Alismail, P. Rander, B. Browning, Visual mapping for natural gas pipe inspection, *Int. J. Robot. Res.* 34 (4–5) (Apr. 2015) 532–558, <https://doi.org/10.1177/0278364914550133>.
- [19] S. Kagami, H. Taira, N. Miyashita, A. Torii, M. Okutomi, 3D pipe network reconstruction based on structure from motion with incremental conic shape detection and cylindrical constraint, in: IEEE International Symposium on Industrial Electronics, Jun. 2020, pp. 1345–1352, <https://doi.org/10.1109/ISIE45063.2020.9152377>, vol. 2020-June.
- [20] A. Oyama, et al., Three-dimensional mapping of pipeline from inside images using earthworm robot equipped with camera, *IFAC-PapersOnLine* 52 (22) (Oct. 2019) 87–90, <https://doi.org/10.1016/j.ifacol.2019.11.053>.

- [21] C.H. Bahnsen, A.S. Johansen, M.P. Philipsen, J.W. Henriksen, K. Nasrollahi, T. B. Moeslund, 3d sensors for sewer inspection: a quantitative review and analysis, *Sensors* 21 (7) (2021) 2553, <https://doi.org/10.3390/s21072553>.
- [22] P. Hansen, H. Alismail, B. Browning, P. Rander, Stereo visual odometry for pipe mapping, in: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep. 2011, pp. 4020–4025, <https://doi.org/10.1109/IROS.2011.6094911>.
- [23] M.E.M. Haertel, T.L.F. da Costa Pinto, A.A. Gonçalves Júnior, Trinocular stereo system with object space oriented correlation for inner pipe inspection, *Measurement* 73 (2015) 162–170, <https://doi.org/10.1016/j.measurement.2015.05.015>.
- [24] S. Beck, B. Froehlich, Volumetric calibration and registration of multiple RGBD-sensors into a joint coordinate system, in: *IEEE Symposium on 3D User Interfaces (3DUI)*, Jun. 2015, pp. 89–96, <https://doi.org/10.1109/3DUI.2015.7131731>.
- [25] C. Chen, et al., Calibrate multiple consumer RGB-D cameras for low-cost and efficient 3D indoor mapping, *Remote Sens.* 10 (2) (2018) 328, <https://doi.org/10.3390/rs10020328>.
- [26] T. Deng, et al., Registration of multiple RGBD cameras via local rigid transformations, in: *IEEE International Conference on Multimedia and Expo (ICME)*, Sep. 2014, pp. 1–6, <https://doi.org/10.1109/ICME.2014.6890122>, vol. 2014-Sept, no. Septmber.
- [27] A. Gunatilake, L. Piyathilaka, A. Tran, V.K. Viswanathan, K. Thiyagarajan, S. Kodagoda, Stereo vision combined with laser profiling for mapping of pipeline internal defects, *IEEE Sensors J.* 21 (10) (May 2020) 11926–11934.
- [28] S. el Kahi, D. Asmar, A. Fakih, J. Nieto, E. Nebot, A vision-based system for mapping the inside of a pipe, in: *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2011, pp. 2605–2611, <https://doi.org/10.1109/ROBIO.2011.6181697>.
- [29] Zhang, Zhao, Hu, Wang, Ai, Li, A 3D reconstruction pipeline of urban drainage pipes based on multiview image matching using low-cost panoramic video cameras, *Water* 11 (10) (Oct. 2019) 2101, <https://doi.org/10.3390/w11102101>.
- [30] R. Summan, et al., Image mosaicing for automated pipe scanning, in: *AIP Conference Proceedings* 1650(1), Apr. 2015, pp. 1334–1342, <https://doi.org/10.1063/1.4914747>.
- [31] S. Esquivel, R. Koch, H. Rehse, Reconstruction of sewer shaft profiles from fisheye-lens camera images, in: *Lecture Notes in Computer Science*, 2009, pp. 332–341, https://doi.org/10.1007/978-3-642-03798-6_34, vol. 5748 LNCS.
- [32] L. Keselman, et al., Intel RealSense stereoscopic depth cameras, in: *IEEE Conference on Computer Vision and Pattern Recognition Workshops* 68, 2017, pp. 1–10, no. January.
- [33] A. Gunatilake, L. Piyathilaka, S. Kodagoda, S. Barclay, D. Vitanage, Real-time 3D profiling with RGB-D mapping in pipelines using stereo camera vision and structured IR laser ring, in: *The 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2019, pp. 916–921, <https://doi.org/10.1109/ICIEA.2019.8834089>.
- [34] Z. Zhang, A flexible new technique for camera calibration, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (11) (Nov. 2000) 1330–1334, <https://doi.org/10.1109/34.888718>.
- [35] E. Fernández-Moral, J. González-Jiménez, P. Rives, V. Arévalo, Extrinsic calibration of a set of range cameras in 5 seconds without pattern, in: *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2014, pp. 429–435, <https://doi.org/10.1109/IROS.2014.6942595>.
- [36] H. Xu, J. Hou, L. Yu, S. Fei, 3D Reconstruction system for collaborative scanning based on multiple RGB-D cameras, *Pattern Recogn. Lett.* 128 (2019) 505–512, <https://doi.org/10.1016/j.patrec.2019.10.020>.
- [37] Z. Liu, Z. Meng, N. Gao, Z. Zhang, Calibration of the relative orientation between multiple depth cameras based on a three-dimensional target, *Sensors* 19 (13) (Jul. 2019) 3008, <https://doi.org/10.3390/s19133008>.
- [38] H. Liu, H. Li, X. Liu, J. Luo, S. Xie, Y. Sun, A novel method for extrinsic calibration of multiple RGB-D cameras using descriptor-based patterns, *Sensors (Switzerland)* 19 (2) (2019) 349, <https://doi.org/10.3390/s19020349>.
- [39] C. Daniel Herrera, J. Kannala, J. Heikkilä, Joint depth and color camera calibration with distortion correction, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (10) (2012) 2058–2064, <https://doi.org/10.1109/TPAMI.2012.125>.
- [40] Q.Y. Zhou, S. Miller, V. Koltun, Elastic fragments for dense scene reconstruction, in: *IEEE International Conference on Computer Vision (ICCV)*, 2013, pp. 473–480, <https://doi.org/10.1109/ICCV.2013.65>.
- [41] S. Beck, B. Froehlich, Sweeping-based volumetric calibration and registration of multiple RGBD-sensors for 3D capturing systems, in: *Virtual Reality (VR)*, Apr. 2017, pp. 167–176, <https://doi.org/10.1109/VR.2017.7892244>.
- [42] X.S. Gao, X.R. Hou, J. Tang, H.F. Cheng, Complete solution classification for the perspective-three-point problem, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (8) (Aug. 2003) 930–943, <https://doi.org/10.1109/TPAMI.2003.1217599>.
- [43] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, M.J. Marín-Jiménez, Automatic generation and detection of highly reliable fiducial markers under occlusion, *Pattern Recogn.* 47 (6) (Jun. 2014) 2280–2292, <https://doi.org/10.1016/j.patcog.2014.01.005>.
- [44] R. Muñoz-Salinas, M.J. Marín-Jiménez, E. Yeguas-Bolivar, R. Medina-Carnicer, Mapping and localization from planar markers, *Pattern Recogn.* 73 (Jan. 2018) 158–171, <https://doi.org/10.1016/j.patcog.2017.08.010>.
- [45] H. Abdi, L.J. Williams, Principal component analysis, in: *Wiley Interdisciplinary Reviews: Computational Statistics* 2(4), John Wiley & Sons, Ltd, Jul. 01, 2010, pp. 433–459, <https://doi.org/10.1002/wics.101>.
- [46] C.L. Lawson, R.J. Hanson, Solving Least Squares Problems, *Society for Industrial and Applied Mathematics (SIAM)*, 1995, <https://doi.org/10.1137/1.9781611971217.bm>.
- [47] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* 24 (6) (Jun. 1981) 381–395, <https://doi.org/10.1145/358669.358692>.
- [48] M.S. Ahn, H. Chae, D. Noh, H. Nam, D. Hong, Analysis and noise modeling of the intel RealSense D435 for mobile robots, in: *16th International Conference on Ubiquitous Robots (UR)*, Jun. 2019, pp. 707–711, <https://doi.org/10.1109/URAL.2019.8768489>.
- [49] P.J. Besl, N.D. McKay, A method for registration of 3-D shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (2) (1992) 239–256, <https://doi.org/10.1109/34.121791>.
- [50] A. Myronenko, X. Song, Point set registration: coherent point drifts, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (12) (2010) 2262–2275, <https://doi.org/10.1109/TPAMI.2010.46>.
- [51] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. Ser. B Methodol.* 39 (1) (1977) 1–22, <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.
- [52] W. Gao, R. Tedrake, FilterReg: robust and efficient probabilistic point-set registration using gaussian filter and twist parameterization, in: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 11095–11104.
- [53] D. Meagher, Geometric modeling using octree encoding, *Comp. Graph. Image Proc.* 19 (2) (1982) 129–147, [https://doi.org/10.1016/0146-664X\(82\)90104-6](https://doi.org/10.1016/0146-664X(82)90104-6).
- [54] D. Shepard, A two-dimensional interpolation function for irregularly-spaced data, in: *The 1968 23rd ACM National Conference*, 1968, pp. 517–524 (SGR: 0014432211).
- [55] S. Choi, Q.Y. Zhou, V. Koltun, Robust reconstruction of indoor scenes, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Oct. 2015, pp. 5556–5565, <https://doi.org/10.1109/CVPR.2015.7299195>, vol. 07-12-June.
- [56] P. Lancaster, K. Salkauskas, Surfaces generated by moving least squares methods, *Math. Comput.* 37 (155) (1981) 141–158, <https://doi.org/10.1090/s0025-5718-1981-0616367-1>.
- [57] J. Park, Q.Y. Zhou, V. Koltun, Colored point cloud registration revisited, in: *Proceedings of the IEEE International Conference on Computer Vision*, Dec. 2017, pp. 143–152, <https://doi.org/10.1109/ICCV.2017.25>, vol. 2017-October.
- [58] M. Labbe, F. Michaud, Appearance-based loop closure detection for online large-scale and long-term operation, *IEEE Trans. Robot.* 29 (3) (2013) 734–745, <https://doi.org/10.1109/TRO.2013.2242375>.
- [59] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, W. Burgard, G2o: a general framework for graph optimization, in: *IEEE International Conference on Robotics and Automation*, 2011, pp. 3607–3613, <https://doi.org/10.1109/ICRA.2011.5979949>.
- [60] B. Curless, M. Levoy, A volumetric method for building complex models from range images, in: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH 1996, 1996, pp. 303–312, <https://doi.org/10.1145/237170.237269>.
- [61] M. Kazhdan, M. Bolitho, H. Hoppe, Poisson surface reconstruction, in: *The fourth Eurographics Symposium on Geometry Processing 7*, Eurographics Association, 2006 (ISBN: 3-905673-36-3).
- [62] N.M. Yatim, R.L.A. Shauri, N. Buniyamin, Automated mapping for underground pipelines: an overview, in: *2014 2nd International Conference on Electrical, Electronics and System Engineering (ICEESE)*, Dec. 2014, pp. 77–82, <https://doi.org/10.1109/ICEESE.2014.7154599>.